SIMULATORE DI RISCHIO

Manuale dell’Utente

Johnathan Mun, Ph.D., MBA, MS, BS, CRM, FRM, CFC, MIFC

Real Options Valuation, Inc.

REAL OPTIONS VALUATION, INC.

Real Options Valuation, Inc.

4101F Dublin Blvd., Ste. 425

Dublin, California 94568 U.S.A.

Phone 925.271.4438 • Fax 925.369.0450

admin@realoptionsvaluation.com

www.risksimulator.com

www.realoptionsvaluation.com

INDICE

1

1. INTRODUZIONE

1.1 Benvenuti al Software SIMULATOR DI RISCHIO

Il **Simulatore di Rischio** è un software di Simulazione Monte Carlo , di Previsione e di Ottimizzazione. Il software è scritto in Microsoft .NET C# e funziona con Excel come componente aggiuntivo (add-in). Il software è anche compatibile e spesso usato con il software Real Options Risolutore di Super Reticoli (SLS) e con il software Toolkit per la Valutazione di Azioni per Dipendenti (ESOV), anche questo sviluppato da Real Options Valuation, Inc. Vi preghiamo di notare che nonostante il nostro sforzo di rendere esauriente questo manuale dell'utente, il manuale non è assolutamente un sostituto per il DVD didattico, per i corsi di formazione dal vivo e per i libri scritti dal creatore del software (per esempio, *Real Options Analysis,* 2nd Edition, Wiley Finance 2005; *Modeling Risk: Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization*, 2nd Edition, Wiley 2010; e *Valuing Employee Stock Options (2004 FAS 123R),* Wiley Finance, 2004), il Dr. Johnathan Mun. Vi preghiamo di visitare il nostro sito web www.realoptionsvaluation.com per maggiori informazioni su questi temi.

Il software ***Simulatore di Rischio*** contiene i seguenti moduli:

- Simulazione Monte Carlo (esegue simulazioni parametriche e non parametriche di 42 distribuzioni di probabilità con differenti profili di simulazione, simulazioni troncate e correlate, distribuzioni personalizzabili, simulazioni con controllo precisione ed errore e molti altri algoritmi)
- Previsione (esegue Box-Jenkins ARIMA, regressioni multiple, estrapolazione non lineari, processi stocastici ed analisi di serie temporali)
- Ottimizzazione sotto Incertezza (esegue ottimizzazioni usando variabili discrete a numero intero e continue per l'ottimizzazione di portafogli e progetti con o senza simulazioni)

- Strumenti Analitici e di Modellazione (esegue analisi Tornado, Ragno e di Sensibilità, così come simulazioni bootstrap, tests di verifica d'ipotesi , adattamenti distribuzionali, ecc.)

Il software ***Real Options SLS*** è usato per calcolare opzioni semplici e complesse e include la capacità di creare modelli di opzioni personalizzabili. Il software ***Real Options SLS*** contiene i seguenti moduli:

- SLS Asset Singolo (per risolvere opzioni di abbandono, chooser, di contrazione, di differimento e di espansione, così come opzioni personalizzate)
- SLS Assets Multipli e Fasi Multiple (per risolvere opzioni sequenziali a fasi multiple, opzioni con fasi e assets sottostanti multipli, combinazioni di opzioni sequenziali di abbandono con fasi multiple, chooser, di contrazione, di differimento e di espansione, e di switching (commutazione); può anche usato per risolvere opzioni personalizzate)
- SLS Multinomiale (per risolvere opzioni trinomiali con ritorno alla media, opzioni quadranomiali con diffusione a salti e opzioni pentanomiali arcobaleno)
- Funzioni dei componenti aggiuntivi (Add-In) di Excel (per risolvere tutte le opzioni di cui sopra, più i modelli a forma chiusa e opzioni personalizzate in un ambiente a base Excel)

1.2 Requisiti di Installazione e Procedure

Per installare il software, seguite le istruzioni visualizzate. I requisiti minimi per questo software sono:

- Processore Pentium IV o successivo (dual core consigliato).
- Windows XP, Vista o Windows 7.
- Microsoft Excel XP, 2003, 2007, 2010, o versioni successive.
- Microsoft .NET Framework 2.0 o successivo (versioni 3.0, 3.5 e così via).
- 500 MB di spazio libero.
- 2GB RAM minimo (2-4GB consigliato).
- Diritti amministrativi per installare il software.

Microsoft .NET Framework 2.0/3.0 è già preinstallato nella maggioranza dei nuovi computers. Tuttavia, se durante l'installazione di Simulatore di Rischio si presenta un messaggio d'errore riguardante la disponibilità di .NET Framework, uscite dall'installazione. Dopo, installate il software pertinente .NET Framework incluso nel CD (scegliete la vostra lingua). Completate l'installazione di .NET, riavviate il computer e poi reinstallate il software Simulatore di Rischio. Il software include di default un file di licenza di prova per 10 giorni. Per ottenere una completa licenza aziendale, siete pregati di contattare Real Options Valuation, Inc. a admin@realoptionsvaluation.com o di chiamare (925) 271-4438 o di visitare il nostro sito web a www.realoptionsvaluation.com. Prego visitare questo sito web e cliccare su DOWNLOAD per ottenere l'ultima versione del software o cliccare sul link FAQ per ottenere eventuali informazioni aggiornate sulla concessione della licenza o sui temi riguardanti l'installazione e le correzioni.

1.3 Concessione della licenza

Se avete installato il software e acquistato una licenza completa per utilizzare il software, dovrete inviarci per e-mail la vostra ID Hardware, per permetterci di generare per voi un file di licenza. Seguite le istruzioni sottostanti:

- Avviate Excel XP/2003/2007/2010, cliccate sull'icona ***Licenza*** o cliccate su ***Simulatore di Rischio*** | ***Licenza***, copiatela ed inviateci per e-mail la vostra ID HARDWARE di 11 a 20 caratteri alfanumerici che inizia con il prefisso "*RS*" (potete anche selezionare la ID Hardware e cliccare con tasto destro del mouse oppure cliccare sul link ID Hardware per e-mail) al seguente indirizzo: admin@realoptionsvaluation.com. Quando riceviamo questa ID, sarà generata una nuova licenza permanente che vi sarà inviata per e-mail. Quando ricevete questo file di licenza, salvatelo semplicemente sul vostro disco fisso (se è un file zippato, decomprimete prima il suo contenuto e salvatelo poi sul vostro disco fisso). Avviate Excel, cliccate su ***Simulatore di Rischio*** | ***Licenza*** oppure cliccate sull'icona ***Licenza*** e poi su ***Installare Licenza***, e puntate su questo nuovo file di licenza. Riavviate Excel e avete terminato. L'intero processo richiede meno di un minuto e sarete in possesso di una licenza completa.

- Una volta terminata l'installazione, avviate Microsoft Excel e, se l'installazione è riuscita, dovreste vedere l'elemento aggiuntivo "*Simulatore di Rischio*" nella barra del menu di Excel XP/2003 o sotto il nuovo gruppo icona in Excel 2007/2010, e una nuova barra delle icone in Excel, come visto nella Figura 1.1. Appare inoltre uno schermo di benvenuto, come visto nella Figura 1.2, indicando che il software è funzionante e caricato in Excel. La Figura 1.3 mostra anche la barra degli strumenti di Simulatore di Rischio. Se questi elementi sono presenti in

Excel, siete pronti per usare il software. Le sezioni seguenti forniscono istruzioni passo passo sull'uso del software.

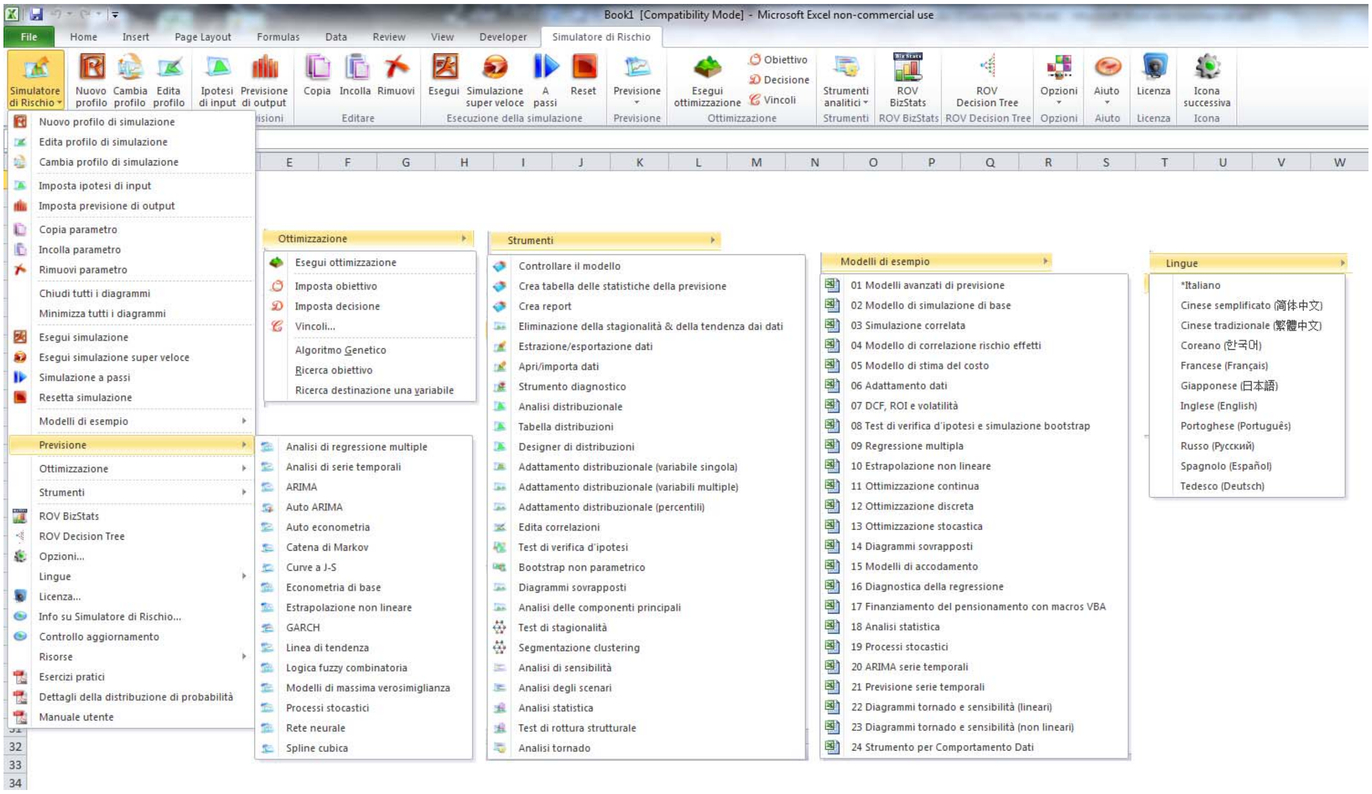


Figura 1.1 – Simulatore di Rischio - Menu e Barra Icone in Excel 2007/2010

Figura 1.2 – Simulatore di Rischio - Schermo di Benvenuto

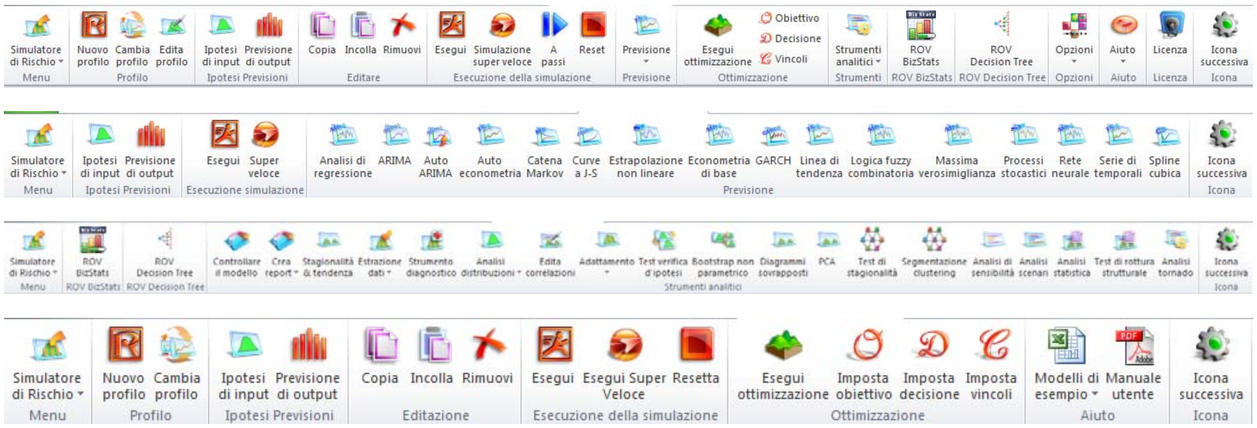


Figura 1.3 – Simulatore di Rischio - Barre degli strumenti delle Icone in 2007/2010

1.4 LE NOVITÀ DELLA VERISONE 2011/2012

Una lista completa delle funzionalità di Simulatore di Rischio

Di seguito sono elencate le principali funzionalità di Simulatore di Rischio. Gli elementi evidenziati indicano le ultime aggiunte alla Versione 2011/2012.

1.4.1 Funzionalità generali

1. Disponibile in 11 lingue — inglese, francese, tedesco, italiano, giapponese, coreano, portoghese, spagnolo, russo, cinese semplificato e cinese tradizionale.
2. ROV Albero decisionale è usato per creare e valutare modelli di alberi decisionali. Sono anche incluse addizionali metodologie e analitiche avanzate.
    - Modelli di alberi decisionali
    - Simulazione di rischio Monte Carlo
    - Analisi di sensibilità
    - Analisi di scenario
    - Analisi Bayesiana (Aggiornamento probabilità congiunte e posteriori)
    - Valore atteso dell'informazione
    - MINIMAXS
    - MAXIMIN
    - Prifili di rischio
3. Libri - la teoria analitica, l'applicazione e i casi di studio sono coadiuvati da 10 libri.
4. Celle commentate — è possibile attivare e disattivare i commenti di cella e decidere se volete mostrare i commenti di cella per tutte le ipotesi di input, le previsioni di output e le variabili decisionali.
5. Modelli dettagliati d'esempio — 24 modelli d'esempio in Simulatore di Rischio e più di 300 modelli nel Modeling Toolkit.
6. Report dettagliati — tutte le analisi sono corredate da report dettagliati.
7. Manuale dell'utente dettagliato

8. Flessibile concessione della licenza — è possibile attivare o disattivare certe funzionalità per permettervi di personalizzare la vostra esperienza di analisi del rischio. Per esempio, se vi interessano solo gli strumenti di previsione di Simulatore di Rischio, potete ottenere una licenza speciale per attivare solo gli strumenti di previsione e mantenere disattivati i restanti moduli. Questo vi permette di risparmiare sul costo del software.

9. Requisti flessibili — funziona con Windows 7, Vista e XP; si integra con Excel 2010, 2007, 2003; funziona con sistemi operativi MAC che eseguono "virtual machines".

10. Colori e diagrammi — inclinazione, 3D, colore, tipo di diagramma e molto altro - completamente personalizzabili.

11. Esercizi pratici — un dettagliato manuale passo-passo su come eseguire Simulatore di Rischio; sono inclusi manuali su come interpretare i risultati.

12. Copia e incolla per celle multiple — permette di copiare ed incollare ipotesi, variabili decisionali e previsioni.

13. Definizione profili — permette la creazione di multipli profili in un singolo modello (è possibile creare, duplicare, modificare ed eseguire differenti scenari di modelli di simulazione in un singolo modello.

14. Icone riviste in Excel 2007/2010 — la barra degli strumenti delle icone è stata completamente rivista, rendendola più intuitiva e facile da usare. Ci saranno quattro gruppi di icone che si adattano alla maggior parte delle risoluzioni dello schermo (1280 x 760 e maggiore).

15. Scelta rapida col tasto destro del mouse — accesso a tutti gli strumenti e i menu di Simulatore di Rischio cliccando sul tasto destro del mouse.

16. Integrazione con i software ROV — funziona bene con altri software ROV incluso Real Options SLS, Modeling Toolkit, Basel Toolkit, ROV Compiler, ROV Extractor e Evaluator, ROV Modeler, ROV Valuator, ROV Optimizer, ROV Dashboard, ESO Valuation Toolkit e altri!

17. Funzioni RS in Excel — è possibile inserire funzioni RS per impostare le ipotesi e le previsioni; supporto mediante l'uso del tasto destro del mouse in Excel.

18. Strumento di risoluzione dei problemi: questo strumento vi permette di riattivare il software, controllare i vostri requisiti di sistema, ottenere la ID Hardware e altro.

19. Analisi a velocità turbo: questa nuova funzionalità esegue previsioni e altri strumenti di analisi ad altissima velocità (potenziato nella versione 5.2). Le analisi e i risultati rimangono uguali, ma sia il calcolo che la generazione dei report avviene molto velocemente.

20. Risorse Web, casi di studio e video — è possibile scaricare modelli gratuiti, video con guide introduttive, casi di studio, white paper e altri materiali dal nostro sito web.

1.4.2 Modulo di Simulazione

21. 6 Generatori di numeri aleatori — ROV Advanced Subtractive Generator, Subtractive Random Shuffle Generator, Long Period Shuffle Generator, Portable Random Shuffle Generator, Quick IEEE Hex Generator, Basic Minimal Portable Generator.

22. 2 Metodi di campionamento — Monte Carlo e Latin Hypercube (Ipercubo latino).

23. 3 Copule di correlazione — applicazione di Copula normale, Copula T e Copula quasi normale per simulazioni correlate.

24. 42 distribuzioni di probabilità — Arcoseno, Bernoulli, Beta, Beta 3, Beta 4, Binomiale, Cauchy, Chi-quadrato, Coseno, Personalizzata, Discreta uniforme, Doppia Log, Erlang, Esponenziale, Esponenziale 2, Distribuzione F, Gamma, Geometrica, Gumbel Max, Gumbel Min, Ipergeometrica, Laplace, Logistica, Lognormale (aritmetica) e Lognormale (Log), Lognormale 3 (aritmetica) e Lognormale 3 (Log), Negativa Binomiale, Normale, Parabolica, Pareto, Pascal, Pearson V, Pearson VI, PERT, Poisson, Potenza, Potenza 3, Rayleigh, T e T2, Triangolare, Uniforme, Weibull, Weibull 3.

25. Parametri alternativi — uso dei percentili come modo alternativo per inserire i parametri.

26. Distribuzione non parametrica personalizzata — è possibile creare le vostre distribuzioni, eseguendo simulazioni storiche e applicando il metodo Delphi.

27. Troncamento delle distribuzioni — attivazione di limiti dei dati.

28. Funzioni Excel — è possibile impostare ipotesi e previsioni usando le funzioni all'interno di Excel.

29. Simulazione multidimensionale — simulazione di parametri di input incerti.

30. Controllo della precisione — determina se il numero di prove di simulazione eseguito è sufficiente.

31. Simulazione Super Veloce — esegue 100000 prove in pochi secondi.

1.4.3 Modulo di Previsione

32. ARIMA — modelli di media mobile integrata autoregressiva ARIMA (P,D,Q).

33. Auto ARIMA — esegue le più comuni combinazioni di ARIMA per trovare il modello col miglior adattamento.

34. Auto-Econometria — esegue migliaia di combinazioni e permutazioni di modelli per ottenere il modello col miglior adattamento per i dati attuali (lineare, non lineare, interagente, sfasamento, lead, tasso, differenza).

35. Econometrica di base — modelli di regressione econometrici e lineari/non lineari e interagenti.

36. Spline cubico — interpolazione ed estrapolazione non lineare.

37. GARCH — proiezioni di volatilità usando modelli di eteroschedasticità condizionata auto regressiva generalizzata: GARCH, GARCH-M, TGARCH, TGARCH-M, EGARCH, EGARCH-T, GJR-GARCH, and GJR-TGARCH.

38. Curva a J — curve a J esponenziali.

39. Variabili dipendenti limitate — Logit, Probit e Tobit.

40. Catene di Markov — due elementi in competizione nel tempo e previsioni di quote di mercato.

41. Regressione multipla — regressione normale lineare e non lineare con metodologie a passi successivi (in avanti, indietro, correlazione, in avanti-indietro).

42. Estrapolazione non lineare — previsione di serie temporali non lineare.

43. Curva a S — curve a S logistiche.

44. Analisi di serie temporali —8 Modelli di scomposizione di serie temporali per predire livelli, tendenza e stagionalità.

45. Linee di tendenza — previsione e adattamento usando lineare, polinomiale non lineare, potenza, logaritmica, esponenziale e media mobile con bontà di adattamento.

46. Previsione a Rete neurale — lineare, logistica, tangente iperbolica, coseno con tangente iperbolica
47. Previsione a Logica fuzzy combinatoria

1.4.4 Modulo di Ottimizzazione

48. Ottimizzazione lineare — ottimizzazione a fasi multiple e ottimizzazione lineare generale.
49. Ottimizzazione non lineare — i risultati dettagliati includono matrici hessiane, funzioni LaGrange e altro.
50. Ottimizzazione statica — esecuzioni veloci di ottimizzazioni continue, a numeri interi e binarie.
51. Ottimizzazione dinamica — simulazione con ottimizzazione.
52. Ottimizzazione stocastica — quadratica, tangenziale, centrale, in avanti, criteri di convergenza.
53. Frontiera efficiente — combinazioni di ottimizzazioni stocastiche e dinamiche su frontiere efficienti multivariate.
54. Algoritmi genetici — usati per una varietà di problemi di ottimizzazione.
55. Ottimizzazione a fasi multiple — esegue test per l'Optimum locale rispetto a quello globale, permettendo un controllo migliore su come è eseguita l'ottimizzazione, e aumenta la precisione e l'affidabilità dei risultati.
56. Percentili e Medie condizionate —statistiche aggiuntive per ottimizzazioni stocastiche, incluso i percentili e le Medie condizionate, che sono cruciali nel calcolare misure del valore a rischio condizionato.
57. Algoritmi di ricerca — algoritmi di ricerca facili, veloci ed efficaci per variabili decisionali singole di base e applicazioni per la ricerca dell'obiettivo.
58. Simulazione Super Veloce nelle Ottimizzazioni dinamiche e stocastiche — esegue una simulazione a super velocità integrata al contempo con una ottimizzazione.

1.4.5 Modulo degli Strumenti analitici

59. Modello di controllo — esamina i vostri modelli per individuare gli errori più comuni.

60. Editore di correlazioni — permette l'inserimento diretto e la modifica di grandi matrici di correlazione.

61. Creazione di report — è possibile automatizzare la generazione di report delle ipotesi e delle previsioni in un modello.

62. Creazione di report statistici — è possibile generare report comparativi di tutte le statistiche della previsione.

63. Diagnostica dei dati —esegue test per l'eteroschedasticità, la micronumerosità, i valori anomali, la non linearità, l'autocorrelazione, la normalità, la sfericità, la non stazionarità, la multicollinearità e le correlazioni.

64. Estrazione ed esportazione dei dati — è possibile estrarre i dati verso Excel o verso file flat di testo e file di tipo Risk Sim, esegue report statistici e report sui risultati della previsione.

65. Apertura ed importazione dei dati — è possibile recuperare i risultati di precedenti esecuzioni di simulazioni.

66. Destagionalizzazione e Rimozione della tendenza — destagionalizza e rimuove la tendenza dai vostri dati.

67. Analisi distribuzionale — calcola esattamente la Funzione di distribuzione della probabilità (PDF), la Funzione di distribuzione cumulativa (CDF) e la Funzione di distribuzione cumulativa inversa (ICDF) di tutte le 42 distribuzioni e generata le tabelle di probabilità.

68. Designer di distribuzioni — è possibile creare le vostre distribuzioni personalizzate.

69. Adattamento distribuzionale (multiplo) — esegue multiple variabili simultaneamente e dà le correlazioni ed il significato delle correlazioni.

70. Adattamento distribuzionale (singolo) — test di Kolmogorov-Smirnov e del Chi-quadrato su distribuzioni continue, completi di report ed ipotesi distribuzionali.

71. Test di verifica d'ipotesi — verifica se due previsioni sono statisticamente simili o differenti.

72. Bootstrap non parametrico — simulazione delle statistiche per ottenere la precisione e l’accuratezza dei risultati.

73. Diagrammi di sovrapposizione — diagrammi di sovrapposizione contestuale delle ipotesi e delle previsioni completamente personalizzabili (diagrammi di tipo CDF, PDF, 2D/3D).

74. Analisi del componente principale — esegue un test per individuare le migliori variabili esplicative e modi di ridurre la serie ordinata dei dati.

75. Analisi dello scenario — centinaia e migliaia di scenari statici bidimensionali.

76. Test di stagionalità — esegue un test per individuare i vari sfasamenti di stagionalità.

77. Clustering di segmentazione — raggruppa i dati in cluster statistici per la segmentazione dei vostri dati.

78. Analisi di sensibilità — sensibilità dinamica (analisi simultanea).

79. Test di rottura strutturale — esegue un test per individuare se i vostri dati di serie temporali presentano rotture strutturali statistiche.

80. Analisi Tornado — perturbazione statica di sensibilità, analisi Spider e Tornado e tabelle di scenari.

1.4.6 Modulo di Statistiche e BizStats

81. Adattamento distribuzionale percentile — uso dei percentili e della ottimizzazione per individuare la distribuzione col miglior adattamento.

82. Diagrammi e tabelle di distribuzioni di probabilità — è possibile eseguire 45 distribuzioni di probabilità, i loro quattro momenti, CDF, ICDF, PDF, diagrammi, sovrapporre multipli diagrammi distribuzionali e generare tabelle di distribuzioni di probabilità.

83. Analisi delle statistiche — statistiche descrittive, adattamento distribuzionale, istogrammi, diagrammi, estrapolazione non lineare, test di normalità, stima dei parametri stocastici, previsioni di serie temporali, trend line proiezioni della linea di tendenza, ecc.

84. ROV BIZSTATS — più di 130 modelli di statistiche aziendali e di analisi:

    Absolute Values, ANOVA: Randomized Blocks Multiple Treatments, ANOVA: Single Factor Multiple Treatments, ANOVA: Two Way Analysis, ARIMA, Auto ARIMA, Autocorrelation and Partial Autocorrelation, Autoeconometrics (Detailed), Autoeconometrics (Quick), Average, Combinatorial Fuzzy Logic Forecasting, Control Chart: C, Control Chart: NP, Control Chart: P, Control Chart: R, Control Chart: U, Control Chart: X, Control Chart: XMR, Correlation, Correlation (Linear, Nonlinear), Count, Covariance, Cubic Spline, Custom Econometric Model, Data Descriptive Statistics, Deseasonalize, Difference, Distributional Fitting, Exponential J Curve, GARCH, Heteroskedasticity, Lag, Lead, Limited Dependent Variables (Logit), Limited Dependent Variables (Probit), Limited Dependent Variables (Tobit), Linear Interpolation, Linear Regression, LN, Log, Logistic S Curve, Markov Chain, Max, Median, Min, Mode, Neural Network, Nonlinear Regression, Nonparametric: Chi-Square Goodness of Fit, Nonparametric: Chi-Square Independence, Nonparametric: Chi-Square Population Variance, Nonparametric: Friedman's Test, Nonparametric: Kruskal-Wallis Test, Nonparametric: Lilliefors Test, Nonparametric: Runs Test, Nonparametric: Wilcoxon Signed-Rank (One Var), Nonparametric: Wilcoxon Signed-Rank (Two Var), Parametric: One Variable (T) Mean, Parametric: One Variable (Z) Mean, Parametric: One Variable (Z) Proportion, Parametric: Two Variable (F) Variances, Parametric: Two Variable (T) Dependent Means, Parametric: Two Variable (T) Independent Equal Variance, Parametric: Two Variable (T) Independent Unequal Variance, Parametric: Two Variable (Z) Independent Means, Parametric: Two Variable (Z) Independent Proportions, Power, Principal Component Analysis, Rank Ascending, Rank Descending, Relative LN Returns, Relative Returns, Seasonality, Segmentation Clustering, Semi-Standard Deviation (Lower), Semi-Standard Deviation (Upper), Standard 2D Area, Standard 2D Bar, Standard 2D Line, Standard 2D Point, Standard 2D Scatter, Standard 3D Area, Standard 3D Bar, Standard 3D Line, Standard 3D Point, Standard 3D Scatter, Standard Deviation (Population), Standard Deviation (Sample), Stepwise Regression (Backward), Stepwise Regression (Correlation), Stepwise Regression (Forward), Stepwise Regression (Forward-Backward), Stochastic Processes (Exponential Brownian Motion), Stochastic Processes (Geometric Brownian Motion), Stochastic Processes (Jump Diffusion), Stochastic Processes (Mean Reversion with Jump Diffusion), Stochastic Processes (Mean Reversion), Structural Break, Sum, Time-Series Analysis (Auto), Time-Series Analysis (Double Exponential Smoothing), Time-Series Analysis (Double Moving Average), Time-Series Analysis (Holt-Winter's Additive), Time-Series Analysis (Holt-Winter's Multiplicative), Time-Series Analysis (Seasonal Additive), Time-Series Analysis (Seasonal Multiplicative), Time-Series Analysis (Single Exponential Smoothing), Time-Series Analysis (Single

Moving Average), Trend Line (Difference Detrended), Trend Line (Exponential Detrended), Trend Line (Exponential), Trend Line (Linear Detrended), Trend Line (Linear), Trend Line (Logarithmic Detrended), Trend Line (Logarithmic), Trend Line (Moving Average Detrended), Trend Line (Moving Average), Trend Line (Polynomial Detrended), Trend Line (Polynomial), Trend Line (Power Detrended), Trend Line (Power), Trend Line (Rate Detrended), Trend Line (Static Mean Detrended), Trend Line (Static Median Detrended), Variance (Population), Variance (Sample), Volatility: EGARCH, Volatility: EGARCH-T, Volatility: GARCH, Volatility: GARCH-M, Volatility: GJR GARCH, Volatility: GJR TGARCH, Volatility: Log Returns Approach, Volatility: TGARCH, Volatility: TGARCH-M, Yield Curve (Bliss), and Yield Curve (Nelson-Siegel).

2

2. SIMULAZIONE MONTE CARLO

La Simulazione Monte Carlo, chiamata così in onore della famosa capitale del gioco di Monaco, è una metodologia molto potente. Per lo specialista, la simulazione offre la possibilità di trovare la soluzione a problemi pratici ma difficili e complessi con grande facilità. Monte Carlo crea futuri artificiali attraverso la generazione di migliaia e anche milioni di percorsi campione di esiti ed esamina le loro caratteristiche prevalenti. Per gli analisti di un'azienda, seguire corsi avanzati di matematica a livello post-universitario non è solo logico o funzionale. Un analista brillante userebbe tutti gli strumenti a sua disposizione per ottenere la stessa risposta nella maniera più facile e pratica possibile. Inoltre, quando modellata correttamente, la Simulazione Monte Carlo offre risposte simili ai metodi più matematicamente eleganti in tutti i tipi di casi. Allora, cos'è la Simulazione Monte Carlo e come funziona?

2.1 Cos'è la Simulazione Monte Carlo?

La Simulazione Monte Carlo, nella sua forma più semplice, è un generatore di numeri casuali, utile in previsioni, in valutazioni e in analisi del rischio. Una simulazione calcola numerosi scenari di un modello mediante la ripetuta selezione di valori da una *distribuzione di probabilità*, predefinita dall'utente, per le variabili incerte e l'utilizzo di questi valori per il modello. Dato che tutti quegli scenari producono risultati associati in un modello, ogni scenario può avere una *previsione*. Le previsioni sono eventi (normalmente con formule o funzioni) che voi definite come outputs importanti del modello. Questi sono di solito eventi come i ricavi totali, il profitto netto o le spese lorde.

Per semplificare, pensate al metodo della Simulazione Monte Carlo come la ripetuta estrazione con rimpiazzo di palline da golf da una grande cesta. La dimensione e la forma della cesta dipendono dall'***ipotesi d'input*** distribuzionale (per esempio, una distribuzione normale con una media di 100 e una deviazione standard di 10, contro una distribuzione uniforme o una distribuzione triangolare), dove alcune ceste sono più profonde o più simmetriche di altre, facendo sì che certe palline siano estratte più frequentemente di altre. Il numero di palline estratte ripetutamente dipende dal numero di ***prove*** simulate. Per un modello grande con multiple ipotesi collegate, immaginate questo modello grande come una cesta molto grande che contiene molte ceste piccole. Ogni piccola cesta ha il suo insieme di palline da golf che vi rimbalzano dentro. Queste ceste piccole sono talvolta collegate tra loro (se esiste una ***correlazione*** tra le variabili) e le palline da golf rimbalzano in tandem, mentre altre rimbalzano indipendentemente una dall'altra. Le palline che sono estratte ogni volta da queste interazioni all'interno del modello (la grande cesta centrale) sono tabulate e registrate, fornendo una ***previsione di output*** come risultato della simulazione.

2.2 Come iniziare con Simulatore di Rischio

2.2.1 SUna visione d'insieme d'alto livello del Software

Il software *Simulatore di Rischio* comprende diverse applicazioni, incluso la Simulazione Monte Carlo, la Previsione, l'Ottimizzazione e l'Analisi del rischio.

- Il *Modulo* di *Simulazione* vi permette di eseguire simulazioni nei vostri modelli esistenti basati su Excel, generare ed estrarre simulazioni di previsioni (distribuzioni di risultati), eseguire adattamenti distribuzionali (trovando automaticamente la distribuzione col miglior adattamento statistico), calcolare le correlazioni (mantenendo relazioni tra variabili casuali simulate), identificare le sensibilità (creando diagrammi Tornado e di sensibilità), testare le ipotesi statistiche (trovando differenze statistiche tra coppie di previsioni), eseguire simulazioni bootstrap (testando la robustezza delle statistiche dei risultati) ed eseguire simulazioni personalizzate e non parametriche (simulazioni usando dati storici senza specificare le distribuzioni o i loro parametri per la previsione senza dati o applicando previsioni basate su opinioni d'esperti).

- Il *Modulo di Previsione* può essere usato per generare automaticamente Previsioni di serie temporali (con e senza stagionalità e tendenza), Regressioni multivariate (modellando relazioni tra variabili), Estrapolazioni non lineari (adattamento alla curva), Processi stocastici (passeggiate casuali, ritorni alla media, diffusione a salti e processi misti), Box-Jenkins ARIMA (previsioni econometriche), Auto-ARIMA, Econometria di base e Auto-Econometria (modellando relazioni e generando previsioni), Curve esponenziali a J, Curve logistiche a S, Modelli GARCH e le multiple variazioni (modellando e prevedendo volatilità), Modelli di massima verosimiglianza per variabili dipendenti limitate (modelli logit, tobit e probit), Catene di Markov, Linee di tendenza, Curve spline ed altro ancora.

- Il *Modulo di Ottimizzazione* è usato per ottimizzare multiple variabili decisionali soggette a vincoli per massimizzare o minimizzare un obiettivo. Può essere eseguito sia come una ottimizzazione statica, dinamica o una ottimizzazione stocastica sotto incertezza insieme con una Simulazione Monte Carlo, oppure come una ottimizzazione stocastica con Simulazioni Super Veloci. Il software può gestire ottimizzazioni lineari e non lineari con variabili binarie, a numero intero e continue. Può anche generare frontiere efficienti di Markowitz.

- Il *Modulo degli Strumenti analitici* vi permette di eseguire la segmentazione clustering, i tests di verifica d'ipotesi, i tests statistici di dati grezzi, la diagnostica dei dati delle ipotesi tecniche di previsione (per esempio, eteroschedasticità, multicollinearità e affini), le analisi di sensibilità e degli scenari, le analisi di diagrammi sovrapposti, i diagrammi Ragno, i diagrammi Tornado e molti altri strumenti potenti.

- Il Real Options Risolutore di Super Reticoli è un altro software autonomo che integra il Simulatore di Rischio. È usato per risolvere problemi semplici e complessi di opzioni reali.

Le sezioni seguenti illustreranno i principi base del Modulo di Simulazione del Simulatore di Rischio, mentre i capitoli futuri scenderanno in maggior dettaglio nelle applicazioni d'altri moduli. Per poter seguire queste spiegazioni, assicuratevi di aver installato il Simulatore di Rischio sul vostro computer prima di procedere. ***In effetti, vi raccomandiamo vivamente di guardare prima sul web i video su come iniziare all'indirizzo www.realoptionsvaluation.com/risksimulator.html o di provare gli esercizi passo passo alla fine di questo capitolo avanti di ritornare e ripassare il testo in questo capitolo.*** Questo perché i video, come anche gli esercizi, vi aiuteranno ad iniziare immediatamente, mentre il testo in questo capitolo si concentra più sulla teoria e sulle spiegazioni dettagliate delle proprietà della simulazione.

2.2.2 Eseguire una Simulazione Monte Carlo

Tipicamente, per eseguire una simulazione nel vostro modello Excel esistente, devono essere compiuti i seguenti passi:

1. Impostate un nuovo profilo di simulazione o aprite un profilo esistente
2. Definite le ipotesi d'input nelle celle pertinenti
3. Definite le previsioni di output nelle celle pertinenti
4. Eseguite la simulazione
5. Interpretate i risultati

Se desiderate, e per esercitarvi, aprite il file d'esempio chiamato ***Modello Simulazione di base*** e seguite gli esempi sottostanti sulla creazione di una simulazione Il file d'esempio si trova o sotto il menu d'avvio ***Start | Real Options Valuation | Simulatore di Rischio | Esempi*** o con accesso diretto tramite ***Simulatore di Rischio | Modelli d'esempio.***

Impostare un Nuovo profilo di simulazione

Per avviare una nuova simulazione dovete prima creare un nuovo profilo di simulazione. Un profilo di simulazione contiene un insieme completo d'istruzioni su come volete che sia eseguita una simulazione, in altre parole tutte le ipotesi, le previsioni, le preferenze d'esecuzione e così via. I profili facilitano la creazione di multipli scenari di simulazioni. In altre parole, usando lo stesso identico modello potete creare vari profili, ciascuno con le sue specifiche proprietà di simulazione ed i suoi requisiti. La stessa persona può creare tests di scenari differenti usando ipotesi ed inputs distribuzionali diversi o più persone possono testare le loro ipotesi e i loro inputs sullo stesso modello.

- **Avviate Excel** e create un nuovo o aprite un modello esistente (potete usare l'esempio del Modello Simulazione di base per seguire)
- Cliccate su Simulatore di Rischio | Nuovo Profilo di simulazione
- Specificate un titolo per la vostra simulazione e tutte le altre informazioni pertinenti (Figura 2.1)

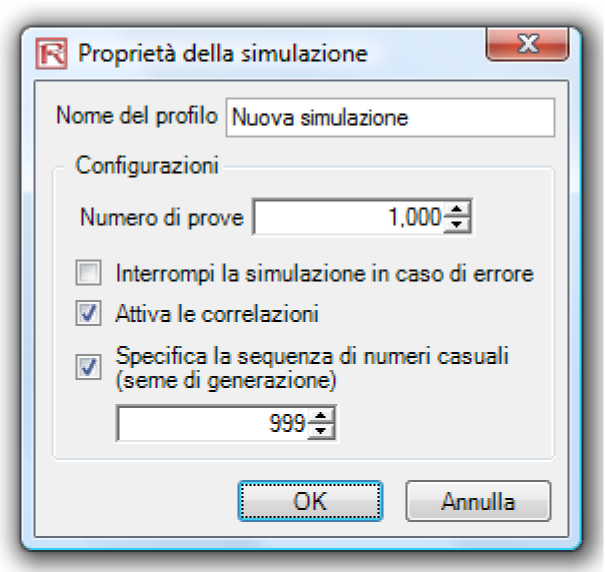


Figura 2.1 – Nuovo Profilo di Simulazione

Titolo:

- Specificare il titolo di una simulazione vi permette di creare multipli profili di simulazione in un singolo modello d'Excel. Questo significa che adesso potete salvare profili differenti di scenari di simulazione all'interno dello stesso modello senza dover cancellare ipotesi esistenti e cambiarle ogni volta che è richiesto un nuovo scenario di simulazione. Potete sempre cambiare il nome del profilo in un secondo momento (***Simulatore di Rischio | Edita profilo***).

Numero di prove:

- Qui s'inserisce il numero richiesto di prove di simulazione. In altre parole, eseguire 1000 prove significa che saranno generate 1000 differenti iterazioni di esiti basato sulle ipotesi d'input. Potete modificare quest'impostazione come richiesto, ma l'input deve essere un numero intero positivo. Il numero di esecuzioni di default è di 1000 prove. Potete usare il Controllo precisione ed errore per aiutarvi a determinare automaticamente il numero di prove di simulazione da eseguire (consultare la sezione sul Controllo precisione ed errore per dettagli).

Interrompi la simulazione in casi d'errore:

- Se questa opzione è attivata, la simulazione si ferma ogni volta che si verifica un errore nel modello Excel. In altre parole, se il vostro modello s'imbatte in un errore di calcolo (per esempio, alcuni valori d'input generati in una prova di simulazione potrebbero dare un errore "dividere per zero" in una delle celle del vostro foglio di lavoro) la simulazione si ferma. Questo è importante per aiutarvi a controllare il vostro modello per assicurarvi che non ci siano errori di calcolo nel vostro modello Excel. Tuttavia, se siete sicuri che il vostro modello funziona, non c'è bisogno di attivare questa preferenza.

Attivare le correlazioni:

- Se si attiva questa opzione, saranno calcolate le correlazioni tra le coppie d'ipotesi d'input. Diversamente le correlazioni saranno tutte impostate su zero e una simulazione sarà eseguita assumendo che non ci sono correlazioni incrociate tra le ipotesi d'input. Ad esempio, l'applicazione delle correlazioni fornirà risultati più accurati se le correlazioni esistono veramente e tenderà a fornire una confidenza di previsione più bassa se ci sono correlazioni negative. Dopo aver attivato qui le correlazioni, potete impostare più tardi i coefficienti di correlazione pertinenti su ciascuna ipotesi generata (consultare la sezione sulle correlazioni per maggiori dettagli).

Specificare la sequenza dei numeri casuali:

- La simulazione, per definizione, fornirà un risultato leggermente diverso ogni volta che viene eseguita una simulazione. Questo accade in virtù della procedura di generazione di numeri casuali nella Simulazione Monte Carlo ed è un fatto teorico in tutti i generatori di numeri casuali. Tuttavia, quando fate una presentazione potreste avere la necessità di mostrare gli stessi risultati (specialmente quando il report che viene presentato mostra un determinato insieme di risultati e volete che la generazione di questi stessi risultati venga visualizzata durante una presentazione dal vivo, o quando condividete i modelli con altri e volete che gli stessi risultati siano ottenuti ogni volta). In tal caso potete attivare questa preferenza ed inserire un numero di seme iniziale. Il numero di seme può essere qualsiasi numero intero positivo. Se usate lo stesso valore di seme iniziale, lo stesso numero di prove e le stesse ipotesi d'input, la simulazione fornirà sempre la stessa sequenza di numeri casuali, garantendo sempre lo stesso insieme finale di risultati.

Prego notare che una volta che avete creato un nuovo profilo di simulazione, potete ritornarvi più tardi e modificare queste selezioni. Per eseguire ciò, assicuratevi che il profilo attivo attuale sia il profilo che volete modificare. Altrimenti cliccate su ***Simulatore di Rischio | Cambia profilo di simulazione***, selezionate il profilo da modificare e cliccate su ***OK*** (Figura 2.2 mostra un esempio dove ci sono multipli profili e indica come attivare il profilo selezionato). Dopo, cliccate su ***Simulatore di Rischio | Edita profilo di simulazione*** ed eseguite le modifiche richieste. Potete anche duplicare o rinominare un profilo esistente. Quando create multipli profili nello stesso modello Excel, assicuratevi di dare un nome univoco a ciascun profilo, in modo da poterli distinguere più tardi. Questi profili sono inoltre archiviati all'interno di settori nascosti del file *.xls di Excel e, quindi, non dovete salvare dei files aggiuntivi. I profili ed i loro contenuti (ipotesi, previsioni e così via) sono automaticamente salvati, quando salvate il file di Excel. Per finire, l'ultimo profilo attivo, quando uscite e salvate il file Excel, è il file che sarà aperto al prossimo avvio del file di Excel.

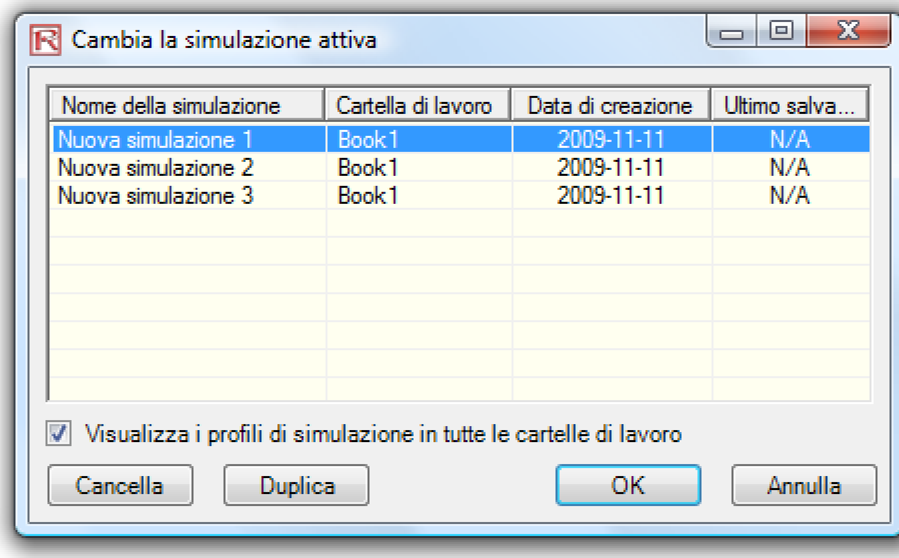


Figura 2.2 – Cambiare la simulazione attiva

Definire le ipotesi d'input

Il passo successivo è di definire le ipotesi d'input nel vostro modello. Prego notare che le ipotesi possono solo essere assegnate a celle senza equazioni o funzioni, vale a dire valori numerici digitati che sono inputs in un modello, mentre le previsioni di output possono solo essere assegnate a celle con equazioni o funzioni, vale a dire outputs di un modello. Ricordatevi che ipotesi e previsioni non possono essere impostate se non esiste già un profilo di simulazione. Impostate le nuove ipotesi d'input nel vostro modello come segue:

- Assicuratevi che esita già un profilo di simulazione o aprite un profilo esistente, oppure iniziate un nuovo profilo (Simulatore di Rischio | Nuovo profilo di simulazione)

- Selezionate la cella sulla quale desiderate impostare un'ipotesi (per esempio, la cella G8 nell'esempio del Modello di Simulazione di base)

- Cliccate su Simulatore di Rischio | Imposta ipotesi d'input o cliccate sull'icona Imposta ipotesi d'input nella Barra degli strumenti di Simulatore di Rischio

- Selezionate la distribuzione pertinente di vostra scelta, inserite i parametri di distribuzione pertinenti (per esempio, distribuzione Triangolare con 1, 2, 2.5 come i valori minimi, più probabili e massimi) e cliccate su OK per inserire le ipotesi d'input nel vostro modello (Figura 2.3)

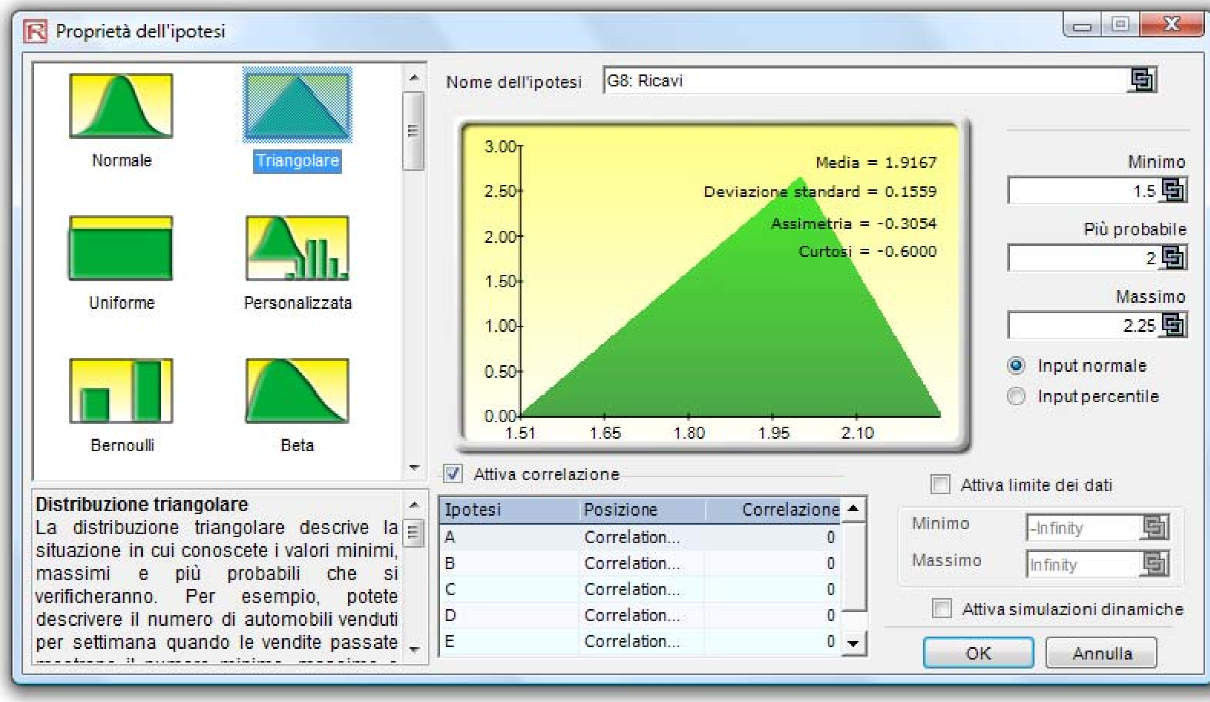


Figura 2.3 – Impostare un'ipotesi d'input

Prego notare che potete anche impostare le ipotesi come segue: selezionate la cella sulla quale desiderate impostare l'ipotesi e, usando il **tasto destro del mouse**, accedete al menu di scelta rapida di Simulatore di Rischio per impostare un'ipotesi d'input. Utenti esperti possono anche impostare le ipotesi d'input usando le **Funzioni SR** (***RS Functions)*** di Simulatore di Rischio: selezionate la cella che desiderate, cliccate in Excel su *Inserisci*, *Funzione* e selezionate *Tutte le categorie* e spostatevi giù fino alla lista delle Funzioni SR (vi consigliamo di non usare le Funzioni SR se non siete un utente esperto). Per gli esempi che seguono, vi consigliamo di seguire le istruzioni di base per accedere ai menu e alle icone.

Prego notare che nella scheda Proprietà dell'ipotesi ci sono alcuni aree fondamentali degne di nota. La Figura 2.4 mostra queste aree:

Nome dell'ipotesi:

- Questa è un'area opzionale che vi permette di inserire nomi univoci per le ipotesi per aiutarvi a distinguere ciò che ogni ipotesi rappresenta. È buona pratica di modellazione usare nomi brevi ma precisi per le ipotesi.
- **Galleria delle Distribuzioni:** Quest'area sulla sinistra mostra tutte le diverse distribuzioni disponibili nel software. Per cambiare le visualizzazioni, cliccate col tasto destro da qualsiasi parte nella galleria e selezionate icone grandi, icone piccole o lista. Sono disponibili più di 42 distribuzioni.

- **Parametri di Input:** I parametri pertinenti saranno visualizzati in base alla distribuzione selezionata. Potete o inserire i parametri direttamente o collegarli a specifiche celle nel vostro foglio di lavoro. L'hardcoding o l'inserimento manuale dei parametri è utile, quando si presume che i parametri d'ipotesi non cambieranno. Il collegamento a celle del foglio di lavoro è utile, quando i parametri d'input devono essere visibili o possono cambiare (cliccate sull'icona Collegamento per collegare un parametro ad una cella del foglio di lavoro).

- **Attiva limite dei data:** Questi limiti non sono normalmente usati dall'analista normale, ma sono presenti per troncare le ipotesi distribuzionali. Per esempio, se viene selezionata una distribuzione normale, i limiti teorici sono tra l'infinto negativo l'infinto positivo. Tuttavia, nella pratica, la variabile simulata esiste solo entro un intervallo più piccolo e quest'intervallo può essere inserito per troncare la distribuzione in modo appropriato.

- **Correlazioni:** Qui si possono assegnare correlazioni accoppiate ad ipotesi d'input. Se le correlazioni sono necessarie, ricordatevi di attivare la preferenza *Attiva correlazioni* cliccando su ***Simulatore di Rischio | Edita profilo di simulazione***. Consultate la sezione sulle correlazioni di seguito in questo capitolo per maggiori dettagli su come assegnare le correlazioni e gli effetti che le correlazioni avranno su un modello. Prego notare che potete o troncare una distribuzione o correlarla con un'altra ipotesi, ma non entrambe le cose.

- **Descrizioni brevi:** Queste sono disponibili per tutte le distribuzioni nella galleria. Le descrizioni brevi spiegano sia quando si usa una determinata distribuzione che i requisiti dei parametri d'input. Consultate la sezione in *Comprendere le Distribuzioni di Probabilità per la Simulazione Monte Carlo* per dettagli su ciascuna delle distribuzioni disponibili nel software.

- **Input normale e Input percentile:** Questa opzione permette all'utente di eseguire un veloce test di "due diligence" sull'ipotesi d'input. Per esempio, se state impostando una distribuzione normale con determinati inputs per la media e la deviazione standard, potete cliccare sull'input percentile per vedere quali sono il 10° e il 90° percentili.

- **Attiva simulazioni dinamiche:** Questa opzione è disattivata di default. Tuttavia, se desiderate eseguire una simulazione multidimensionale (vale a dire, se collegate i parametri d'input dell'ipotesi con un'altra cella che è essa stessa un'ipotesi, state simulando gli inputs o simulando la simulazione), ricordatevi di attivare questa opzione. La simulazione dinamica non funzionerà se gli inputs non sono collegati ad altre ipotesi d'input mutevoli.

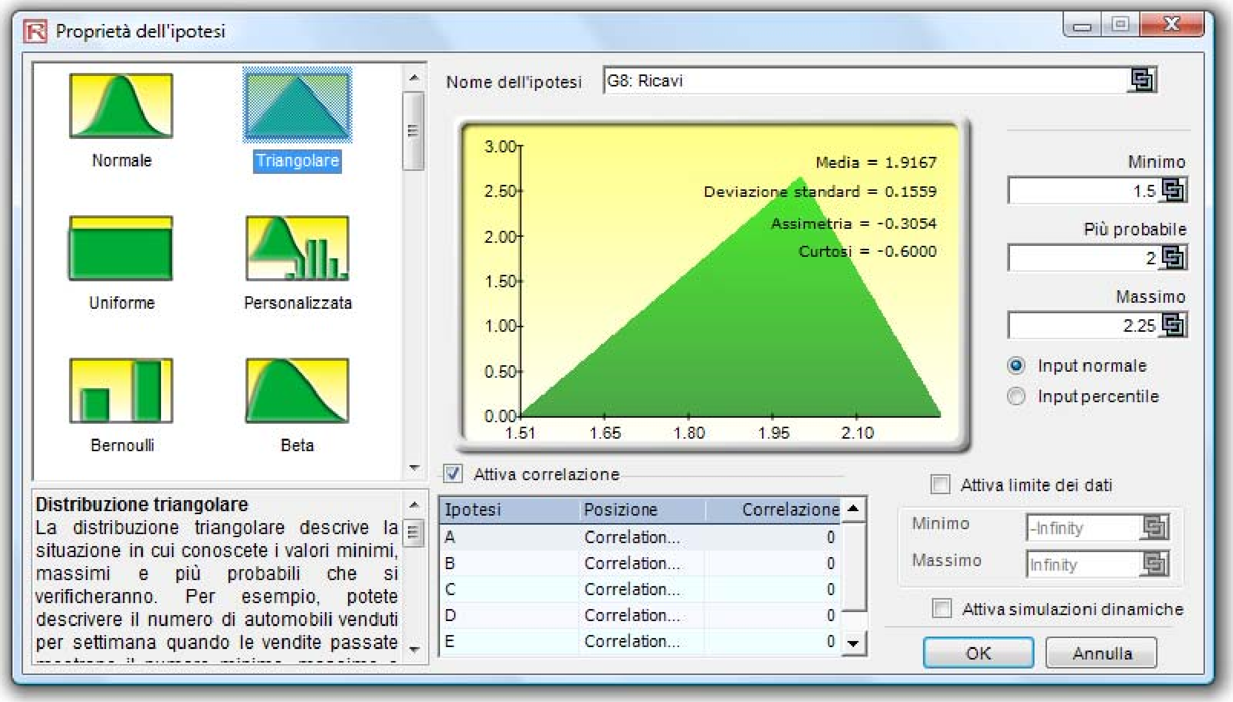


Figura 2.4 –Proprietà dell'ipotesi

Nota: Se state seguendo quest'esempio, continuate impostando un'altra ipotesi sulla cella ***G9***. Questa volta usate la distribuzione ***Uniforme*** con un valore minimo di ***0.9*** e un valore massimo di ***1.1***. Dopo procedete con la definizione delle previsioni di output nel passo successivo.

Definire le previsioni di output

Il passo successivo e di definire le previsioni di output nel modello. Le previsioni possono essere definite solo su celle di output con equazioni o funzioni. Di seguito descriviamo il processo d'impostazione della previsione :

- Selezionate la cella sulla quale desiderate impostare una previsione (per esempio, la cella ***G10*** nell'esempio del Modello di Simulazione di base)
- Cliccate su ***Simulatore di Rischio*** e selezionate ***Imposta previsione di output*** oppure cliccate sull'icona Imposta previsione di output nella Barra degli strumenti delle icone di Simulatore di Rischio (Figura 1.3)
- Inserite l'informazione pertinente e cliccate su ***OK***

Prego notare che potete anche impostare le previsioni di output come segue: selezionate la cella sulla quale volete impostare la previsione e, usando il **tasto destro del mouse**, accedete al menu di scelta rapida di Simulatore di Rischio per impostare una previsione di output.

La Figura 2.5 illustra le proprietà dell'impostazione della previsione.

Nome della previsione:

- Specificate il nome della cella di previsione. Questo è importante perché quando avete un modello grande con molte celle di previsione, il denominare individualmente le celle di previsione vi permette di accedere rapidamente ai risultati giusti. Non sottostimate l'importanza di questo semplice passo. E' buona pratica di modellazione usare nomi brevi ma precisi per la previsione.

Precisione della previsione:

- Invece di affidarvi ad un calcolo approssimativo su quante prove eseguire nella vostra simulazione, potete impostare i controlli di precisione ed errore. Quando una determinata combinazione errore-precisione è stata raggiunta nella simulazione, la simulazione si fermerà e v'informerà sulla precisione raggiunta. Questo rende il numero di prove di simulazione un processo automatizzato e non richiede congetture sul numero richiesto di prove da simulare. Consultate la sezione sui controlli errore e precisione per maggiori dettagli specifici.

Mostra la finestra della previsione:

- Permette all'utente di visualizzare o no una determinata finestra di previsione. Di default viene sempre visualizzato il diagramma della previsione.

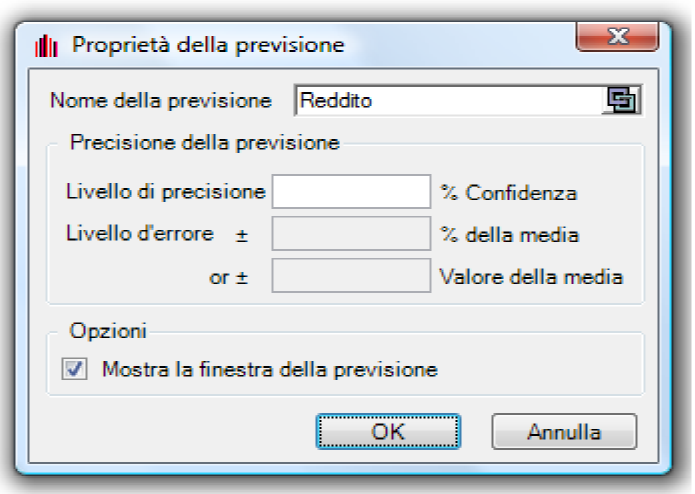


Figura 2.5 – Impostare la previsione di output

Eseguire la simulazione

Quando tutto sembra a posto, cliccate semplicemente su ***Simulatore di Rischio | Esegui simulazione*** o cliccate sull'icona ***Esegui*** nella Barra degli strumenti di Simulatore di Rischio e la simulazione inizierà. Potete anche resettate una simulazione dopo la sua esecuzione per eseguirla di nuovo (***Simulatore di Rischio | Resetta simulazione*** o cliccate sull'icona Resetta simulazione nella Barra degli strumenti), o fermarla durante un'esecuzione. In più, la funzione *passi successivi* (***Simulatore di Rischio | Simulazione a passi*** o l'icona Simulazione a passi nella Barra degli strumenti) vi permette di simulare una singola prova un passo alla volta. Questo è utile per insegnare ad altri il concetto della simulazione (in pratica potete mostrare che ad ogni prova, tutti i valori nelle celle d'ipotesi sono sostituiti e l'intero modello è calcolato ogni volta di nuovo). Potete anche accedere al menu Esegui simulazione cliccando col tasto destro del mouse da qualsiasi parte nel modello e selezionando Esegui simulazione.

Simulatore di Rischio vi permette anche di eseguire la simulazione a velocità molto elevate, chiamato Super Veloce. Per attuare ciò, cliccate su ***Simulatore di Rischio | Esegui Simulazione Super Veloce*** o usate l'icona Esegui Super Veloce. Notate con quale velocità superiore viene eseguita la Simulazione Super Veloce. Per esercitarvi: cliccate su ***Resetta simulazione***, poi su ***Edita profilo simulazione*** e cambiate il ***Numero di prove*** a ***100000*** e infine su ***Esegui Super Veloce***. L'esecuzione dovrebbe richiedere solo pochi secondi. Tuttavia, siate consapevoli che la Simulazione Super Veloce non sarà eseguita se il modello contiene errori, VBA (Visual Basic for Applications) o collegamenti a sorgenti dati o applicazioni esterne. In tali situazioni sarete notificati e sarà eseguita invece la simulazione a velocità normale. Le simulazioni a velocità normale sono sempre eseguibili anche con errori, VBA o collegamenti esterni.

Interpretare i risultati della previsione

Il passo finale in una simulazione Monte Carlo è di interpretare i risultanti diagrammi di previsione. Le Figure 2.6 a 13 mostrano un diagramma di previsione e le corrispondenti statistiche generate dopo l'esecuzione della simulazione. Tipicamente, i seguenti elementi sono importanti nell'interpretare i risultati di una simulazione:

Diagramma della previsione:

- Il diagramma della previsione mostrato nella Figura 2.6 è un istogramma di probabilità che mostra il conteggio della frequenza dei valori che accadono nel numero totale di prove simulate. Le barre verticali mostrano la frequenza di un determinato valore *x* che avviene nel corso di un numero totale di prove, mentre la frequenza cumulativa (linea smussale) mostra le probabilità totali che avvengono nella previsione di tutti i valori fino a e sotto lo *x*.

Statistiche della previsione:

- Le statistiche della previsione mostrate nella Figura 2.7 riassumono la distribuzione dei valori della previsione in termini dei quattro momenti di una distribuzione. Consultare la sezione su *Comprendere le statistiche della previsione* per maggiori dettagli sul significato di alcune di queste statistiche. Potete alternare tra le linguette Istogramma e Statistiche premendo la barra spaziatrice.

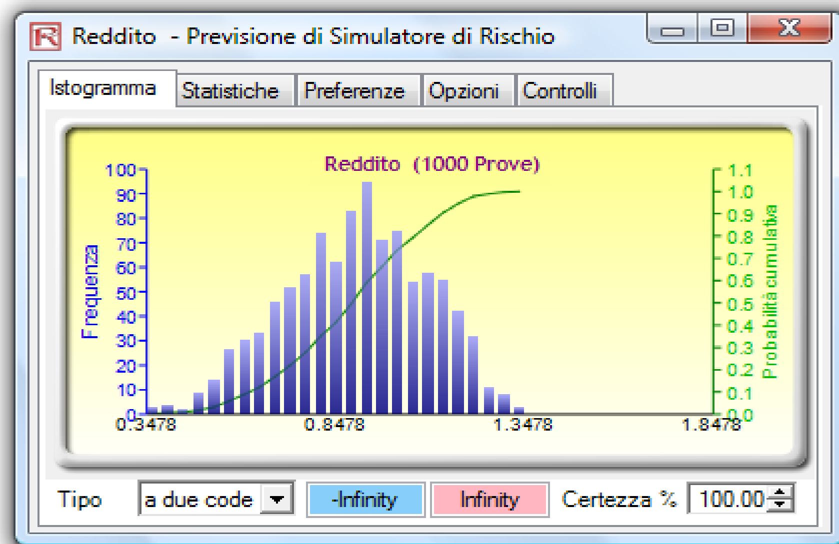
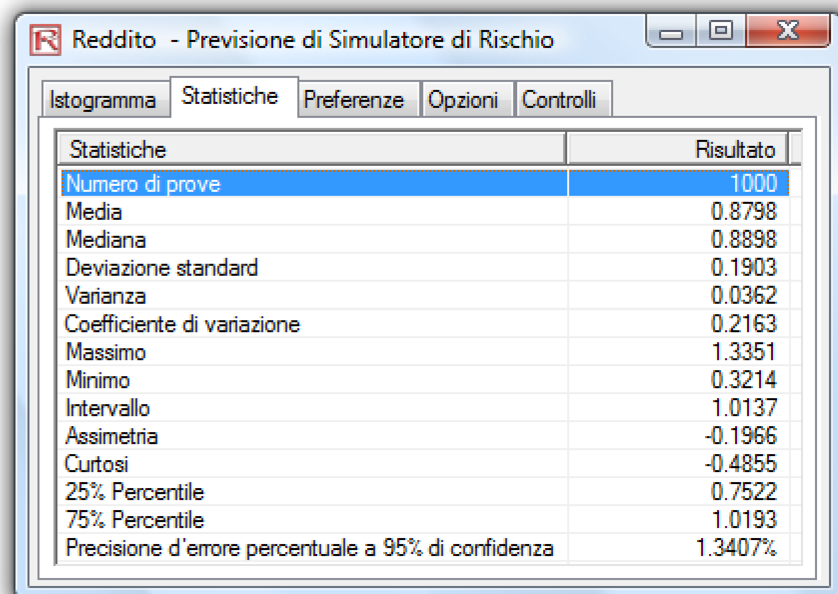


Figura 2.6 – Diagramma della previsione



Statistiche	Risultato
Numero di prove	1000
Media	0.8798
Mediana	0.8898
Deviazione standard	0.1903
Varianza	0.0362
Coefficiente di variazione	0.2163
Massimo	1.3351
Minimo	0.3214
Intervallo	1.0137
Assimetria	-0.1966
Curtosi	-0.4855
25% Percentile	0.7522
75% Percentile	1.0193
Precisione d'errore percentuale a 95% di confidenza	1.3407%

Figura 2.7 – Statistiche della previsione

Preferenze:

- La linguetta Preferenze nel diagramma della previsione vi permette di cambiare l'aspetto e la manualità dei diagrammi. Per esempio, se è selezionata l'opzione ***Sempre in primo piano***, i diagrammi della previsione saranno sempre visibili a prescindere dagli altri software in esecuzione sul vostro computer. ***Risoluzione dell'istogramma*** vi permette di cambiare il numero di "bins" dell'istogramma, da 5 a 100 "bins". La sezione ***Aggiorna dati*** vi permette di controllare la velocità della simulazione contro la frequenza d'aggiornamento del diagramma della previsione. In altre parole, se volete vedere un aggiornamento del diagramma della previsione quasi ad ogni prova, questo rallenterà la simulazione, dato che più memoria viene allocata all'aggiornamento del diagramma che all'esecuzione della simulazione. Questa è solamente una preferenza dell'utente e non cambia in nessun modo i risultati della simulazione, solo la velocità di completamento della simulazione. Per aumentare ulteriormente la velocità della simulazione, potete minimizzare Excel durante l'esecuzione della simulazione, riducendo così la memoria richiesta per aggiornare visivamente il foglio di lavoro di Excel e liberando la memoria per l'esecuzione della simulazione. I bottoni ***Chiudi tutto*** e ***Minimizza tutto*** controllano tutti i diagrammi della previsione aperti.

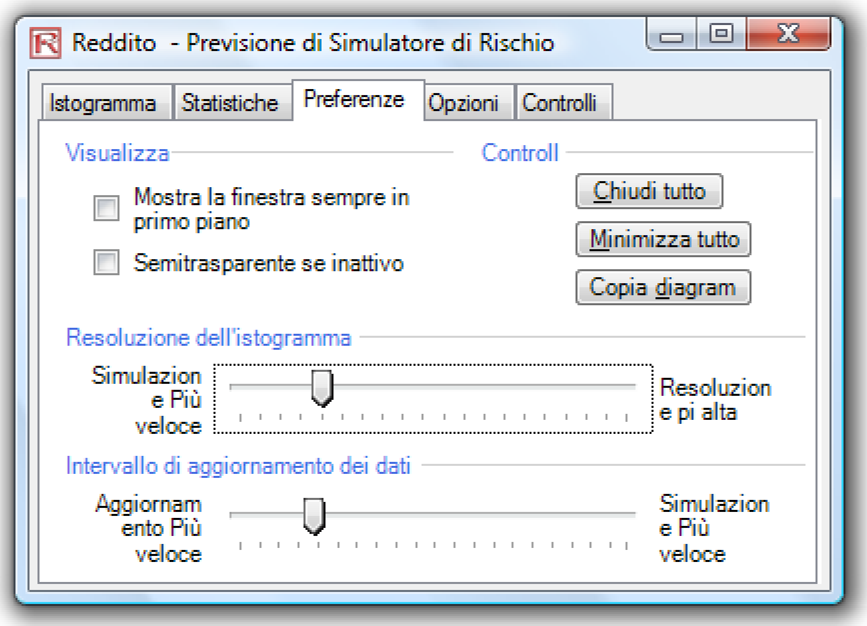


Figura 2.8 –Preferenze del diagramma della previsione

Opzioni:

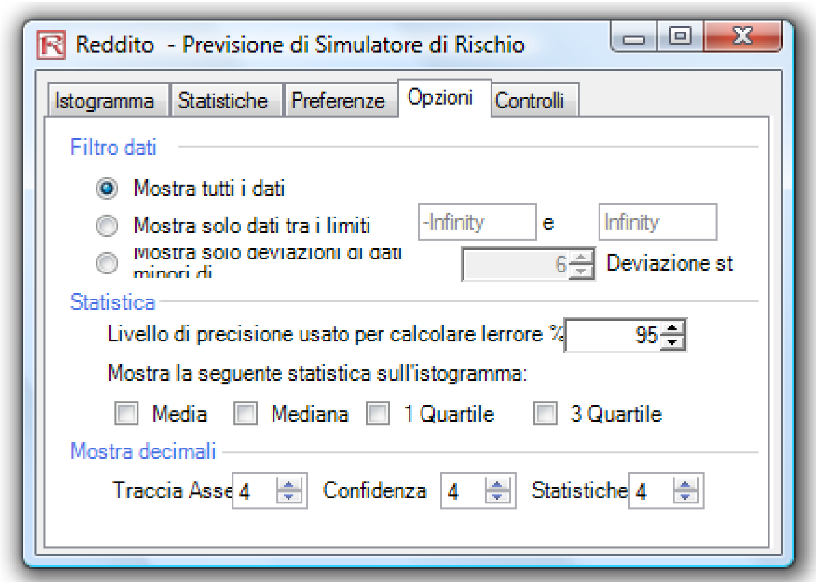
- L'opzione Diagramma della previsione vi permette di visualizzare tutti i dati della previsione oppure di filtrare dentro/fuori i valori che cadono entro uno specifico intervallo da voi scelto o entro una deviazione standard da voi scelta. Qui si può inoltre impostare il livello di precisione per questa specifica previsione per mostrare i livelli d'errore nella visualizzazione delle statistiche. Consultare la sezione sul Controllo errore e precisione per maggiori dettagli. *Mostra le seguenti statistiche* è una preferenza dell'utente per indicare se la media, la mediana e la linea del 1° quartile e del 4° quartile (25° e 75° percentile) devono essere visualizzati nel diagramma della previsione.

Controlli:

- Questa linguetta contiene tutte le funzionalità che vi permettono di modificare il tipo, il colore, la dimensione, lo zoom, l'inclinazione, il 3D ed altro nel diagramma della previsione. Fornisce anche i diagrammi sovrapposti (PDF, CDF) ed esegue l'adattamento distribuzionale sui vostri dati di previsione (consultare le sezioni Adattamento dati per maggiori dettagli su questa metodologia).

Usare i diagrammi della previsione e gli Intervalli di confidenza

- Nei diagrammi della previsione potete determinare la probabilità di avvenimento, denominato intervalli di confidenza. Vale a dire, dato due valori, quali sono le probabilità che gli esiti cadranno tra questi due valori? La Figura 2.10 illustra che c'è una probabilità del 90% che l'esito finale (in questo caso, il livello del reddito) sarà tra \$0.5307 e \$1.1739. L'intervallo di confidenza a due code può essere ottenuto selezionando prima ***a due code*** come il tipo, inserendo poi il valore di certezza desiderato (per esempio, ***90***) e premendo infine il tasto ***TAB*** sulla tastiera. A questo punto saranno visualizzati i due valori calcolati che corrispondono al valore di certezza. In questo esempio c'è una probabilità del 5% che il reddito sarà sotto \$0.5307 e un'altra probabilità del 5% che il reddito sarà sopra \$1.1739. In altre parole, l'intervallo di confidenza a due code è un intervallo simmetrico centrato sulla mediana o il 50° valore percentile. Pertanto entrambe le code avranno la stessa probabilità.



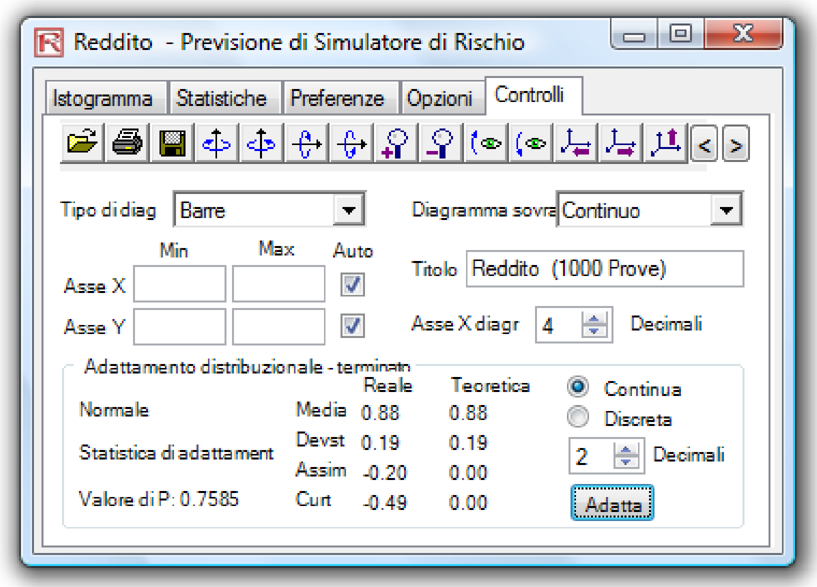


Figura 2.9 – Diagramma della previsione - Opzioni e Controlli

Figura 2.10 – Diagramma della previsione - Intervallo di confidenza a due code

In alternativa si può calcolare la probabilità ad una coda. La Figura 2.11 mostra la selezione di una Coda sinistra ad una confidenza del 95% (vale a dire, scegliete ***Coda sinistra ≤*** come tipo, inserite ***95*** come livello di certezza e premete ***TAB*** sulla tastiera). Questo significa che c'è una probabilità del 95% che il reddito sarà sotto $1.1739 o una probabilità del 5% che il reddito sarà sopra $1.1739. Questo corrisponde perfettamente ai risultati visti nella Figura 2.10.

Figura 2.11 – Diagramma della previsione - Intervallo di confidenza ad una coda

Oltre alla valutazione dell'intervallo di confidenza (cioè, dato un livello di probabilità trovare i valori pertinenti del reddito), potete determinare la probabilità di un determinato valore del reddito. Per esempio, qual è la probabilità che il reddito sarà minore di o uguale a $1? Per eseguire ciò, selezionate il tipo di probabilità ***Coda sinistra ≤***, inserite ***1*** nel riquadro del valore d'input e premete ***TAB***. Sarà ora calcolata la certezza corrispondente (in questo caso, c'è una probabilità del 74.30% che il reddito sarà di o sotto $1).

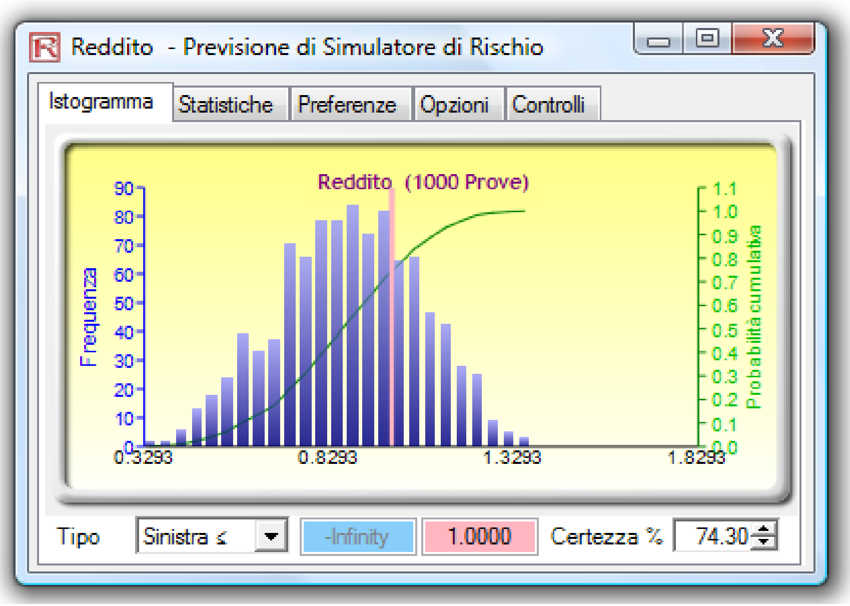


Figura 2.12 – Diagramma della previsione - Valutazione della Probabilità

Nell'interesse della completezza, potete selezionare il tipo di probabilità ***Coda destra >*** probabilità, inserire il valore ***1*** nel riquadro del valore d'input e premere ***TAB***. La probabilità risultante indica che la probabilità della coda destra è oltre il valore di 1, vale a dire, la probabilità che il reddito superi $1 (in questo caso vediamo che c'è una probabilità del 25.70% che il reddito superi $1). La somma di 74.30% e 25.70% è naturalmente 100%, la probabilità totale sotto la curva.

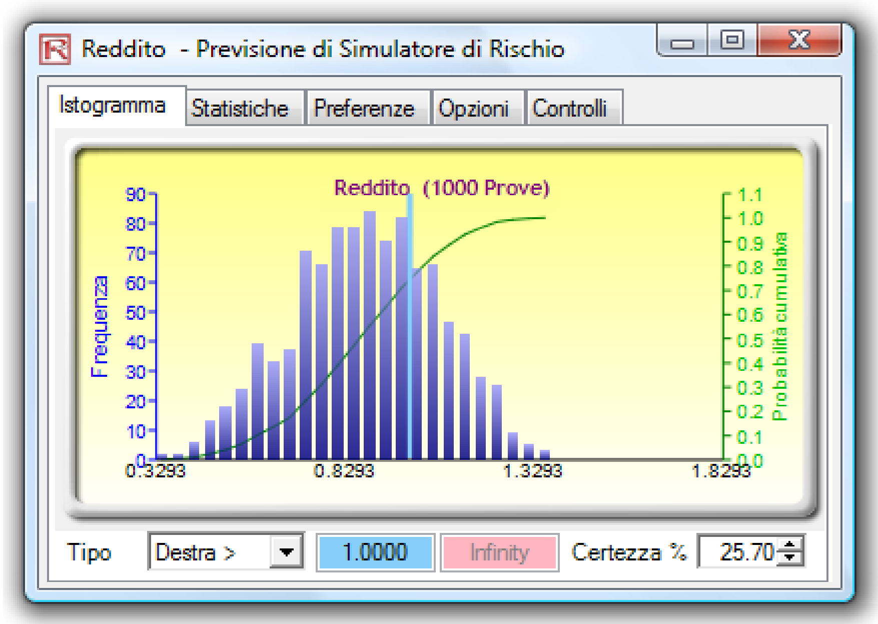


Figura 2.13 – Diagramma della previsione - Valutazione della Probabilità

SUGGERIMENTI:

- Si può ridimensionare la finestra della previsione cliccando su e trascinando l'angolo inferiore destro della finestra della previsione.
- Prima di rieseguire la simulazione, consigliamo di resettare (***Simulatore di Rischio | Resetta simulazione***) la simulazione attiva.
- Ricordatevi che dovete premere ***TAB*** sulla tastiera per aggiornare il diagramma e i risultati quando digitate i valori della certezza o i valori delle due code (destra e sinistra).
- Potete anche premere ripetutamente la ***Barra spaziatrice*** sulla tastiera per passare da linguetta a linguetta (Istogramma, Statistiche, Preferenze, Opzioni e Controlli.
- In aggiunta, se cliccate su ***Simulatore di Rischio | Opzioni***, potete accedere a varie opzioni di Simulatore di Rischio. Potete impostare Simulatore di Rischio per avviarsi ogni volta che si avvia Excel oppure solo quando lo volete voi (andate a ***Start | Programmi | Real Options Valuation | Simulatore di Rischio | Simulatore di Rischio***). Potete modificare i ***colori delle celle*** delle ipotesi e delle previsioni e potete attivare/disattivare i ***commenti cella*** (i commenti cella vi permetto di vedere sia quali celle sono ipotesi d'input e quali sono previsioni di output che i loro rispettivi parametri d'input e i nomi). Prendete del tempo per familiarizzarvi con i diagrammi di previsioni di output e con le varie "campane e fischietti", specialmente la linguetta ***Controlli***.

2.3 Correlazioni e Controllo Precisione

2.3.1 I principi base delle correlazioni

Il coefficiente di correlazione è una misura della forza e della direzione della relazione tra due variabili e può prendere qualsiasi valore tra –1.0 e +1.0. In altre parole, il coefficiente di correlazione può essere scomposto nel suo segno (relazione positiva o negativa tra due variabili) e nella magnitudine o forza della relazione (tanto più alto è il valore assoluto del coefficiente di correlazione, tanto più forte è la relazione).

Il coefficiente di correlazione può essere calcolato in vari modi. Il primo metodo è di calcolare manualmente la correlazione r di due variabili x e y usando:

$$r_{x,y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$$

Il secondo metodo è di usare la funzione *CORREL* di Excel. Per esempio, se i 10 punti dati per *x* e *y* sono elencati nelle celle A1:B10, allora la funzione Excel da usare è *CORREL (A1:A10, B1:B10)*.

Il terzo metodo è di eseguire lo Strumento *Adattamento multiplo (Multi-Fit) di Simulatore di Rischio*. Sarà calcolata e visualizzata la risultante matrice di correlazione.

È importante notare che la correlazione non implica la causalità. Due variabili aleatorie completamente irrelate potrebbero manifestare una qualche correlazione, ma questo non implica nessun rapporto di causa ed effetto tra le due (per esempio, l'attività delle macchie solari e gli eventi del mercato azionario sono correlati, ma non esiste una causalità tra i due).

Ci sono due tipi generali di correlazioni: correlazioni parametriche e non parametriche. Il Coefficiente di correlazione di Pearson è la misura di correlazione più comune ed è di solito chiamato semplicemente Coefficiente di correlazione. Tuttavia, la correlazione di Pearson è una misura parametrica. Questo significa che ha due presupposti: entrambe le variabili correlate devono avere una sottostante distribuzione normale e la relazione tra le due variabili deve essere lineare. Quando queste due condizioni sono violate, cosa che succede spesso nelle simulazioni Monte Carlo, diventano più importanti le controparti non parametriche. La Correlazione per ranghi di Spearman ed il Tau di Kendall sono due alternative. La correlazione di Spearman è usata più comunemente ed è più appropriata se applicata nel contesto di una Simulazione Monte Carlo – non c'è subordinazione a distribuzioni normali o linearità, il che significa che si possono applicare correlazioni tra variabili differenti con distribuzioni diverse. Per calcolare la correlazione di Spearman, ordinate per rango prima tutti i valori delle variabili first *x* e *y* e applicate poi il calcolo della correlazione di Pearson.

Nel caso di Simulatore di Rischio, la correlazione usata è la più robusta e non parametrica Correlazione per ranghi di Spearman. Tuttavia, per semplificare il processo di simulazione e per essere coerenti con la funzione di Correlazione di Excel, gli inputs richiesti per la correlazione sono i Coefficienti di correlazione di Pearson. Simulatore di Rischio applicherà poi i propri algoritmi per convertirli in una Correlazione per ranghi di Spearman, semplificando così il processo. In ogni caso, per semplificare l'interfaccia utente, permettiamo agli utenti di inserire la più comune correlazione prodotto-momento di Pearson (per esempio calcolata usando la funzione *CORREL* di Excel). Invece per quanto riguarda i codici matematici, convertiamo queste semplici correlazioni in correlazioni di Spearman basati sul rango per le simulazioni distribuzionali.

2.3.2 Applicare le correlazioni in Simulatore di Rischio

Le correlazioni possono essere applicate in Simulatore di Rischio in vari modi:

- Quando definite le ipotesi (***Simulatore di Rischio*** | ***Imposta ipotesi di input***), inserite semplicemente le correlazioni nella griglia di matrice di correlazione nella Galleria delle Distribuzioni.
- Con dati esistenti, eseguite lo strumento Adattamento multiplo (Multi-Fit) (***Simulatore di Rischio*** | ***Strumenti*** | **Adattamento *distribuzionale*** | ***Variabili multiple***) per eseguire l'adattamento distribuzionale e per ottenere la matrice di correlazione tra variabili accoppiate. Se esiste un profilo di simulazione, le ipotesi adattate conteranno automaticamente i valori di correlazione pertinenti.
- Con ipotesi esistenti, potete cliccare su ***Simulatore di Rischio*** | ***Strumenti*** | ***Edita correlazioni*** per inserire le correlazioni accoppiate di tutte le ipotesi direttamente in un'unica interfaccia utente.

Prego notare che la matrice di correlazione deve essere positiva definita. In altre parole, la correlazione deve essere matematicamente valida. Per esempio, supponiamo che stiate tentando di correlare tre variabili: i voti di dottorandi in un determinato anno, il numero di birre che consumano in una settimana e il numero d'ore di studio in una settimana. Si potrebbe presumere che esistano le seguenti relazioni di correlazione:

Voti e Birre: – Tanto più bevono, tanto più bassi i voti (assenze agli esami)

Voti e Studio: + Tanto più studiano, tanto più alti sono i voti

Birre e Studio: – Tanto più bevono, tanto meno studiano (sempre ubriachi e festaioli)

Se inserite però una correlazione negativa tra Voti e Studio, e assumendo che i coefficienti di correlazione sono di grandi magnitudini, la matrice di correlazione sarà non positiva definita. Questo andrebbe contro la logica, i requisiti della correlazione e la matematica delle matrici. Tuttavia, coefficienti più piccoli possono talvolta funzionare nonostante la cattiva logica. Quando viene inserita una matrice di correlazione non positiva o errata, il Simulatore di Rischio v'informerà automaticamente e offrirà di correggere queste correlazioni in qualcosa di semi-positivo definito, mantenendo comunque la struttura generale della relazione di correlazione (gli stessi segni e le stesse forza relative).

2.3.3 Gli effetti delle correlazioni in una Simulazione Monte Carlo

Anche se i calcoli richiesti per correlare le variabili in una simulazione sono complessi, gli effetti risultanti sono abbastanza chiari. La Figura 2.14 presenta un semplice modello di correlazione (Modello Effetti della correlazione nella cartella degli Esempi Il calcolo per Ricavi è semplicemente prezzo moltiplicato per quantità. Lo stesso modello è replicato per senza correlazione, per correlazione positiva (+0.8) e per correlazione negativa (–0.8) tra prezzo e quantità.

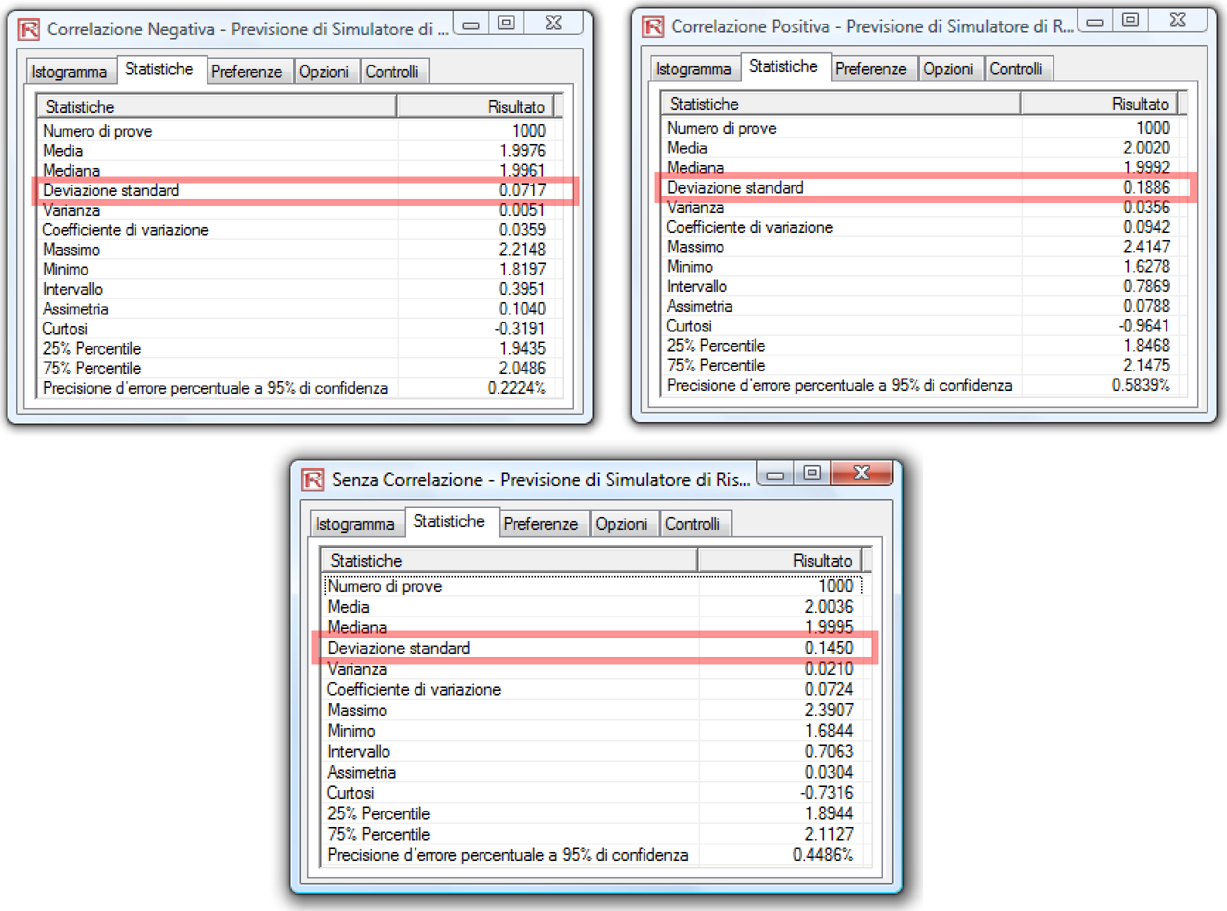
Modello di correlazione

	Senza correlazione	Correlazione positiva	Correlazione negativa
Prezzo	$2.00	$2.00	$2.00
Quantità	1.00	1.00	1.00
Ricavi	$2.00	$2.00	$2.00

Per replicare questo modello, usate le seguenti ipotesi:
Prezzi sono impostati come Distribuzioni triangolari (1.8, 2.0, 2.2), mentre le Quantità sono impostate come Distribuzioni uniformi (0.9, 1.1) con correlazioni impostate a 0,0, +0,8, -0,8 con 1000 prove con valore di seme di generazione di 123456.

Figura 2.14 –Modello semplice di correlazione

Le statistiche risultanti sono mostrate nella Figura 2.15. Prego notare che la deviazione standard del modello senza correlazioni è 0.1450, paragonato a 0.1886 per la correlazione positiva e 0.0717 per la correlazione negativa. In altre parole, per modelli semplici, le correlazioni negative tendono a ridurre la dispersione media della distribuzione e a creare una distribuzione della previsione stretta e più concentrata rispetto a correlazioni positive con dispersioni medie più grandi. La media rimane comunque relativamente stabile. Questo implica che le correlazioni fanno poco per cambiare il valore atteso di progetti, ma può ridurre o aumentare il rischio di un progetto.



Statistiche	Risultato
Numero di prove	1000
Media	1.9976
Mediana	1.9961
Deviazione standard	0.0717
Varianza	0.0051
Coefficiente di variazione	0.0359
Massimo	2.2148
Minimo	1.8197
Intervallo	0.3951
Assimetria	0.1040
Curtosi	-0.3191
25% Percentile	1.9435
75% Percentile	2.0486
Precisione d'errore percentuale a 95% di confidenza	0.2224%

Statistiche	Risultato
Numero di prove	1000
Media	2.0020
Mediana	1.9992
Deviazione standard	0.1886
Varianza	0.0356
Coefficiente di variazione	0.0942
Massimo	2.4147
Minimo	1.6278
Intervallo	0.7869
Assimetria	0.0788
Curtosi	-0.9641
25% Percentile	1.8468
75% Percentile	2.1475
Precisione d'errore percentuale a 95% di confidenza	0.5839%

Statistiche	Risultato
Numero di prove	1000
Media	2.0036
Mediana	1.9995
Deviazione standard	0.1450
Varianza	0.0210
Coefficiente di variazione	0.0724
Massimo	2.3907
Minimo	1.6844
Intervallo	0.7063
Assimetria	0.0304
Curtosi	-0.7316
25% Percentile	1.8944
75% Percentile	2.1127
Precisione d'errore percentuale a 95% di confidenza	0.4486%

Figura 2.15 – Risultati della correlazione

La Figura 2.16 illustra i risultati dopo l'esecuzione di una simulazione, l'estrazione dei dati grezzi delle ipotesi ed il calcolo delle correlazioni tra le variabili. La Figura mostra che le ipotesi d'input sono recuperate nella simulazione. Vale a dire, voi inserite le correlazioni +0.8 and –0.8 e i valori simulati risultanti hanno le stesse correlazioni.

Prezzo Correlazione positiva	***Quantità Correlazione positiva***			***Prezzo Correlazione negativa***	***Quantità Correlazione negativa***		
1.95	*0.91*			*1.89*	*1.06*		
1.92	*0.95*			*1.98*	*1.05*		
2.02	*1.04*	*Correlazione di Pearson:*		*1.89*	*1.09*	*Correlazione di Pearson:*	
2.04	*1.03*			*1.88*	*1.04*		
1.89	*0.91*		***0.80***	*1.96*	*0.93*		***-0.80***
1.98	*1.05*			*2.02*	*0.93*		
2.05	*1.03*			*2.00*	*1.02*		
1.87	*0.91*			*1.86*	*1.04*		
1.84	*0.91*			*1.96*	*1.02*		

Figura 2.16 –Correlazioni recuperate

2.3.4 Controllo precisione ed errore

Uno strumento molto potente nella Simulazione Monte Carlo è il controllo precisione. Per esempio, quante prove sono ritenute sufficienti da eseguire in un modello complesso? Il Controllo precisione elimina le congetture nello stimare il numero pertinente di prove, permettendo alla simulazione di fermarsi se viene raggiunto il livello prespecificato di precisione.

La funzionalità del Controllo precisione vi permette di impostare il vostro livello desiderato di precisione della previsione. Generalmente parlando, tante più prove sono calcolate, tanto più si restringe l'intervallo di confidenza e le statistiche diventano più accurate. La funzione di Controllo precisione in Simulatore di Rischio utilizza la caratteristica degli intervalli di confidenza per determinare quando viene raggiunta la precisione specificata di una statistica. Potete specificare lo specifico intervallo di confidenza per il livello di precisione per ciascuna previsione.

Assicuratevi di non confondere tre termini molto diversi tra loro: errore, precisione e confidenza. Anche se sembrano simili, i concetti sono significativamente differenti tra loro. Una semplice illustrazione è appropriata. Supponiamo che siate un produttore di gusci per taco ("taco shell") e siate interessati a scoprire quanti gusci rotti ci sono in media in una scatola di 100 gusci. Un modo di fare ciò è di raccogliere un campione di scatole preconfezionate di 100 gusci, aprirli e contare quanti di questi gusci sono effettivamente rotti. Voi producete 1 milione di scatole il giorno (questa è la vostra *popolazione*), ma aprite a caso solo 10 scatole (questa è la dimensione del vostro campione, noto anche come il numero di *prove* in una simulazione). Il numero di gusci rotti in ciascuna scatola è come segue: 24, 22, 4, 15, 33, 32, 4, 1, 45 e 2. Il numero medio calcolato di gusci rotti è di 18.2. Basato su questi 10 campioni o prove, la media è di 18.2 unità; mentre basato sul campione, l'intervallo di confidenza del 80 percento è tra 2 e 33 unità (vale a dire, 80 percento delle volte, il numero di gusci rotti è tra 2 e 33 *basato su questa dimensione del campione o su questo numero di prove eseguite*). Quanto siete sicuri, però, che 18.2 è la media corretta? Bastano 10 prove per stabilire ciò? L'intervallo di confidenza tra 2 e 33 è troppo ampio e troppo variabile. Supponiamo che abbiate bisogno di un valore medio più preciso, dove l'errore è ±2 gusci il 90 percento delle volte – questo significa che se aprite *tutta* la produzione di 1 milione di scatole il giorno, 900000 di queste scatole avranno gusci rotti mediamente con un'unità media di ±2 gusci. Quante scatole di gusci dovreste ancora esaminare (o eseguire prove) per ottenere questo livello di precisione? In questo caso, i 2 gusci sono il livello d'errore, mentre il 90 percento è il livello di precisione. Se si eseguono un numero sufficiente di prove, allora l'intervallo di confidenza del 90 percento sarà identico al livello di precisione del 90 percento, dove si ottiene una misura più precisa per la media, così che il 90 percento delle volte l'errore, e quindi la confidenza, sarà di

±2 gusci. Ad esempio, se diciamo che la media è di 20 unità, allora l'intervallo di confidenza del 90 percento sarà tra 18 e 22 unità, dove quest'intervallo è preciso il 90 percento delle volte: se apriamo tutta la produzione di 1 milione di scatole, 900000 di queste scatole conteranno tra 18 e 22 gusci rotti. Il numero di prove richiesto per raggiungere questa precisione è basato sulla seguente equazione d'errore di campionamento $\bar{x} \pm Z\frac{s}{\sqrt{n}}$, dove $Z\frac{s}{\sqrt{n}}$ è l'errore di 2 gusci, $\bar{x}$ è la media del campione, Z è il punteggio Z standard normale ottenuto dal livello di precisione del 90 percento, s è la deviazione standard del campione e n è il numero di prove richieste per raggiungere questo livello d'errore con la precisione specificata. Le Figure 2.17 e 2.18 illustrano come si può eseguire il controllo precisione sulle multiple previsioni simulate in Simulatore di Rischio. Questa caratteristica impedisce che l'utente debba decidere sul numero di prove da eseguire in una simulazione ed elimina la necessità di congetture. La Figura 2.17 illustra il diagramma di previsione con impostato un livello di precisione del 95%. Questo valore può essere modificato e il cambiamento sarà riportato nella linguetta delle Statistiche, come mostrato nella Figura 2.18.

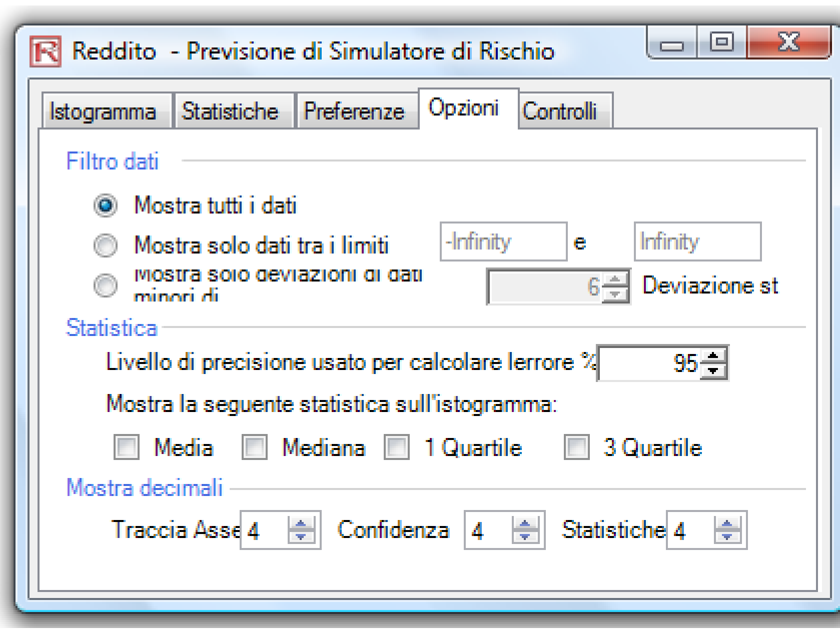
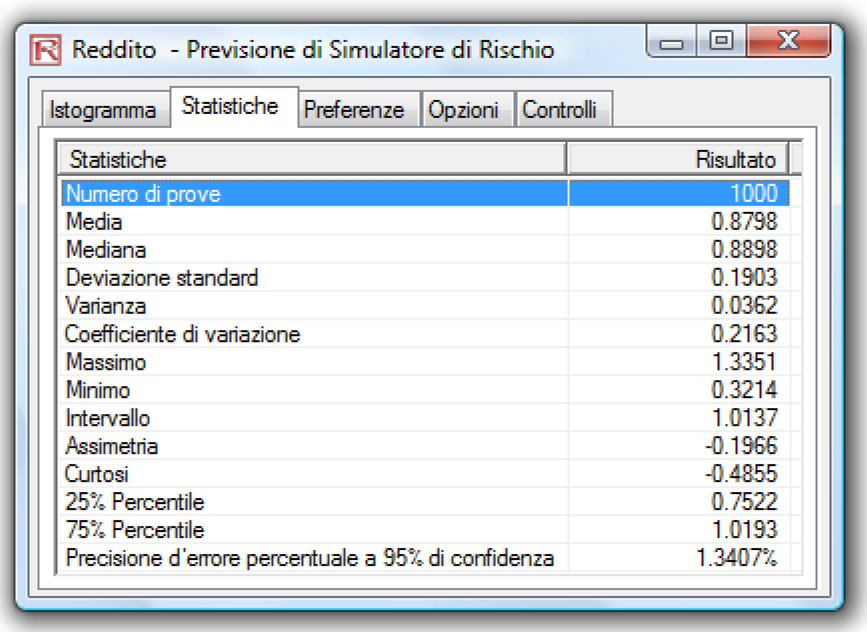


Figura 2.17 – Impostare il livello di precisione della previsione



Statistiche	Risultato
Numero di prove	1000
Media	0.8798
Mediana	0.8898
Deviazione standard	0.1903
Varianza	0.0362
Coefficiente di variazione	0.2163
Massimo	1.3351
Minimo	0.3214
Intervallo	1.0137
Assimetria	-0.1966
Curtosi	-0.4855
25% Percentile	0.7522
75% Percentile	1.0193
Precisione d'errore percentuale a 95% di confidenza	1.3407%

Figura 2.18 – Calcolare l'errore

2.3.5 Comprendere le Statistiche della previsione

La maggioranza delle distribuzioni può essere definita entro quattro momenti. Il primo momento descrive la sua posizione o tendenza centrale (ricavi attesi). Il secondo momento descrive la sua ampiezza o dispersione (rischi). Il terzo momento descrive la sua asimmetria direzionale (gli eventi più probabili). Infine il quarto momento descrive la sua "punta" o lo spessore delle code (perdite catastrofiche o ricavi). Tutti i quattro momenti dovrebbero essere calcolati nella pratica e poi interpretati per fornire una visione esaustiva del progetto sotto analisi. Il Simulatore di Rischio fornisce i risultati di tutti quattro momenti nella scheda delle *Statistiche* dei diagrammi di previsione.

Misurare il centro della distribuzione — il primo momento

Il primo momento di una distribuzione misura il tasso atteso di rendimento di un determinato progetto. Misura la posizione degli scenari del progetto e i possibili esiti nella media. Le statistiche comuni per il primo momento comprendono la media (valore medio), la mediana (centro di una distribuzione) e la moda (il valore che accade più comunemente). La Figura 2.19 illustra il primo momento — dove, in questo caso, il primo momento di questa distribuzione è misurata dalla media *(μ)* o valore medio.

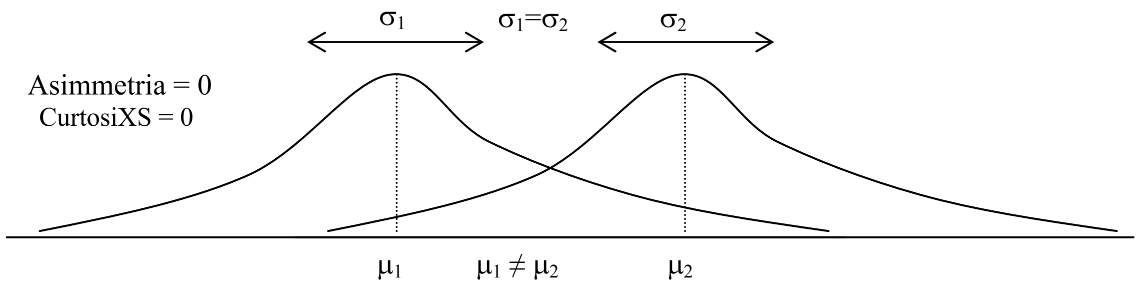


Figura 2.19 – Primo Momento

Misurare la dispersione (spread) della distribuzione — il secondo momento

Il secondo momento misura la dispersione di una distribuzione, che è una misura del rischio. La dispersione o ampiezza di una distribuzione misura la variabilità di una variabile, vale a dire, il potenziale che la variabile cadi in aree differenti della distribuzione — in altre parole, gli scenari potenziali degli esiti. La Figura 2.20 illustra due distribuzioni con primi momenti identici (medie identiche), ma con secondi momenti o rischi molto differenti. La visualizzazione diventa più chiara nella Figura 2.21. Ad esempio, supponiamo che ci siano due titoli e che i movimenti del primo titolo (illustrati dalla linea più scura) con una fluttuazione più piccola siano confrontati con i movimenti del secondo titolo (illustrati dalla linea punteggiata) con una fluttuazione di prezzo molto maggiore. Un investitore ovviamente valuterebbe il titolo con la fluttuazione più forte come più rischioso, perché gli esiti del titolo più rischioso sono relativamente meno noti che nel caso del titolo meno rischioso. L'asse verticale nella Figura 2.21 misura i prezzi del titolo; di conseguenza il titolo più rischioso ha un campo di variazione di potenziali esiti più ampio. Questo campo è riportato nell'ampiezza di una distribuzione (l'asse orizzontale) nella Figure 2.20, dove la distribuzione più ampia rappresenta l'asset più rischioso. L'ampiezza o la dispersione di una distribuzione misura perciò il rischio di una variabile.

Prego notare nella Figura 2.20 che, anche se entrambe le distribuzioni hanno primi momenti identici o tendenze centrali identiche, le distribuzioni sono chiaramente molto differenti. Questa differenza nell'ampiezza distribuzionale è misurabile. L'ampiezza o il rischio di una variabile può essere misurata(o) matematicamente e statisticamente mediante varie differenti statistiche, compreso il campo di variazione, la deviazione standard (σ), la varianza, il coefficiente di variazione e i percentili.

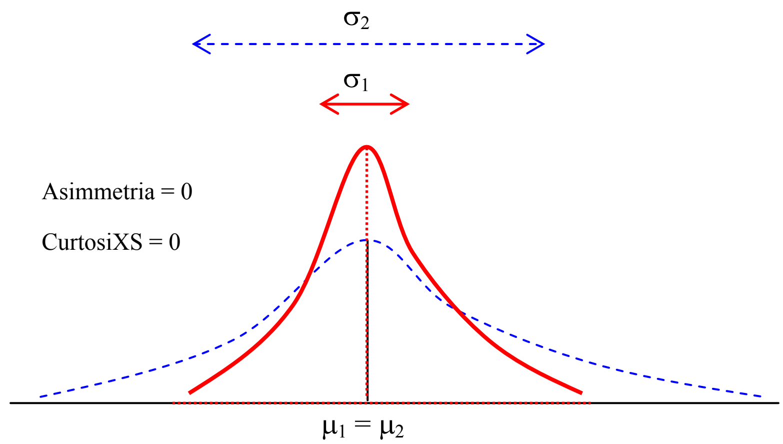


Figura 2.20 – Secondo Momento

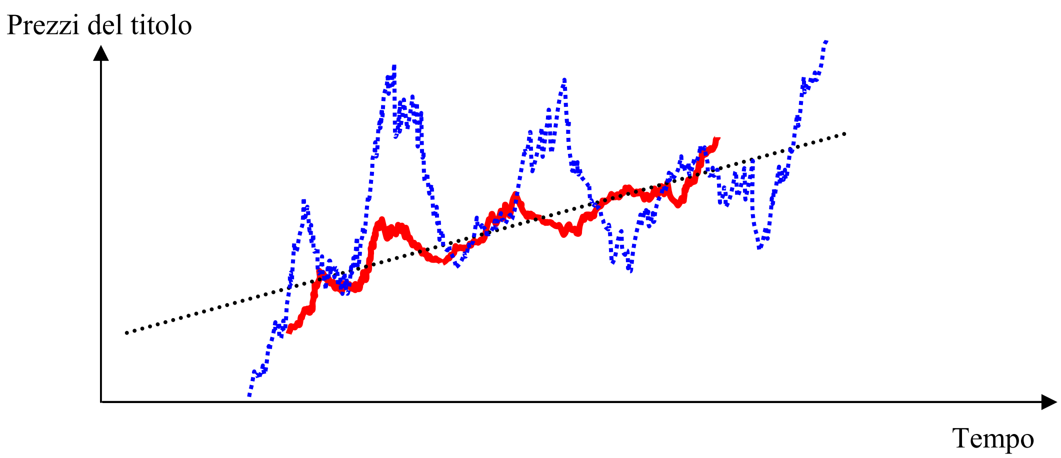


Figura 2.21 –Fluttuazioni nel prezzo del titolo

Misurare l'asimmetria di una distribuzione — il terzo momento

Il terzo momento misura l'asimmetria di una distribuzione, vale a dire, il modo in cui la distribuzione è tirata da una o dall'altra parte. La Figura 2.22 illustra un'asimmetria negativa o a sinistra (la coda della distribuzione punta verso sinistra) e la Figura 2.23 illustra un'asimmetria positiva o a destra (la coda della distribuzione punta verso destra). La media è sempre asimmetrica verso la coda della distribuzione, mentre la mediana rimane costante. Un altro punto di vista è che la media si muove, ma la deviazione standard , la varianza o l'ampiezza possono comunque rimanere costanti. Se il terzo momento non viene preso in considerazione e si guarda solo ai ricavi attesi (per esempio, mediana o media) e al rischio (deviazione standard), si potrebbe erroneamente scegliere un progetto con asimmetria positiva! Per esempio, se l'asse orizzontale rappresenta i ricavi netti di un progetto, allora è chiaro che una distribuzione con asimmetria negativa o a sinistra potrebbe essere preferibile, dato che c'è una probabilità più alta di rendimenti maggiori (Figura 2.22) paragonato ad una probabilità più alta di rendimenti a livelli inferiori (Figura 2.23). Pertanto, in una distribuzione asimmetrica, la mediana è una misura migliore del rendimento: dato che le mediane d'entrambe le Figure 2.22 e 2.23 sono identiche, i rischi sono identici e, quindi, un progetto con una distribuzione negativamente asimmetrica dei ricavi netti è una scelta migliore. La mancata considerazione dell'asimmetria distribuzionale di un progetto potrebbe significare la scelta di un progetto sbagliato (per esempio, due progetti potrebbero avere identici primi e secondi momenti, cioè identici rendimenti e profili di rischio, ma le loro asimmetrie distribuzionali potrebbero essere molto differenti).

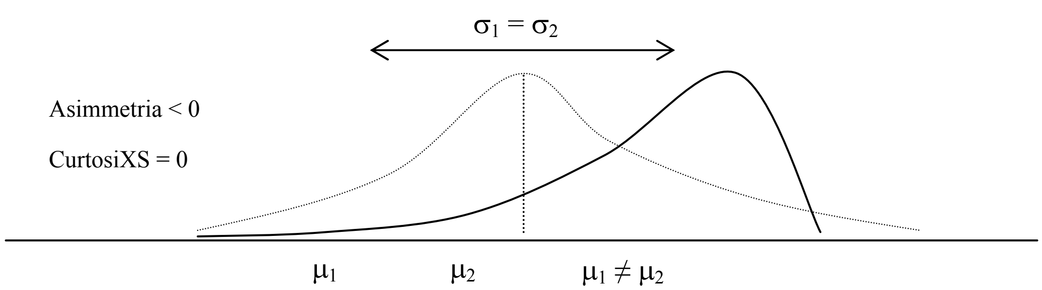


Figura 2.22 – Terzo Momento (Asimmetria a sinistra)

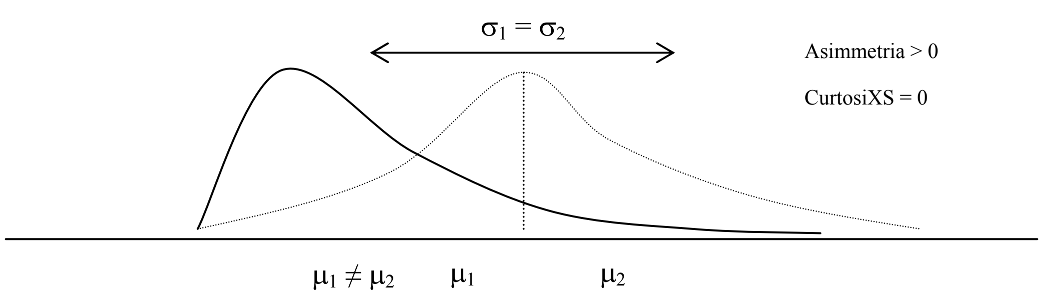


Figura 2.23 – Terzo Momento (Asimmetria a destra)

Misurare gli eventi di coda catastrofici in una distribuzione — il quarto momento

Il quarto momento o curtosi misura la "punta" di una distribuzione. La Figura 2.24 illustra quest'effetto. Il contesto (rappresentato dalla linea punteggiata) è una distribuzione normale con una curtosi di 3.0, o una curtosi in eccesso (CurtosiXS) di 0.0. I risultati di Simulatore di Rischio mostrano il valore di CurtosiXS, usando 0 come il livello normale di curtosi. Questo significa che una CurtosiXS negativa indica code più piatte (distribuzioni platicurtiche come la distribuzione uniforme), mentre i valori positivi indicano code più spesse (distribuzioni leptocurtiche come la distribuzione T di Student o la distribuzione lognormale). La distribuzione raffigurata dalla linea in grassetto ha una curtosi in eccesso più alta e quindi l'area sotto la curva è più spessa alle code con meno area nel corpo centrale. Questa condizione ha un impatto rilevante sull'analisi del rischio dato che, come si vede per le due distribuzioni nella Figura 2.24, i primi tre momenti (media, deviazione standard e asimmetria) possono essere identici, ma il quarto momento (curtosi) è differente. Questa condizione significa che sebbene i rendimenti e i rischi siano identici, le probabilità che accadano eventi estremi e catastrofici (potenziali grosse perdite o grandi ricavi) sono più alte per una distribuzione con una curtosi alta (per esempio, i rendimenti del mercato azionario sono leptocurtici o hanno una curtosi alta). Ignorare la curtosi di un progetto potrebbe essere dannoso. Un valore più alto di curtosi in eccesso indica tipicamente che i rischi di svantaggio sono più alti (per esempio, il Valore a rischio di un progetto potrebbe essere significativo).

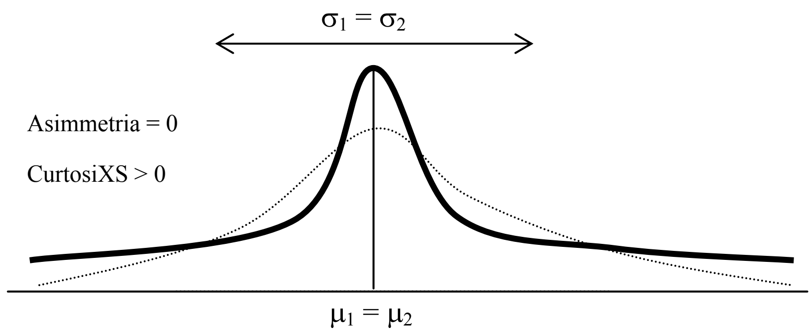


Figura 2.24 – Quarto Momento

Le funzioni dei momenti

Vi siete mai chiesti perché queste statistiche del rischio sono chiamate momenti? Nel gergo matematico, il momento significa elevato alla potenza di un qualche valore. In altre parole, il terzo momento implica che, in un'equazione, tre è più probabilmente la potenza più alta. Le equazioni sottostanti illustrano le funzioni e applicazioni matematiche di alcuni momenti di una statistica campione. Notate, per esempio, che la potenza più alta per la media del primo momento è uno, la deviazione standard del secondo momento è due, l'asimmetria del terzo momento è tre, e la potenza più alta per il quarto momento è quattro.

Primo Momento: Media Aritmetica o Media Semplice (Campione)

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$ La funzione equivalente di Excel è AVERAGE (Media)

Secondo Momento: Deviazione Standard (Campione)

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$ La funzione equivalente di Excel è STDEV (Deviazione Standard)

per una deviazione standard del campione

La funzione equivalente di Excel è STDEV (Deviazione Standard) per una deviazione standard della popolazione

Terzo Momento: Asimmetria

$$skew = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^3}{s}$$

La funzione equivalente di Excel è SKEW (Asimmetria)

Quarto Momento: Curtosi

$$kurtosis = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^4}{s} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

La funzione equivalente di Excel è KURT (Curtosi)

2.3.6 Comprendere le Distribuzioni di probabilità per una Simulazione Monte Carlo

Questa sezione dimostra la potenza della Simulazione Monte Carlo, ma per poter iniziare con la simulazione è prima necessario comprendere il concetto delle distribuzioni di probabilità. Per iniziare a comprendere la probabilità, considerate il seguente esempio. Volete esaminare la distribuzione dei salari non esenti all'interno di un reparto di una grande azienda. Prima raccogliete i dati grezzi —in questo caso i salari di tutti gli impiegati non esenti nel reparto. Poi organizzate i dati in un formato significativo e tracciate i dati nella forma di una distribuzione di frequenza su un diagramma. Per creare una distribuzione di frequenza, dividete i salari in intervalli di gruppo ed elencate questi intervalli sull'asse orizzontale del diagramma. Elencate poi il numero o la frequenza degli impiegati in ciascun intervallo sull'asse verticale del diagramma. Ora potete facilmente vedere la distribuzione dei salari non esenti all'interno del reparto.

Uno sguardo al diagramma illustrato nella Figura 2.25 rivela che la maggior parte degli impiegati (circo 60 su un totale di 180) guadagna tra $7.00 e $9.00 l'ora.

Potete tracciare questi dati come una distribuzione di probabilità. Una distribuzione di probabilità mostra il numero degli impiegati in ciascun intervallo come una frazione del numero totale degli impiegati. Per creare una distribuzione di probabilità, dividete il numero degli impiegati in ciascun intervallo per il numero totale degli impiegati ed elencate i risultati sull'asse verticale del diagramma.

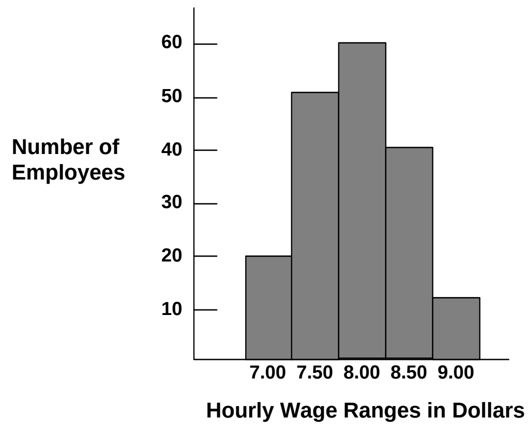


Figura 2.25 – Istogramma di frequenza I

Il diagramma nella Figura 2.26 mostra il numero degli impiegati in ciascun gruppo salariale come una frazione di tutti gli impiegati; potete stimare la verosimiglianza o la probabilità che un impiegato estratto a sorte dall'intero gruppo guadagni un salario all'interno di un dato intervallo. Per esempio, se si assume che esistano le stesse condizioni del momento di rilevazione del campione, esiste una probabilità dello 0.33 (una probabilità su tre) che un impiegato estratto a sorte dall'intero gruppo guadagni tra $8.00 e $8.50 l'ora.

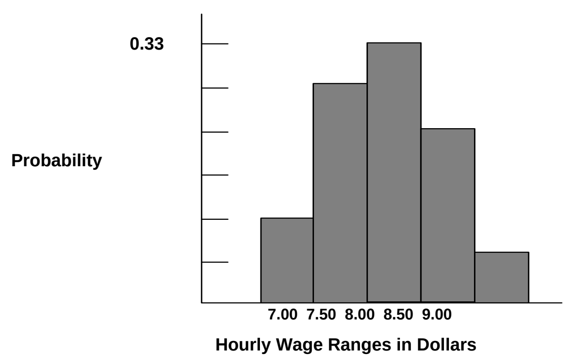


Figura 2.26 – Istogramma di frequenza II

Le distribuzioni di probabilità sono o discrete o continue. Le *distribuzioni discrete di probabilità* descrivono valori distinti, normalmente di numeri interi, senza valori intermedi e sono mostrate come una serie di barre verticali. Una distribuzione discreta potrebbe, per esempio, descrivere il numero di teste per quattro lanci di una moneta come 0, 1, 2, 3 or 4. Le *distribuzioni continue* sono in realtà astrazioni matematiche perché assumono l'esistenza di tutti i possibili valori intermedi tra due numeri. In altre parole, una distribuzione continua assume che ci sia un numero infinito di valori tra qualsiasi due punti di una distribuzione. Ciò nonostante, potete usare una distribuzione continua in molte situazioni per approssimare una distribuzione discreta, anche se il modello continuo non descrive per forza la situazione con esattezza.

Selezionare la giusta Distribuzione di Probabilità

Tracciare o diagrammare i dati è una guida nella selezione della distribuzione di probabilità. I seguenti passi forniscono un altro processo per la selezione delle distribuzioni di probabilità che descrivono al meglio le variabili incerte nei vostri fogli di lavoro.

Seguite le istruzioni sottostanti per selezionare la giusta distribuzione di probabilità:

- Esaminate la variabile in questione. Elencate tutto ciò che sapete riguardo alle condizioni che circondano questa variabile. Forse potrete raccogliere informazioni di valore sulla variabile incerta da dati storici. Se dati storici non sono disponibili, usate il vostro giudizio valutativo e, basato sulla vostra esperienza, elencate tutto ciò che sapete su questa variabile incerta.
- Analizzate le descrizioni delle distribuzioni di probabilità.
- Selezionate la distribuzione che caratterizza questa variabile. Una distribuzione caratterizza una variabile, quando le condizioni della distribuzione corrispondono a quelle della variabile.

Simulazione Monte Carlo

La Simulazione Monte Carlo nella sua forma più semplice è un generatore di numeri casuali utile per la previsione, la valutazione e l'analisi del rischio. Una simulazione calcola numerosi scenari di un modello mediante la ripetuta estrazione di valori per le variabili incerte da una *distribuzione di probabilità* predefinita dall'utente e l'uso di questi valori per il modello. Dato che tutti questi scenari producono risultati associati in un modello, ogni scenario può avere una *previsione*. Le previsioni sono eventi (normalmente con formule o funzioni) che voi definite come outputs importanti del modello. Questi sono normalmente eventi come i totali, il profitto netto o le spese lorde.

Per semplificare, pensate al metodo della Simulazione Monte Carlo come alla ripetuta estrazione con rimpiazzo di palline da golf da una grande cesta. La dimensione e la forma della cesta dipendono dall'***ipotesi*** distribuzionale (per esempio, una distribuzione normale con una media di 100 e una deviazione standard di 10, contro una distribuzione uniforme o una distribuzione triangolare), dove alcune ceste sono più profonde o più simmetriche di altre, permettendo che certe palline siano estratte più frequentemente di altre. Il numero di palline estratte ripetutamente dipende dal numero di ***prove*** simulate. Per un modello grande con multiple ipotesi collegate, immaginate il modello grande come una cesta molto grande che contiene molte ceste piccole. Ogni piccola cesta ha il suo insieme di palline da golf che vi rimbalzano dentro. Queste ceste piccole sono talvolta collegate tra loro (se esiste una ***correlazione*** tra le variabili) e le palline da golf rimbalzano in tandem, mentre altre rimbalzano indipendentemente una dall'altra. Le palline che sono estratte ogni volta da queste interazioni all'interno del modello (la grande cesta centrale) sono tabulate e registrate, fornendo il risultato di una ***previsione*** della simulazione.

Con la Simulazione Monte Carlo, il Simulatore di Rischio genera valori casuali totalmente indipendenti per la distribuzione di probabilità di ciascun'ipotesi. In altre parole, il valore casuale selezionato per una prova non ha nessun effetto sul successivo valore casuale generato. Usate la campionatura Monte Carlo quando volere simulare scenari reali del tipo "cosa succederebbe se" per il vostro modello di foglio di lavoro.

2.4 Distribuzioni Discrete

Segue una lista dettagliata di ciascun tipo diverso di distribuzione di probabilità che può essere usata in una Simulazione Monte Carlo. Questa lista è compresa nell'Appendice per consultazione.

Distribuzione Bernoulli o Si/No

La distribuzione Bernoulli è una distribuzione discreta con due esiti (per esempio, testa o croce, successo o insuccesso, 0 o 1). La distribuzione Bernoulli è la distribuzione binomiale con una prova e può essere usata per simulare condizioni di Si/No o Successo/Insuccesso. Questa distribuzione è la componente fondamentale di altre distribuzioni più complesse. Per esempio:

- Distribuzione binomiale: è una distribuzione Bernoulli con un numero più alto di *n* prove totali. Calcola la probabilità di *x* successesi all'interno questo numero totale di prove.
- Distribuzione geometrica: è una distribuzione Bernoulli con un numero più alto di prove totali. Calcola il numero d'insuccessi necessari prima che avvenga il primo successo.
- Distribuzione negativa binomiale: è una distribuzione Bernoulli con un numero più alto di prove totali. Calcola il numero d'insuccessi, prima che avvenga lo *X*-esimo successo.

I costrutti matematici della distribuzione Bernoulli sono come segue:

$$P(n) = \begin{cases} 1-p & \text{per } x = 0 \\ p & \text{per } x = 1 \end{cases}$$

o

$$P(n) = p^x (1-p)^{1-x}$$

$$\text{media} = p$$

$$\text{deviazione standard} = \sqrt{p(1-p)}$$

$$\text{asimmetria} = \frac{1-2p}{\sqrt{p(1-p)}}$$

$$\text{curtosi in eccesso} = \frac{6p^2 - 6p + 1}{p(1-p)}$$

La probabilità di successo (p) è il solo parametro distribuzionale. È anche importante notare che esiste solo una prova nella distribuzione Bernoulli e che il risultante valore simulato è o 0 o 1.

Requisiti d'input:

Probabilità di successo > 0 e < 1 (vale a dire, $0.0001 \leq p \leq 0.9999$)

Distribuzione Binomiale

La distribuzione binomiale descrive il numero di volte del verificarsi di un particolare evento in un numero fissato di prove, come il numero di teste in 10 lanci di una moneta o il numero d'oggetti difettosi tra 50 oggetti scelti.

Condizioni

Le tre condizioni sottostanti della distribuzione binomiale sono:

- Per ciascuna prova sono possibili solo due esiti che si escludono a vicenda.
- Le prove sono indipendenti - quello che succede nella prima prova non influisce sulla prova successiva.
- La probabilità che si verifichi un evento rimane la stessa di prova in prova.

I costrutti matematici della distribuzione binomiale sono come segue:

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)} \quad \text{per } n > 0; x = 0, 1, 2, \ldots n; \text{e } 0 < p < 1$$

$$\text{media} = np$$

$$\text{deviazione standard} = \sqrt{np(1-p)}$$

$$\text{asimmetria} = \frac{1-2p}{\sqrt{np(1-p)}}$$

$$\text{curtosi in eccesso} = \frac{6p^2 - 6p + 1}{np(1-p)}$$

La probabilità di successo (p) e il numero intero di prove totali (n) sono i parametri distribuzionali. Il numero di prove di successo è denominato x. È importante notare che le probabilità di successo (p) di 0 o 1 sono condizioni insignificanti e non richiedono una simulazione, quindi, non sono permesse nel software.

Requisiti d'input:

Probabilità di successo > 0 e < 1 (vale a dire, $0.0001 \leq p \leq 0.9999$)

Numero di prove ≥ 1 o numeri interi positivi e ≤ 1000 (per prove più grandi usate la distribuzione normale con le pertinenti calcolate media e deviazione standard binomiali come i parametri della distribuzione normale).

Distribuzione Discreta Uniforme

La distribuzione discreta uniforme è conosciuta anche come la distribuzione degli *esiti ugualmente possibili*, dove se la distribuzione ha un insieme di *N* elementi, allora ogni elemento ha la stessa probabilità. Questa distribuzione è collegata alla distribuzione uniforme, ma i suoi elementi sono discreti e non continui.

I costrutti matematici della distribuzione discreta uniforme sono come segue:

$$P(x) = \frac{1}{N}$$

media = $\frac{N+1}{2}$ valore di classifica

deviazione standard = $\sqrt{\frac{(N-1)(N+1)}{12}}$ valore di classifica

asimmetria = 0 (vale a dire, la distribuzione è perfettamente simmetrica)

curtosi in eccesso = $\frac{-6(N^2+1)}{5(N-1)(N+1)}$ valore di classifica

Requisiti d'input:

Minimo < Massimo ed entrambi devono essere numeri interi (numeri interi negativi e zero sono permessi)

Distribuzione Geometrica

La distribuzione geometrica descrive il numero di prove fino al primo avvenimento di successo, come il numero di volte che uno deve girare la ruota di una roulette prima di vincere.

Condizioni

Le tre condizioni sottostanti della distribuzione geometrica sono:

- Il numero di prove non è fisso.
- Le prove continuano fino al primo successo.
- La probabilità di successo è la stessa da prova in prova.

I costrutti matematici della distribuzione geometrica sono come segue:

$$P(x) = p(1-p)^{x-1} \quad \text{per } 0 < p < 1 \text{ e } x = 1, 2, ..., n$$

$$\text{media} = \frac{1}{p} - 1$$

$$\text{deviazione standard} = \sqrt{\frac{1-p}{p^2}}$$

$$\text{asimmetria} = \frac{2-p}{\sqrt{1-p}}$$

$$\text{curtosi in eccesso} = \frac{p^2 - 6p + 6}{1-p}$$

La probabilità di successo (p) è il solo parametro distribuzionale. Il numero di prove di successo simulate è denominato x, che deve essere un numero intero positivo.

Requisiti d'input:

La probabilità di successo > 0 e < 1 (vale a dire, $0.0001 \leq p \leq 0.9999$). È importante notare che le probabilità di successo (p) di 0 o 1 sono condizioni insignificanti e non richiedono una simulazione, quindi, non sono permesse nel software.

Distribuzione Ipergeometrica

The distribuzione ipergeometrica è simile alla distribuzione binomiale in quanto entrambe descrivono il numero di volte che accade un determinato evento in un numero fissato di prove. La differenza è che le prove della distribuzione binomiale sono indipendenti, mentre le prove della distribuzione ipergeometrica cambiano la probabilità per ciascuna prova successiva e sono chiamate "prove senza rimpiazzo". Per esempio, supponiamo che si sappia che una scatola di componenti fabbricati contiene alcuni componenti difettosi. Voi scegliete un componente dalla scatola, lo trovate difettoso e lo togliete dalla scatola. Se scegliete un altro componente dalla scatola, la probabilità che sia difettoso è un po' più piccola che per il primo componente perché avete già rimosso un componente difettoso. Se aveste riposto il componente difettoso nella scatola, le probabilità sarebbero rimaste le stesse ed il processo avrebbe soddisfatto le condizioni per una distribuzione binomiale.

Condizioni

Le tre condizioni sottostanti della distribuzione ipergeometrica sono:

- Il numero totale di componenti o elementi (la dimensione della popolazione) è un numero fisso, una popolazione finita, e la dimensione della popolazione deve essere minore di o uguale a 1750.
- La dimensione del campione (il numero di prove) rappresenta una parte della popolazione.

- La probabilità iniziale nota di successo nella popolazione cambia dopo ogni prova.

I costrutti matematici della distribuzione ipergeometrica sono come segue:

$$P(x) = \frac{\dfrac{(N_x)!}{x!(N_x - x)!}\dfrac{(N - N_x)!}{(n - x)!(N - N_x - n + x)!}}{\dfrac{N!}{n!(N - n)!}} \quad \text{per } x = Max(n - (N - N_x), 0), \ldots, Min(n, N_x)$$

media = $\dfrac{N_x n}{N}$

deviazione standard = $\sqrt{\dfrac{(N - N_x)N_x n(N - n)}{N^2(N - 1)}}$

asimmetria = $\sqrt{\dfrac{N - 1}{(N - N_x)N_x n(N - n)}}$

curtosi in eccesso = funzione complessa

Il numero d'elementi nella popolazione o la Dimensione della popolazione (*N*), le prove campionate o la Dimensione del campione (*n*) e il numero d'elementi nella popolazione che hanno la caratteristica di successo o Successi della popolazione (N_x) sono i parametri distribuzionali. Il numero di prove di successo è denominato *x*.

Requisiti d'input:

Dimensione della popolazione ≥ 2 e numero intero

Dimensione del campione > 0 e numero intero

Successi della popolazione > 0 e numero intero

Dimensione della popolazione > Successi della popolazione

Dimensione del campione < Successi della popolazione

Dimensione della popolazione < 1750

Distribuzione Negativa Binomiale

La distribuzione negativa binomiale è utile per modellare la distribuzione del numero di prove addizionali richieste sopra il numero d'accadimenti di successo richiesti (*R*). Per esempio, per poter portare a termine un totale di dieci opportunità di vendita, quante chiamate di vendita addizionali devono essere fatte oltre le dieci chiamate, data una determinata probabilità di successo per ogni chiamata? L'asse x mostra il numero di chiamate addizionali richieste o il numero di chiamate senza successo. Il numero di prove non è fisso, le prove continuano fino al *R-es*imo successo e la probabilità di successo rimane la stessa di prova in prova. La probabilità di successo (*p*) e il numero di successi richiesti (*R*) sono i parametri distribuzionali. Si tratta essenzialmente di una *super distribuzione* delle distribuzioni geometriche e binomiali. Questa distribuzione mostra le probabilità di ciascun numero di prove sopra *R* per produrre il successo richiesto *R*.

Condizioni

Le tre condizioni sottostanti della distribuzione negativa binomiale sono:

- Il numero di prove non è fisso.
- Le prove continuano fino al *R-es*imo successo.
- La probabilità di successo rimane la stessa di prova in prova.

I costrutti matematici della distribuzione negativa binomiale sono come segue:

$$P(x) = \frac{(x+r-1)!}{(r-1)!x!} p^r (1-p)^x \quad \text{for } x = r, \text{r} + 1, ...; \text{and } 0 < p < 1$$

$$\text{media} = \frac{r(1-p)}{p}$$

$$\text{deviazione standard} = \sqrt{\frac{r(1-p)}{p^2}}$$

$$\text{asimmetria} = \frac{2-p}{\sqrt{r(1-p)}}$$

$$\text{curtosi in eccesso} = \frac{p^2 - 6p + 6}{r(1-p)}$$

La probabilità di successo (*p*) e i successi richiesti (*R*) sono i parametri distribuzionali.

Requisiti d'input:

Successi richiesti devono essere numeri interi > 0 e < 8000.

Probabilità di successo > 0 e < 1 (vale a dire, 0.0001 ≤ *p* ≤ 0.9999). È importante notare che le probabilità di successo (*p*) di 0 o 1 sono condizioni insignificanti e non richiedono una simulazione, quindi, non sono permesse nel software.

Distribuzione Pascal

La distribuzione Pascal è utile per modellare la distribuzione del numero di prove totali necessarie per ottenere il numero richiesto di accadimenti con esito favorevole. Per esempio, per portare a buon fine 10 opportunità di vendita, quale è il numero totale di chiamate di vendita che bisognerebbe fare, data una determinata probabilità di successo per ciascuna chiamata? L'asse x mostra il numero totale di chiamate necessarie, che includono sia quelle con esito favorevole che quelle fallite. Il numero di prove non è fissato, le prove continuano fino al Resimo successo e le probabilità di successo sono le stesse da prova a prova. La distribuzione Pascal è imparentata con la distribuzione negativa binomiale. La distribuzione negativa binomiale calcola il numero di eventi necessari al di sopra del numero di successi richiesti, data una determinata probabilità (in altre parole, il numero totale di fallimenti), mentre la distribuzione Pascal calcola il numero totale di eventi necessari (in altre parole, la somma dei fallimenti e dei successi) per ottenere i successi richiesti, data una determinata probabilità. Successi richiesti e probabilità sono i parametri distribuzionali.

I costrutti matematici della distribuzione Pascal sono come segue:

$$f(x) = \begin{cases} \dfrac{(x-1)!}{(x-s)!(s-1)!} p^S (1-p)^{X-S} & \text{for all } x \geq s \\ 0 \text{ otherwise} \end{cases}$$

$$F(x) = \begin{cases} \displaystyle\sum_{x=1}^{k} \dfrac{(x-1)!}{(x-s)!(s-1)!} p^S (1-p)^{X-S} & \text{for all } x \ \geq \ s \\ 0 \text{ otherwise} \end{cases}$$

$$\text{media} = \frac{s}{p}$$

$$\text{deviazione standard } = \sqrt{s(1-p)p^2}$$

$$\text{asimmetria } = \frac{2-p}{\sqrt{r(1-p)}}$$

$$\text{curtosi in eccesso } = \frac{p^2 - 6p + 6}{r(1-p)}$$

Requisiti di Input:

Successi richiesti > 0 ed è un numero intero

0 ≤ Probabilità ≤ 1

Distribuzione Poisson

La distribuzione Poisson descrive il numero di volte che si verifica un evento in un determinato intervallo, come il numero di chiamate telefoniche per minuto o il numero d'errori per pagina in un documento.

Condizioni

Le tre condizioni sottostanti della distribuzione Poisson sono:

- Il numero di possibili accadimenti in qualsiasi intervallo è illimitato.
- Gli accadimenti sono indipendenti. Il numero d'accadimenti in un intervallo non influisce sul numero d'accadimenti in altri intervalli.
- Il numero medio d'accadimenti deve rimanere lo stesso d'intervallo in intervallo.

I costrutti matematici della distribuzione Poisson sono come segue:

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad \text{per } x \text{ e } \lambda > 0$$

$$\text{media} = \lambda$$

deviazione standard = $\sqrt{\lambda}$

asimmetria = $\dfrac{1}{\sqrt{\lambda}}$

curtosi in eccesso = $\dfrac{1}{\lambda}$

Tasso o Lambda (λ) è il solo parametro distribuzionale.

Requisiti d'input:

Tasso > 0 e ≤ 1000 (vale a dire, 0.0001 ≤ tasso ≤ 1000)

2.5 Distribuzioni Continue

Distribuzione Arcoseno

La distribuzione arcoseno è a forma di U ed è un caso speciale della distribuzione Beta, quando sia la forma che la scala sono uguali a 0.5. Valori vicini al minimo ed al massimo hanno alte probabilità del verificarsi, mentre valori tra questi due estremi hanno probabilità del verificarsi molto piccole. Minimo e massimo sono i parametri distribuzionali.

I costrutti matematici della distribuzione Arcoseno sono come segue:

$$f(x) = \begin{cases} \dfrac{1}{\pi\sqrt{x(1-x)}} & \textit{for } 0 \le x \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{2}{\pi}\sin^{-1}(\sqrt{x}) & \textit{for } 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

$$\text{media} = \frac{Min + Max}{2}$$

$$\text{deviazione standard} = \sqrt{\frac{(Max - Min)^2}{8}}$$

asimmetria = 0

curtosi in eccesso = 1.5

Requisiti di Input:

Minimo < Massimo

Distribuzione Beta

La distribuzione beta è molto flessibile ed è comunemente usata per rappresentare la variabilità nel corso di un campo. Una delle applicazioni più importanti della distribuzione beta è il suo uso come distribuzione coniugata per il parametro di una distribuzione Bernoulli. In quest'applicazione, la distribuzione beta è usata per rappresentare l'incertezza nella probabilità dell'accadimento di un evento. È anche usata per dati empirici e predire il comportamento casuale di percentuali e frazioni, dato che l'intervallo degli esiti è normalmente tra 0 e 1.

Il valore della distribuzione beta sta nella grande varietà di forme che prende, quando si variano i due parametri, alfa e beta. Se i parametri sono uguali, la distribuzione è simmetrica. Se uno dei parametri è 1, mentre l'altro parametro è maggiore di 1, la distribuzione è con forma a J. Se alfa è minore di beta, la distribuzione è detta essere positivamente asimmetrica (la maggior parte dei valori è vicina al valore minimo). Se alfa è maggiore di beta, la distribuzione è negativamente asimmetrica (la maggior parte dei valori è vicina al valore massimo).

I costrutti matematici della distribuzione beta sono come segue:

$$f(x) = \frac{(x)^{(\alpha-1)}(1-x)^{(\beta-1)}}{\left[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right]} \quad \text{per } \alpha > 0; \beta > 0; x > 0$$

$$\text{media} = \frac{\alpha}{\alpha+\beta}$$

$$\text{deviazione standard} = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}}$$

$$\text{asimmetria} = \frac{2(\beta-\alpha)\sqrt{1+\alpha+\beta}}{(2+\alpha+\beta)\sqrt{\alpha\beta}}$$

$$\text{curtosi in eccesso} = \frac{3(\alpha+\beta+1)[\alpha\beta(\alpha+\beta-6)+2(\alpha+\beta)^2]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)} - 3$$

Alfa (α) e beta (β) sono i due parametri di forma distribuzionali e Γ è la funzione Gamma.

Condizioni

Le due condizioni sottostanti della distribuzione beta sono:

- La variabile incerta è un valore casuale tra 0 e un valore positivo.
- La forma della distribuzione può essere specificata usando due valori positivi.

Requisiti d'input:

Alfa e beta entrambi > 0 e possono essere qualsiasi valore positivo

Distribuzione Moltiplicativa Beta Spostato

La distribuzione beta è molto flessibile ed è comunemente usata per rappresentare la variabilità nel corso di un determinato intervallo. E' usata per descrivere dati empirici e predire il comportamento casuale di percentuali e frazioni, dato che l'intervallo degli esiti è tipicamente tra 0 e 1. Il valore della distribuzione beta consiste nell'ampia varietà di forme che può assumere quando si variano i due parametri, alfa and beta. La distribuzione beta spostato è tipicamente ottenuta moltiplicando la distribuzione beta per un fattore e spostando i risultati di un certo parametro di posizione, per permettere all'intervallo degli esisti di espandersi oltre i suoi limiti naturali di 0 e 1, con un punto iniziale diverso da 0. Alfa, Beta, Posizione e Fattore sono i parametri di input.

Requisiti di Input:

Alfa > 0

Beta > 0

Posizione può essere qualsiasi valore positivo o negativo incluso lo zero

Fattore > 0

Distribuzione Cauchy o Lorentziana o Breit-Wigner

La distribuzione Cauchy, denominata anche la distribuzione Lorentziana o la distribuzione Breit-Wigner, è una distribuzione continua che descrive il comportamento di risonanza. Descrive anche la distribuzione di distanze orizzontali alle quali un segmento di linea inclinato in un angolo casuale taglia l'asse x.

I costrutti matematici della distribuzione Cauchy o Lorentziana sono come segue:

$$f(x) = \frac{1}{\pi} \frac{\gamma / 2}{(x-m)^2 + \gamma^2 / 4}$$

La distribuzione Cauchy è un caso speciale: non ha nessun momento teorico (media, deviazione standard, asimmetria e curtosi), dato che sono tutti non definiti.

Posizione modale (α) e scala (β) sono i soli due parametri in questa distribuzione. Il parametro di posizione specifica la punta o moda della distribuzione, mentre il parametro di scala specifica la semi-ampiezza al semi-massimo della distribuzione. Oltre a ciò, la media e la varianza di una distribuzione Cauchy o Lorentziana sono non definiti.

In aggiunta, la distribuzione Cauchy è la distribuzione T di Student con solo 1 grado di libertà. Questa distribuzione può anche essere costruita prendendo il rapporto ratio di due distribuzioni standard normali (distribuzioni normali con una media di zero e una varianza di uno) che sono indipendenti tra loro.

Requisiti d'input:

Posizione Alfa può essere qualsiasi valore

Scala Beta > 0 e può essere qualsiasi valore positivo

Distribuzione Chi-Quadrato

La distribuzione Chi-Quadrato è una distribuzione di probabilità usata principalmente nel test di verifica d'ipotesi ed è collegata alla distribuzione gamma e alla distribuzione standard normale. Per esempio, le somme di distribuzioni normali indipendenti sono distribuite come un Chi-Quadrato (χ^2) con *k* gradi di libertà:

$$Z_1^2 + Z_2^2 + \ldots + Z_k^2 \overset{d}{\sim} \chi_k^2$$

I costrutti matematici della distribuzione Chi-Quadrato sono come segue:

$$f(x) = \frac{0.5^{-k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2} \text{ per tutti } x > 0$$

media = k

deviazione standard = $\sqrt{2k}$

asimmetria = $2\sqrt{\frac{2}{k}}$

curtosi in eccesso = $\frac{12}{k}$

Γ è la funzione gamma. Gradi di libertà *k* è il solo parametro distribuzionale.

La distribuzione Chi-Quadrato può anche essere modellata usando la distribuzione gamma con le seguenti impostazioni:

parametro di forma = $\frac{k}{2}$ e scala = $2S^2$ dove S è la scala.

Requisiti d'input:

Gradi di libertà > 1 e deve essere un numero intero < 300

Distribuzione Coseno

La distribuzione coseno ha le sembianze di una distribuzione logistica, dove il valore mediano tra il minimo ed il massimo ha il picco o valore modale più alto, comportando la massima probabilità del verificarsi, mentre le code estreme vicine ai valori minimi e massimi hanno probabilità più basse. Minimo e massimo sono i parametri distribuzionali.

I costrutti matematici della distribuzione Cosine sono come segue:

$$f(x) = \begin{cases} \frac{1}{2b}\cos\left[\frac{x-a}{b}\right] & for \min \le x \le \max \\ 0 & otherwise \end{cases}$$

$$where\ a = \frac{\min + \max}{2}\ and\ b = \frac{\max - \min}{\pi}$$

$$F(x) = \begin{cases} \frac{1}{2}\left[1 + \sin\left(\frac{x-a}{b}\right)\right] & for \min \le x \le \max \\ 1 & for\ x > max \end{cases}$$

$$\text{media} = \frac{Min + Max}{2}$$

$$\text{deviazione standard} = \sqrt{\frac{(Max - Min)^2(\pi^2 - 8)}{4\pi^2}}$$

asimmetria = 0

$$\text{curtosi in eccesso} = \frac{6(90 - \pi^4)}{5(\pi^2 - 6)^2}$$

Requisiti di Input:

Minimo < Massimo

Distribuzione Erlang

La distribuzione Erlang è uguale alla distribuzione Gamma, con il requisito che Alfa, o il parametro di forma, sia un numero intero positivo. Un esempio di applicazione della distribuzione Erlang è la calibratura del tasso di transizione di elementi attraverso un sistema di compartimenti. Tali sistemi trovano largo impiego in biologia ed ecologia (p.e., in epidemiologia, una persona potrebbe progredire ad un tasso esponenziale dal essere sano al diventare un portatore di malattia, e continuare esponenzialmente dal essere portatore a essere contagioso). Alfa (noto anche come forma) e Beta (noto anche come scala) sono i parametri distribuzionali.

I costrutti matematici della distribuzione Erlang sono come segue:

$$f(x) = \begin{cases} \dfrac{\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}}{\beta(\alpha-1)} & for\ x \ge 0 \\ 0 & otherwise \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-x/\beta}\sum_{i=0}^{\alpha-1}\frac{(x/\beta)^i}{i!} & for\ x \ge 0 \\ 0 & otherwise \end{cases}$$

media = $\alpha\beta$

deviazione standard = $\sqrt{\alpha\beta^2}$

asimmetria = $\frac{2}{\sqrt{\alpha}}$

curtosi in eccesso = $\frac{6}{\alpha} - 3$

Requisiti di Input:

Alfa (Forma) > 0 ed è un numero intero

Beta (Scala) > 0

Distribuzione Esponenziale

La distribuzione esponenziale è largamente usata per descrivere eventi che si verificano in punti casuali nel tempo, come il tempo tra eventi come la rottura di dispositivi elettronici o il tempo tra arrivi ad uno sportello d'assistenza. È collegata alla distribuzione Poisson che descrive il numero degli accadimenti di un evento durante un determinato intervallo di tempo. Una caratteristica importante della distribuzione esponenziale è la sua proprietà *senza memoria*, il che significa che la vita futura di un determinato oggetto ha la stessa distribuzione a prescindere dal tempo della sua esistenza.

I costrutti matematici della distribuzione esponenziale sono come segue:

$f(x) = \lambda e^{-\lambda x}$ per $x \geq 0; \lambda > 0$

media = $\frac{1}{\lambda}$

deviazione standard = $\frac{1}{\lambda}$

asimmetria = 2 (questo valore vale per tutti gli inputs del tasso di successo λ)

curtosi in eccesso = 6 (questo valore vale per tutti gli inputs del tasso di successo λ)

Il tasso di successo (λ) è il solo parametro distribuzionale. Il numero di prove di successo è denominato x.

Condizioni

La condizione sottostante della distribuzione esponenziale è:

- La distribuzione esponenziale descrive il periodo tra accadimenti.

Requisiti d'input: Tasso > 0

Distribuzione Esponenziale Spostato

La distribuzione esponenziale è largamente usata per descrivere eventi che si verificano in punti casuali nel tempo, come il tempo tra eventi come la rottura di dispositivi elettronici o il tempo tra arrivi ad uno sportello di assistenza. È collegata alla distribuzione Poisson che descrive il numero degli accadimenti di un evento durante un determinato intervallo di tempo. Una caratteristica importante della distribuzione esponenziale è la sua proprietà senza memoria, che significa che la vita futura di un determinato oggetto ha la stessa distribuzione a prescindere dal tempo della sua esistenza. In altre parole, il tempo non ha nessun effetto sugli esiti futuri. Il tasso di successo (λ) è il solo parametro distribuzionale.

Requisiti di Input:

Tasso Lambda > 0

Posizione può essere qualsiasi valore positivo o negativo incluso lo zero

Distribuzione del Valore Estremo o Distribuzione Gumbel

La distribuzione del valore estremo (Tipo 1) è comunemente usata per descrivere il valore più grande di una risposta nel corso di un periodo di tempo, per esempio, durante inondazioni, precipitazioni e terremoti. Altre applicazioni includono il carico di rottura di materiali, la progettazione edile e carichi e tolleranze d'aeromobili. La distribuzione del valore estremo è nota anche come la distribuzione Gumbel.

I costrutti matematici della distribuzione del valore estremo sono come segue:

$$f(x) = \frac{1}{\beta} z e^{-Z} \text{ dove } z = e^{\frac{x-\alpha}{\beta}} \quad \text{per } \beta > 0; \text{e qualsiasi valore di } x \text{ e } \alpha$$

media = $\alpha + 0.577215\beta$

deviazione standard = $\sqrt{\frac{1}{6}\pi^2\beta^2}$

asimmetria = $\frac{12\sqrt{6}(1.2020569)}{\pi^3} = 1.13955$ (questo vale per tutti i valori di moda e scala)

curtosi in eccesso = 5.4 (questo vale per tutti i valori di moda e scala)

Moda (α) e scala (β) sono i parametri distribuzionali.

Calcolare i Parametri

Ci sono due parametri standard per la distribuzione del valore estremo: moda e scala. Il parametro di moda è il valore più probabile della variabile (il punto più alto sulla distribuzione di probabilità). Dopo aver selezionato il parametro di moda, potete stimare il parametro di scala. Il parametro di scala è un numero maggiore di 0. Tanto più grande è il parametro di scala, tanto più grandi sono le varianze.

Requisiti d'input:

Moda Alfa può essere qualsiasi valore

Scala Beta > 0

Distribuzione F o Distribuzione Fisher-Snedecor

La distribuzione F, conosciuta anche come la distribuzione Fisher-Snedecor, è un'altra distribuzione continua usata principalmente per il test di verifica d'ipotesi. In modo specifico, è usata per testare la differenza statistica tra due varianze nei test d'analisi della varianza e nei test del rapporto di verosimiglianza. La distribuzione F, con il numeratore grado di libertà n e il denominatore grado di libertà m, è collegata alla distribuzione Chi-Quadrato in quanto:

$$\frac{\chi_n^2 / n}{\chi_m^2 / m} \overset{d}{\sim} F_{n,m}$$

media = $\dfrac{m}{m-2}$

deviazione standard = $\dfrac{2m^2(m+n-2)}{n(m-2)^2(m-4)}$ per tutti $m > 4$

asimmetria = $\dfrac{2(m+2n-2)}{m-6}\sqrt{\dfrac{2(m-4)}{n(m+n-2)}}$

curtosi in eccesso =

$$\frac{12(-16+20m-8m^2+m^3+44n-32mn+5m^2n-22n^2+5mn^2}{n(m-6)(m-8)(n+m-2)}$$

Il numeratore grado di libertà n e il denominatore grado di libertà m sono gli unici parametri distribuzionali.

Requisiti d'input:

Gradi di libertà numeratore e gradi di libertà denominatore entrambi numeri interi > 0

Distribuzione Gamma (Distribuzione Erlang)

La distribuzione gamma è applicabile ad un ampio campo di quantità fisiche ed è collegata ad altre distribuzioni: lognormale, esponenziale, Pascal, Erlang, Poisson e Chi-Quadrato. È usata nei processi meteorologici per rappresentare concentrazioni d'inquinanti e quantità di precipitazioni. La distribuzione gamma è anche usata per misurare il tempo tra gli accadimenti di eventi, quando il processo dell'evento non è completamente casuale. Altre applicazioni della distribuzione gamma includono il controllo dell'inventario, la teoria economica e la teoria del rischio assicurativo.

Condizioni

La distribuzione gamma è usata più spesso come la distribuzione del tempo fino al *r*-esimo accadimento di un evento in un processo Poisson. Quando si usa in questa maniera, le tre condizioni sottostanti della distribuzione gamma sono:

- Il numero di possibili accadimenti in una qualsiasi unità di misura non è limitato ad un numero fissato.
- Gli accadimenti sono indipendenti. Il numero d'accadimenti in un'unità di misura non influisce sul numero d'accadimenti in altre unità.
- Il numero medio d'accadimenti deve rimanere lo stesso d'unità in unità.

I costrutti matematici della distribuzione gamma sono come segue:

$$f(x) = \frac{\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta} \text{ con qualsiasi valore di } \alpha > 0 \text{ e } \beta > 0$$

media = $\alpha\beta$

deviazione standard = $\sqrt{\alpha\beta^2}$

asimmetria = $\frac{2}{\sqrt{\alpha}}$

curtosi in eccesso = $\frac{6}{\alpha}$

Il parametro di forma alfa (α) e il parametro di scala beta (β) sono i parametri distribuzionali e Γ è la funzione Gamma.

Quando il parametro alfa è un numero intero positivo, la distribuzione gamma è denominata la distribuzione Erlang, usata per predire i tempi d'attesa in un sistema d'accodamento, dove la distribuzione Erlang è la somma di variabili casuali indipendenti e analogamente distribuite, ciascuna con una distribuzione esponenziale senza memoria. Impostando *n* come il numero di queste variabili casuali, il costrutto matematico della distribuzione Erlang è:

$$f(x) = \frac{x^{n-1}e^{-x}}{(n-1)!} \text{ per tutti } x > 0 \text{ e tutti i numeri interi positivi di } n$$

Requisiti d'input:

Scala beta > 0 e può essere qualsiasi valore positivo

Forma alfa ≥ 0.05 e può essere qualsiasi valore positivo

Posizione può essere qualsiasi valore

Distribuzione Laplace

La distribuzione Laplace è talvolta anche chiamata la distribuzione esponenziale doppia, perché può essere costruita con due distribuzioni esponenziali (con un parametro di posizione addizionale) attaccate in sequenza, creando un picco inconsueto nel mezzo peak. La funzione di densità di probabilità della distribuzione Laplace richiama alla mente la distribuzione normale. Tuttavia, mentre la distribuzione normale viene espressa in termini di differenza quadrata dalla media, la densità Laplace è espressa in termini di differenza assoluta dalla media. Questo rende le code della distribuzione Laplace più ampie di quelle della distribuzione normale. Quando il parametro di posizione è impostato a zero, la variabile casuale della distribuzione Laplace è distribuita esponenzialmente con un inverso del parametro di scala. Alfa (noto anche come posizione) e Beta (noto anche come scala) sono i parametri distribuzionali.

I costrutti matematici della distribuzione Laplace sono come segue:

$$f(x) = \frac{1}{2\beta}\exp\left(-\frac{|x-\alpha|}{\beta}\right)$$

$$F(x) = \begin{cases} \frac{1}{2}\exp\left[\frac{x-\alpha}{\beta}\right] & \text{when } x < \alpha \\ 1-\frac{1}{2}\exp\left[-\frac{x-\alpha}{\beta}\right] & \text{when } x \geq \alpha \end{cases}$$

$$\text{media} = \alpha$$

$$\text{deviazione standard} = 1.4142\beta$$

asimmetria = 0

curtosi in eccesso = 3

Requisiti di Input:

Alfa (Posizione) può essere qualsiasi valore positivo o negativo incluso lo zero

Beta (Scala) > 0

Distribuzione Log Doppia

La distribuzione log doppia assomiglia alla distribuzione Cauchy, dove la tendenza centrale è "a picco" e ha la massima densità di probabilità, ma che diminuisce tanto più velocemente quanto più si allontana dal centro, creando una distribuzione simmetrica con un picco estremo tra i valori minimi e massimi. Minimo e massimo sono i parametri distribuzionali.

I costrutti matematici della distribuzione Log Doppia sono come segue:

$$f(x) = \begin{cases} \dfrac{-1}{2b} \ln\left(\dfrac{|x-a|}{b}\right) & \textit{for } \min \le x \le \max \\ 0 & \textit{otherwise} \end{cases}$$

$$\textit{where } a = \frac{\min + \max}{2} \textit{ and } b = \frac{\max - \min}{2}$$

$$F(x) = \begin{cases} \dfrac{1}{2} - \left(\dfrac{|x-a|}{2b}\right)\left[1 - \ln\left(\dfrac{|x-a|}{b}\right)\right] & \textit{for } \min \le x \le a \\ \dfrac{1}{2} + \left(\dfrac{|x-a|}{2b}\right)\left[1 - \ln\left(\dfrac{|x-a|}{b}\right)\right] & \textit{for } a \le x \le \max \end{cases}$$

media = $\dfrac{Min + Max}{2}$

deviazione standard = $\sqrt{\dfrac{(Max - Min)^2}{36}}$

asimmetria = 0

Requisiti di Input:

Minimo < Massimo

Distribuzione Logistica

La distribuzione logistica è comunemente usata per descrivere la crescita, e cioè la dimensione della popolazione espressa come una funzione di una variabile temporale. Può anche essere usata per descrivere reazioni chimiche e il corso della crescita di una popolazione o di un individuo.

I costrutti matematici della distribuzione logistica sono come segue:

$$f(x) = \frac{e^{\frac{\alpha - x}{\beta}}}{\beta\left[1 + e^{\frac{\alpha - x}{\beta}}\right]^2} \quad \text{per qualsiasi valore di } \alpha \text{ e } \mu$$

$$\text{media} = \alpha$$

$$\text{deviazione standard} = \sqrt{\frac{1}{3}\pi^2\beta^2}$$

asimmetria = 0 (questo vale per tutti gli inputs di media e scala)

curtosi in eccesso = 1.2 (questo vale per tutti gli inputs di media e scala)

Media (α) e scala (β) sono i parametri distribuzionali.

Calcolare i Parametri

Ci sono due parametri standard per la distribuzione logistica: media e scala. Il parametro di media è il valore medio, che per questa distribuzione è lo stesso della moda, dato che questa è una distribuzione simmetrica. Dopo aver selezionato il parametro di media, potete stimare il parametro di scala. Il parametro di scala è un numero maggiore di 0. Tanto più grande è il parametro di scala, tanto più grandi sono le varianze.

Requisiti d'input:

Scala Beta > 0 e può essere qualsiasi valore positivo

Media Alfa può essere qualsiasi valore

Distribuzione Lognormale

La distribuzione lognormale è largamente usata in situazioni dove i valori sono positivamente asimmetrici, per esempio, in analisi finanziarie per la valutazione di titoli o nel campo immobiliare per la valutazione di proprietà, e dove i valori non possono scendere sotto zero.

I prezzi di titoli sono di norma positivamente asimmetrici, piuttosto che normalmente (simmetricamente) distribuiti. I prezzi di titoli presentano questa tendenza perché non possono scendere sotto il limite inferiore di zero, ma possono salire di prezzo senza limiti. Analogamente, i prezzi immobiliari illustrano un'asimmetria positiva dato che i prezzi immobiliari non possono essere negativi.

Condizioni

Le tre condizioni sottostanti della distribuzione lognormale sono:

- La variabile incerta può salire senza limiti, ma non può scendere sotto zero.
- La variabile incerta è positivamente asimmetrica, con la maggior parte dei valori vicini al limite inferiore.
- Il logaritmo naturale della variabile incerta fornisce una distribuzione normale.

Generalmente, se il coefficiente di variabilità è maggiore del 30 percento, è appropriato usare una distribuzione lognormale. Altrimenti, usate la distribuzione normale.

I costrutti matematici della distribuzione lognormale sono come segue:

$$f(x) = \frac{1}{x\sqrt{2\pi}\ln(\sigma)} e^{-\frac{[\ln(x)-\ln(\mu)]^2}{2[\ln(\sigma)]^2}} \quad \text{per } x > 0; \mu > 0 \text{ e } \sigma > 0$$

$$\text{media} = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

deviazione standard = $\sqrt{\exp(\sigma^2 + 2\mu)[\exp(\sigma^2) - 1]}$

asimmetria = $\left\lfloor\sqrt{\exp(\sigma^2) - 1}\right\rfloor(2 + \exp(\sigma^2))$

curtosi in eccesso = $\exp(4\sigma^2) + 2\exp(3\sigma^2) + 3\exp(2\sigma^2) - 6$

Media (μ) e deviazione standard (σ) sono i parametri distribuzionali.

Requisiti d'input:

Media e deviazione standard entrambe > 0 e possono essere qualsiasi valore positivo

Lognormale - Insiemi di Parametri

Per default, la distribuzione lognormale usa la media e la deviazione standard aritmetica. Per applicazioni per le quali sono disponibili dati storici, è più appropriato usare o la media e la deviazione standard logaritmica o la media e la deviazione standard geometrica.

Distribuzione Lognormale Spostata

La distribuzione lognormale è ampiamente usata in situazioni dove i valori sono positivamente asimmetrici, per esempio, nell'analisi finanziaria per la valutazione delle garanzie o nel campo immobiliare per la valutazione di proprietà, dove i valori non possono scendere sotto lo zero. I prezzi dei titoli sono generalmente positivamente asimmetrici anziché essere normalmente (simmetricamente) distribuiti. I prezzi dei titoli esibiscono questo trend perché non possono scendere sotto il limite inferiore di zero, ma potrebbero aumentare fino a qualsiasi prezzo senza limiti. Per contrasto, la distribuzione lognormale spostata è uguale alla distribuzione lognormale, ma è spostata in modo che il risultante valore possa assumere valori negativi. Media (valore medio), Deviazione standard e Spostamento sono i parametri distribuzionali.

Requisiti di Input: Media > 0. Deviazione standard > 0

Spostamento può assumere qualsiasi valore positivo o negativo incluso lo zero

Distribuzione Normale

La distribuzione normale è la distribuzione più importante nella teoria della probabilità, perché descrive molti fenomeni naturali, come il Q.I. o l'altezza di persone. I decisori possono usare la distribuzione normale per descrivere variabili incerte come il tasso d'inflazione o il prezzo futuro della benzina.

Condizioni

Le tre condizioni sottostanti della distribuzione normale sono:

- Un qualche valore della variabile incerta è quello più probabile (la media della distribuzione).
- La variabile incerta potrebbe essere sia sopra che sotto la media (simmetrica attorno alla media).
- La variabile incerta sarà più probabilmente vicina che lontana dalla media.

I costrutti matematici della distribuzione normale sono come segue:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{per tutti i valori di } x \text{ e } \mu; \text{ mentre } \sigma > 0$$

$$\text{media} = \mu$$
$$\text{deviazione standard} = \sigma$$

asimmetria = 0 (questo vale per tutti gli inputs di media e deviazione standard)

curtosi in eccesso = 0 (questo vale per tutti gli inputs di media e deviazione standard)

Media (μ) e deviazione standard (σ) sono i parametri distribuzionali.

Requisiti d'input:

Deviazione standard > 0 e può essere qualsiasi valore positivo

Media può essere qualsiasi valore

Distribuzione Parabolica

La distribuzione parabolica è un caso speciale della distribuzione di beta quando la Forma = la Scala = 2. I valori chiudono al minimo e al massimo ha delle probabilità basse di evento poiché i valori tra questi due estremi hanno delle più alti probabilità o l'evento. Il minimo ed il massimo sono i parametri distribuzionali.

I costrutti matematici della distribuzione parabolica sono come segue:

$$f(x) = \frac{(x)^{(\alpha-1)}(1-x)^{(\beta-1)}}{\left[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right]} \quad \text{for } \alpha > 0; \beta > 0; x > 0$$

Where the functional form above is for a Beta distribution, and for a Parabolic function, we set Alpha = Beta = 2 and a shift of location in Minimum, with a multiplicative factor of (Maximum – Minimum).

media = $\frac{Min + Max}{2}$

deviazione standard = $\sqrt{\frac{(Max - Min)^2}{20}}$

asimmetria = 0

curtosi in eccesso = *–0.8571*

Requisiti di Input:

Minimo < Massimo

Distribuzione Pareto

La distribuzione Pareto è largamente usata per l'analisi di distribuzioni associate a fenomeni empirici come la dimensione delle popolazioni di città, la ricorrenza di risorse naturali, la dimensione di aziende, i redditi personali, le fluttuazioni del prezzo di un titolo ed il raggruppamento di errori in circuiti di comunicazione.

I costrutti matematici della distribuzione Pareto sono come segue:

$$f(x) = \frac{\beta L^{\beta}}{x^{(1+\beta)}} \quad \text{per } x > L$$

$$\text{media} = \frac{\beta L}{\beta - 1}$$

$$\text{deviazione standard} = \sqrt{\frac{\beta L^2}{(\beta - 1)^2 (\beta - 2)}}$$

$$\text{asimmetria} = \sqrt{\frac{\beta - 2}{\beta}} \left[\frac{2(\beta + 1)}{\beta - 3} \right]$$

$$\text{curtosi in eccesso} = \frac{6(\beta^3 + \beta^2 - 6\beta - 2)}{\beta(\beta - 3)(\beta - 4)}$$

Forma (α) e Posizione (β) sono i parametri distribuzionali.

Calcolare i Parametri

Ci sono due parametri standard per la distribuzione Pareto: posizione e forma. Il parametro di posizione è il limite inferiore della variabile. Dopo aver selezionato il parametro di posizione, potete stimare il parametro di forma. Il parametro di forma è un numero maggiore di 0, normalmente maggiore di 1. Tanto più grande è il parametro di forma, tanto più piccola è la varianza e tanto più spessa è la coda destra della distribuzione.

Requisisti d'input:

Posizione > 0 e può essere qualsiasi valore positivo

Forma ≥ 0.05

Distribuzione Pearson V

La distribuzione Pearson V è imparentata con la distribuzione gamma inversa. E' il reciproco della variabile distribuita secondo la distribuzione Gamma. La distribuzione Pearson V è anche usata per modellare i ritardi temporali, quando c'è una quasi certezza di un minimo di ritardo e il ritardo massimo è senza limiti (p.e., il ritardo nell'arrivo dei servizi di emergenza o il tempo richiesto per riparare un macchinario). Alfa (noto anche come forma) e Beta (noto anche come scala) sono i parametri distribuzionali.

I costrutti matematici della distribuzione Pearson V sono come segue:

$$f(x) = \frac{x^{-(\alpha+1)}e^{-\beta/x}}{\beta^{-\alpha}\Gamma(\alpha)}$$

$$F(x) = \frac{\Gamma(\alpha, \beta / x)}{\Gamma(\alpha)}$$

$$\text{media} = \frac{\beta}{\alpha - 1}$$

$$\text{deviazione standard} = \sqrt{\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}}$$

$$\text{asimmetria} = \frac{4\sqrt{\alpha-2}}{\alpha-3}$$

$$\text{curtosi in eccesso} = \frac{30\alpha - 66}{(\alpha-3)(\alpha-4)} - 3$$

Requisiti di Input:

Alfa (Forma) > 0

Beta (Scala) > 0

Distribuzione Pearson VI

La distribuzione Pearson VI è imparentata con la distribuzione Gamma. E' la funzione razionale di due variabili distribuite secondo le due distribuzioni Gamma. Alfa 1 (noto anche come forma 1), Alfa 2 (noto anche come forma 2) e Beta (noto anche come scala) sono i parametri distribuzionali.

I costrutti matematici della distribuzione Pearson VI sono come segue:

$$f(x) = \frac{(x/\beta)^{\alpha_1 - 1}}{\beta \ \mathrm{B}(\alpha_1, \alpha_2)[1 + (x/\beta)]^{\alpha_1 + \alpha_2}}$$

$$F(x) = F_B\left(\frac{x}{x+\beta}\right)$$

$$\text{media} = \frac{\beta\alpha_1}{\alpha_2 - 1}$$

$$\text{deviazione standard} = \sqrt{\frac{\beta^2\alpha_1(\alpha_1+\alpha_2-1)}{(\alpha_2-1)^2(\alpha_2-2)}}$$

$$\text{asimmetria} = 2\sqrt{\frac{\alpha_2-2}{\alpha_1(\alpha_1+\alpha_2-1)}}\left[\frac{2\alpha_1+\alpha_2-1}{\alpha_2-3}\right]$$

$$\text{curtosi in eccesso} = \frac{3(\alpha_2-2)}{(\alpha_2-3)(\alpha_2-4)}\left[\frac{2(\alpha_2-1)^2}{\alpha_1(\alpha_1+\alpha_2-1)} + (\alpha_2+5)\right] - 3$$

Requisiti di Input:

Alfa 1 (Forma 1) > 0

Alfa 2 (Forma 2) > 0

Beta (Scala) > 0

Distribuzione PERT

La distribuzione PERT è molto usata nel management di progetto e di programma per definire gli scenari del caso peggiore, del caso nominale e del caso migliore dei tempi di completamento di un progetto. E' imparentata con le distribuzioni Beta e Triangolari. La distribuzione PERT può essere usata per identificare i rischi nei modelli di progetto e di costo basato sulla probabilità di raggiungere i target e gli scopi nell'ambito di qualsiasi numero di componenti del progetto usando i valori minimi, massimi e più probabili, ma è stata progettata per generare una distribuzione che assomigli più da vicino alle distribuzioni di probabilità realistiche. La distribuzione PERT può fornire una disposizione vicina alle distribuzioni normali o lognormali. Come la distribuzione triangolare, la distribuzione PERT enfatizza il valore "più probabile" sopra le stime minime e massime. Tuttavia, diversamente dalla distribuzione triangolare, la distribuzione PERT costruisce una curva "morbida" che pone progressivamente più enfasi sui valori attorno (vicino) al valore più probabile, a scapito dei valori attorno alle estremità. In pratica, questo significa che ci "fidiamo" della stima per il valore più probabile e crediamo che, anche se non fosse esattamente corretta (come le stime quasi mai lo sono), abbiamo l'aspettativa che il valore risultante sarà vicino a quella stima. Se si presume che molti fenomeni del mondo reale siano normalmente distribuiti, l'attrattiva della distribuzione PERT sta nel fatto che produce una curva simile in forma alla curva normale, senza dover sapere i precisi parametri della relativa curva normale. Minimo, Più probabile e Massimo sono i parametri distribuzionali.

I costrutti matematici della distribuzione PERT sono come segue:

$$f(x) = \frac{(x-\min)^{A1-1}(\max - x)^{A2-1}}{B(A1, A2)(\max - \min)^{A1+A2-1}}$$

$$where\ A1 = 6\left[\frac{\frac{\min + 4(\text{likely}) + \max}{6} - \min}{\max - \min}\right]\ and\ A2 = 6\left[\frac{\max - \frac{\min + 4(\text{likely}) + \max}{6}}{\max - \min}\right]$$

$$and\ B\ is\ the\ Beta\ \ function$$

$$Mean = \frac{Min + 4Mode + Max}{6}$$

$$Standard\ Deviation = \sqrt{\frac{(\mu - Min)(Max - \mu)}{7}}$$

$$Skew = \sqrt{\frac{7}{(\mu - Min)(Max - \mu)}}\left(\frac{Min + Max - 2\mu}{4}\right)$$

Requisiti di Input:

Minimo ≤ Più probabile ≤ Massimo e possono essere valori positivi, valori negativi o zero

Distribuzione Potenza

La distribuzione Potenza è imparentata alla distribuzione esponenziale nel senso che la probabilità di piccoli esiti è grande, ma decresce esponenzialmente col crescere del valore dell'esito. Alfa (noto anche come forma) è il solo parametro distribuzionale.

I costrutti matematici della distribuzione potenza sono come segue:

$$f(x) = \alpha x^{\alpha - 1}$$

$$F(x) = x^{\alpha}$$

$$\text{media} = \frac{\alpha}{1+\alpha}$$

$$\text{deviazione standard} = \sqrt{\frac{\alpha}{(1+\alpha)^2(2+\alpha)}}$$

$$\text{asimmetria} = \sqrt{\frac{\alpha+2}{\alpha}}\left(\frac{2(\alpha-1)}{\alpha+3}\right)$$

Requisiti di Input:

Alfa (Forma) > 0

Distribuzione Moltiplicativa Potenza Spostato

La distribuzione Potenza è imparentata alla distribuzione esponenziale nel senso che la probabilità di piccoli esiti è grande, ma decresce esponenzialmente col crescere del valore dell'esito. Alfa (noto anche come forma) è il solo parametro distribuzionale.

Requisiti di Input:

Alfa (Forma) > 0

Posizione può essere qualsiasi valore positivo o negativo incluso lo zero

Fattore > 0

Distribuzione t di Student

La distribuzione t di Student è la distribuzione più largamente usata nel test di verifica d'ipotesi. Questa distribuzione è usata per stimare la media di una popolazione distribuita normalmente, quando la dimensione del campione è piccola ed è usata per testare la significatività statistica della differenza tra le medie di due campioni o gli intervalli di confidenza per campioni di dimensioni piccole.

I costrutti matematici della distribuzione t sono come segue:

$$f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{r\pi}\,\Gamma[r/2]}(1+t^2/r)^{-(r+1)/2}$$

media = 0 (questo vale per tutti i gradi di libertà r tranne se la distribuzione viene spostata ad un'altra posizione centrale di "non zero")

$$\text{deviazione standard} = \sqrt{\frac{r}{r-2}}$$

asimmetria = 0 (questo vale per tutti i gradi di libertà r)

$$\text{curtosi in eccesso} = \frac{6}{r-4} \text{ per tutti } r > 4$$

dove $t = \dfrac{x - \bar{x}}{s}$ e Γ è la funzione gamma.

Gradi di libertà *r* è il solo parametro distribuzionale.

La distribuzione t è collegata alla distribuzione F come segue: il quadrato di un valore di t con r gradi di libertà è distribuito come F con 1 e r gradi di libertà. La forma complessiva della funzione di densità di probabilità della distribuzione t assomiglia anche alla forma a campana di una variabile distribuita normalmente con media 0 e varianza 1, tranne che è un po' più bassa e ampia o leptocurtica (code spesse agli estremi e centro a picco). Col aumentare del numero di gradi di libertà (diciamo, sopra 30), la distribuzione t si avvicina alla distribuzione normale con media 0 and varianza 1.

Requisiti d'input:

Gradi di libertà ≥ 1 e deve essere un numero intero

Distribuzione Triangolare

La distribuzione triangolare descrive la situazione in cui conoscete i valori minimi, massimi e più probabili che si verificheranno. Per esempio, potete descrivere il numero d'automobili vendute per settimana, quando le vendite passate mostrano il numero minimo, massimo e normale d'automobili vendute.

Condizioni

Le tre condizioni sottostanti della distribuzione triangolare sono:

- Il numero minimo di elementi è fisso.
- Il numero massimo di elementi è fisso.
- Il numero più probabile di elementi cade tra i valori minimi e massimi, formando una distribuzione a forma di triangolo che mostra che è meno probabile che si verifichino i valori vicini al minimo e al massimo che quelli vicino al valore più probabile.

I costrutti matematici della distribuzione triangolare sono come segue:

$$f(x) = \begin{cases} \dfrac{2(x - Min)}{(Max - Min)(Likely - \min)} & \text{per } Min < x < Likely \\ \dfrac{2(Max - x)}{(Max - Min)(Max - Likely)} & \text{for } Likely < x < \text{Max} \end{cases}$$

(Likely = Più Probabile)

media = $\dfrac{1}{3}(Min + Likely + Max)$

deviazione standard =

$$\sqrt{\frac{1}{18}(Min^2 + Likely^2 + Max^2 - Min\,Max - Min\,Likely - Max\,Likely)}$$

asimmetria =

$$\frac{\sqrt{2}(Min + Max - 2Likely)(2Min - Max - Likely)(Min - 2Max + Likely)}{5(Min^2 + Max^2 + Likely^2 - MinMax - MinLikely - MaxLikely)^{3/2}}$$

curtosi in eccesso = –0.6 (questo vale per tutti gli inputs di *Min*, *Max* e *Likely*)

Valore minimo (*Min*), valore più probabile (*Likely*) e valore massimo (*Max*) sono i parametri distribuzionali.

Requisiti d'input:

Min ≤ Most Likely ≤ Max e può essere qualsiasi valore

Tuttavia, Min < Max e può essere qualsiasi valore

Distribuzione Uniforme

Con la distribuzione uniforme tutti i valori cadono tra il minimo e il massimo e si verificano con la stessa probabilità.

Condizioni

Le tre condizioni sottostanti della distribuzione uniforme sono:

- Il valore minimo e fisso.
- Il valore massimo e fisso.
- Tutti i valori tra il minimo e il massimo si verificano con la stessa probabilità.

I costrutti matematici della distribuzione uniforme sono come segue:

$$f(x) = \frac{1}{Max - Min} \quad \text{per tutti i valori così che } Min < Max$$

$$\text{media} = \frac{Min + Max}{2}$$

$$\text{deviasione standard} = \sqrt{\frac{(Max - Min)^2}{12}}$$

asimmetria = 0 (questo vale per tutti gli inputs di *Min* e *Max*)

curtosi in eccesso = –1.2 (questo vale per tutti gli inputs di *Min* e *Max*)

Valore massimo (*Max*) e valore minimo (*Min*) sono i parametri distribuzionali.

Requisiti d'input: Min < Max e può essere qualsiasi valore

Distribuzione Weibull (Distribuzione Rayleigh)

La distribuzione Weibull descrive i dati che provengono da prove di durata e di fatica. E' comunemente usata per descrivere il tempo di fallimento in studi d'affidabilità come anche i carichi di rottura di materiali in prove d'affidabilità e di controllo qualità. Le distribuzioni Weibull sono anche usate per rappresentare varie quantità fisiche come la velocità del vento.

La distribuzione Weibull è una famiglia di distribuzioni che possono assumere le proprietà di varie altre distribuzioni. Per esempio, in base al parametro di forma che definite, la distribuzione Weibull può essere usata per modellare, tra le altre, le distribuzioni esponenziali e Rayleigh. La distribuzione Weibull è molto flessibile. Quando il parametro di forma della Weibull è uguale a 1,0, la distribuzione Weibull è identica alla distribuzione esponenziale. Il parametro di posizione della Weibull vi permette di impostare una distribuzione esponenziale che inizia da una posizione diversa da 0,0. Quando il parametro di forma è minore di 1,0, la distribuzione Weibull diventa una curva fortemente declinante. Un produttore potrebbe trovare quest'effetto utile nella descrizione di guasti di componenti durante un periodo di "burn-in".

I costrutti matematici della distribuzione Weibull sono come segue:

$$f(x) = \frac{\alpha}{\beta}\left[\frac{x}{\beta}\right]^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

$$\text{media} = \beta\,\Gamma(1+\alpha^{-1})$$
$$\text{deviazione standard} = \beta^2\left[\Gamma(1+2\alpha^{-1}) - \Gamma^2(1+\alpha^{-1})\right]$$

$$\text{asimmetria} = \frac{2\Gamma^3(1+\beta^{-1}) - 3\Gamma(1+\beta^{-1})\Gamma(1+2\beta^{-1}) + \Gamma(1+3\beta^{-1})}{\left[\Gamma(1+2\beta^{-1}) - \Gamma^2(1+\beta^{-1})\right]^{3/2}}$$

curtosi in eccesso =

$$\frac{-6\Gamma^4(1+\beta^{-1}) + 12\Gamma^2(1+\beta^{-1})\Gamma(1+2\beta^{-1}) - 3\Gamma^2(1+2\beta^{-1}) - 4\Gamma(1+\beta^{-1})\Gamma(1+3\beta^{-1}) + \Gamma(1+4\beta^{-1})}{\left[\Gamma(1+2\beta^{-1}) - \Gamma^2(1+\beta^{-1})\right]^2}$$

Forma (α) e scala di posizione centrale (β)sono i parametri distribuzionali e Γ è la funzione Gamma.

Requisiti d'input: Forma Alfa $\geq$ 0.05. Scala Beta > 0 e può essere qualsiasi valore positivo

Distribuzioni Moltiplicativa Weibull e Rayleigh Spostato

La distribuzione Weibull descrive i dati che provengono da prove di durata e di fatica. E' comunemente usata per descrivere il tempo di fallimento in studi di affidabilità come anche i carichi di rottura di materiali in prove di affidabilità e di controllo qualità. Le distribuzioni Weibull sono anche usate per rappresentare varie quantità fisiche come la velocità del vento. La distribuzione Weibull è una famiglia di distribuzioni che possono assumere le proprietà di varie altre distribuzioni. Per esempio, in base al parametro di forma che definite, la distribuzione Weibull può essere usata per modellare, tra le altre, le distribuzioni esponenziali e Rayleigh. La distribuzione Weibull è molto flessibile. Quando il parametro di forma Weibull è uguale a 1,0, la distribuzione Weibull è identica alla distribuzione esponenziale. La scala di posizione centrale o parametro beta vi permette di impostare una distribuzione esponenziale che inizia da una posizione diversa da 0,0. Quando il parametro di forma è minore di 1,0, la distribuzione Weibull diventa una curva fortemente declinante. Un produttore potrebbe trovare questo effetto utile nella descrizione di guasti di componenti durante un periodo di burn-in. Forma (α) e scala (β) sono i parametri distribuzionali.

Requisiti di Input: Forma Alfa $\geq$ 0,05. Scala di posizione centrale o Beta > 0 e può avere qualsiasi valore positivo. Posizione può essere qualsiasi valore positivo o negativo incluso lo zero

Fattore > 0

3

3. PREVISIONE

La Previsione è l'atto di predire il futuro, che sia basato su dati storici o sulla congettura del futuro quando non esistono dati storici. Quando sono disponibili dati storici, è meglio un metodo quantitativo o statistico, ma se i dati storici non esistono, allora di solito è probabile che un metodo qualitativo o basato sul giudizio sia l'unico ricorso. La Figura 3.1 elenca le metodologie più comuni di previsione.

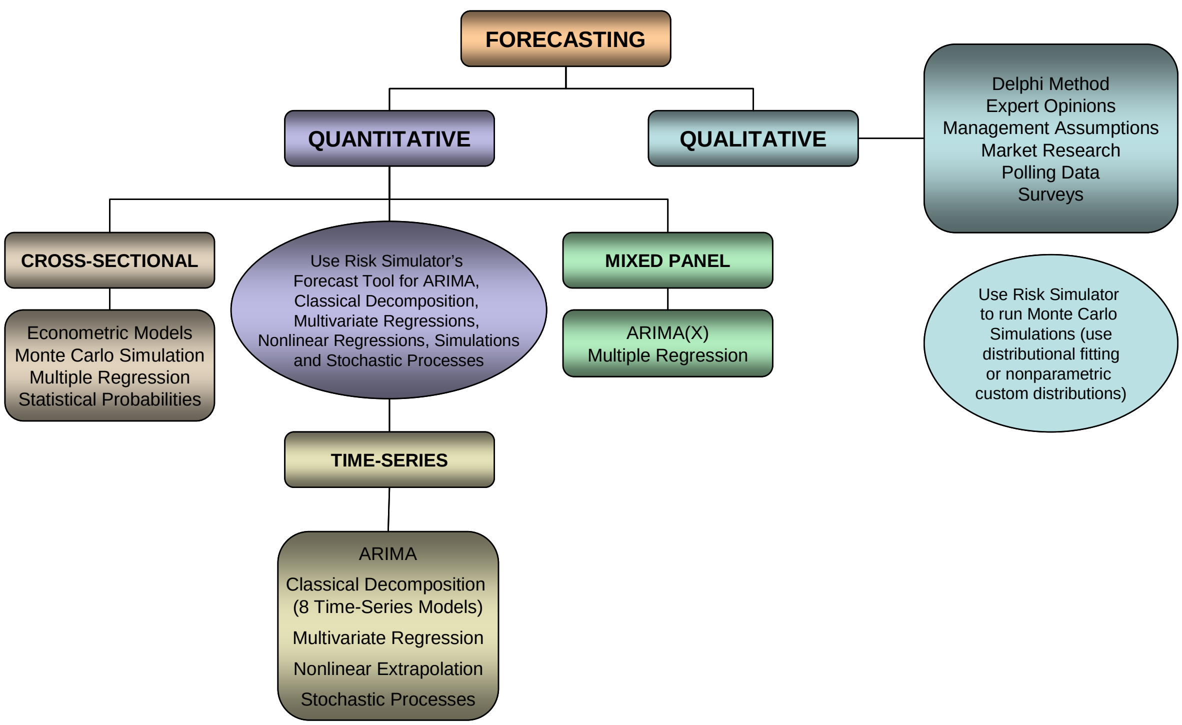


Figura 3.1 – Metodi di Previsione

3.1 Tipi diversi di tecniche di Previsione

La previsione può generalmente essere divisa in quantitativa e qualitativa. La previsione qualitativa è usata quando non ci sono o ci sono pochi dati storici, dati contemporanei o dati comparabili affidabili. Esistono diversi metodi qualitativi come il Metodo Delphi o il Metodo delle opinioni d'esperti (una previsione ad approccio consensuale da parte d'esperti del settore, esperti di marketing o membri interni dello staff), il Metodo delle ipotesi del management (obiettivi di tassi di crescita fissati dall'alta dirigenza), come anche la ricerca di mercato o dati esterni o sondaggi e rilevamenti (dati ottenuti da fonti esterne, dall'industria e dagli indici del settore o dalla ricerca attiva di mercato). Queste stime possono essere o stime a punto singolo (un consenso medio) o un insieme di valori di previsione (una distribuzione di previsioni). Questa ultima può essere inserita in *Simulatore di Rischio* come una distribuzione personalizzata e le risultanti previsioni possono essere simulate. Vale a dire, una simulazione non parametrica usando i punti dati stimati stessi come la distribuzione.

Per quanto riguarda il tipo quantitativo di previsione, i dati disponibili o i dati che devono essere previsti possono essere divisi in dati di serie temporali (valori che contengono un elemento di tempo, come i ricavi in anni diversi, i tassi d'inflazione, i tassi d'interesse, la quota di mercato, il tasso dei guasti e così via), in dati di sezioni trasversali (valori che non dipendono dal tempo, come la media dei voti degli studenti della decima classe di tutto il paese in un determinato anno, dato i livelli dei punteggi degli esami pre-universitari, il Q.I. ed il numero di consumazioni alcoliche a settimana), o in dati misti di panel (un misto di dati di serie temporali e di panel; per esempio, predire le vendite nei prossimi 10 anni, dato le spese di marketing messe in bilancio e le proiezioni della quota di mercato: questo significa che i dati di vendita sono di tipo serie temporali, ma sono presenti variabili esogene come le spese di marketing e la quota di mercato per aiutare nella modellazione delle predizioni della previsione).

Il software *Simulatore di Rischio* fornisce all'utente varie metodologie di previsione:

1. ARIMA (Media mobile integrata autoregressiva)
2. Auto-ARIMA
3. Econometria di base
4. Auto-Econometria
5. Distribuzioni Personalizzate
6. Logica fuzzy combinatoria

7. GARCH (Eteroschedasticità condizionale autoregressiva generalizzata)
8. Curve a J
9. Catene di Markov
10. Rete neurale
11. Massima verosimiglianza (Logit, Probit, Tobit)
12. Regressione Multivariata
13. Estrapolazione non lineare
14. Curve a S
15. Spline Cubico
16. Previsione con Processi Stocastici
17. Analisi di serie temporali e Scomposizione
18. Linee di tendenza

I dettagli analitici di ciascun metodo di previsione non rientrano nello scopo di questo Manuale dell'Utente. Per maggiori dettagli prego consultare *Modeling Risk,* 2nd Edition,: *Applying Monte Carlo Simulation, Real Options Analysis, Stochastic Forecasting and Portfolio Optimization* (Wiley Finance 2010) di Dr. Johnathan Mun, che è anche il creatore del software Simulatore di Rischio. Ciò nonostante, di seguito illustriamo alcuni dei metodi più comuni. Tutti gli altri metodi di previsione sono abbastanza semplici da applicare in Simulatore di Rischio.

Il seguito fornisce una veloce analisi di ciascuna metodologia e alcuni esempi veloci per aiutarvi ad iniziare ad usare software. Descrizioni più dettagliate e modelli d'esempio di ciascuna di queste tecniche si trovano ovunque in questo e nel prossimo capitolo.

ARIMA

La Media mobile integrata autoregressiva (ARIMA, nota anche come Box-Jenkins ARIMA) è una tecnica avanzata di modellazione econometrica. ARIMA esamina dati storici di serie temporali ed esegue processi di ottimizzazione d'adattamento all'indietro per rendere conto dell'autocorrelazione storica (la relazione di un valore contro un altro nel tempo) e della stabilità dei dati per correggere le caratteristiche non stazionarie dei dati. Questo modello predittivo impara col tempo mediante la correzione degli errori di previsione. Normalmente è richiesta una conoscenza avanzata dell'econometria per costruire validi modelli predittivi usando questo metodo.

Auto ARIMA Il modulo Auto-ARIMA automatizza una parte della modellazione ARIMA tradizionale, testando automaticamente multiple permutazioni di specifiche del modello, e fornisce il modello con miglior adattamento. Eseguire Auto-ARIMA è simile all'esecuzione di previsioni ARIMA normali. La differenza è che non sono più richiesti gli inputs di P, D e Q e che combinazioni differenti di questi inputs sono eseguite e confrontate automaticamente.

Econometria di base L'Econometria si riferisce ad un ramo dell'analitica economica (business analytics): tecniche di modellazione e previsione per modellare il comportamento di o per prevedere determinate variabili nel campo del business, dell'economia, della finanza, della fisica, della produzione, delle operazioni ed altre ancora. Eseguire i modelli dell'Econometria di base è simile all'esecuzione di una normale analisi di regressione, tranne che le variabili dipendenti ed indipendenti possono essere modificate prima dell'esecuzione di una regressione.

Auto-Econometria Simile all'Econometria, ma qui vengono eseguite automaticamente migliaia di variabili lineari, non lineari, interagenti, sfasate e miste sui vostri dati, per determinare il modello econometrico col miglior adattamento che meglio descrive il comportamento della variabile dipendente. Questa tecnica è utile per modellare gli effetti delle variabili e per prevedere esiti futuri, tuttavia non richiede che l'analista sia un econometrico esperto.

Distribuzioni Personalizzate Usando il Simulatore di Rischio, si possono raccogliere opinioni d'esperti e generare così una distribuzione personalizzata. Questa tecnica di previsione è utile se l'insieme di dati è piccolo o se la bontà dell'adattamento non è valida quando applicata ad una procedura d'adattamento distribuzionale.

Logica fuzzy combinatoria Il termine logica fuzzy deriva dalla teoria degli insiemi fuzzy per affrontare ragionamenti che sono approssimativi anziché accurati — contrapposta alla logica "crisp" (incisiva), dove insiemi binari hanno logica binaria, le variabili della logica fuzzy possono avere un valore di verità che varia tra 0 e 1 e non è vincolato ai due valori di verità della classica logica preposizionale. Questo schema di ponderazione fuzzy è usato insieme con un metodo combinatorio per fornire risultati di previsioni di serie temporali in Simulatore di Rischio.

GARCH Il modello GARCH (eteroschedasticità condizionale autoregressiva generalizzata) è usato per modellare livelli storici di volatilità e prevedere livelli futuri di volatilità di valori negoziabili (per esempio, prezzi di titoli, prezzi di commodities, prezzi del petrolio e così via). L'insieme dei dati deve essere una serie temporale di livelli di prezzi grezzi. GARCH converte prima i prezzi in rendimenti relativi ed esegue poi un'ottimizzazione interna per adattare i dati storici in una struttura a termine della volatilità con ritorno alla media, assumendo al contempo che la volatilità è di natura eteroschedastica (ovvero che cambia col tempo secondo determinate caratteristiche econometriche). Diverse variazioni di questa metodologia sono disponibili in Simulatore di Rischio, comprese EGARCH, EGARCH-T, GARCH-M, GJR-GARCH, GJR-GARCH-T, IGARCH e T-GARCH.

Curve a J

La curva a J o curva di crescita esponenziale è una curva dove la crescita del periodo successivo dipende dal livello del periodo attuale e l'aumento è esponenziale. Questo significa che, col tempo, i valori aumenteranno in modo rilevante da un periodo all'altro. Questo modello è tipicamente usato nella previsione della crescita biologica e delle reazioni chimiche col passare del tempo.

Catene di Markov

Una Catena di Markov esiste, quando la probabilità di uno stato futuro dipende da uno stato precedente e quando sono collegati formano una catena che ritorna ad un livello d'equilibrio stazionario di lungo termine. Questo metodo è tipicamente usato per prevedere la quota di mercato di due concorrenti. Gli inputs richiesti sono la probabilità iniziale che un cliente del primo negozio (il primo stato) ritornerà allo stesso negozio nel periodo successivo contro la probabilità che vada al negozio del concorrente nello stato successivo.

Rete neurale

Il termine Rete neurale è spesso usato per indicare una rete o un circuito of neuroni biologici, mentre l'uso moderno del termine si riferisce a reti neurali artificiali che consistono di neuroni o nodi artificiali ricreati in un ambiente software. Questa metodologia tenta di imitare il cervello o i neuroni umani per quanto riguarda i modi di pesare e di identificare gli schemi, e nel nostro caso, di identificare gli schemi con lo scopo di eseguire previsioni di dati di serie temporali.

Massima Verosimiglianza su Logit, Probit e Tobit

La Stima di massima verosimiglianza (MLE) è usata per prevedere la probabilità che qualcosa accada, dato alcune variabili indipendenti. MLE è usata, per esempio, per predire se una linea di credito o un debito resterà inadempiuta(o), date le caratteristiche del debitore (30 anni, celibe/nubile, remunerazione annua di $100000 e con un debito totale di carte di credito di $10000); o la probabilità che un paziente svilupperà un cancro polmonare se questa persona è un maschio d'età tra 50 e 60 anni, fuma 5 pacchetti di sigarette per mese e così via. In queste circostanze, la variabile dipendente è limitata (per esempio, limitata ad essere binaria 1 e 0 per inadempienza/morte e non inadempienza/vita, o limitata a valori di numeri interi come 1, 2, 3 e così via) e l'esito desiderato del modello è di predire la probabilità che si verifichi un evento. L'analisi di regressione tradizionale non funziona in queste situazioni: la probabilità predetta è normalmente minore di zero o maggiore di uno; molte delle ipotesi di regressione richieste, come l'indipendenza e la normalità degli errori, sono violate; e gli errori saranno abbastanza grandi).

Regressione Multivariata

La Regressione Multivariata è usata per modellare la struttura e le caratteristiche relazionali di una certa variabile per come dipende da altre variabili esogene indipendenti. Usando la relazione modellata, possiamo predire i valori futuri della variabile dipendente. Si possono anche determinare la precisione e la bontà di adattamento per questo modello. È possibile adattare modelli lineari e non lineari nell'analisi di regressione multipla.

Estrapolazione non lineare

Si presume che la struttura sottostante dei dati da prevedere sia non lineare nel corso del tempo. Per esempio, un insieme di dati come 1, 4, 9, 16, 25 è considerato essere non lineare (questi punti dati provengono da una funzione al quadrato).

Curve a S

La curva a S o curva di crescita logistica inizia come una curva a J, con tassi di crescita esponenziali. Col tempo, l'ambiente si satura (per esempio, saturazione del mercato, concorrenza, sovraffollamento), la crescita rallenta ed il valore previsto si ferma infine ad un livello massimo o di saturazione. Questo modello è tipicamente usato nella previsione delle quote di mercato o della crescita delle vendite di un prodotto nuovo dalla sua introduzione nel mercato fino alla sua maturità e diminuzione, delle dinamiche delle popolazioni e d'altri fenomeni che accadono in natura.

Spline Cubico

Talvolta ci sono dati mancanti nell'insieme di dati di una serie temporale. Per esempio, sono disponibili i tassi d'interesse per gli anni 1 a 3, seguiti dai tassi per gli anni 5 a 8 e poi per l'anno 10. Le curve spline possono essere usate per interpolare i valori dei tassi d'interesse per gli anni mancanti, basato sui dati esistenti. Le curve spline possono anche essere usate per prevedere od estrapolare valori di periodi temporali futuri oltre il periodo temporale dei dati disponibili. I dati possono essere lineari o non lineari.

Previsione con Processi Stocastici

Le variabili sono talvolta stocastiche e non possono essere facilmente predette usando mezzi tradizionali. Queste variabili si chiamano stocastiche. Ciononostante, la maggior parte dei fenomeni finanziari, economici e naturali (per esempio, il movimento delle molecole nell'aria) seguono una nota legge o relazione matematica. Anche se i valori risultanti sono incerti, la struttura matematica sottostante è nota e può essere simulata usando la simulazione del rischio Monte Carlo. I processi supportati in Simulatore di Rischio comprendono la Passeggiata casuale di moto Browniano, il Ritorno alla media, la Diffusione a salti e processi misti, utili nella previsione di variabili non stazionarie di serie temporali.

Analisi di serie temporali e Scomposizione

In dati di serie temporali con un "buon comportamento" data (esempi tipici comprendono i ricavi da vendite e le strutture dei costi di grandi aziende), i valori tendono ad avere fino a tre elementi: un valore base, la tendenza e la stagionalità. L'Analisi di serie temporali usa questi dati storici, li scompone in questi tre elementi e li ricompone in previsioni del futuro. In altre parole, questo metodo di previsione, come alcuni degli altri che abbiamo descritto, esegue prima un adattamento all'indietro ("previsione all'indietro") dei dati storici, avanti di fornire stime di valori futuri (previsioni).

3.2 Eseguire lo strumento di previsione in Simulatore di Rischio

In generale, per creare delle previsioni sono richiesti alcuni passi veloci:

- Avviate Excel ed inserite od aprite i vostri dati storici esistenti
- Selezionate i dati e cliccate su *Simulazione* e selezionate *Previsione*
- Selezionate le sezioni pertinenti (ARIMA, Regressione Multivariata, Estrapolazione non lineare, Previsione stocastica, Analisi di serie temporali) ed inserite gli inputs attinenti

La Figura 3.2 illustra lo strumento di *Previsione* e le varie metodologie.

Figura 3.2 – Simulatore di Rischio - Metodi di Previsione

Di seguito forniamo una veloce analisi di ciascuna metodologia ed alcuni esempi veloci per aiutarvi ad iniziare ad usare il software. Il file degli esempi si trova o nel menu d'avvio sotto ***Avvio | Real Options Valuation | Simulatore di Rischio | Esempi*** o con accesso diretto sotto ***Simulatore di Rischio | Modelli d'esempio.***

3.3 Analisi di serie temporali

Teoria:

La Figura 3.3 elenca gli otto modelli più comuni di serie temporali, divisi per stagionalità e tendenza. Per esempio, se la variabile dati non ha tendenza trend o stagionalità, allora è sufficiente un modello con media mobile singola o un modello di graduazione esponenziale singola. Tuttavia, se esiste stagionalità, ma non è presente una tendenza rilevabile, è meglio usare un modello con stagionalità additiva o stagionalità moltiplicativa, e così via.

	No Seasonality	**With Seasonality**
No Trend	*Single Moving Average*	*Seasonal Additive*
	Single Exponential Smoothing	*Seasonal Multiplicative*
With Trend	*Double Moving Average*	*Holt-Winter's Additive*
	Double Exponential Smoothing	*Holt-Winter's Multiplicative*

Figura 3.3 – Gli otto più comuni metodi di serie temporali

Procedura:

Avviate Excel e, se necessario, aprite i vostri dati storici (l'esempio sottostante utilizza il file ***Previsione di serie temporali*** nella cartella degli esempi)

- Selezionate i dati storici (i dati devono essere elencati in una singola colonna)
- Selezionate Simulatore di Rischio | Previsione | Analisi di serie temporali
- Scegliete il modello da applicare, inserite le ipotesi pertinenti e cliccate su **OK**

Ricavi storici da vendite

Anno	Trimestre	Periodo	Vendite
2006	1	1	$684.20
2006	2	2	$584.10
2006	3	3	$765.40
2006	4	4	$892.30
2007	1	5	$885.40
2007	2	6	$677.00
2007	3	7	$1,006.60
2007	4	8	$1,122.10
2008	1	9	$1,163.40
2008	2	10	$993.20
2008	3	11	$1,312.50
2008	4	12	$1,545.30
2009	1	13	$1,596.20
2009	2	14	$1,260.40
2009	3	15	$1,735.20
2009	4	16	$2,029.70
2010	1	17	$2,107.80
2010	2	18	$1,650.30
2010	3	19	$2,304.40
2010	4	20	$2,639.40

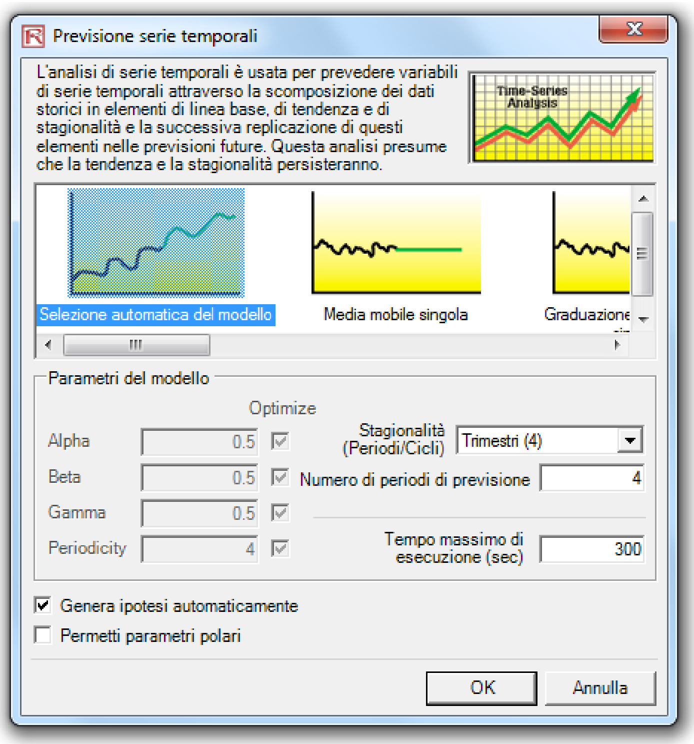


Figura 3.4 – Analisi di serie temporali

Interpretare i risultati:

La Figura 3.5 illustra i risultati generati usando lo strumento di *Previsione*. Il modello utilizzato è stato il Modello Moltiplicativo di Holt-Winters. Prego notare che nella Figura 3.5, il diagramma di previsione e adattamento del modello indica che la tendenza e la stagionalità sono rilevate bene dal Modello Moltiplicativo di Holt-Winters. Il report dell'analisi di serie temporali fornisce i pertinenti parametri ottimizzati di alfa, beta e gamma, le misure d'errore, i dati adattati, i valori di previsione e il diagramma della previsione adattata. I parametri sono puramente per consultazione. Alfa cattura l'effetto memoria dei cambiamenti del livello base nel tempo, beta è il parametro di tendenza che misura la robustezza della tendenza, mentre gamma misura la forza della stagionalità dei dati storici. L'analisi scompone i dati storici in questi tre elementi e poi li ricompone per prevedere il futuro. I dati adattati illustrano sia i dati adattati che i dati storici usando il modello ricomposto e mostra la vicinanza al passato delle previsioni (una tecnica chiamata "previsione all'indietro"). I valori di previsione sono o stime di punti singoli o ipotesi (se si seleziona l'opzione Genera automaticamente le ipotesi e se esiste un profilo di simulazione). Il diagramma illustra questi valori storici, adattati e di previsione. Il diagramma è uno strumento potente di comunicazione e visualizzazione per vedere la bontà del modello di previsione.

Note:

Il modulo dell'analisi di serie temporali contiene gli otto modelli visti nella Figura 3.3. Potete scegliere il modello specifico da eseguire, basato sui criteri della tendenza e della stagionalità, o scegliere l'opzione Selezione automatica del modello, che eseguirà ripetutamente gli otto metodi, ottimizzerà i parametri e troverà il modello col miglior adattamento per i vostri dati. In alternativa, se scegliete uno degli otto modelli, potete anche deselezionare la casella di controllo *Ottimizza* ed inserire i vostri parametri di alfa, beta e gamma. Consultare il libro *Modeling Risk,* 2nd Edition*: Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization* (Wiley, 2010) di Dr. Johnathan Mun per maggiori dettagli sulle specifiche tecniche di questi parametri. In aggiunta, se scegliete l'opzione di selezione automatica del modello o qualsiasi dei modelli stagionali, dovete inserire i periodi pertinenti di stagionalità. L'input di stagionalità deve essere un numero intero positivo (per esempio, se i dati sono trimestrali, inserite 4 come il numero di stagioni o cicli per anno, o inserite 12 se i dati sono mensili). Dopo, inserite il numero di periodi di previsione. Anche questo valore deve essere un numero intero**Fehler! Textmarke nicht definiert.** positivo. Il tempo massimo di esecuzione è impostato su 300 secondi. Normalmente non sono richieste modifiche. Tuttavia, quando eseguite una previsione con una quantità rilevante di dati storici, l'analisi potrebbe durare leggermente di più e se il tempo di elaborazione supera questo tempo di esecuzione, il processo sarà terminato. Potete anche optare di fare generare le ipotesi automaticamente alla previsione. In altre parole, invece di stime di punti singoli, le previsioni saranno ipotesi. Per finire, l'opzione Parametri polari vi permette di ottimizzare i parametri alfa, beta e gamma per includere zero e uno. Certi software di previsione permettono parametri polari, mentre altri non li permettono. Il Simulatore di Rischio vi permette di scegliere quali usare. Normalmente non è necessario usare i parametri polari.

Holt-Winters Moltiplicativa

Sommario delle statistiche

Alfa, Beta, Gamma	RMSE	Alfa, Beta, Gamma	RMSE
0.00, 0.00, 0.00	914.824	0.00, 0.00, 0.00	914.824
0.10, 0.10, 0.10	415.322	0.10, 0.10, 0.10	415.322
0.20, 0.20, 0.20	187.202	0.20, 0.20, 0.20	187.202
0.30, 0.30, 0.30	118.795	0.30, 0.30, 0.30	118.795
0.40, 0.40, 0.40	101.794	0.40, 0.40, 0.40	101.794
0.50, 0.50, 0.50	102.143		

L'analisi è stata eseguita con alfa = 0.2429, beta = 1.0000, gamma = 0.7797 e stagionalità = 4

Sommario dell'analisi serie temporali

Quando esistono sia stagionalità che trend sono necessari modelli più avanzati per scomporre i dati nei loro elementi di base: un livello di caso-base (L) ponderato dal parametro alfa; un componente di trend (b) ponderato dal parametro beta; un componente di stagionalità (S) ponderato dal parametro gamma. Esistono diversi metodi ma i due più comuni sono i metodi della stagionalità additiva di Holt-Winters e della stagionalità moltiplicativa di Holt-Winters. Nel modello additivo di Holt-Winters si sommano il livello di caso-base, la stagionalità ed il trend per ottenere l'adattamento della previsione.

Il test di miglior adattamento per la previsione con media mobile usa la radice dell'errore quadratico medio (RMSE). La radice dell'errore quadratico medio (RMSE) calcola la radice quadrata delle deviazioni medie al quadrato dei valori adattati contro i punti dati reali.

L'errore quadratico medio (MSE) è una misura d'errore assoluto che eleva al quadrato gli errori (la differenza tra i dati storici reali e i dati adattati alla previsione pronosticati dal modello) per evitare che gli errori positivi e negativi si annullino a vicenda. Questa misura tende anche ad esagerare gli errori grandi, dato che pondera gli errori grandi più pesantemente degli errori piccoli elevandoli al quadrato. Questo può essere d'aiuto se si confrontano modelli di serie temporali differenti. La radice dell'errore quadratico medio (RMSE) è la radice quadrata dell'errore quadratico medio (MSE) ed è la misura d'errore più comune, nota anche come la funzione di perdita quadratica. La radice dell'errore quadratico medio (RMSE) può essere definita come la media dei valori assoluti degli errori della previsione ed il suo uso è molto appropriato se il costo degli errori della previsione è proporzionale alla dimensione assoluta dell'errore della previsione. La radice dell'errore quadratico medio (RMSE) è usata come criterio di selezione per il modello di serie temporali col miglior adattamento.

L'errore percentuale assoluto medio (MAPE) è una statistica d'errore relativa misurata come errore percentuale medio di punti dati storici. È adatto quando il costo dell'errore della previsione è più strettamente collegato all'errore percentuale che alla dimensione numerica dell'errore. Per finire, la statistica U di Theil è una misura associata che misura l'ingenuità della previsione del modello. Vale a dire, se la statistica U di Theil è meno di 1.0, allora il metodo di previsione usato fornisce una stima che è statisticamente migliore di una congettura.

Periodo	Reale	Adattamento previsione
1	684.20	
2	584.10	
3	765.40	
4	892.30	
5	885.40	684.20
6	677.00	667.55
7	1006.60	935.45
8	1122.10	1198.09
9	1163.40	1112.48
10	993.20	887.95
11	1312.50	1348.38
12	1545.30	1546.53
13	1596.20	1572.44
14	1260.40	1299.20
15	1735.20	1704.77
16	2029.70	1976.23
17	2107.80	2026.01
18	1650.30	1637.28
19	2304.40	2245.93
20	2639.40	2643.09
Previsione21		2713.69
Previsione22		2114.79
Previsione23		2900.42
Previsione24		3293.81

Misure d'errore	
Radice dell'errore quadratico medio (RMSE)	71.8132
Errore quadratico medio (MSE)	5157.1348
Deviazione assoluta media (MAD)	53.4071
Errore percentuale assoluto medio (MAPE)	4.50%
U di Theil	0.3054

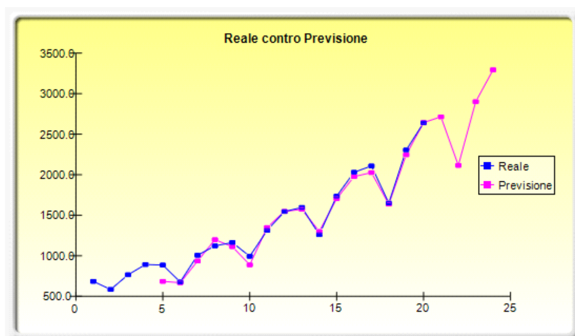


Figura 3.5 – Esempio del Report di una Previsione Holt-Winters

3.4 Regressione Multivariata

Teoria: Si presume che l'utente sia sufficientemente informato sui principi fondamentali dell'analisi di regressione. L'equazione generale della regressione lineare bivariata prende la forma di $Y = \beta_0 + \beta_1 X + \varepsilon$, dove β_0 è l'intercetta, β_1 è la pendenza e ε è il termine d'errore. È bivariata dato che ci sono solo due variabili, una variabile dipendente *Y* e una variabile indipendente *X*, dove *X* è anche nota come il regressore (una regressione bivariata è talvolta conosciuta anche come una regressione univariata, visto che c'è solo una singola variabile indipendente *X*). La variabile dipendente è denominata così, dato che *dipende* dalla variabile indipendente. Per esempio, i ricavi da vendite dipendono dalla somma dei costi di marketing spesi per la pubblicità e la promozione di un prodotto: vendite è la variabile dipendente e i costi di marketing è la variabile indipendente. Un esempio di una regressione bivariata è visto come il semplice inserimento della linea col miglior adattamento attraverso un insieme di punti dati in un piano bidimensionale, come visto nel panello sinistro nella Figura 3.6. In altri casi, si può eseguire una regressione multivariata, quando ci sono multiple o un numero *n* di variabili indipendenti *X*: ora l'equazione generale di regressione prenderà la forma di $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 ... + \beta_n X_n + \varepsilon$. In questo caso, la linea col miglior adattamento sarà dentro un piano dimensionale *n + 1*.

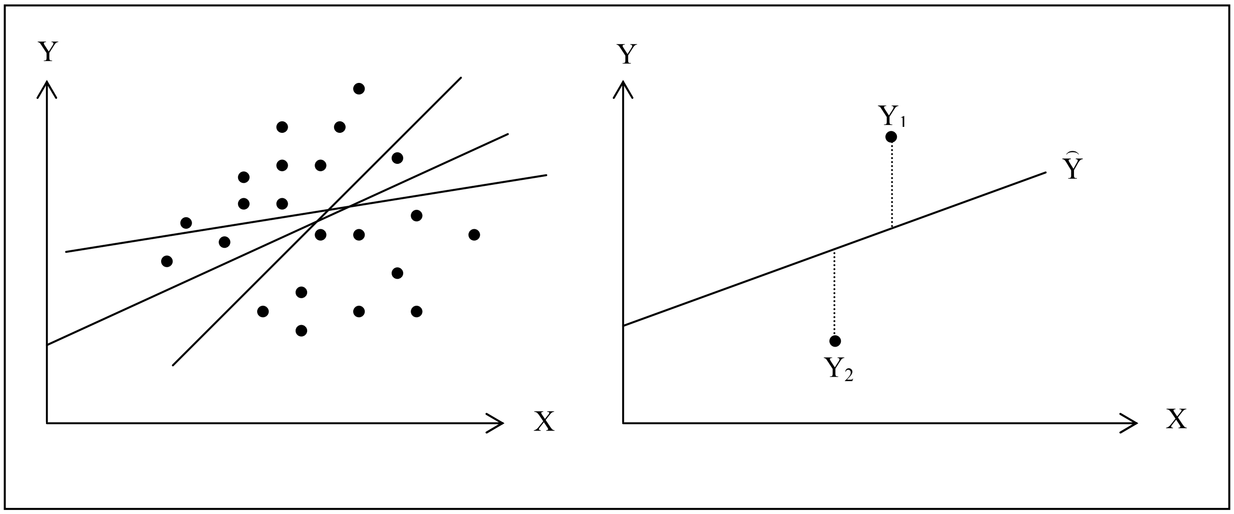


Figura 3.6 –Regressione Bivariata

Tuttavia, adattare una linea attraverso un insieme di punti dati in un diagramma a dispersione, come nella Figura 3.6, può risultare in numerose linee possibili. La linea col miglior adattamento è definita come l'unica singola line che minimizza gli errori totali verticali. In altre parole, la somma delle distanze assolute tra i punti dati effettivi (Y_i) e la linea stimata ($\hat{Y}$), come mostrato nel panello destro della Figura 3.6. Per poter trovare la linea col miglior adattamento che minimizza gli errori, è richiesto un metodo più sofisticato, vale a dire, l'analisi di regressione. L'analisi di regressione , dunque, trova l'unica linea col miglior adattamento mediante la minimizzazione degli errori totali, o calcolando

$$Min \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

dove solo un'unica linea minimizza questa somma degli errori quadrati. Gli errori (le distanze verticali tra i dati effettivi e la linea predetta) sono elevati al quadrato per evitare che gli errori negativi neutralizzino gli errori positivi. La soluzione di questo problema di minimizzazione rispetto alla pendenza e l'intercetta richiede prima il calcolo di un primo derivativo e la loro impostazione su zero:

$$\frac{d}{d\beta_0}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = 0 \text{ e } \frac{d}{d\beta_1}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = 0$$

che fornisce le equazioni dei minimi quadrati della regressione bivariata

$$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{\sum_{i=1}^{n}X_iY_i - \frac{\sum_{i=1}^{n}X_i\sum_{i=1}^{n}Y_i}{n}}{\sum_{i=1}^{n}X_i^2 - \frac{\left(\sum_{i=1}^{n}X_i\right)^2}{n}}$$

$$\beta_0 = \overline{Y} - \beta_1\overline{X}$$

Per la regressione multivariata, l'analogia è allargata per tenere conto delle multiple variabili indipendenti, dove $Y_i = \beta_1 + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i$ e la pendenza stimate possono essere calcolate mediante:

$$\hat{\beta}_2 = \frac{\sum Y_i X_{2,i} \sum X_{3,i}^2 - \sum Y_i X_{3,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$
$$\hat{\beta}_3 = \frac{\sum Y_i X_{3,i} \sum X_{2,i}^2 - \sum Y_i X_{2,i} \sum X_{2,i} X_{3,i}}{\sum X_{2,i}^2 \sum X_{3,i}^2 - \left(\sum X_{2,i} X_{3,i}\right)^2}$$

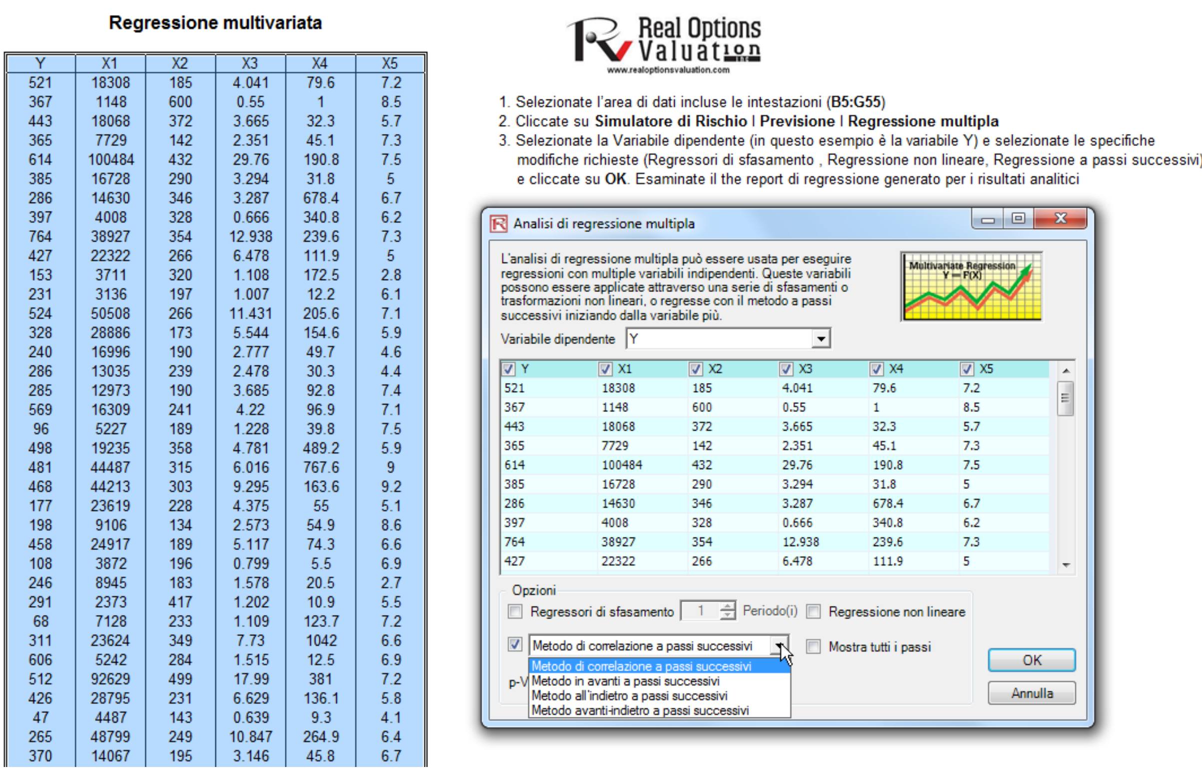
Nell'esecuzione di regressioni multivariate, bisogna prestare molta attenzione nell'impostare ed interpretare i risultati. Per esempio è richiesta una buona comprensione della modellazione econometrica (per esempio, poter identificare le trappole della regressione come le rotture strutturali, la multicollinearità, l'eteroschedasticità, l'autocorrelazione, i tests delle specifiche, le non linearità e così via) prima di poter costruire un buon modello. Consultate il libro *Modeling Risk,* 2nd Edition*: Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization* (Wiley, 2010) di Dr. Johnathan Mun per un'analisi e una discussione più dettagliata della regressione multivariata e su come identificare queste trappole della regressione.

Procedura:

- Avviate Excel e aprite se richiesto i vostri dati storici (la figura sottostante usa il file ***Regressione Multipla*** nella cartella degli esempi)
- Assicuratevi che i dati siano disposti in colonne, selezionate l'intera area di dati, incluso il nome della variabile, e selezionate ***Simulatore di Rischio | Previsione | Regressione Multipla***
- Selezionate la variabile dipendente e attivate le opzioni attinenti (sfasamenti, regressione a passi, regressione non lineare e così via) e cliccate su **OK**

Interpretare i risultati:

La Figura 3.8 illustra un esempio del report dei risultati di una regressione multivariata. Il report contiene tutti i risultati della regressione, l'analisi dei risultati di varianza, il diagramma adattato e i risultati del test di verifica d'ipotesi. I dettagli tecnici per l'interpretazione di questi risultati non rientrano nello scopo di questo Manuale dell'Utente. Consultare il libro *Modeling Risk,* 2nd Edition*: Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization* (Wiley 2010) di Dr. Johnathan Mun per un'analisi e una discussione più dettagliata sulla regressione multivariata e su come interpretare i reports della regressione.



Regressione multivariata

Y	X1	X2	X3	X4	X5
521	18308	185	4.041	79.6	7.2
367	1148	600	0.55	1	8.5
443	18068	372	3.665	32.3	5.7
365	7729	142	2.351	45.1	7.3
614	100484	432	29.76	190.8	7.5
385	16728	290	3.294	31.8	5
286	14630	346	3.287	678.4	6.7
397	4008	328	0.666	340.8	6.2
764	38927	354	12.938	239.6	7.3
427	22322	266	6.478	111.9	5
153	3711	320	1.108	172.5	2.8
231	3136	197	1.007	12.2	6.1
524	50508	266	11.431	205.6	7.1
328	28886	173	5.544	154.6	5.9
240	16996	190	2.777	49.7	4.6
286	13035	239	2.478	30.3	4.4
285	12973	190	3.685	92.8	7.4
569	16309	241	4.22	96.9	7.1
96	5227	189	1.228	39.8	7.5
498	19235	358	4.781	489.2	5.9
481	44487	315	6.016	767.6	9
468	44213	303	9.295	163.6	9.2
177	23619	228	4.375	55	5.1
198	9106	134	2.573	54.9	8.6
458	24917	189	5.117	74.3	6.6
108	3872	196	0.799	5.5	6.9
246	8945	183	1.578	20.5	2.7
291	2373	417	1.202	10.9	5.5
68	7128	233	1.109	123.7	7.2
311	23624	349	7.73	1042	6.6
606	5242	284	1.515	12.5	6.9
512	92629	499	17.99	381	7.2
426	28795	231	6.629	136.1	5.8
47	4487	143	0.639	9.3	4.1
265	48799	249	10.847	264.9	6.4
370	14067	195	3.146	45.8	6.7

Figura 3.7 – Eseguire una Regressione Multivariata

Report dell'analisi di regressione

Statistiche della regressione

R-Quadrato (Coefficiente di determinazione)	0.3272
R- Quadrato corretto	0.2508
R-Multiplo (Coefficiente di correlazione multiplo)	0.5720
Errore standard delle stime (SEy)	149.6720
Numero delle osservazioni	50

Il R-Quadrato, o Coefficiente di determinazione, indica che 0.33 della variazione nella variabile dipendente può essere spiegata e giustificata dalle variabili indipendenti in questa analisi di regressione. Tuttavia in una regressione multipla, il R-Quadrato corretto tiene conto dell'esistenza di variabili indipendenti o regressori addizionali e corregge il valore di R-Quadrato per ottenere una visione più accurata del potere esplicativo della regressione. Perciò solo 0.25 della variazione nella variabile dipendente può essere spiegato dai regressori.

Il Coefficiente di correlazione multiplo (R-Multiplo) misura la correlazione tra la variabile dipendente reale (Y) e quella stimata o adattata (Y) basato sull'equazione di regressione. Questo è anche la radice quadrata del Coefficiente di determinazione (R-Quadrato).

L'Errore standard delle Stime (SEy) descrive la the dispersione dei punti dati sopra e sotto la linea od il piano di regressione. Questo valore è usato come parte del calcolo per ottenere più tardi il livello di confidenza delle stime.

Risultati della regressione

	Intercetta	X1	X2	X3	X4	X5
Coefficienti	57.9555	-0.0035	0.4644	25.2377	-0.0086	16.5579
Errore standard	108.7901	0.0035	0.2535	14.1172	0.1016	14.7996
tatistica di t	0.5327	-1.0066	1.8316	1.7877	-0.0843	1.1188
Valore di p	0.5969	**0.3197**	**0.0738**	**0.0807**	**0.9332**	**0.2693**
5% inferiore	-161.2966	-0.0106	-0.0466	-3.2137	-0.2132	-13.2687
95% superiore	277.2076	0.0036	0.9753	53.6891	0.1961	46.3845

Gradi di libertà

Gradi di libertà per la regressione	5
Gradi di libertà per il residuo	44
Gradi di libertà totali	49

Test di verifica d'ipotesi

Statistica di t critica (99% confidenza con gradi di libertà 44)	2.6923
Statistica di t critica (95% confidenza con gradi di libertà 44)	2.0154
Statistica di t critica (90% confidenza con gradi di libertà 44)	1.6802

I coefficienti forniscono l'intercetta stimata e le pendenze stimate della regressione. Per esempio, i coefficienti sono stime dei veri valori b di popolazione nella seguente equazione di regressione Y = b0 + b1X1 + b2X2 + ... + bnXn. L'errore standard misura l'accuratezza dei coefficienti predetti e le statistiche di t sono i rapporti di ciascun coefficiente predetto con il suo errore standard.

La statistica di t è usata nel test di verifica d'ipotesi, dove impostiamo l'ipotesi nulla (Ho) in modo tale che la media reale del coefficiente sia = 0 e l'ipotesi alternativa (Ha) in modo tale che la media reale del coefficiente non sia uguale a 0. Viene eseguito un t-test e la statistica di t calcolata viene paragonata con i valori critici ai pertinenti gradi di libertà per il residuo. Il t-test è molto importante dato che calcola se ciascuno dei coefficienti è statisticamente significativo in presenza degli altri regressori. Questo significa che il t-test verifica statisticamente se un regressore o una variabile indipendente dovrebbe restare nella regressione o se dovrebbe essere rimosso o rimossa.

Il coefficiente è statisticamente significativo se la sua statistica di t calcolata supera la statistica di t critica ai pertinenti gradi di libertà (df). I tre principali livelli di confidenza usati per testare la significatività sono 90%, 95% e 99%. Se la statistica di t di un coefficiente supera il livello critico, è considerato statisticamente significativo. Alternativamente, il valore di p calcola la probabilità di occorrenza di ciascuna statistica di t, il che significa che tanto più piccolo è il valore di p, tanto più significativo è il coefficiente. I livelli di significatività normali per il valore di p sono 0,01, 0,05 e 0,10, che corrispondono ai livelli di confidenza di 99%, 95% e 99%.
I coefficienti di cui i valori di p sono evidenziati in blu indicano che sono statisticamente significativi al 90% livello di confidenza o al livello alfa di 0,10. Invece quelli evidenziati in rosso indicano che non sono statisticamente significativi a nessun altro livello di alfa.

Analisi di Varianza

	Somma dei quadrati	Media dei quadrati	Statistica di F	Valore di p	Test di verifica d'ipotesi	
Regressione	479388.49	95877.70	4.28	**0.0029**	Statistica critica di F (99% confidenza con gradi di libertà 5 e 44)	3.4651
Residuo	985675.19	22401.71			Statistica critica di F (95% confidenza con gradi di libertà 5 e 44)	2.4270
Totale	1465063.68				Statistica critica di F (90% confidenza con gradi di libertà 5 e 44)	1.9828

La tabella dell'analisi di varianza (ANOVA) fornisce un F-test della significatività statistica generale del modello di regressione. Invece di esaminare i regressori individuali come accade nel t-test, il F-test esamina tutte le proprietà statistiche dei coefficienti stimati. La statistica di F è calcolata come il rapporto della Media dei quadrati della regressione con la Media dei quadrati del residuo. Il numeratore misura quanto della regressione è spiegato, mentre il denominatore misura quanto resta non spiegato. Per cui tanto più grande è la statistica di F, tanto più significativo è il modello. Il corrispondente valore di p è calcolato per testare l'ipotesi nulla (Ho), dove tutti i coefficienti sono simultaneamente uguali a zero, contro l'ipotesi alternativa (Ha), dove sono tutti simultaneamente diversi da zero, indicando un modello di regressione generale significativo. Se il valore di p è più piccolo della significatività alfa di 0,01, 0,05 o 0,10, allora la regressione è significativa. Lo stesso metodo può essere applicato alla statistica di F paragonando la statistica di F calcolata con i valori critici di F a vari livelli di significatività.

Forecasting

Periodo	Reale (Y)	Previsione (F)	Errore (E)
1	521.0000	299.5124	221.4876
2	367.0000	487.1243	(120.1243)
3	443.0000	353.2789	89.7211
4	365.0000	276.3296	88.6704
5	614.0000	776.1336	(162.1336)
6	385.0000	298.9993	86.0007
7	286.0000	354.8718	(68.8718)
8	397.0000	312.6155	84.3845
9	764.0000	529.7550	234.2450
10	427.0000	347.7034	79.2966
11	153.0000	266.2526	(113.2526)
12	231.0000	264.6375	(33.6375)
13	524.0000	406.8009	117.1991
14	328.0000	272.2226	55.7774
15	240.0000	231.7882	8.2118
16	286.0000	257.8862	28.1138
17	285.0000	314.9521	(29.9521)
18	569.0000	335.3140	233.6860
19	96.0000	282.0356	(186.0356)
20	498.0000	370.2062	127.7938
21	481.0000	340.8742	140.1258
22	468.0000	427.5118	40.4882
23	177.0000	274.5298	(97.5298)
24	198.0000	294.7795	(96.7795)

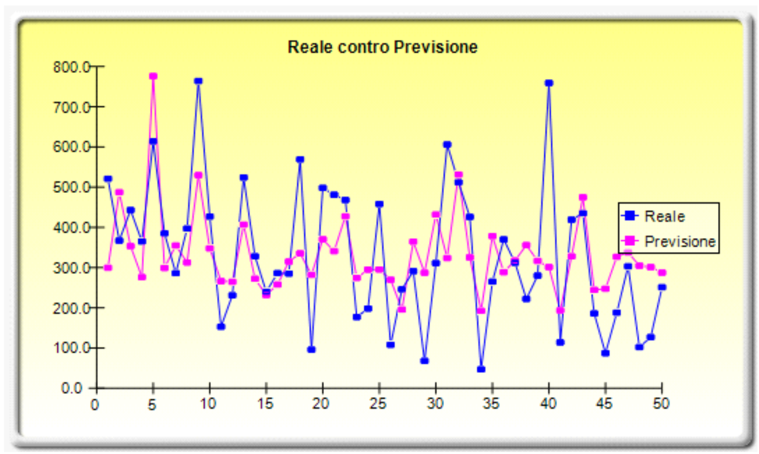


Figura 3.8 – Risultati di una Regressione Multivariata

3.5 Previsione Stocastica

Teoria:

Un processo stocastico non è altro che un'equazione matematicamente definita che può creare una serie di esiti nel tempo, esiti che non sono di natura deterministica. Vale a dire, un'equazione o un processo che non segue una regola facilmente rilevabile come, per esempio, il prezzo aumenterà di *X* percento ogni anno o i ricavi aumenteranno di questo fattore di *X* più *Y* percento. Un processo stocastico è, per definizione, non deterministico. Inoltre è possibile inserire numeri nell'equazione del processo stocastico e ottenere ogni volta dei risultati differenti. Per esempio, il percorso del prezzo di un titolo è di natura stocastica e non è possibile predire il prezzo di un titolo in modo affidabile e con certezza. Tuttavia, l'evoluzione del prezzo nel tempo è racchiusa in un processo che genera questi prezzi. Il processo è fisso e and predeterminato, ma gli esiti non lo sono. Quindi, mediante la simulazione stocastica noi creiamo multipli percorsi di prezzi, otteniamo un campionamento statistico di queste simulazioni e facciamo delle deduzioni sui percorsi potenziali che i prezzi effettivi potrebbero prendere, dato la natura e i parametri del processo stocastico usato per generare le serie temporali. Tre processi stocastici di base sono inclusi bello strumento di *Previsione* di *Simulatore di Rischio*, compreso il Moto Browniano Geometrico o passeggiata casuale. Questo è il processo più comune e prevalentemente usato a causa della sua semplicità e dell'ampio campo delle sue applicazioni. Gli altri due processi stocastici sono il Processo con ritorno alla media e il Processo di diffusione a salti.

L'aspetto interessante della simulazione con processo stocastico è che non sono necessariamente richiesti i dati storici. In altre parole, il modello non deve adattarsi ad un qualsiasi determinato insieme di dati storici. Basta semplicemente calcolare i rendimenti attesi e la volatilità dei dati storici oppure stimarli usando dati esterni paragonabili oppure fare delle ipotesi riguardo questi valori. Consultare il libro *Modeling Risk: Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization,* 2nd Edition (Wiley 2010) di Dr. Johnathan Mun per maggiori dettagli su come sono calcolati ciascuno di questi inputs (per esempio, il tasso di ritorno alla media, le probabilità del salto, la volatilità e così via).

Procedura:

- Avviate il modulo selezionando Simulatore di Rischio | Previsione | Processi Stocastici
- Selezionate il processo desiderato, inserite gli inputs richiesti e cliccate varie volte su Aggiorna diagramma, per assicuravi che il processo si stia comportando come previsto. Dopo cliccate su **OK** (Figura 3.9)

Interpretare i Risultati:

La Figura 3.10 mostra i risultati di un esempio di un processo stocastico. Il diagramma mostra l'esempio di un insieme di iterazioni, mentre il report spiega i principi base dei processi stocastici. Sono inoltre forniti i valori della previsione (media e deviazione standard) per ciascun periodo temporale. Usando questi valori, potete decidere quale periodo temporale è attinente alla vostra analisi e impostare le ipotesi, basato su questi valori della media e della deviazione standard, usando la distribuzione normale. Queste ipotesi possono poi essere simulate nel vostro modello personalizzato.

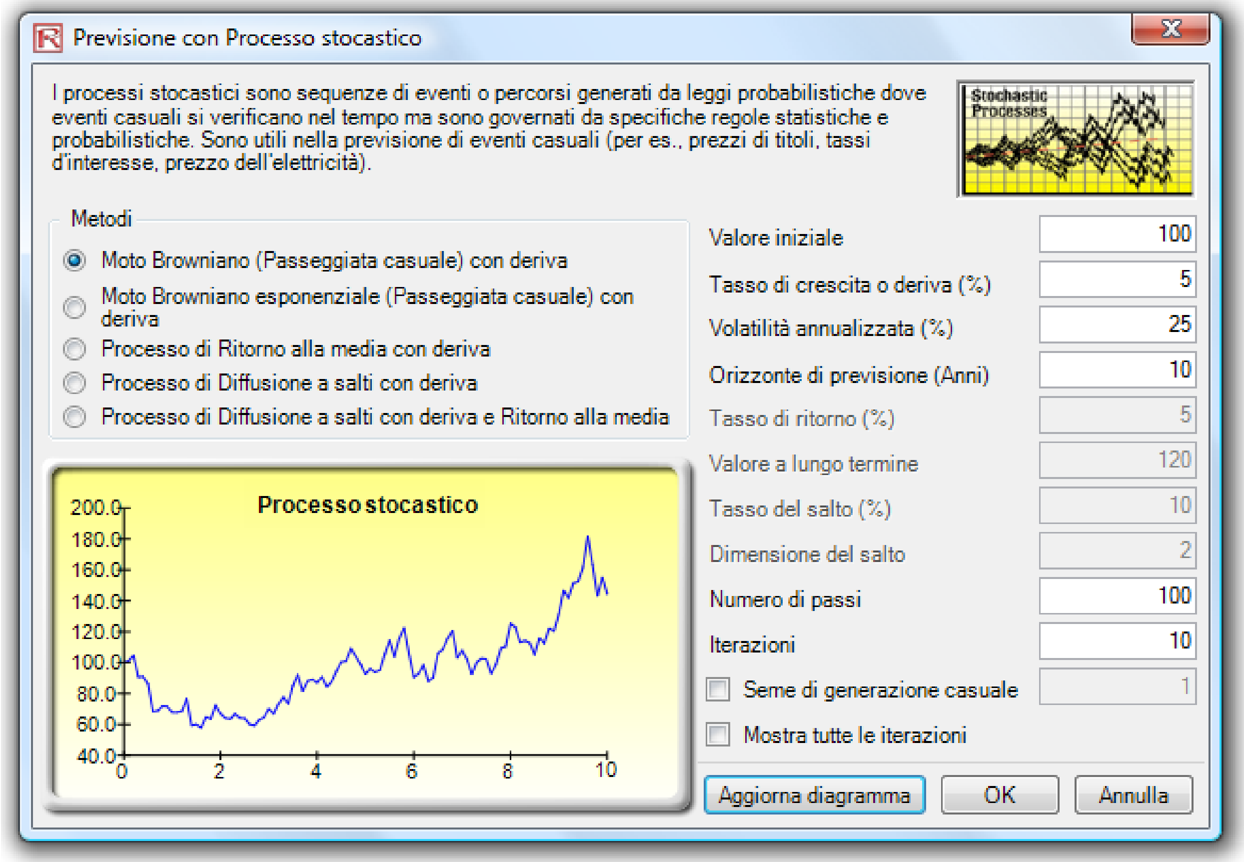


Figura 3.9 – Previsione con Processi Stocastici

Previsione con processi stocastici

Sommario statistico

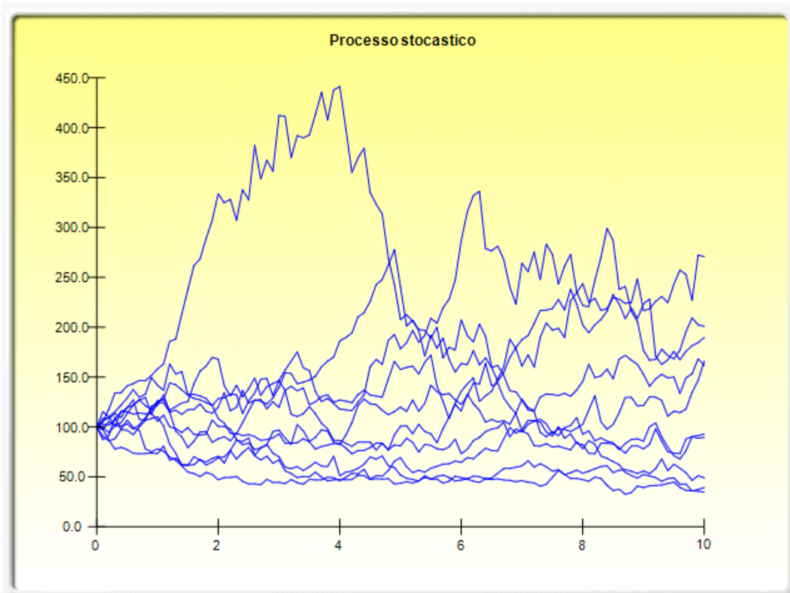
Un processo stocastico è una sequenza di eventi o percorsi generati da leggi probabilistiche. In altre parole, eventi casuali possono verificarsi nel tempo, ma essi sono governati da specifiche regole statistiche e probabilistiche. I principali processi stocastici includono Passeggiata casuale o Moto Browniano, Ritorno alla media e Diffusione a salti. Questi processi possono essere usati per prevedere una molteplicità di variabili che sembrano seguire tendenze casuali, ma che sono tuttavia limitate da leggi probabilistiche.

Il processo di Passeggiata casuale di Moto Browniano può essere usato per prevedere i prezzi di titoli, i prezzi di commodities ed altri dati di serie temporali, dato una deriva o tasso di crescita e una volatilità intorno al percorso della deriva. Il processo di Ritorno alla media può essere usato per ridurre le fluttuazioni del processo di Passeggiata casuale permettendo al percorso di stabilire come obiettivo un valore a lungo termine. Questo lo rende utile nella previsione di variabili di serie temporali che hanno un tasso a lungo termine, come i tassi di interesse e dell'inflazione (questi sono tassi di obiettivo a lungo termine delle autorità di regolamentazione o del mercato). Il processo di Diffusione a salti è utile nella previsione di dati di serie temporali quando la variabile può occasionalmente manifestare salti casuali, come nel caso dei prezzi del petrolio o dell'elettricità (gli shock di un evento esogeno discreto possono fare balzare i prezzi in alto o in basso). In conclusione, questi tre processi stocastici possono essere combinati e composti come richiesto.

I risultati sulla destra indicano la media e la deviazione standard di tutte le iterazioni generate ad ogni passo temporale. Se è stata selezionata l'opzione the Mostra tutte le iterazioni, il percorso di ogni iterazione sarà mostrato in un foglio di lavoro separato. Il grafico generato sottostante mostra un insieme campione di percorsi di iterazione.

Processo stocastico: Moto Browniano (Passeggiata casuale) con deriva

Valore iniziale	100	Passi	100.00	Tasso del salto	N/A		
Tasso di deriva	5.00%	Iterazioni	10.00	Dimensione del salto	N/A		
Volatilità	25.00%	Tasso di ritorno	N/A	Seme di generazione casuale	937426889		
Orizzonte	10	Valore a lungo termine	N/A				



Tempo	Media	Devst
0.0000	100.00	0.00
0.1000	100.94	8.29
0.2000	101.29	11.00
0.3000	104.86	15.36
0.4000	107.60	15.21
0.5000	112.11	17.61
0.6000	112.79	21.78
0.7000	112.37	21.02
0.8000	111.68	22.90
0.9000	114.07	25.63
1.0000	116.43	26.12
1.1000	118.10	26.51
1.2000	115.88	39.15
1.3000	112.36	39.91
1.4000	111.20	48.60
1.5000	110.13	54.12
1.6000	114.75	60.38
1.7000	117.71	62.63
1.8000	118.77	68.60
1.9000	121.05	73.93
2.0000	123.38	81.50
2.1000	120.71	78.04
2.2000	118.79	78.00
2.3000	118.03	72.11
2.4000	120.09	81.10
2.5000	121.91	77.54
2.6000	127.23	95.91
2.7000	123.50	85.90
2.8000	126.75	90.45
2.9000	125.12	87.54
3.0000	133.57	104.14
3.1000	132.72	106.61
3.2000	126.78	95.39
3.3000	129.74	101.59
3.4000	127.99	100.66
3.5000	126.69	101.20
3.6000	127.11	106.54
3.7000	129.05	113.80
3.8000	126.57	106.06
3.9000	127.73	115.17
4.0000	126.86	118.57
4.1000	123.77	105.92
4.2000	120.40	93.48
4.3000	128.23	97.73
4.4000	132.36	100.16
4.5000	130.54	90.72
4.6000	131.49	90.20
4.7000	130.69	87.93
4.8000	129.77	83.86
4.9000	133.62	81.86
5.0000	125.03	68.10
5.1000	121.69	63.65
5.2000	122.10	66.63
5.3000	117.38	61.94
5.4000	118.11	59.32

Figura 3.10 – Risultati della Previsione Stocastica

3.6 Estrapolazione Non lineare

Teoria:

L'estrapolazione consiste nel fare proiezioni statistiche usando trend storici che sono proiettati nel futuro per un periodo specifico. È utilizzata solo per previsioni di serie temporali. Per dati in sezione trasversale o dati panel misti (serie temporali con dati in sezione), è più adatta una regressione multivariata. Questa metodologia è utile, quando non sono previsti cambiamenti rilevanti, vale a dire, è previsto che i fattori causali rimangano costanti, o quando i fattori causali di una situazione non sono chiaramente compresi. Aiuta anche a scoraggiare l'inserimento di preconcetti personali nel processo. L'estrapolazione è abbastanza affidabile, relativamente semplice ed economica. Tuttavia l'estrapolazione, che presume che sia le tendenze recenti che quelle storiche perdureranno, produce grandi errori di previsione se si verificano delle discontinuità entro il periodo temporale proiettato. In altre parole, la pura estrapolazione di serie temporali presume che tutto ciò che dobbiamo sapere è contenuto nei valori storici della serie che stiamo prevedendo. Se presumiamo che il comportamento passato è un buon predittore del comportamento futuro, l'estrapolazione è interessante. Questo la rende un metodo utile, quando sono richieste solamente molte previsioni a breve termine.

Questa metodologia esegua una stima della funzione f(x) per qualsiasi valore arbitrario di x. interpolando una curva regolare non lineare attraverso tutti i valori di x, ed estrapola valori futuri di x oltre l'insieme di dati storici usando questa curva regolare. La metodologia utilizza o la forma funzionale polinomiale o la forma funzionale razionale (un rapporto di due polinomiali). La forma funzionale polinomiale è normalmente sufficiente per dati con un "buon comportamento", tuttavia le forme funzionali razionali sono talvolta più precise (specialmente con funzioni polari, vale a dire funzioni con denominatori che si avvicinano a zero).

Procedura:

- Avviate Excel e aprite se richiesto I vostri dati storici (la figura mostrata di seguito usa il file ***Estrapolazione non lineare*** dalla cartella degli esempi)
- Selezionate i dati di serie temporali e selezionate ***Simulatore di Rischio | previsione | Estrapolazione non lineare***
- Selezionate il tipo di estrapolazione (selezione automatica, polinomiale o funzione razionale), inserite il numero desiderato di periodi di previsione (Figure 3.11) e cliccate su **OK**

Interpretare i Risultati:

Il report dei risultati mostrato nella Figura 3.12 indica i valori estrapolati della previsione, le misure d'errore e la rappresentazione grafica dei risultati dell'estrapolazione. Le misure d'errore sono da usare per controllare la validità della previsione e sono particolarmente importanti, se usate per confrontare la qualità della previsione e la precisione dell'estrapolazione rispetto all'analisi di serie temporali.

Note: Quando i dati storici sono graduati e seguono certi schemi e certe curve non lineari, l'estrapolazione è più adatta dell'analisi di serie temporali. Tuttavia, quando gli schemi dei dati seguono cicli stagionali ed una tendenza, l'analisi di serie temporali fornirà risultati migliori.

Ricavi da vendite storici
Tassi di crescita polinomiali

Anno	*Mese*	*Periodo*	*Vendite*
2004	*1*	*1*	*$1.00*
2004	*2*	*2*	*$6.73*
2004	*3*	*3*	*$20.52*
2004	*4*	*4*	*$45.25*
2004	*5*	*5*	*$83.59*
2004	*6*	*6*	*$138.01*
2004	*7*	*7*	*$210.87*
2004	*8*	*8*	*$304.44*
2004	*9*	*9*	*$420.89*
2004	*10*	*10*	*$562.34*
2004	*11*	*11*	*$730.85*
2004	*12*	*12*	*$928.43*

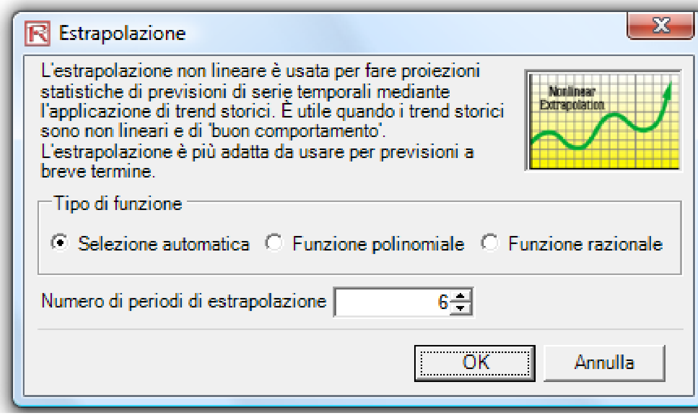


Figura 3.11 – Eseguire una Estrapolazione non lineare

Estrapolazione non lineare

Sommario statistico

L'estrapolazione consiste nel fare proiezioni statistiche usando trend storici che sono proiettati nel futuro per un periodo specifico. È utilizzata solo per previsioni di serie temporali. Per dati in sezione o dati panel misti (serie temporali con dati in sezione) è più adatta una regressione multivariata. Questa metodologia è utile quando non sono previsti cambiamenti rilevanti, vale a dire, è previsto che i fattori causali rimangano costanti, o quando i fattori causali di una situazione non sono chiaramente compresi. Aiuta anche a scoraggiare l'inserimento di preconcetti personali nel processo. L'estrapolazione è abbastanza affidabile, relativamente semplice ed economica. Tuttavia l'estrapolazione, che presume che sia i trend recenti che quelli storici proseguiranno, produce grandi errori di previsione, se si verificano delle discontinuità entro il periodo temporale proiettato. In altre parole, la pura estrapolazione di serie temporali presume che tutto ciò che dobbiamo sapere è contenuto nei valori storici della serie che stiamo prevedendo. Se presumiamo che il comportamento passato è un buon predittore del comportamento futuro, l'estrapolazione è interessante. Questo la rende un metodo utile, quando sono richieste solamente molte previsioni a breve termine.

Questa metodologia esegua una stima della funzione f(x) per qualsiasi arbitrario valore di x interpolando una curva regolare non lineare attraverso tutti i valori di x ed estrapola futuri valori di x oltre l'insieme di dati storici usando questa curva regolare. La metodologia utilizza o la forma funzionale polinomiale o la forma funzionale razionale (un rapporto di due polinomiali). La forma funzionale polinomiale è normalmente sufficiente per dati con 'comportamento buono', tuttavia le forme funzionali razionali sono talvolta più precise (specialmente con funzioni polari, vale a dire, funzioni con denominatori che si avvicinano a zero).

Periodo	Reale	Adattamento previsione
1	1.00	
2	6.73	1.00
3	20.52	-1.42
4	45.25	99.82
5	83.59	55.92
6	138.01	136.71
7	210.87	211.96
8	304.44	304.43
9	420.89	420.89
10	562.34	562.34
11	730.85	730.85
12	928.43	928.43
Previsione 13		1157.03

Misure d'errore	
Radice dell'errore quadratico medio (RMSE)	19.6799
Errore quadratico medio (MSE)	387.2974
Deviazione assoluta media (MAD)	10.2095
Errore percentuale assoluto medio (MAPE)	31.56%
U di Theil	1.1210

Tipo di funzione: Razionale

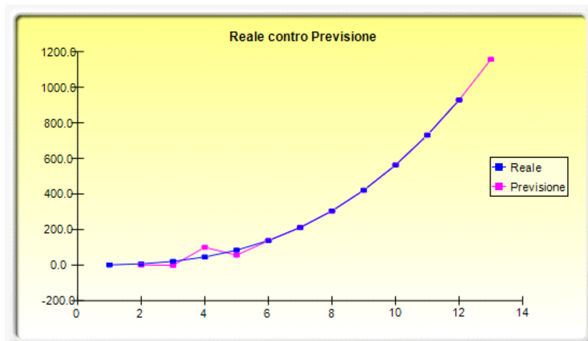


Figura 3.12 – Risultati di una Estrapolazione non lineare

3.7 ARIMA Box-Jenkins Serie temporali avanzate

Teoria:

Uno strumento molto potente di previsioni avanzate di serie temporali è il metodo ARIMA o *Media mobile integrata autoregressiva*. La previsione ARIMA raccoglie tre strumenti separati in un modello generale. Il primo segmento di strumento è il termine autoregressiva o "AR", che corrisponde al numero di valori sfasati del residuo nel modello non condizionato di previsione. Sostanzialmente, il modello cattura la variazione storica di dati reali rispetto ad un modello di previsione e usa questa variazione o residuo per creare un miglior modello predittivo. Il secondo segmento di strumento è il termine di ordine d'integrazione o "I". Questo termine d'integrazione corrisponde al numero di differenziazione a cui è sottoposta la serie temporale che deve essere prevista. Questo elemento rende conto degli eventuali tassi di crescita non lineari che esistono nei dati. Il terzo segmento di strumento è il termine media mobile o "MA", che è sostanzialmente la media mobile degli errori sfasati di previsione. Incorporando questi errori sfasati di previsione, il modello essenzialmente impara dai suoi errori di previsione e li corregge attraverso un calcolo di media mobile. Il modello ARIMA segue la metodologia Box-Jenkins, dove ogni termine rappresenta i passi intrapresi nella costruzione del modello fino a che rimane solo il rumore bianco. Inoltre, la modellazione ARIMA usa le tecniche di correlazione nella generazione delle previsioni. ARIMA può essere usata per modellare schemi che potrebbero non essere visibili in dati diagrammati. I modelli ARIMA possono anche essere combinati con variabili esogene, ma assicuratevi che le variabili esogene abbiano abbastanza punti dati da coprire il numero addizionale di periodi da prevedere. Per finire, siate consapevoli, che a causa della complessità dei modelli, questo modulo richiede più tempo per la sua esecuzione.

Ci sono molte ragioni per la superiorità del modello ARIMA rispetto alle più comuni analisi di serie temporali e regressioni multivariate. La conclusione comune nell'analisi di serie temporali e nella regressione multivariata è che i residui degli errori sono correlati con i propri valori sfasati. Questa correlazione seriale viola l'ipotesi standard della teoria della regressione, ovvero che le perturbazioni non sono correlate con altre perturbazioni. I problemi principali associati alla correlazione seriale sono:

- L'analisi di regressione e l'analisi base di serie temporali non sono più efficienti tra i diversi stimatori lineari. Tuttavia, dato che i residui degli errori possono aiutare a predire i residui attuali degli errori, possiamo utilizzare queste informazioni per costruire una predizione migliore della variabile dipendente usando ARIMA.

- Gli errori standard calcolati usando la formula di regressione e la formula di serie temporali non sono giusti e sono generalmente sottovalutati. Inoltre, se ci sono variabili sfasate dipendenti impostate come i regressori, le stime della regressione sono distorte e incongruenti, ma possono essere corrette usando ARIMA.

I modelli della Media mobile integrata autoregressiva o ARIMA(p,d,q) sono l'estensione del modello AR che usa tre componenti per modellare la correlazione seriale nei dati di serie temporali. Il primo componente è il termine autoregressivo (AR). Il modello AR(p) usa gli sfasamenti p delle serie temporali nell'equazione. Un modello AR(p) ha la forma: $y_t = a_1 y_{t-1} + ... + a_p y_{t-p} + e_t$. Il secondo componente è il termine ordine d'integrazione (d). Ciascun ordine d'integrazione corrisponde ad una differenziazione delle serie temporali. I(1) significa differenziare i dati una volta. I(d) significa differenziare i dati d volte. Il terzo componente è il termine media mobile (MA). Il modello MA(q) usa gli sfasamenti q degli errori della previsione per migliorare la previsione. Un modello MA(q) ha la forma $y_t = e_t + b_1 e_{t-1} + ... + b_q e_{t-q}$. Infine, un modello ARMA(p,q) ha la forma combinata: $y_t = a_1\, y_{t-1} + ... + a_p\, y_{t-p} + e_t + b_1\, e_{t-1} + ... + b_q\, e_{t-q}$.

Procedura:

- Avviate Excel ed inserite I vostri dati o aprite un foglio di lavoro con dati storici da prevedere (la mostrata di seguito usa il file d'esempio ***ARIMA***) ***Serie Temporali***
- Selezionate i dati di serie temporali e selezionate ***Simulatore di Rischio | Previsione | ARIMA***
- Inserite i pertinenti parametri per P, D e Q (solo numeri interi positivi) ed il numero desiderato di periodi di previsione e cliccate su ***OK***

Nota per ARIMA and AUTO-ARIMA:

- Per ARIMA e Auto-ARIMA, potete modellare e prevedere periodi futuri usando solo la variabile dipendente (*Y*), vale a dire, la ***Variabile di Serie temporali*** da sola, oppure potete aggiungere addizionali Variabili Esogene ($X_1, X_2, \ldots, X_n$) esattamente come in un'analisi di regressione, dove ci sono multiple variabili indipendenti. Potete eseguire quanti periodi di previsione desiderate, se usate solo la variabile di serie temporali (*Y*). Tuttavia, se aggiungete variabili esogene (*X*), prego notare che i vostri periodi di previsione sono limitati al numero dei periodi di dati delle variabili esogene meno i periodi di dati della variabile di serie temporali. Per esempio, potete prevedere solo fino a 5 periodi se avete dati storici di serie temporali per 100 periodi e solo se avete variabili esogene per 105 periodi (100 periodi storici per uguagliare la variabile di serie temporali e 5 periodi futuri aggiuntivi di variabili indipendenti esogene per prevedere la variabile dipendente di serie temporali).

Interpretare i Risultati:

Nell'interpretare i risultati di un modello ARIMA, la maggior parte delle specifiche è identica come nell'analisi di regressione multivariata (consultare il libro *Modeling Risk*, 2nd Edition di Dr. Johnathan Mun per maggiori dettagli sull'interpretazione dell'analisi di regressione multivariata e dei modelli ARIMA). Ci sono tuttavia vari addizionali insiemi di risultati specifici all'analisi ARIMA, come visto nella Figura 3.14. Il primo è l'aggiunta del Criterio informativo di Akaike (AIC) e del Criterio di Schwarz (SC), usati spesso nella selezione e identificazione del modello ARIMA. Vale a dire, AIC e SC sono usati per determinare se un particolare modello con un insieme specifico di parametri p, d e q è un buon adattamento statistico. SC impone una penalità maggiore per coefficienti addizionali di AIC, ma generalmente si deve scegliere il modello con i valori più bassi di AIC e SC. Per finire, nel report ARIMA viene fornito un insieme addizionale di risultati, denominati le statistiche di autocorrelazione (AC) e di autocorrelazione parziale (PAC).

Per esempio, l'autocorrelazione AC(1) è non-zero, significa che la serie è correlata serialmente di primo ordine. Se AC(k) si esaurisce più o meno geometricamente con uno sfasamento crescente, questo implica che la serie segue un processo autoregressivo di basso ordine. Se AC(k) scende a zero dopo un numero limitato di sfasamenti, questo implica che la serie segue un processo a media mobile di basso ordine. Per contrasto, PAC misura la correlazione di valori che sono separati da k periodi dopo la rimozione della correlazione dagli sfasamenti intercorsi. Se lo schema di autocorrelazione può essere catturato da una autoregressione di ordine minore di k, allora l'autocorrelazione parziale allo sfasamento k sarà vicino a zero. Vengono fornite anche le statistiche di Q di Ljung-Box ed i loro valori di p allo sfasamento k, dove l'ipotesi nulla che viene testata è tale, che non esiste autocorrelazione fino all'ordine *k*. Le linee punteggiate nei tracciati delle autocorrelazioni sono i due approssimativi estremi standard d'errore. Se l'autocorrelazione si trova entro questi estremi, essa non è significativamente diversa da zero al livello di significatività (approssimativo) di 5%. Trovare il giusto modello ARIMA richiede pratica ed esperienza. AC, PAC, SC e AIC sono strumenti diagnostici molto utili nell'aiutare a identificare le giuste specifiche del modello.

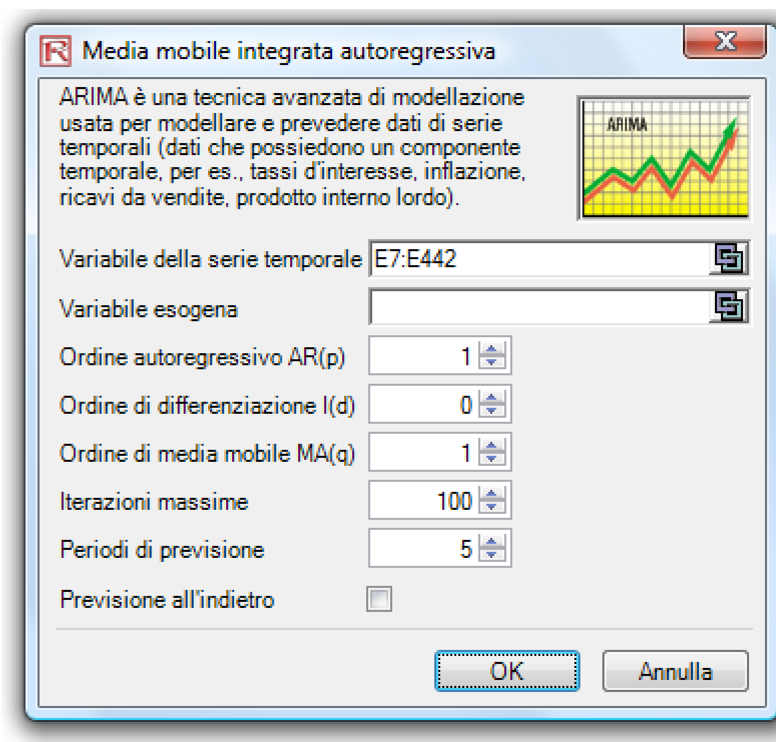


Figura 3.13A - Strumento di Previsione ARIMA Box Jenkins

ARIMA (Media mobile integrata autoregressiva)

Statistiche della regressione

R-Quadrato (Coefficiente di determinazione)	0.9999	Criterio informativo di Akaike (AIC)	4.6213
R- Quadrato corretto	0.9999	Criterio di Schwarz (SC)	4.6632
R-Multiplo (Coefficiente di correlazione multiplo)	1.0000	Probabilità Log	-1005.13
Errore standard delle stime (SEy)	297.52	Statistica di Durbin-Watson (DW)	1.8588
Numero di osservazioni	435	Numero di iterazioni	5

I modelli a media mobile integrata autoregressiva o ARIMA(p,d,q) sono l'estensione dei modelli AR che usano tre componenti per modellare la correlazione seriale nei dati di serie temporali. Il primo componente è il termine autoregressivo (AR). Il modello AR(p) usa gli sfasamenti p delle serie temporali nell'equazione. Un modello AR(p) ha la forma: y(t)=a(1)*y(t-1)+...+a(p)*y(t-p)+e(t).Il secondo componente è il termine ordine d'integrazione (d). Ciascun ordine d'integrazione corrisponde ad una differenziazione delle serie temporali. I(1) significa differenziare i dati una volta. I(d) significa differenziare i dati d volte. Il terzo componente è il termine media mobile (MA). Il modello MA(q) usa gli sfasamenti q degli errori della previsione per migliorare la previsione. Un modello MA(q) ha la forma: y(t)=e(t)+b(1)*e(t-1)+...+b(q)*e(t-q). Infine, un modello ARMA(p,q) ha la forma combinata: y(t)=a(1)*y(t-1)+...+a(p)*y(t-p)+e(t)+b(1)*e(t-1)+...+b(q)*e(t-q).

Il R-Quadrato, o Coefficiente di determinazione, indica la variazione percentuale nella variabile dipendente che può essere spiegata e giustificata dalle variabili indipendenti in quest'analisi di regressione. Tuttavia in una regressione multipla, il R-Quadrato corretto tiene conto dell'esistenza di variabili indipendenti o regressori addizionali e corregge il valore di R-Quadrato per ottenere una visione più accurata del potere esplicativo della regressione. Ma in alcune situazioni di modellazione ARIMA (per esempio con modelli non convergenti), il R-Quadrato tende ad essere inaffidabile.

Il Coefficiente di correlazione multiplo (R-Multiplo) misura la correlazione tra la variabile dipendente reale (Y) e quella stimata o adattata (Y) basato sull'equazione di regressione. Questo è anche la radice quadrata del Coefficiente di determinazione (R-Quadrato).

L'Errore standard delle Stime (SEy) descrive la the dispersione dei punti dati sopra e sotto la linea od il piano di regressione. Questo valore è usato come parte del calcolo per ottenere più tardi il livello di confidenza delle stime.

AIC e SC sono usati spesso nella selezione del modello. SC impone una penalità maggiore per coefficienti addizionali. Generalmente l'utente dovrebbe scegliere un modello con il valore più basso di AIC e SC.

La statistica di Durbin-Watson misura la correlazione seriale nei residui. Generally, Una statistica DW minore di 2 implica generalmente una correlazione seriale positiva

Risultati della regressione

	Intercetta	AR(1)	MA(1)
Coefficienti	-0.0626	1.0055	0.4936
Errore standard	0.3108	0.0006	0.0420
Statistica di t	-0.2013	1691.1373	11.7633
Valore di p	0.8406	**0.0000**	**0.0000**
5% inferiore	0.4498	1.0065	0.5628
Upper 95% superiore	-0.5749	1.0046	0.4244

Gradi di libertà		**Test di verifica d'ipotesi**	
Gradi di libertà per la regressione	2	Statistica critica di t (99% confidenza con gradi di libertà 4	2.5873
Gradi di libertà per il residuo	432	Statistica critica di t (95% confidenza con gradi di libertà 4	1.9655
Gradi di libertà totali	434	Statistica critica di t (90% confidenza con gradi di libertà 4	1.6484

I coefficienti forniscono l'intercetta stimata e le pendenze stimate della regressione. Per esempio, i coefficienti sono stime dei veri valori b di popolazione nella seguente equazione di regressione Y = b0 + b1X1 + b2X2 + ... + bnXn. L'errore standard misura l'accuratezza dei coefficienti predetti e le statistiche di t sono i rapporti di ciascun coefficiente predetto con il suo errore standard.

La statistica di t è usata nel test di verifica d'ipotesi, dove impostiamo l'ipotesi nulla (Ho) in modo tale che la media reale del coefficiente sia = 0 e l'ipotesi alternativa (Ha) in modo tale che la media reale del coefficiente non sia uguale a 0. Viene eseguito un t-test e la statistica di t calcolata viene paragonata con i valori critici ai pertinenti gradi di libertà per il residuo. Il t-test è molto importante dato che calcola se ciascuno dei coefficienti è statisticamente significativo in presenza degli altri regressori. Questo significa che il t-test verifica statisticamente se un regressore o una variabile indipendente dovrebbe restare nella regressione o se dovrebbe essere rimosso o rimossa.

Il coefficiente è statisticamente significativo se la sua statistica di t calcolata supera la statistica di t critica ai pertinenti gradi di libertà (df). I tre principali livelli di confidenza usati per testare la significatività sono 90%, 95% e 99%. Se la statistica di t di un coefficiente supera il livello critico, è considerato statisticamente significativo. Alternativamente, il valore di p calcola la probabilità di occorrenza di ciascuna statistica di t, il che significa che tanto più piccolo è il valore di p, tanto più significativo è il coefficiente. I livelli di significatività normali per il valore di p sono 0,01, 0,05 e 0,10, che corrispondono ai livelli di confidenza di 99%, 95% e 99%.

I coefficienti di cui i valori di p sono evidenziati in blu indicano che sono statisticamente significativi al 90% livello di confidenza o al livello alfa di 0,10. Invece quelli evidenziati in rosso indicano che non sono statisticamente significativi a nessun altro livello di alfa.

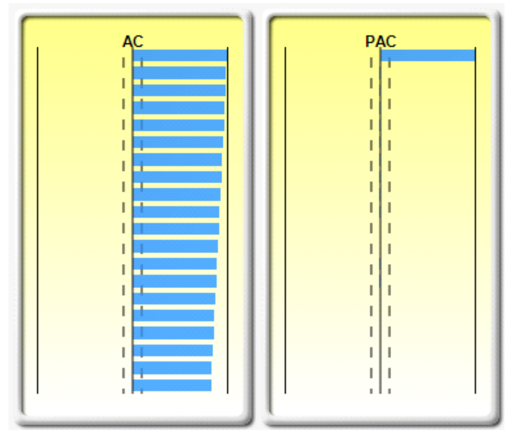
Analisi di Varianza

	Somma dei quadrati	**Media dei quadrati**	**Statistica di F**	**Valore di p**	**Test di verifica d'ipotesi**	
Regressione	38415447.53	19207723.76	3171851.1	**0.0000**	Statistica critica di F (99% confidenza con gradi di libertà 2	4.6546
Residuo	2616.05	6.06			Statistica critica di F (95% confidenza con gradi di libertà 2	3.0166
Totale	38418063.58				Statistica critica di F (90% confidenza con gradi di libertà 2	2.3149

La tabella dell'analisi di varianza (ANOVA) fornisce un F-test della significatività statistica generale del modello di regressione. Invece di esaminare i regressori individuali come accade nel t-test, il F-test esamina tutte le proprietà statistiche dei coefficienti stimati. La statistica di F è calcolata come il rapporto della Media dei quadrati della regressione con la Media dei quadrati del residuo. Il numeratore misura quanto della regressione è spiegato, mentre il denominatore misura quanto resta non spiegato. Per cui tanto più grande è la statistica di F, tanto più significativo è il modello. Il corrispondente valore di p è calcolato per testare l'ipotesi nulla (Ho), dove tutti i coefficienti sono simultaneamente uguali a zero, contro l'ipotesi alternativa (Ha), dove sono tutti simultaneamente diversi da zero, indicando un modello di regressione generale significativo. Se il valore di p è più piccolo della significatività alfa di 0,01, 0,05 o 0,10, allora la regressione è significativa. Lo stesso metodo può essere applicato alla statistica di F paragonando la statistica di F calcolata con i valori critici di F a vari livelli di significatività.

Autocorrelazione

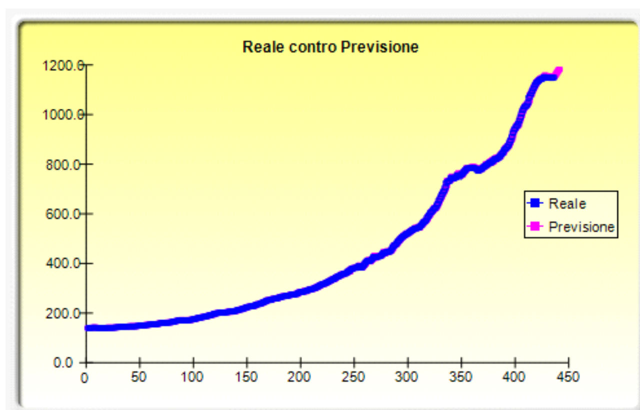
Periodo	AC	PAC	Limite inferiore	mite superiore	Statistica di Q	Prob
1	0.9921	0.9921	(0.0958)	0.0958	431.1216	-
2	0.9841	(0.0105)	(0.0958)	0.0958	856.3037	-
3	0.9760	(0.0109)	(0.0958)	0.0958	1,275.4818	-
4	0.9678	(0.0142)	(0.0958)	0.0958	1,688.5499	-
5	0.9594	(0.0098)	(0.0958)	0.0958	2,095.4625	-
6	0.9509	(0.0113)	(0.0958)	0.0958	2,496.1572	-
7	0.9423	(0.0124)	(0.0958)	0.0958	2,890.5594	-
8	0.9336	(0.0147)	(0.0958)	0.0958	3,278.5669	-
9	0.9247	(0.0121)	(0.0958)	0.0958	3,660.1152	-
10	0.9156	(0.0139)	(0.0958)	0.0958	4,035.1192	-
11	0.9066	(0.0049)	(0.0958)	0.0958	4,403.6117	-
12	0.8975	(0.0068)	(0.0958)	0.0958	4,765.6032	-
13	0.8883	(0.0097)	(0.0958)	0.0958	5,121.0697	-
14	0.8791	(0.0087)	(0.0958)	0.0958	5,470.0032	-
15	0.8698	(0.0064)	(0.0958)	0.0958	5,812.4256	-
16	0.8605	(0.0056)	(0.0958)	0.0958	6,148.3694	-
17	0.8512	(0.0062)	(0.0958)	0.0958	6,477.8620	-
18	0.8419	(0.0038)	(0.0958)	0.0958	6,800.9622	-
19	0.8326	(0.0003)	(0.0958)	0.0958	7,117.7709	-
20	0.8235	0.0002	(0.0958)	0.0958	7,428.3952	-



Se l'autocorrelazione AC(1) è non-zero, significa che la serie è correlata serialmente di primo ordine. Se AC(k) si esaurisce più o meno geometricamente con uno sfasamento crescente, questo implica che la serie segue un processo autoregressivo di basso ordine. Se AC(k) scende a zero dopo un numero limitato di sfasamenti, questo implica che la serie segue un processo a media mobile di basso ordine. Una correlazione parziale PAC(k) misura la correlazione di valori che sono separati da k periodi dopo la rimozione della correlazione dagli sfasamenti intercorsi. Se lo schema di autocorrelazione può essere catturato da una autoregressione di ordine minore di k, allora l'autocorrelazione parziale allo sfasamento k sarà vicino a zero. Le statistiche di Q di Ljung-Box ed i loro valori di p allo sfasamento k hanno l'ipotesi nulla che non esiste nessuna autocorrelazione fino all'ordine k. Le linee punteggiate nei tracciati delle autocorrelazioni sono i due approssimativi limiti standard d'errore. Se l'autocorrelazione è entro questi estremi, essa non è significativamente diversa da zero al (approssimativo) livello di significatività di 5%.

Previsione

Periodo	Reale (Y)	Previsione (F)	Errore (E)
2	139.4000	139.6056	(0.2056)
3	139.7000	140.0069	(0.3069)
4	139.7000	140.2586	(0.5586)
5	140.7000	140.1343	0.5657
6	141.2000	141.6948	(0.4948)
7	141.7000	141.6741	0.0259
8	141.9000	142.4339	(0.5339)
9	141.0000	142.3587	(1.3587)
10	140.5000	141.0466	(0.5466)
11	140.4000	140.9447	(0.5447)
12	140.0000	140.8451	(0.8451)
13	140.0000	140.2946	(0.2946)
14	139.9000	140.5663	(0.6663)
15	139.8000	140.2823	(0.4823)
16	139.6000	140.2726	(0.6726)
17	139.6000	139.9775	(0.3775)
18	139.6000	140.1232	(0.5231)
19	140.2000	140.0513	0.1487
20	141.3000	140.9862	0.3138
21	141.2000	142.1738	(0.9738)
22	140.9000	141.4377	(0.5377)
23	140.9000	141.3513	(0.4513)
24	140.7000	141.3939	(0.6939)
25	141.1000	141.0731	0.0270
26	141.6000	141.8311	(0.2311)
27	141.9000	142.2065	(0.3065)
28	142.1000	142.4709	(0.3709)
29	142.7000	142.6402	0.0598
30	142.9000	143.4561	(0.5561)
31	142.9000	143.3532	(0.4532)
32	143.5000	143.4040	0.0960
33	143.8000	144.2784	(0.4784)
34	144.1000	144.2966	(0.1966)
35	144.8000	144.7374	0.0626
36	145.2000	145.5692	(0.3692)
37	145.2000	145.7582	(0.5582)
38	145.7000	145.6649	0.0351
39	146.0000	146.4605	(0.4605)
40	146.4000	146.5176	(0.1176)
41	146.8000	147.0891	(0.2891)



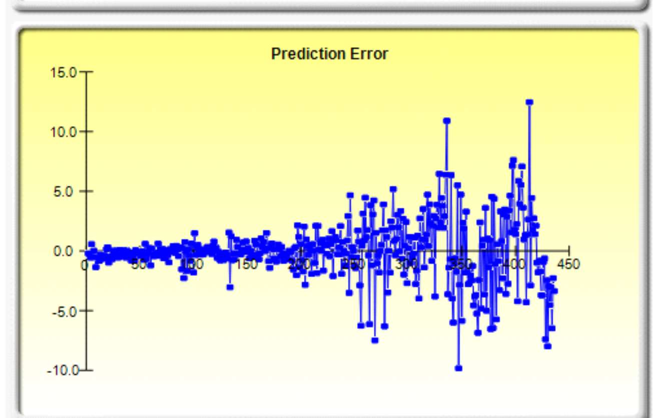


Figura 3.13B - Report della Previsione ARIMA Box Jenkins

3.8 AUTO-ARIMA (ARIMA Box-Jenkins Serie Temporali avanzate)

Teoria:

Questo strumento fornisce analisi identiche al modulo ARIMA, con la differenza che il modulo Auto-ARIMA automatizza alcuni processi della modellazione ARIMA tradizionale, testando automaticamente multiple permutazioni di specifiche del modello, e fornisce il modello col miglior adattamento. L'esecuzione di Auto-ARIMA è simile all'esecuzione delle normali previsioni ARIMA. La differenza è che non sono più richiesti gli inputs di P, D e Q e che combinazioni diverse di questi inputs sono automaticamente eseguite e confrontate.

Procedura:

- Avviate Excel ed inserite i vostri dati o aprite un foglio di lavoro esistente con dati storici da prevedere (l'illustrazione mostrata nella Figura 3.14 usa il file d'esempio ***Modelli avanzati di previsione*** nel menu ***Esempi*** di Simulatore di Rischio)
- Nel foglio di lavoro ***Auto ARIMA***, selezionate ***Simulatore di Rischio | Previsione | AUTO-ARIMA.*** Potete anche accedere a questo metodo attraverso la barra delle icone Previsione o cliccando col tasto destro del mouse in qualsiasi parte del modello e selezionando il menu di previsione a scelta.
- Cliccate sull'icona Collegamento e collegate ai dati esistenti di serie temporali. Poi inserite il numero desiderato di periodi di previsione e cliccate su ***OK***

Nota per ARIMA and AUTO-ARIMA:

- Per ARIMA e Auto-ARIMA, potete modellare e prevedere periodi futuri usando solo la variabile dipendente (*Y*), vale a dire, la ***Variabile di Serie temporali*** da sola, oppure potete aggiungere addizionali Variabili Esogene (X_1, X_2,…, X_n) esattamente come in un'analisi di regressione, dove ci sono multiple variabili indipendenti. Potete eseguire quanti periodi di previsione desiderate, se usate solo la variabile di serie temporali (*Y*). Tuttavia, se aggiungete variabili esogene (*X*), prego notare che i vostri periodi di previsione sono limitati al numero dei periodi di dati delle variabili esogene meno i periodi di dati della variabile di serie temporali. Per esempio, potete prevedere solo fino a 5 periodi se avete dati storici di serie temporali per 100 periodi e solo se avete variabili esogene per 105 periodi (100 periodi storici per uguagliare la variabile di serie temporali e 5 periodi futuri aggiuntivi di variabili indipendenti esogene per prevedere la variabile dipendente di serie temporali).

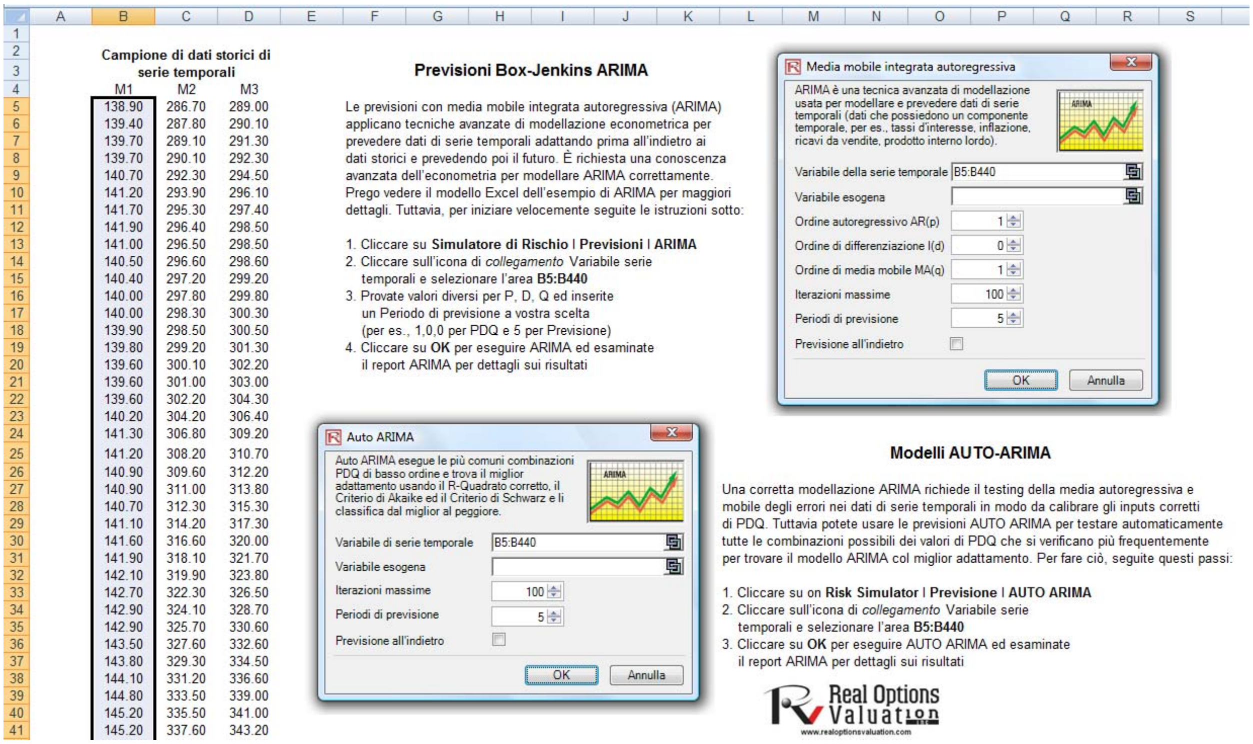


Figura 3.14 - Modulo AUTO-ARIMA

3.9 Econometria di base

Teoria: L'Econometria si riferisce ad un ramo dell'analitica economica (business analytics): tecniche di modellazione e previsione per modellare il comportamento di o per prevedere determinate variabili nel campo del business e dell'economia. Eseguire i modelli dell'Econometria di base è simile all'esecuzione di una normale analisi di regressione, tranne che le variabili dipendenti ed indipendenti possono essere modificate prima dell'esecuzione di una regressione. Il report generato è lo stesso di quello mostrato in precedenza nella sezione sulla Regressione Multipla e l'interpretazione è identica a quelle già descritte.

Procedura:

- Avviate Excel ed inserite i vostri dati o aprite un foglio di lavoro esistente con dati storici da prevedere (l'illustrazione mostrata nella Figura 3.15 usa il file d'esempio ***Modelli avanzati di previsione*** nel menu ***Esempi*** di Simulatore di Rischio)
- Selezionate i dati nel foglio di lavoro Econometria di base e selezionate Simulatore di Rischio | Previsione | Econometria di base
- Inserite le variabili dipendenti e indipendenti desiderate (vedere la Figura 3.15 per esempi) e cliccate su ***OK*** per eseguire il modello e il report, o

cliccate su ***Mostra risultati*** per vedere i risultati prima di generare il report, nel caso dovete apportare delle modifiche al modello

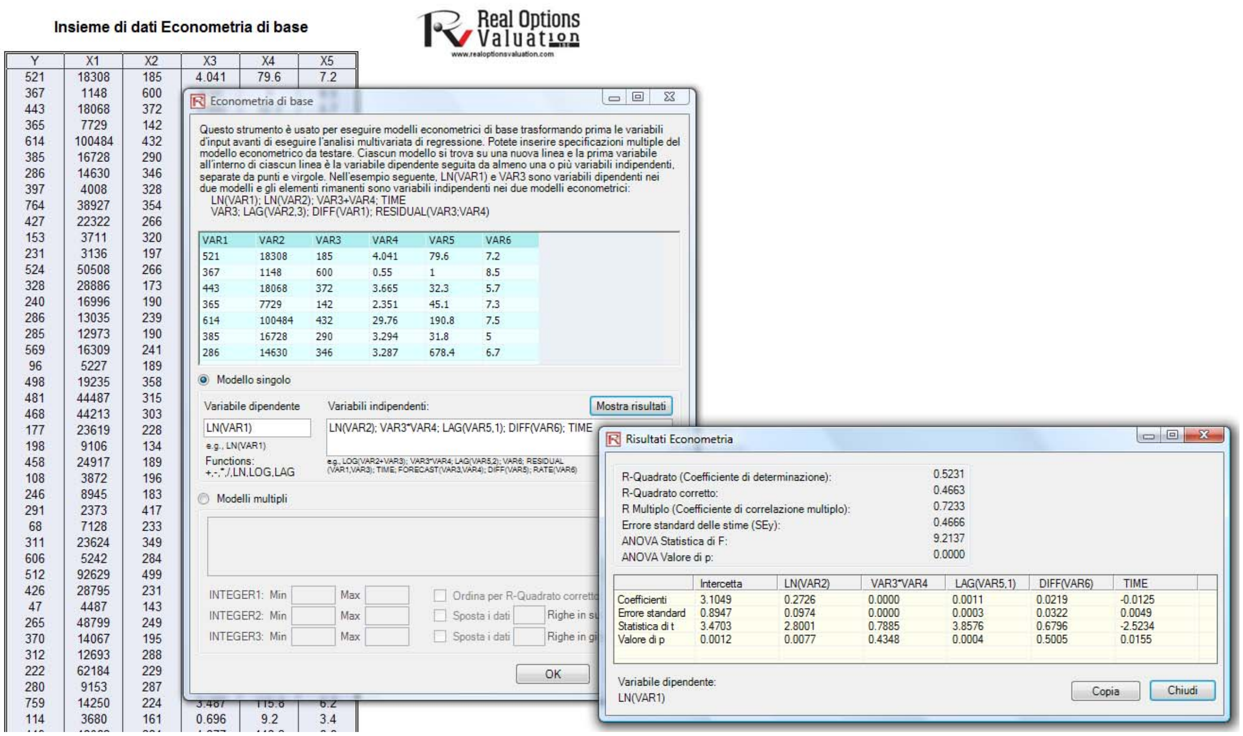


Figura 3.15 - Modulo di Econometrica di base

Note:

- Consultate il Capitolo 9 per dettagli su come interpretare gli outputs della regressione e, per estensione, gli di un'analisi di econometria di base.

- Per eseguire un modello econometrico, selezionate semplicemente i dati (B5:G55), incluso le intestazioni, e cliccate su ***Simulatore di Rischio | Previsione | Econometria di base***. Dopo potete inserire le variabili e le loro modifiche per le variabili dipendenti e indipendenti (Figura 8.15). Prego notare che è permessa una sola variabile come la Variabile Dipendente (*Y*), mentre nella sezione Variabili Indipendenti (*X*) sono permesse multiple variabili, separate da un punto e virgola ";". Si possono usare funzioni matematiche di base (per esempio, LN, LOG, LAG, +, -, /, *, TIME, RESIDUAL, DIFF). Cliccate su ***Mostra risultati*** per vedere in anteprima il modello calcolato e cliccate su ***OK*** per generare il report del modello econometrico.

- Potete anche generare automaticamente Modelli Multipli, inserendo un modello d'esempio e usando sia la variabile predefinita ***INTEGER(N)*** che ripetutamente ***Spostando i dati*** su o giù di specifiche righe. Per esempio, se usate la variabile *LAG(VAR1, INTEGER1)* e impostate *INTEGER1* per essere tra *MIN* = 1 e *MAX* = 3, allora saranno eseguiti i seguenti tre modelli: *LAG(VAR1,1)*, poi *LAG(VAR1,2)* e infine *LAG(VAR1,3)*. Potreste talvolta anche voler testare se i dati di serie temporali hanno spostamenti strutturali o se il comportamento del modello è coerente nel tempo, spostando i dati e rieseguendo poi lo stesso modello. Per esempio, se avete 100 mesi di dati elencati in ordine cronologico, potete spostarli giù 3 mesi alla volta per 10 volte (in altre parole, il modello sarà eseguito per i mesi 1-100, 4-100, 7-100 e così via). Usando questa sezione ***Modelli Multipli*** nell'Econometria di base, potete eseguire centinaia di modelli semplicemente inserendo una singola equazione di modello, se usate questi metodi delle variabili predefinite a numero intero e dello spostamento.

3.10 Previsioni Curve J-S

Teoria:

La curva a J o curva di crescita esponenziale è una curva dove la crescita del periodo successivo dipende dal livello del periodo attuale e l'aumento è esponenziale. Questo significa che, col tempo, i valori aumenteranno in modo rilevante da un periodo all'altro. Questo modello è tipicamente usato nella previsione della crescita biologica e delle reazioni chimiche col passare del tempo.

Procedura:

- Avviate Excel e selezionate Simulatore di Rischio | Previsione | Curve a JS.
- Selezionate il tipo di curva J o S, inserite le ipotesi d'input richieste (vedere le Figure 3.16 e 3.17 per esempi) e cliccate su ***OK*** per eseguire il modello e il report.

La curva a S o curva di crescita logistica inizia come una curva a J, con tassi di crescita esponenziali. Col tempo, l'ambiente si satura (per esempio, saturazione del mercato, concorrenza, sovraffollamento), la crescita rallenta ed il valore previsto si ferma infine ad un livello massimo o di saturazione. Questo modello è tipicamente usato nella previsione delle quote di mercato o della crescita delle vendite di un prodotto nuovo dalla sua introduzione nel mercato fino alla sua maturità e diminuzione, delle dinamiche delle popolazioni, della crescita di colture batteriche e di altri fenomeni che accadono in natura. La Figura 3.17 illustra un esempio di curva a S.

Curva a J Curve esponenziali di accrescimento

In matematica, una quantità che cresce esponenzialmente è una quantità la cui curva di accrescimento è sempre proporzionale alla sua dimensione attuale. Si dice che tale crescita segue una legge esponenziale. Questo implica che per qualsiasi quantità che cresce esponenzialmente, tanto più grande diventa la quantità, tanto più velocemente essa cresce. Ma questo implica anche che la relazione tra la dimensione della variabile dipendente ed il suo tasso di crescita è governata da una legge rigorosa, del tipo più semplice: la proporzione diretta. Il principio generale della crescita esponenziale è che tanto più grande diventa un numero, tanto più velocemente esso cresce. Un numero con crescita esponenziale diventerà alla fine più grande di qualsiasi altro numero che cresce solamente ad un tasso costante nello stesso lasso di tempo. Questo metodo di previsione è

Per generare una previsione con curva a J, seguite le istruzioni sotto:

1. Cliccate su **Simulatore di Rischio | Previsione | Curve a J-S**
2. Selezionate la **Curva esponenziale a J** ed inserite gli inputs desiderati (per es., Valore iniziale di **100**, Tasso di crescita del **5** percento, Periodo finale di **100**)
3. Cliccate su **OK** per eseguire la previsione e prendete del tempo per esaminare il report della previsione

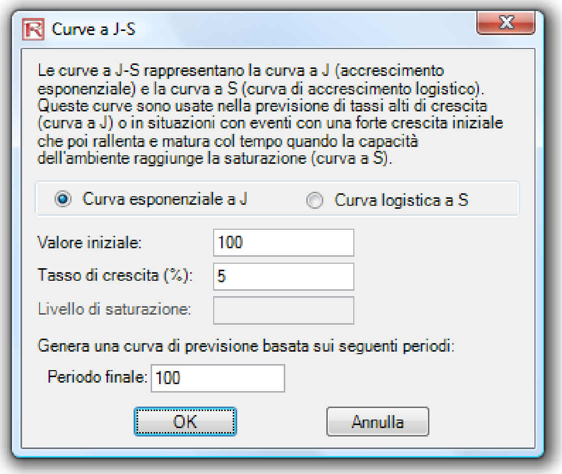


Figura 3.16 - Previsione Curve a J

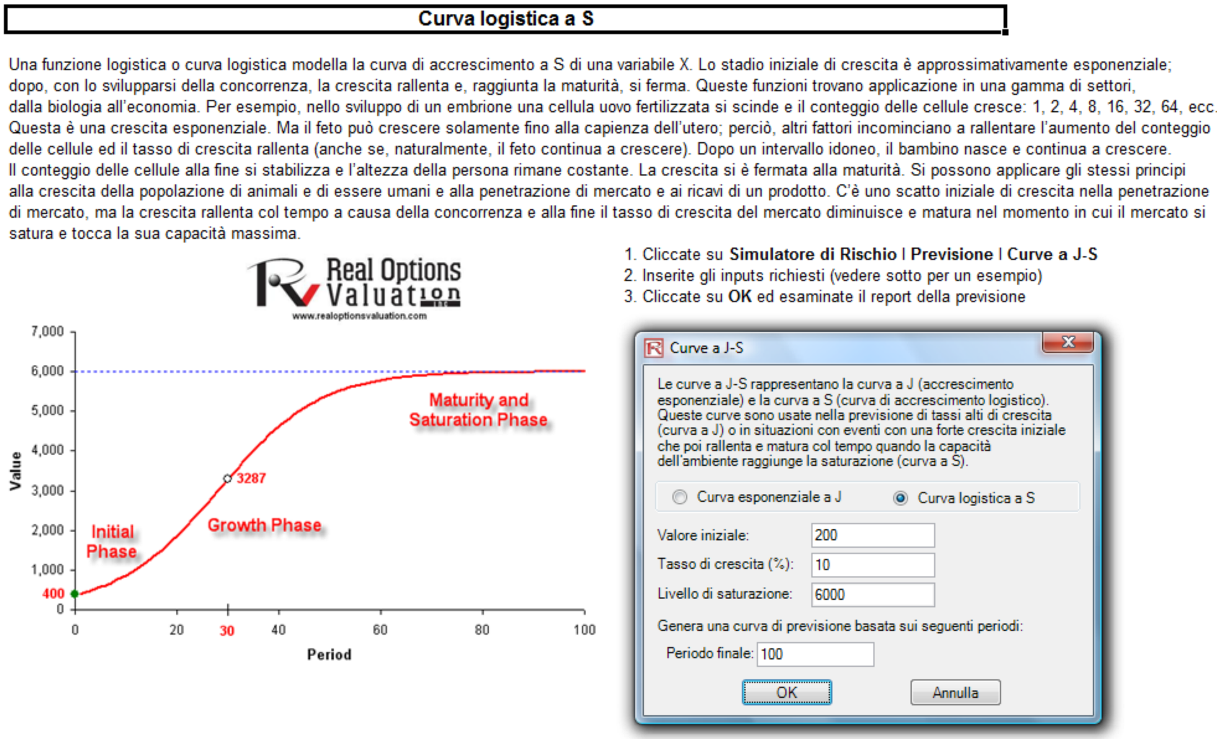


Figura 3.17 - Previsione Curve a S

3.11 Previsioni di Volatilità GARCH

Teoria: Il modello GARCH (eteroschedasticità condizionale autoregressiva generalizzata) è usato per modellare livelli storici di volatilità e prevedere livelli futuri di volatilità di valori negoziabili (per esempio, prezzi di titoli, prezzi di commodities, prezzi del petrolio e così via). L'insieme dei dati deve essere una serie temporale di livelli di prezzi grezzi. GARCH converte prima i prezzi in rendimenti relativi ed esegue poi una ottimizzazione interna per adattare i dati storici ad una struttura a termine della volatilità con ritorno alla media, assumendo al contempo che la volatilità è di natura eteroschedastica (ovvero che cambia col tempo secondo determinate caratteristiche econometriche).

Le specifiche teoriche del modello GARCH non rientrano nello scopo di questo Manuale dell'Utente. Per maggiori dettagli sui modelli GARCH, prego consultare il libro "Advanced Analytical Models" di Dr. Johnathan Mun (Wiley 2008).

Procedura:

- Avviate Excel e aprite il file d'esempio ***Modelli avanzati di previsione***, andate al foglio di lavoro ***GARCH*** e selezionate ***Simulatore di Rischio | Previsione | GARCH***.

- Cliccate sull'icona Collegamento, selezionate la *Posizione dei dati*, inserite le ipotesi d'input richieste (vedere la Figura 3.18) e cliccate su ***OK*** per eseguire il modello e il report.

Nota: La tipica situazione di previsione della volatilità richiede i seguenti parametri: P = 1; Q = 1; Periodicità = numero di periodi per anno (12 per dati mensili, 52 per dati settimanali, 252 o 365 per dati giornalieri); Base = minimo di 1 e fino al valore della periodicità; Periodi di previsione = numero di previsioni annualizzate di volatilità che desiderate ottenere. Diverse variazioni di questa metodologia sono disponibili in Simulatore di Rischio, comprese EGARCH, EGARCH-T, GARCH-M, GJR-GARCH, GJR-GARCH-T, IGARCH e T-GARCH. Consultare il libro Modeling Risk, Second Edition (Wiley 2010) sulla modellazione GARCH per maggiori dettagli sullo scopo di ciascuna specifica.

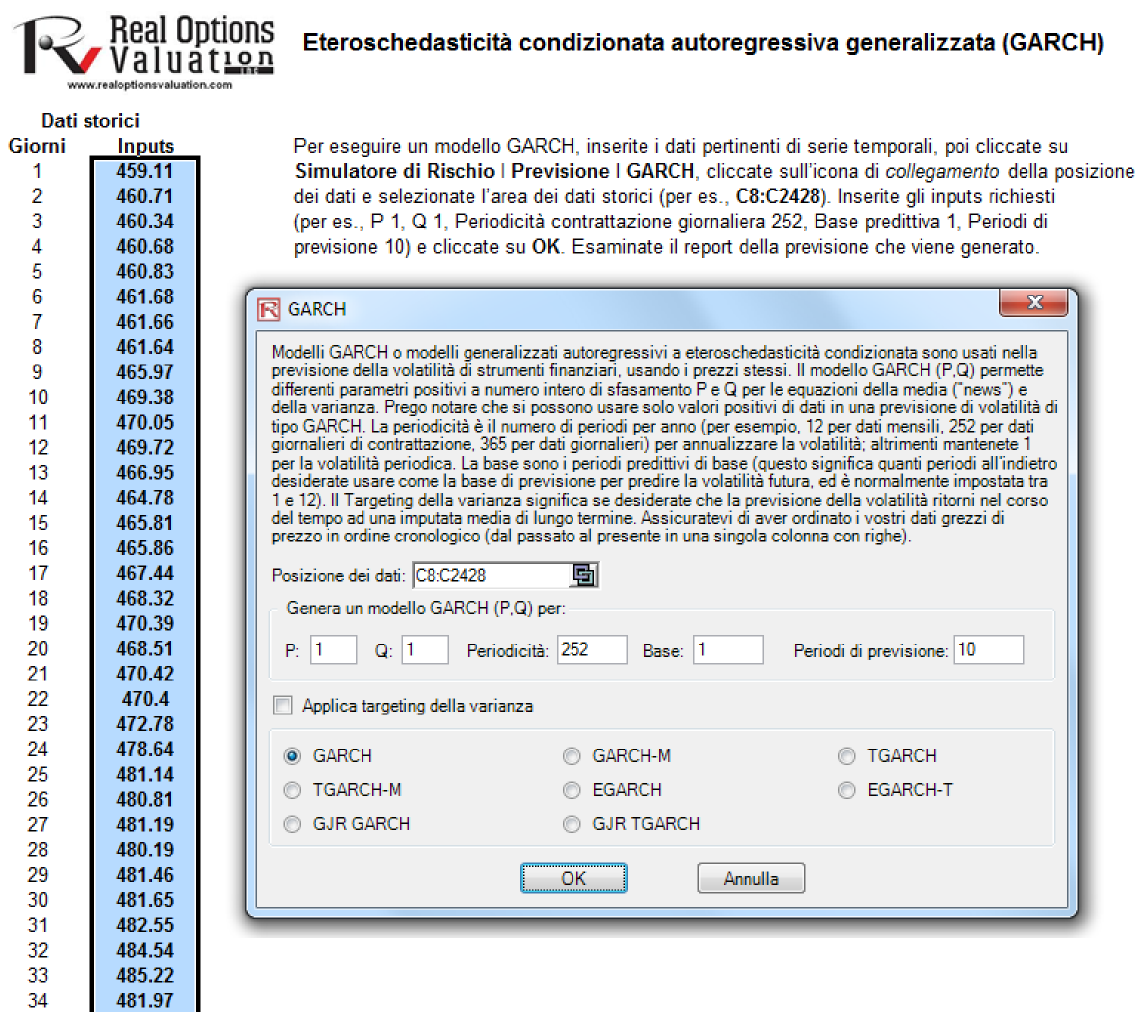


Figura 3.18 - Previsione di Volatilità GARCH

	$z_t \sim$ Normal	$z_t \sim$ T
GARCH-M	$y_t = c + \lambda\sigma_t^2 + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	$y_t = c + \lambda\sigma_t^2 + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
GARCH-M	$y_t = c + \lambda\sigma_t + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	$y_t = c + \lambda\sigma_t + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
GARCH-M	$y_t = c + \lambda \ln(\sigma_t^2) + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	$y_t = c + \lambda \ln(\sigma_t^2) + \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
GARCH	$y_t = x_t\gamma + \varepsilon_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$	$y_t = \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
EGARCH	$y_t = \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\ln\left(\sigma_t^2\right) = \omega + \beta \cdot \ln\left(\sigma_{t-1}^2\right) +$ $\alpha\left[\left\lvert\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right\rvert - E(\lvert\varepsilon_t\rvert)\right] + r\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ $E(\lvert\varepsilon_t\rvert) = \sqrt{\frac{2}{\pi}}$	$y_t = \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\ln\left(\sigma_t^2\right) = \omega + \beta \cdot \ln\left(\sigma_{t-1}^2\right) +$ $\alpha\left[\left\lvert\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right\rvert - E(\lvert\varepsilon_t\rvert)\right] + r\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ $E(\lvert\varepsilon_t\rvert) = \frac{2\sqrt{\nu-2}\,\Gamma((\nu+1)/2)}{(\nu-1)\Gamma(\nu/2)\sqrt{\pi}}$
GJR-GARCH	$y_t = \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 +$ $r\varepsilon_{t-1}^2 d_{t-1} + \beta\sigma_{t-1}^2$ $d_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{otherwise} \end{cases}$	$y_t = \varepsilon_t$ $\varepsilon_t = \sigma_t z_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 +$ $r\varepsilon_{t-1}^2 d_{t-1} + \beta\sigma_{t-1}^2$ $d_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{otherwise} \end{cases}$

3.12 Catene di Markov

Teoria: Una Catena di Markov esiste, quando la probabilità di uno stato futuro dipende da uno stato precedente e quando sono collegati formano una catena che ritorna ad un livello di equilibrio stazionario di lungo termine. Questo metodo è tipicamente usato per prevedere la quota di mercato di due concorrenti. Gli inputs richiesti sono la probabilità iniziale che un cliente del primo negozio (il primo stato) ritornerà allo stesso negozio nel periodo successivo contro la probabilità che vada al negozio del concorrente nello stato successivo.

Procedura:

- Avviate Excel e selezionate Simulatore di Rischio | Previsione | Catena di Markov.
- Inserite le ipotesi d'input richieste (vedere la Figura 3.19 per un esempio) e cliccate su ***OK*** per eseguire il modello e il report.

Nota: Impostate entrambe le probabilità su 10% e rieseguite la Catena di Markov; vedrete molto chiaramente gli effetti dei comportamenti di commutazione nel diagramma risultante.

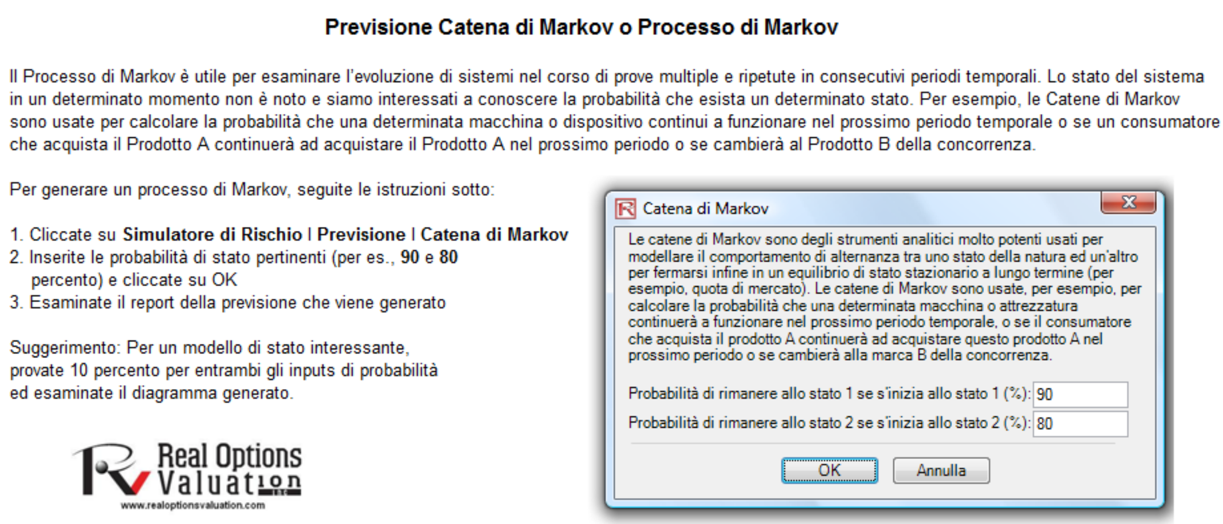


Figura 3.19 Catene di Markov (Regimi di Commutazione)

3.13 Modelli di Massima Verosimiglianza (MLE) su Logit, Probit e Tobit

Teoria: Le Variabili dipendenti limitate descrivono la situazione dove la variabile dipendente contiene dati che sono limitati nello scopo e nella portata, come le risposte binarie (0 o 1), e dati troncati, ordinati o censurati. Per esempio, dato un insieme di variabili indipendenti (per es., età, reddito, grado d'istruzione di detentori di carte di credito o prestiti ipotecari), possiamo modellare la probabilità d'inadempienza usando la stima di massima verosimiglianza (MLE). La risposta o la variabile dipendente Y è binaria, vale a dire che può avere solo due esiti possibili che denominiamo 1 e 0 (per es., Y può rappresentare la presenza/assenza di una certa condizione, l'inadempienza/non inadempienza di prestiti precedenti, il successo/insuccesso di un certo dispositivo, la riposta si/no in un sondaggio, ecc.). Abbiamo inoltre un vettore dei regressori delle variabili indipendenti X, che si presume condizionino l'esito Y. Il tipico normale Metodo della regressione dei minimi quadrati non è valido perché gli errori della regressione sono eteroschedastici e non normali, e le risultanti stime della probabilità stimata forniranno valori senza senso sopra 1 o sotto 0. L'analisi MLE gestisce questi problemi usando una routine iterativa d'ottimizzazione per massimizzare una funzione di verosimiglianza log quando le variabili dipendenti sono limitate.

La regressione Logit o Logistica è usata per predire la probabilità di occorrenza di un evento mediante l'adattamento dei dati ad una curva logistica. Questo è un modello lineare generalizzato usato per la regressione binomiale e, come molte forme di analisi della regressione, fa uso di alcuni predittori (variabili esplicative) che possono essere numerici o categoriali. La stima di massima verosimiglianza (MLE) su un'analisi logistica multivariata binaria è usata per modellare variabili dipendenti per determinare la probabilità attesa del successo di appartenere ad un determinato gruppo. I coefficienti stimati per il modello Logit sono i rapporti logaritmici di probabilità e non possono essere interpretati direttamente come probabilità. Prima è richiesto un veloce calcolo. Il metodo è semplice.

Nello specifico, il modello Logit è indicato come Stima di Y = LN[Pi/(1–Pi)] o viceversa, Pi = EXP(Stima di Y)/(1+EXP(Stima di Y)) e i coefficienti βi sono i rapporti logaritmici di probabilità. Per cui se si prende l'antilog o EXP(βi) otteniamo il rapporto di probabilità di Pi/(1–Pi). Questo significa che con l'aumento in una unità di βi il rapporto logaritmico di probabilità aumenta di questa quantità. Per finire, il tasso di cambiamento nella probabilità dP/dX = βiPi(1–Pi). L'Errore standard misura l'accuratezza dei Coefficienti predetti e le Statistiche di t sono i rapporti di ciascun Coefficiente predetto con il suo Errore standard. Sono usati nel tipico test di verifica d'ipotesi della regressione per la significatività di ciascun parametro stimato. Per stimare la probabilità di successo dell'appartenenza ad un determinato gruppo (per esempio, predire se un fumatore svilupperà complicazioni polmonari data una certa quantità di sigarette fumate l'anno), calcolate semplicemente il valore stimato di Y usando i coefficienti della Stima di massima verosimiglianza (MLE). Per esempio, se il modello è Y = 1.1 + 0.005 (sigarette), allora chi fuma 100 pacchetti l'anno ha uno Y stimato di 1,1 + 0,005(100) = 1,6. Poi calcolate l'antilog inverso del rapporto di probabilità come segue: EXP(Y stimato)/[1 + EXP(Y stimato)] = EXP(1,6)/(1+ EXP(1,6)) = 0,8320. Tale persona ha perciò una probabilità del 83.20% di sviluppare complicazioni polmonari durante la sua vita.

Un modello Probit (conosciuto talvolta anche come modello Normit) è una comune specificazione alternativa per un modello a riposta binaria che utilizza una funzione probit stimata usando la stima di massima verosimiglianza (MLE). Questo metodo è denominato regressione probit. I modelli di regressione Probit e Logistica tendono a produrre previsioni molto simili, laddove i parametri stimati in una regressione logistica tendono ad essere da 1.6 a 1.8 volte più alti che nel corrispondente modello Probit. La scelta di usare un Probit o Logit sta solamente nella convenienza e la distinzione principale è che la distribuzione logistica ha una curtosi più alta (code più ampie) per rendere conto dei valori estremi. Per esempio, supponiamo che la decisione da modellare sia il possesso di un immobile, che questa risposta sia binaria (comprare l'immobile o non comprare l'immobile) e che dipenda da una serie di variabili indipendenti Xi come reddito, età, e così via. In questo modo Ii = β0 + β1X1 +...+ βnXn, dove tanto più alto è il valore di Ii, tanto più alta è la probabilità del possesso di un immobile. Per ciascuna famiglia esiste una soglia critica I* che, se superata, porta all'acquisto di un immobile, altrimenti l'immobile non viene acquistato. Si presume che la probabilità dell'esito (P) sia distribuita normalmente, così che Pi = CDF(I) usando una funzione di distribuzione cumulativa (CDF) normale standard. Perciò usate i coefficienti stimati esattamente come in un modello di regressione e, usando il valore stimato di Y, applicate una distribuzione normale standard (potete usare la funzione NORMSDIST di Excel o lo Strumento di Analisi Distribuzionale di Simulatore di Rischio, selezionando una distribuzione Normale ed impostando la media a 0 e la deviazione standard a 1). Per finire, per ottenere un Probit o una misura di unità della probabilità impostate Ii su + 5 (questo perché ogni volta che la probabilità Pi < 0.5, il Ii stimato è negativo dovuto al fatto che la distribuzione normale è simmetrica attorno ad una media di zero).

Il Modello Tobit (Tobit censurato) è un metodo di modellazione econometrica e biometrica usato per descrivere la relazione tra una variabile dipendente non negativa Yi e una o più variabili indipendenti Xi. Un modello Tobit è un modello econometrico nel quale la variabile dipendente è censurata; vale a dire, la variabile dipendente è censurata perché i valori sotto lo zero non sono osservati. Il modello Tobit assume che ci sia una variabile Y* latente non osservabile. Questa variabile è linearmente dipendente dalle variabili Xi attraverso un vettore di coefficienti βi che determina le loro inter-relazioni. In aggiunta, c'è un termine d'errore normalmente distribuito Ui per catturare le influenze casuali su questa relazione. La variabile Yi osservabile è definita come uguale alle variabili latenti ogni volta che le variabili latenti sono sopra lo zero; altrimenti si presume che Yi sia zero. In altre parole, Yi = Y* if Y* > 0 e Yi = 0 se Y* = 0. Se il parametro di relazione βi è stimato usando una normale regressione dei minimi quadrati di Yi osservato su Xi, i risultanti stimatori della regressione sono inconsistenti e forniscono coefficienti di pendenza con una distorsione verso il basso e una intercetta con una distorsione verso l'alto. Solo MLE sarà consistente per un modello Tobit. Nel modello Tobit c'è una statistica ausiliaria denominata sigma, che è equivalente all'errore standard della stima in una normale regressione standard dei minimi quadrati e i coefficienti stimati sono usati nello stesso modo che in una analisi di regressione.

Procedura:

- Avviate Excel, aprite il file d'esempio ***Modelli avanzati di previsione***, andate al foglio di lavoro ***MLE***, selezionate l'insieme di dati, incluso le intestazioni, e cliccate su ***Simulatore di Rischio | Previsione | Massima verosimiglianza***.
- Selezionate la variabile dipendente dal menu a tendina (vedere la Figura 3.20) e cliccate su ***OK*** per eseguire il modello e il report.

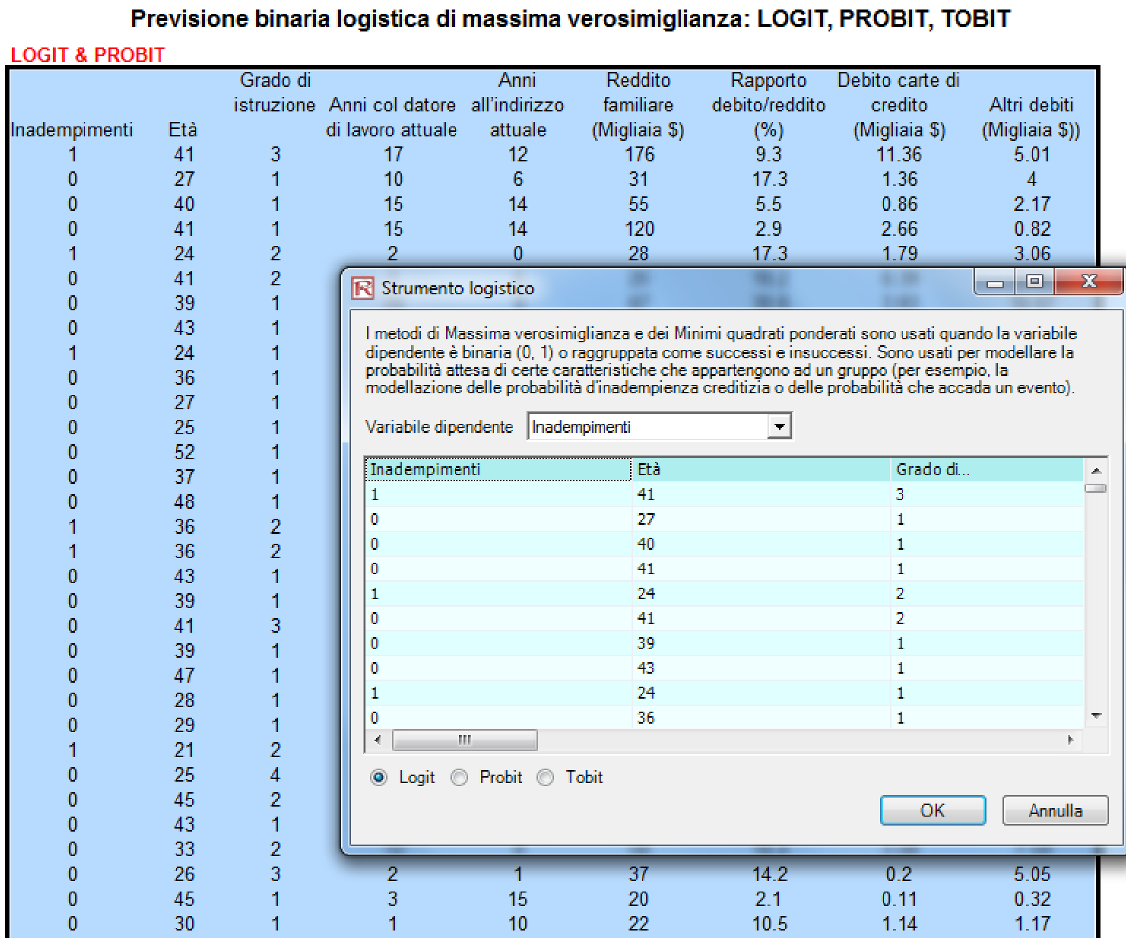


Figura 3.20 - Modulo di Massima Verosimiglianza

3.14 Spline (Spline Cubico - Interpolazione ed Estrapolazione)

Teoria:

Talvolta ci sono dati mancanti nell'insieme di dati di una serie temporale. Per esempio, sono disponibili i tassi d'interesse per gli anni 1 a 3, seguiti dai tassi per gli anni 5 a 8 e poi per l'anno 10. Le curve spline possono essere usate per interpolare i valori dei tassi d'interesse per gli anni mancanti, basato sui dati esistenti. Le curve spline possono anche essere usate per prevedere od estrapolare valori di periodi temporali futuri oltre il periodo temporale dei dati disponibili. I dati possono essere lineari o non lineari. La Figura 3.21 illustra come si esegue uno spline cubico e la Figura 3.22 mostra il report della previsione derivante da questo modulo. I valori Noti di *X* rappresentano i valori sull'asse *x* di un diagramma (nel nostro esempio, questi valori sono gli Anni dei tassi d'interesse noti; l'asse *x* rappresenta, inoltre, valori noti in anticipo, come il tempo o gli anni) e i valori Noti di *Y* rappresentano i valori sull'asse *y* (nel nostro esempio, i tassi d'interesse noti). La variabile dell'asse *y* è tipicamente la variabile di cui desiderate interpolare valori mancanti o estrapolare i valori nel futuro.

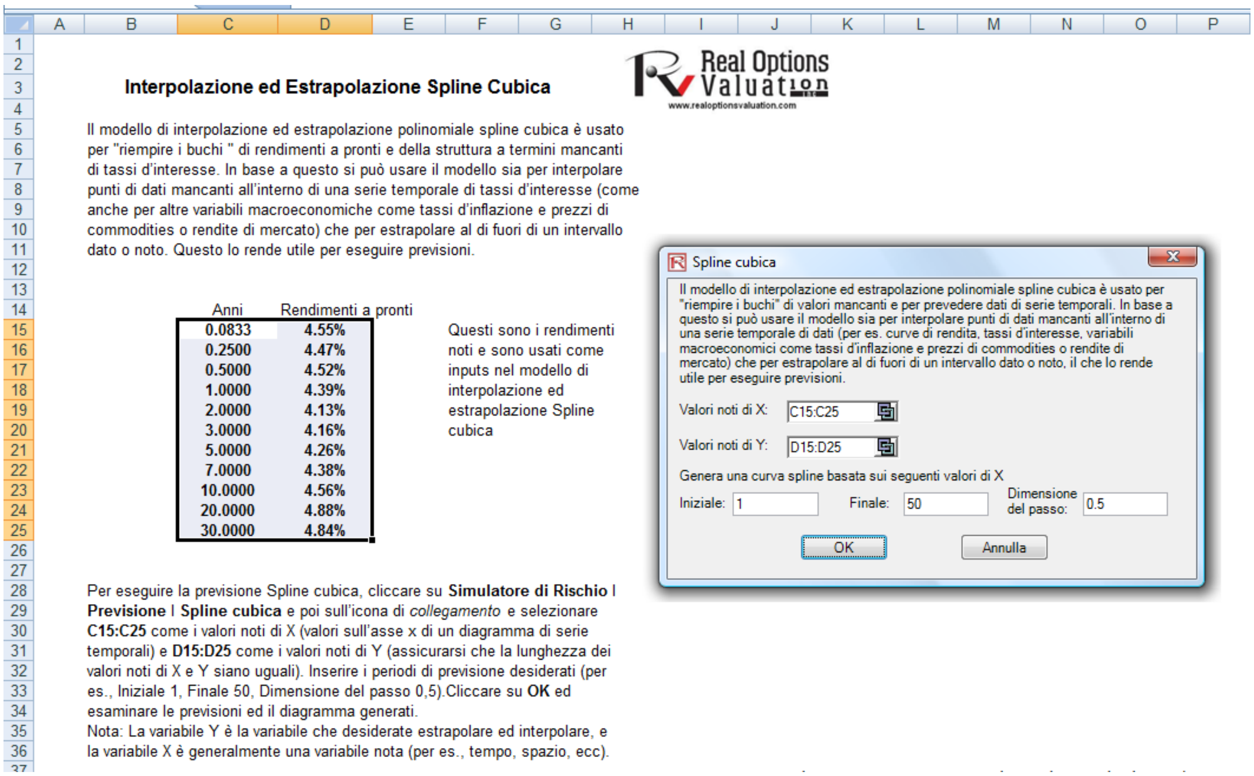


Figura 3.21 - Modulo Spline Cubico

Procedura:

- Avviate Excel, aprite il file d'esempio Modelli avanzati di previsione, andate al foglio di lavoro Spline Cubico, selezionate l'insieme di dati, incluso le intestazioni, e cliccate su Simulatore di Rischio | Previsione | Spline Cubico.
- La posizione dei dati è automaticamente inserita nell'interfaccia utente se selezionate prima i dati. Potete altrimenti cliccare manualmente sull'icona Collegamento e collegare i valori Noti di X e i valori Noti di Y (vedere la Figura 3.21 per un esempio). Ora inserite i valori richiesti Iniziali e Finali da estrapolare ed interpolare, come anche la debita Dimensione del passo tra questi valori iniziali e finali. Cliccate su OK per eseguire il modello e il report (vedere la Figura 3.22).

Previsioni Spline Cubica

Il modello di interpolazione ed estrapolazione polinomiale spline cubica è usato per "riempire i buchi" di valori mancanti e per prevedere dati di serie temporali. In base a questo si può usare il modello sia per interpolare punti di dati mancanti all'interno di una serie temporale di dati (per es. curve di rendita, tassi d'interesse, variabili macroeconomici come tassi d'inflazione e prezzi di commodities o rendite di mercato) che per estrapolare al di fuori di un intervallo dato o noto, il che lo rende utile per eseguire previsioni.

Risultati dell'interpolazione ed estrapolazione Spline

X	Y adattato	Note
1.0	4.39%	Interpolare
1.5	4.21%	Interpolare
2.0	4.13%	Interpolare
2.5	4.13%	Interpolare
3.0	4.16%	Interpolare
3.5	4.19%	Interpolare
4.0	4.22%	Interpolare
4.5	4.24%	Interpolare
5.0	4.26%	Interpolare
5.5	4.29%	Interpolare
6.0	4.32%	Interpolare
6.5	4.35%	Interpolare
7.0	4.38%	Interpolare
7.5	4.41%	Interpolare
8.0	4.44%	Interpolare
8.5	4.47%	Interpolare
9.0	4.50%	Interpolare
9.5	4.53%	Interpolare
10.0	4.56%	Interpolare
10.5	4.59%	Interpolare
11.0	4.61%	Interpolare
11.5	4.64%	Interpolare
12.0	4.66%	Interpolare
12.5	4.68%	Interpolare
13.0	4.70%	Interpolare

Questi sono i valori d'input noti per il modello d'interpolazione ed estrapolazione spline cubica:

Osservazione	X noto	Y noto
1	0.0833	4.55%
2	0.2500	4.47%
3	0.5000	4.52%
4	1.0000	4.39%
5	2.0000	4.13%
6	3.0000	4.16%
7	5.0000	4.26%
8	7.0000	4.38%
9	10.0000	4.56%
10	20.0000	4.88%
11	30.0000	4.84%

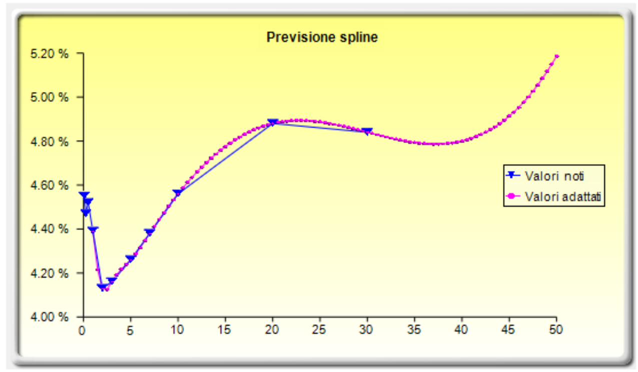


Figura 3.22 - Risultati della Previsione Spline

4

4. OTTIMIZZAZIONE

Questa sezione esamina più approfonditamente il processo e le metodologie di ottimizzazione per quanto riguarda l'uso di Simulatore di Rischio. Queste metodologie comprendono l'utilizzo dell'ottimizzazione continua contro quella discreta a numero intero, come anche delle ottimizzazioni statiche contro quelle dinamiche e stocastiche.

4.1 Metodologie di Ottimizzazione

Esistono molti algoritmi per eseguire una ottimizzazione e ci sono molte procedure differenti, quando l'ottimizzazione è accoppiata alla Simulazione Monte Carlo. Il Simulatore di Rischio contiene tre procedure distinte e tre tipi distinti di ottimizzazione, così come diversi tipi di variabili decisionali. Per esempio, il Simulatore di Rischio può gestire sia ***Variabili decisionali continue*** (1.2535, 0.2215 e così via) che ***Variabili decisionali a numero intero*** (per esempio, 1, 2, 3, 4 e così via), ***Variabili decisionali binarie*** (1 e 0 per decisioni "go/no-go") e ***Variabili decisionali miste*** (variabili sia a numero intero che continue). Oltre a ciò, il Simulatore di Rischio può gestire sia ***Ottimizzazioni lineari*** (vale a dire, quando sia l'obiettivo che i vincoli sono tutte equazioni e funzioni lineari) che ***Ottimizzazioni non lineari*** (vale a dire, quando l'obiettivo e i vincoli sono un misto di funzioni ed equazioni sia lineari che non lineari).

Ottimizzazione discreta

Per quanto riguarda il processo di ottimizzazione, il Simulatore di Rischio può essere usato per eseguire una ***Ottimizzazione discreta***, vale a dire, una ottimizzazione eseguita su un modello discreto o statico dove non sono eseguite simulazioni. In altre parole, tutti gli inputs nel modello sono statici e immutabili. Questo tipo di ottimizzazione è applicabile, quando si presume che il modello sia noto e che non ci siano incertezze. Una ottimizzazione discreta può inoltre essere eseguita per prima cosa per determinare il portafoglio ottimale e la sua corrispondente allocazione ottimale delle variabili decisionali, prima di applicare procedure di ottimizzazione più avanzate. Per esempio, avanti di eseguire un problema di ottimizzazione stocastica si esegue prima una ottimizzazione discreta per determinare se esistono delle soluzioni al problema di ottimizzazione, avanti di eseguire un'analisi più approfondita.

Ottimizzazione dinamica

Continuando, l'***Ottimizzazione dinamica*** è applicata se la Simulazione Monte Carlo è usata insieme all'ottimizzazione. Un altro nome per tale procedura è ***Simulazione-Ottimizzazione***. Vale a dire, prima viene eseguita una simulazione, poi i risultati della simulazione sono applicati nel modello Excel e dopo viene applicata una ottimizzazione ai valori simulati. In altre parole, prima viene eseguita una simulazione per N prove e dopo viene eseguito un processo di ottimizzazione per M iterazioni fino ad ottenere i risultati ottimali o un insieme impraticabile. Vale a dire, se usate il modulo d'ottimizzazione di Simulatore di Rischio, potete scegliere quali previsioni e statistiche d'ipotesi usare e sostituire nel modello dopo l'esecuzione della simulazione. Queste statistiche di previsione possono poi essere applicate nel processo d'ottimizzazione. Questo metodo è utile quando avete un modello grande con molte ipotesi e previsioni interagenti e quando alcune delle statistiche di previsione sono richieste nell'ottimizzazione. Per esempio, se la deviazione standard di una ipotesi o previsione è richiesta nel modello di ottimizzazione (per esempio, calcolare il Rapporto di Sharpe in problemi di allocazione degli assets e di ottimizzazione, dove abbiamo la media divisa per la deviazione standard del portafoglio), allora si dovrebbe usare questo metodo.

Ottimizzazione stocastica

Il processo di ***Ottimizzazione stocastica***, per contrasto, è simile alla procedura di ottimizzazione dinamica con l'eccezione che l'intero processo d'ottimizzazione dinamica è ripetuto T volte. Vale a dire, prima viene eseguita una simulazione con *N* prove e dopo viene eseguita una ottimizzazione per M iterazioni per ottenere i risultati ottimali. Il processo viene poi replicato *T* volte. I risultati sono diagrammi di previsione per ciascuna variabile decisionale con T valori. In altre parole, viene eseguita una simulazione e le statistiche della previsione o dell'ipotesi sono usate nel modello di ottimizzazione per trovare l'allocazione ottimale delle variabili decisionali. Dopo viene eseguita un'altra simulazione che genera statistiche di previsione differenti e questi nuovi valori aggiornati sono poi ottimizzati e così via. Ciascuna variabile decisionale finale avrà quindi il suo diagramma di previsione, indicando il campo delle variabili decisionali ottimali. Per esempio, invece di ottenere stime di singoli punti in una procedura di ottimizzazione dinamica, potete ora ottenere una distribuzione delle variabili decisionali e, quindi, un campo dei valori ottimali per ciascuna variabile decisionale, nota anche come ottimizzazione stocastica.

Per finire, una procedura di ottimizzazione di tipo Frontiera Efficiente applica i concetti degli incrementi marginali e del prezzo ombra nelle ottimizzazioni. In altre parole, cosa succederebbe ai risultati dell'ottimizzazione se uno dei vincoli fosse leggermente mitigato? Diciamo, per esempio, che il vincolo per il budget è fissato a $1 milione. Cosa succederebbe all'esito e alle decisioni ottimali del portafoglio, se il vincolo fosse ora di $1.5 milioni o di $2 milioni e così via. Questo è il concetto delle frontiere efficienti di Markowitz negli investimenti finanziari: se si permette alla deviazione standard del portafoglio di aumentare leggermente, quali rendimenti addizionali potrà il portafoglio generare? Questo processo è simile al processo dell'ottimizzazione dinamica, tranne che *uno* dei vincoli può essere cambiato e, con ogni modifica, si esegue la simulazione e il processo di ottimizzazione. La procedura migliore è di applicare questo processo manualmente usando il Simulatore di Rischio. In altre parole, eseguite una ottimizzazione dinamica o stocastica, poi rieseguite un'altra ottimizzazione con un vincolo e ripetete questa procedura varie volte. Questo processo manuale è importante, dato che mediante la modifica del vincolo, l'analista può determinare se i risultati sono simili o differenti e, quindi, se vale la pena eseguire delle analisi aggiuntive, o può determinare quanto grande dovrebbe essere l'aumento marginale nel vincolo per ottenere una modifica rilevante nell'obiettivo e nelle variabili decisionali .

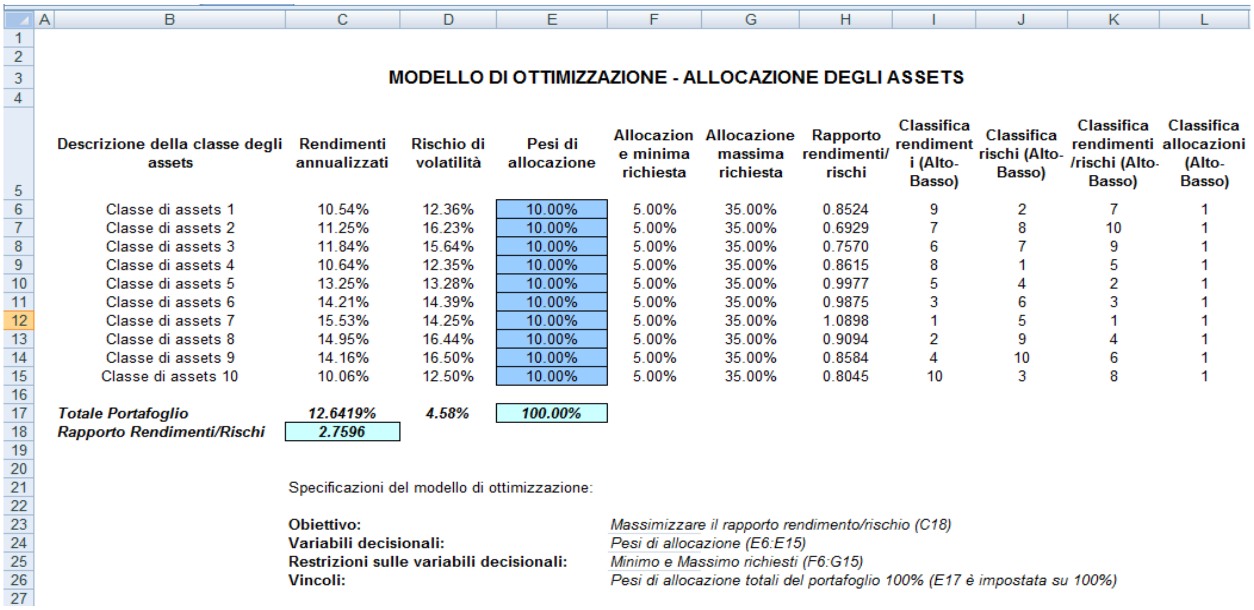
Un aspetto degno di considerazione. Esistono altri prodotti software che apparentemente eseguono le ottimizzazioni stocastiche, ma che in realtà non lo fanno. Per esempio, dopo l'esecuzione di una simulazione, viene generata poi *una* iterazione del processo di ottimizzazione, dopo viene eseguita un'altra simulazione e poi viene generata una *seconda* iterazione dell'ottimizzazione e così via. Questo metodo è semplicemente una perdita di tempo e di risorse. Vale a dire, in una ottimizzazione, il modello è sottoposto ad un insieme rigoroso di algoritmi, dove sono richieste multiple iterazioni (da decine a migliaia di iterazioni) per ottenere i risultati ottimali. Quindi, generare *una* iterazione alla volta è una perdita di tempo e di risorse. Lo stesso portafoglio può essere risolto con Simulatore di Rischio in meno di un minuto, paragonato alle molte ore richieste se si usa un metodo così retrogrado. Inoltre, questo metodo di simulazione-ottimizzazione fornirà tipicamente dei risultati non validi e non è un metodo di ottimizzazione stocastica. Prestate molta attenzione a tali metodologie, quando applicate l'ottimizzazione ai vostri modelli.

Seguono due esempi di problemi di ottimizzazione. Uno utilizza le variabili decisionali continue, mentre l'altro usa le variabili decisionali discrete a numero intero. In entrambi i modelli si possono applicare l'ottimizzazione discreta, l'ottimizzazione dinamica, l'ottimizzazione stocastica o anche le frontiere efficienti con prezzi ombra. Si possono usare qualsiasi di questi metodi per questi due esempi. Per semplicità illustreremo, perciò, solo l'impostazione del modello e la decisione su quale processo di ottimizzazione dovrà essere eseguito viene lasciata all'utente. Il modello continuo utilizza il metodo di ottimizzazione non lineare (questo perché il rischio calcolato del portafoglio è una funzione non lineare e l'obiettivo è una funzione non lineare dei rendimenti del portafoglio diviso per i rischi del portafoglio). Viceversa, il secondo esempio di ottimizzazione a numero intero è un esempio di un modello di ottimizzazione lineare (il suo obiettivo e i suoi vincoli sono lineari). Per queste ragioni, i due esempi racchiudono tutte le procedure citate.

4.2 Ottimizzazione con Variabili Decisionali Continue

La Figura 4.1 illustra l'esempio di un modello di ottimizzazione continua. Questo esempio utilizza il file ***Ottimizzazione Continua*** che si trova nel menu di avvio sotto ***Avvio | Real Options Valuation | Simulatore di Rischio | Esempi*** o tramite accesso diretto sotto ***Simulatore di Rischio | Modelli d'esempio.*** In quest'esempio ci sono 10 distinte classi di assets (per esempio, differenti tipi di fondi comuni d'investimento, di titoli o di assets), dove l'idea è di trovare l'allocazione più efficace e più valida del portafoglio per ottenere il miglior rendimento ("bang-for-the-buck"). Vale a dire, per generare il migliore rendimento possibile del portafoglio, dato i rischi inerenti di ciascuna classe di assets. Per poter veramente comprendere il concetto dell'ottimizzazione, dobbiamo esaminare più approfonditamente questo modello d'esempio per capire come meglio applicare il processo d'ottimizzazione.

Il modello mostra le 10 classi di assets. Ciascuna classe di assets ha il suo set di rendimenti annualizzati e volatilità annualizzate. Queste misure di rendimento e rischio sono valori annualizzate così da poterli confrontare coerentemente con classi di assets differenti. I rendimenti sono calcolati usando la media geometrica dei rendimenti relativi, mentre i rischi sono calcolati usando il metodo logaritmico dei rendimenti relativi di titoli. Consultate l'Appendice di questo capitolo per dettagli su come calcolare la volatilità annualizzata e i rendimenti annualizzati di un titolo o di una classe di assets.



MODELLO DI OTTIMIZZAZIONE - ALLOCAZIONE DEGLI ASSETS

Descrizione della classe degli assets	Rendimenti annualizzati	Rischio di volatilità	Pesi di allocazione	Allocazione minima richiesta	Allocazione massima richiesta	Rapporto rendimenti/ rischi	Classifica rendimenti (Alto-Basso)	Classifica rischi (Alto-Basso)	Classifica rendimenti /rischi (Alto-Basso)	Classifica allocazioni (Alto-Basso)
Classe di assets 1	10.54%	12.36%	10.00%	5.00%	35.00%	0.8524	9	2	7	1
Classe di assets 2	11.25%	16.23%	10.00%	5.00%	35.00%	0.6929	7	8	10	1
Classe di assets 3	11.84%	15.64%	10.00%	5.00%	35.00%	0.7570	6	7	9	1
Classe di assets 4	10.64%	12.35%	10.00%	5.00%	35.00%	0.8615	8	1	5	1
Classe di assets 5	13.25%	13.28%	10.00%	5.00%	35.00%	0.9977	5	4	2	1
Classe di assets 6	14.21%	14.39%	10.00%	5.00%	35.00%	0.9875	3	6	3	1
Classe di assets 7	15.53%	14.25%	10.00%	5.00%	35.00%	1.0898	1	5	1	1
Classe di assets 8	14.95%	16.44%	10.00%	5.00%	35.00%	0.9094	2	9	4	1
Classe di assets 9	14.16%	16.50%	10.00%	5.00%	35.00%	0.8584	4	10	6	1
Classe di assets 10	10.06%	12.50%	10.00%	5.00%	35.00%	0.8045	10	3	8	1
Totale Portafoglio	*12.6419%*	*4.58%*	*100.00%*							
Rapporto Rendimenti/Rischi	*2.7596*									

Specificazioni del modello di ottimizzazione:

Obiettivo: *Massimizzare il rapporto rendimento/rischio (C18)*
Variabili decisionali: *Pesi di allocazione (E6:E15)*
Restrizioni sulle variabili decisionali: *Minimo e Massimo richiesti (F6:G15)*
Vincoli: *Pesi di allocazione totali del portafoglio 100% (E17 è impostata su 100%)*

Figura 4.1 - Modello di Ottimizzazione Continua

I Pesi di allocazione nella colonna E contengono le variabili decisionali. Queste sono le variabili che devono essere "aggiustate" e testate in modo che il peso totale sia limitato a 100% (cella E17). Tipicamente, per avviare l'ottimizzazione si impostano queste celle su un valore uniforme: in questo caso le celle da E6 ad E15 sono impostate ciascuna al 10%. In aggiunta, ciascuna variabile decisionale può avere specifiche restrizioni nel suo campo consentito. In quest'esempio, le allocazioni inferiori e superiori consentite sono 5% e 35%, come mostrato nelle colonne F e G. Questa impostazione significa che ciascuna classe di assets può avere i suoi limiti di allocazione. Proseguendo, la colonna H mostra il rapporto rendimenti/rischi, che è semplicemente la percentuale dei rendimenti divisa per la percentuale dei rischi, dove tanto più alto è questo valore, tanto più alto è il rendimento ("bang-for-the-buck"). Il modello restante mostra la classifica delle individuali classi di assets per rischio, rapporto rendimenti/rischi e allocazione. In altre parole, questa classifica mostra in un'occhiata quale classe di assets ha il rischio più basso o il rendimento più alto e così via.

I rendimenti totali del portafoglio nella cella C17 è *SUMPRODUCT(C6:C15, E6:E15)*, vale a dire, la somma dei pesi di allocazione moltiplicata per i rendimenti annualizzati di ciascuna classe di assets. In altre parole, abbiamo $R_P = \omega_A R_A + \omega_B R_B + \omega_C R_C + \omega_D R_D$, dove R_P è il rendimento del portafoglio, $R_{A,B,C,D}$ sono i rendimenti individuali dei progetti, e $\omega_{A,B,C,D}$ sono i rispettivi pesi o le rispettive allocazioni di capitale per ciascun progetto.

Oltre a ciò, il rischio diversificato del portafoglio nella cella D17 è calcolato eseguendo $\sigma_P = \sqrt{\sum_{i=1}^{i} \omega_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{m} 2\omega_i \omega_j \rho_{i,j} \sigma_i \sigma_j}$. Qui, $\rho_{i,j}$ sono le rispettive correlazioni incrociate tra le classi di assets — quindi, se le correlazioni incrociate sono negative, esistono effetti di diversificazione del rischio e il rischio del portafoglio diminuisce. Tuttavia, per semplificare i calcoli in quest'esempio, presumiamo zero correlazioni tra le classi di assets in questo calcolo del rischio del portafoglio. Viceversa presumiamo che ci siano correlazioni, quando applichiamo la simulazione sui rendimenti, come mostreremo di seguito. Quindi, invece di applicare correlazioni statiche tra questi differenti rendimenti di assets, applichiamo le correlazioni nelle ipotesi stesse della simulazione, creando così una relazione più dinamica tra i valori di rendimento simulati.

Per finire, viene calcolato il rapporto rendimenti/rischio o Rapporto di Sharpe Ratio del portafoglio. Questo valore è mostrato cella C18 e rappresenta l'obiettivo da massimizzare in questo esercizio di ottimizzazione. Per riassumere, abbiamo le seguenti specifiche in questo modello d'esempio:

Obiettivo: Massimizzare il rapporto rendimento/rischio (C18)

Variabili Decisionali: Pesi di allocazione (E6:E15)

Restrizioni sulle variabili decisionali: Minimo e Massimo richiesti (F6:G15)

Vincoli: Pesi di allocazione totali del portafoglio 100% (E17)

Procedura:

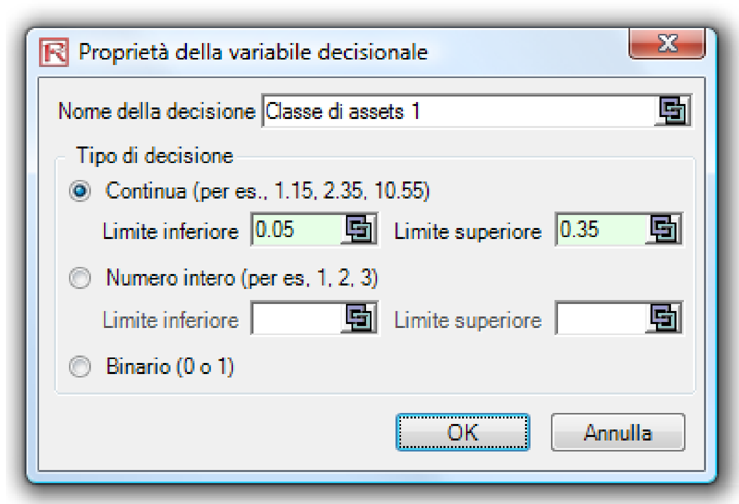
- Aprite il file d'esempio, avviate un nuovo profilo cliccando su ***Simulatore di Rischio | Nuovo profilo*** e dategli un nome.
- Il primo passo in una ottimizzazione è di impostare le variabili decisionali. Selezionate la cella E6, impostate la prima variabile decisionale (***Simulatore di Rischio | Ottimizzazione | Imposta decisione***) e cliccate sull'icona di collegamento per selezionare il nome della cella (B6), e i valori dei limiti inferiore e superiore nelle celle F6 e G6. Poi, usando la funzione Copia di Simulatore di Rischio, copiate questa variabile decisionale della cella E6 e incollate la variabile decisionale nelle celle rimanenti da E7 ad E15.
- Il secondo passo in una ottimizzazione è di impostare il vincolo. Qui c'è solo un vincolo, vale a dire, l'allocazione totale del portafoglio deve sommarsi a 100%. Per cui, cliccate su ***Simulatore di Rischio |***

Ottimizzazione* | *Vincoli… e selezionate ***AGGIUNGI*** per aggiungere un nuovo vincolo. Dopo di questo, selezionate la cella E17 e rendetela uguale (=) a 100%. Cliccate su **OK** quando avete finito.

- L'ultimo passo in una ottimizzazione è di impostare la funzione dell'obiettivo e di avviare l'ottimizzazione: selezionate la cella obiettivo C18, poi ***Simulatore di Rischio* | *Ottimizzazione* | *Esegui ottimizzazione*** e scegliete l'ottimizzazione che desiderate (Ottimizzazione statica, Ottimizzazione dinamica o Ottimizzazione stocastica). Per iniziare, selezionate ***Ottimizzazione Statica.*** Assicuratevi che la cella dell'obiettivo sia impostata su C18 e selezionate ***Massimizza***. Ora, se necessario, potete esaminare le variabili decisionali e i vincoli, o cliccare su **OK** per eseguire l'ottimizzazione statica.

- Una volta terminata l'ottimizzazione, potete selezionare ***Ritorna*** per ritornare ai valori originali delle variabili decisionali e dell'obiettivo, o selezionare ***Sostituisci*** per applicare le variabili decisionali ottimizzate. Tipicamente viene scelto Sostituisci, dopo che l'ottimizzazione è terminata.

La Figura 4.2 mostra le visualizzazioni del display per i passi della procedura citata sopra. Potete aggiungere ipotesi di simulazione sui rendimenti e rischi del modello (colonne C e D) e applicare una ottimizzazione dinamica e una ottimizzazione stocastica per esercitarvi ulteriormente.



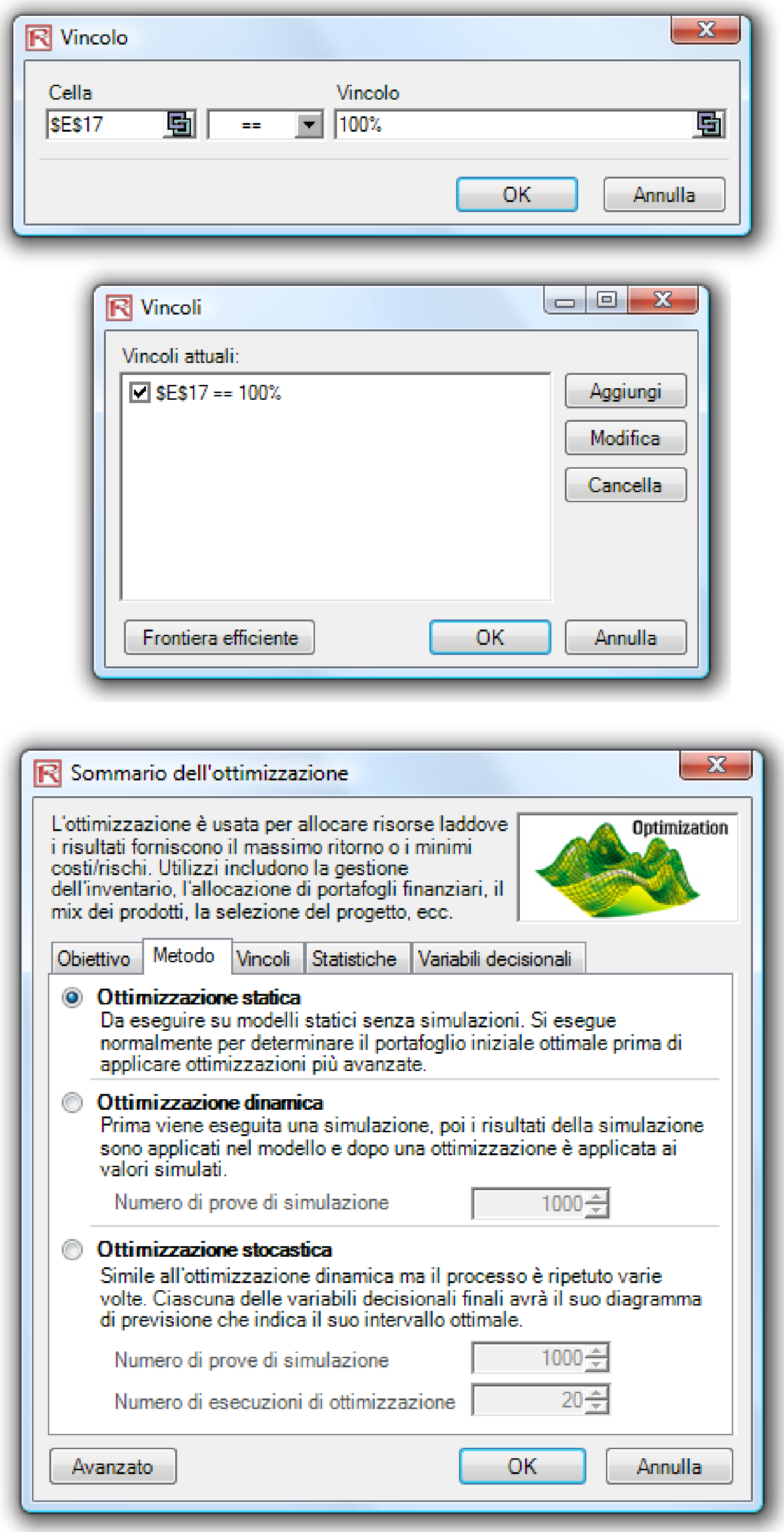


Figura 4.2 - Eseguire una Ottimizzazione Continua in Simulatore di Rischio

Interpretare i Risultati:

I risultati finali dell'ottimizzazione sono mostrati nella Figura 4.3, dove vediamo l'allocazione ottimale degli assets del portafoglio nelle celle E6:E15. Vale a dire, dato le restrizioni, che ciascun asset deve fluttuare tra 5% e 35% e che la somma dell'allocazione deve essere uguale a 100%, l'allocazione che massimizza il rapporto rendimenti/rischio è indicato nella Figura 4.3.

È importante notare alcuni aspetti rilevanti, quando si esaminano i risultati e le procedure di ottimizzazione eseguite fino ad ora:

- Il modo giusto di eseguire una ottimizzazione è di massimizzare il rendimento ("bang-for-the-buck") o il Rapporto rendimenti/rischio di Sharpe, come noi abbiamo fatto.
- Se massimizziamo invece i rendimenti totali del portafoglio, l'allocazione ottimale sarà insignificante e non richiede una ottimizzazione per essere trovata. In altre parole, basta semplicemente allocare il 5% (il minimo permesso) agli 8 assets più bassi, il 35% (il massimo permesso) all'asset col rendimento più alto, ed il resto (25%) all'asset col secondo rendimento più alto. Non è necessaria una ottimizzazione. Tuttavia, quando allocate il portafoglio in questa maniera, il rischio è molto più alto rispetto a quando massimizzate il rapporto rendimenti/rischio, anche se i rendimenti stessi del portafoglio sono più alti.
- Per contrasto, si può minimizzare il rischio totale del portafoglio, ma così i rendimenti saranno minori.

La Tabella 4.1 illustra i risultati dei tre diversi obiettivi da ottimizzare:

Obiettivo:	Rendimenti Portafoglio	Rischio Portafoglio	Rapporto Rendimenti/Rischio Portafoglio
Massimizzare Rapporto Rendimenti/Rischio	12.69%	4.52%	2.8091
Massimizzare Rendimenti	13.97%	6.77%	2.0636
Minimizzare Rischio	12.38%	4.46%	2.7754

Tabella 4.1 - Risultati dell'Ottimizzazione

Dalla Tabella vediamo che il metodo migliore è di massimizzare il rapporto rendimenti/rischio. Vale a dire, questa allocazione, a fronte della stessa quantità di rischio, fornisce la quantità più alta di rendimento. Viceversa, per la stessa quantità di rendimento, questa allocazione fornisce la quantità più bassa possibile di rischio. Questo metodo di rendimento alto (“bang-for-the-buck”) o rapporto rendimenti/rischio è il caposaldo della frontiera efficiente di Markowitz nella moderna teoria del portafoglio. In altre parole, se sottoponiamo a vincoli i livelli totali di rischio del portafoglio e li aumentiamo successivamente nel tempo, otteniamo differenti allocazioni efficienti del portafoglio per differenti caratteristiche di rischio. In questo modo possiamo ottenere differenti allocazioni efficienti del portafoglio per individui differenti con differenti preferenze di rischio.

MODELLO DI OTTIMIZZAZIONE - ALLOCAZIONE DEGLI ASSETS

Descrizione della classe degli assets	Rendimenti annualizzati	Rischio di volatilità	Pesi di allocazione	Allocazione minima richiesta	Allocazione massima richiesta	Rapporto rendimenti/ rischi
Classe di assets 1	10.54%	12.36%	11.09%	5.00%	35.00%	0.8524
Classe di assets 2	11.25%	16.23%	6.86%	5.00%	35.00%	0.6929
Classe di assets 3	11.84%	15.64%	7.78%	5.00%	35.00%	0.7570
Classe di assets 4	10.64%	12.35%	11.23%	5.00%	35.00%	0.8615
Classe di assets 5	13.25%	13.28%	12.09%	5.00%	35.00%	0.9977
Classe di assets 6	14.21%	14.39%	11.04%	5.00%	35.00%	0.9875
Classe di assets 7	15.53%	14.25%	12.30%	5.00%	35.00%	1.0898
Classe di assets 8	14.95%	16.44%	8.90%	5.00%	35.00%	0.9094
Classe di assets 9	14.16%	16.50%	8.37%	5.00%	35.00%	0.8584
Classe di assets 10	10.06%	12.50%	10.35%	5.00%	35.00%	0.8045
Totale Portafoglio	***12.6919%***	***4.52%***	***100.00%***			
Rapporto Rendimenti/Rischi	***2.8091***					

Figura 4.3 - Risultati di una Ottimizzazione Continua

4.3 Ottimizzazione con Variabili Discrete a Numero Intero

Talvolta le variabili decisionali non sono continue ma discrete a numeri interi (per esempio, 0 e 1). Vale a dire che possiamo usare queste ottimizzazioni come commutatori "on-off" o decisioni "go/no-go". La Figura 4.4 illustra un modello di selezione del progetto, dove sono elencati 12 progetti. L'esempio qui utilizza il file ***Ottimizzazione discreta*** che si trova nel menu di avvio sotto ***Avvio | Real Options Valuation | Simulatore di Rischio |Esempi*** o tramite accesso diretto sotto ***Simulatore di Rischio | Modelli d'esempio.*** Come prima, ciascun progetto ha i suoi rendimenti (ENPV e NPV per valore netto attuale espanso e per valore netto attuale —ENPV è semplicemente NPV più eventuali valori di opzioni reali strategiche), costi d'implementazione, rischi e così via. Se richiesto, questo modello può essere modificato per includere obbligatorie equivalenze a tempo pieno (FTE) e altre risorse di varie funzioni. Si possono inoltre impostare vincoli aggiuntivi su queste risorse addizionali. Gli inputs per questo modello sono tipicamente collegati da altri modelli di fogli di lavoro. Per esempio, ciascun progetto avrà il suo cash flow scontato o modello di redditività sull'investimento. L'applicazione qui è di massimizzare il Rapporto di Sharpe del portafoglio, soggetto ad una determinata allocazione del budget. Si possono creare molte altre versioni di questo modello. Per esempio, massimizzare i rendimenti del portafoglio o minimizzare i rischi o aggiungere nuovi vincoli, come il numero totale dei progetti scelti non può superare 6, e così via. Tutti questi elementi possono essere eseguiti usando questo modello.

Procedura:

- Aprite il file d'esempio, avviate un nuovo profilo cliccando su ***Simulatore di Rischio | Nuovo profilo*** e dategli un nome.
- Il primo passo in una ottimizzazione è di impostare le variabili decisionali. Impostate la prima variabile decisionale selezionando la cella J4 e selezionate ***Simulatore di Rischio | Ottimizzazione | Imposta decisione***. Poi cliccate sull'icona di collegamento per selezionare il nome della cella (B4) e selezionate la variabile **Binaria**. Poi, usando la funzione Copia di Simulatore di Rischio, copiate questa variabile decisionale della cella J4 e incollate la variabile decisionale nelle celle rimanenti da J5 a J15. Questo è il metodo migliore se avete solo poche variabili decisionali e potete nominare ogni variabile decisionale con un nome univoco per poterla identificare in seguito.
- Il secondo passo in una ottimizzazione è di impostare il vincolo. Qui ci sono due vincoli, vale a dire, l'allocazione totale del budget nel portafoglio

deve essere meno di $5000 e il numero totale dei progetti non deve superare 6. Per cui, cliccate su on ***Simulatore di Rischio | Ottimizzazione | Vincoli…*** e selezionate ***AGGIUNGI*** per aggiungere un nuovo vincolo. Dopo, selezionate la cella D17 e rendetela minore di o uguale a (<=)5000. Ripetete la procedura impostando la cella J17 su <= 6.

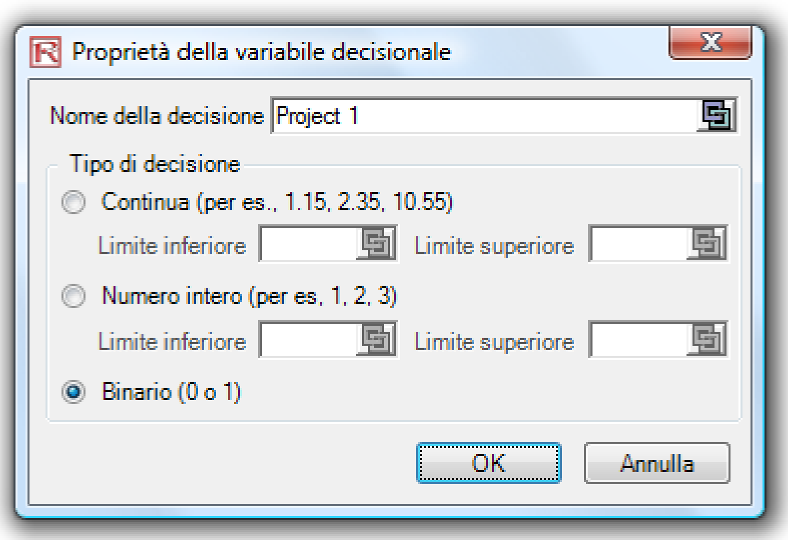
- L'ultimo passo in una ottimizzazione è di impostare la funzione dell'obiettivo e di avviare l'ottimizzazione: selezionate la cella obiettivo C19 e poi ***Simulatore di Rischio | Ottimizzazione | Imposta obiettivo***. Ora eseguite l'ottimizzazione usando ***Simulatore di Rischio | Ottimizzazione | Esegui ottimizzazione*** e scegliete l'ottimizzazione che desiderate (Ottimizzazione statica, Ottimizzazione dinamica o Ottimizzazione stocastica). Per iniziare, selezionate ***Ottimizzazione Statica.*** Assicuratevi che la cella dell'obiettivo sia o il Rapporto di Sharpe o il Rapporto rendimenti/rischio del portafoglio e selezionate ***Massimizza***. Ora, se necessario, potete esaminare le variabili decisionali e i vincoli, o cliccare su **OK** per eseguire l'ottimizzazione statica.

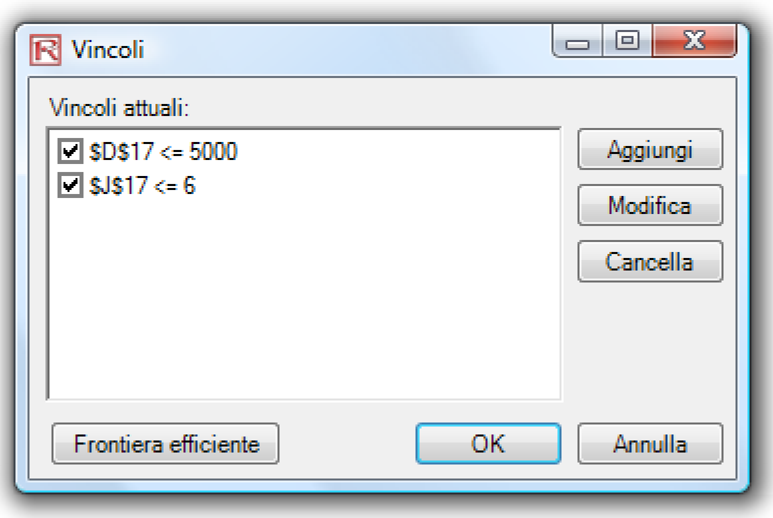
La Figura 4.2 mostra le visualizzazioni del display per i passi della procedura citata sopra. Potete aggiungere ipotesi di simulazione sull'ENPV e sui rischi del modello (colonne C ed E) e applicare una ottimizzazione dinamica e una ottimizzazione stocastica per esercitarvi ulteriormente.

Progetti	ENPV	Costo	Rischio $	Rischio %	Rapporto Rendimenti /Rischi	Indice di Profittabilità	Selezione
Project 1	$458.00	$1,732.44	$54.96	12.00%	8.33	1.26	1.0000
Project 2	$1,954.00	$859.00	$1,914.92	98.00%	1.02	3.27	1.0000
Project 3	$1,599.00	$1,845.00	$1,551.03	97.00%	1.03	1.87	1.0000
Project 4	$2,251.00	$1,645.00	$1,012.95	45.00%	2.22	2.37	1.0000
Project 5	$849.00	$458.00	$925.41	109.00%	0.92	2.85	1.0000
Project 6	$758.00	$52.00	$560.92	74.00%	1.35	15.58	1.0000
Project 7	$2,845.00	$758.00	$5,633.10	198.00%	0.51	4.75	1.0000
Project 8	$1,235.00	$115.00	$926.25	75.00%	1.33	11.74	1.0000
Project 9	$1,945.00	$125.00	$2,100.60	108.00%	0.93	16.56	1.0000
Project 10	$2,250.00	$458.00	$1,912.50	85.00%	1.18	5.91	1.0000
Project 11	$549.00	$45.00	$263.52	48.00%	2.08	13.20	1.0000
Project 12	$525.00	$105.00	$309.75	59.00%	1.69	6.00	1.0000
Totale	$17,218.00	$8,197.44	$7,007	40.70%			12.00
Scopo:	MAX	< =$5000					<=6
Rapporto di Sharpe	2.4573						

ENPV è il valore attuale netto (NPV) atteso per ciascun investimento o progetto, mentre Costo può essere il costo totale dell'investimento e Rischio è il Coefficiente di variazione del ENPV del progetto.

Figura 4.4 - Modello di Ottimizzazione Discreta a Numeri Interi





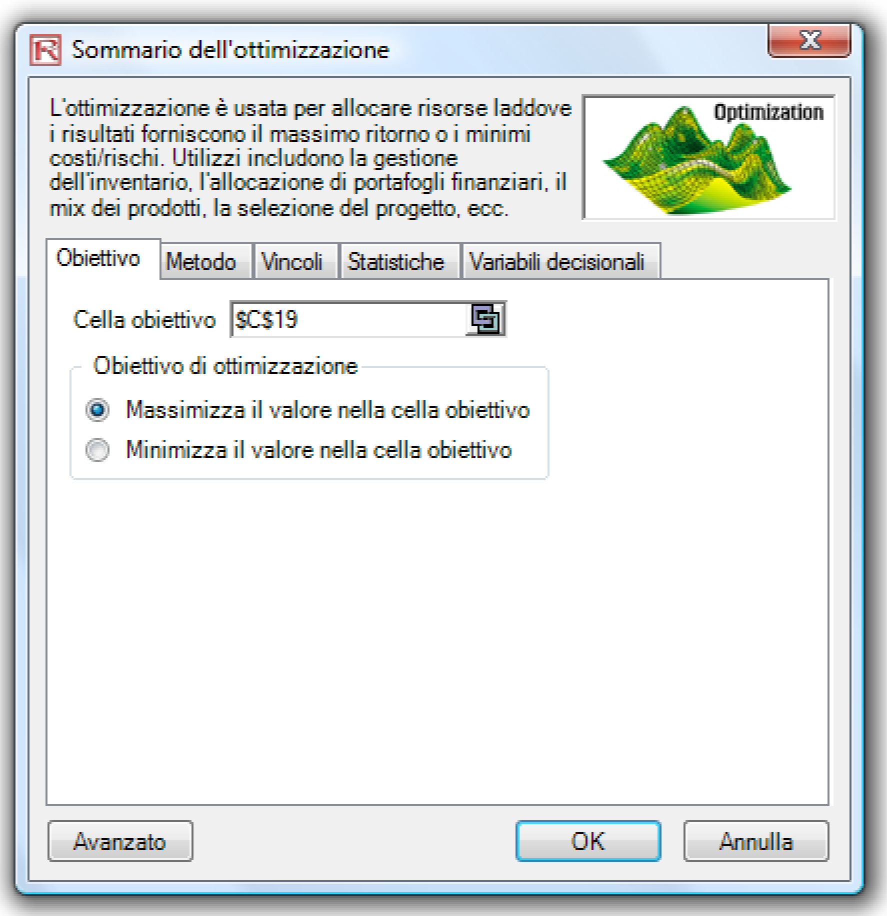


Figura 4.5 - Eseguire una Ottimizzazione Discreta a Numeri Interi in Simulatore di Rischio

Interpretare i Risultati:

La Figura 4.6 mostra l'esempio di una selezione ottimale dei progetti che massimizzano il Rapporto di Sharpe. Per contrasto, uno può sempre massimizzare i ricavi totali, ma, come notato in precedenza, questo è un processo insignificante e implica semplicemente scegliere il progetto col maggior rendimento e poi continuare giù nella lista fino a finire i fondi disponibili o a superare il vincolo del budget. Agire in questa maniera fornirà progetti teoricamente non desiderabili, dato che i progetti col miglior rendimento hanno anche tipicamente i maggiori rischi. A questo punto, se desiderato, potete replicare l'ottimizzazione usando una ottimizzazione stocastica o una ottimizzazione dinamica mediante l'aggiunta di ipotesi nell'ENPV e/o nei valori dei costi e/o dei rischi.

Progetti	ENPV	Costo	Rischio $	Rischio %	Rapporto Rendimenti /Rischi	Indice di Profittabilità	Selezione
Project 1	$458.00	$1,732.44	$54.96	12.00%	8.33	1.26	1.0000
Project 2	$1,954.00	$859.00	$1,914.92	98.00%	1.02	3.27	0.0000
Project 3	$1,599.00	$1,845.00	$1,551.03	97.00%	1.03	1.87	0.0000
Project 4	$2,251.00	$1,645.00	$1,012.95	45.00%	2.22	2.37	1.0000
Project 5	$849.00	$458.00	$925.41	109.00%	0.92	2.85	0.0000
Project 6	$758.00	$52.00	$560.92	74.00%	1.35	15.58	1.0000
Project 7	$2,845.00	$758.00	$5,633.10	198.00%	0.51	4.75	0.0000
Project 8	$1,235.00	$115.00	$926.25	75.00%	1.33	11.74	1.0000
Project 9	$1,945.00	$125.00	$2,100.60	108.00%	0.93	16.56	0.0000
Project 10	$2,250.00	$458.00	$1,912.50	85.00%	1.18	5.91	0.0000
Project 11	$549.00	$45.00	$263.52	48.00%	2.08	13.20	1.0000
Project 12	$525.00	$105.00	$309.75	59.00%	1.69	6.00	1.0000
Totale	$5,776.00	$3,694.44	$1,539	26.64%			6.00
Scopo:	MAX	< =$5000					<=6
Rapporto di Sharpe	3.7543						

ENPV è il valore attuale netto (NPV) atteso per ciascun investimento o progetto, mentre Costo può essere il costo totale dell'investimento e Rischio è il Coefficiente di variazione del ENPV del progetto.

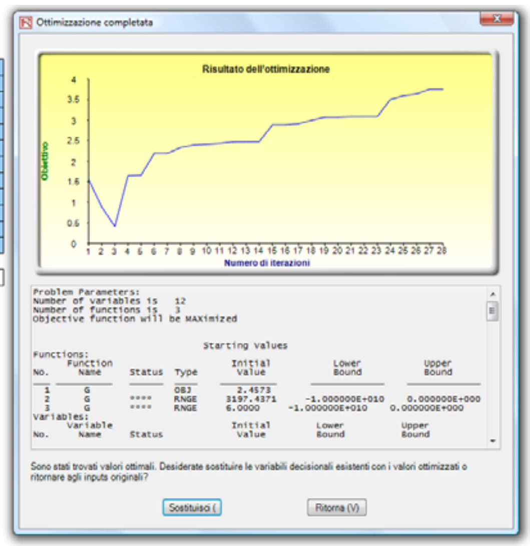


Figura 4.6 - Selezione Ottimale dei Progetti che Massimizzano il Rapporto di Sharpe

Per addizionali esempi pratici dell'ottimizzazione in azione, consultate il caso di studio nel Capitolo 11 su *Integrated Risk Analysis* nel libro *Real Options Analysis: Tools and Techniques*, 2nd Edition (Wiley Finance, 2005). Questo caso di studio spiega come generare una frontiera efficiente e come la previsione, la simulazione, l'ottimizzazione e le opzioni reali possono essere combinate in un processo analitico perfetto.

4.4 Frontiera Efficiente e Impostazioni avanzate dell'Ottimizzazione

Il secondo diagramma nella Figura 4.5 mostra i Vincoli per l'ottimizzazione. Qui, se avete cliccate sul bottone ***Frontiera Efficiente*** dopo aver impostato dei vincoli, potete rendere questi vincoli mutevoli. In altre parole, ciascuno dei vincoli può essere creato per intervenire tra determinati valori minimi e massimi. Ad esempio, si può impostare il vincolo nella cella J17 <= 6 per essere eseguito tra 4 e 8 (Figura 4.7). In altre parole, saranno eseguite cinque ottimizzazioni, ciascuna con i seguenti vincoli: J17 <= 4, J17 <= 5, J17 <= 6, J17 <= 7 and J17 <= 8. I risultati ottimali saranno poi tracciati come una frontiera efficiente e sarà generato un report (Figura 4.8). Di seguito illustriamo in modo specifico i passi richiesti per creare un vincolo mutevole:

- In un modello di ottimizzazione (vale a dire, un modello con Obiettivo, Variabili Decisionali e Vincoli già impostati), cliccate su ***Simulatore di Rischio | Ottimizzazione | Vincoli*** e poi su ***Frontiera Efficiente***.
- Selezionate il vincolo che volete rendere mutevole o che deve intervenire a passi (per esempio, J17), inserite i parametri per Min, Max e Dimensione passo (Figura 4.7), cliccate su ***AGGIUNGI*** e poi su ***OK*** e di nuovo su ***OK***. Dovete disattivare il vincolo D17 <= 5000 prima dell'esecuzione.
- Eseguite l'Ottimizzazione come sempre (***Simulatore di Rischio | Ottimizzazione | Esegui ottimizzazione***). Potete scegliere statica, dinamica o stocastica.
- I risultati saranno visualizzati come un'interfaccia utente (Figura 4.8). Cliccate su ***Crea Report*** per generare un report sotto forma di foglio di lavoro con tutti i dettagli delle esecuzioni dell'ottimizzazione.

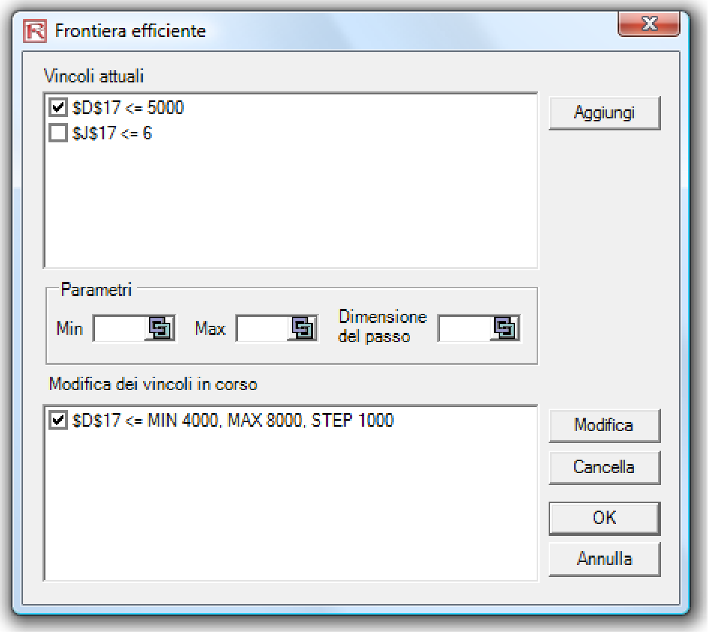
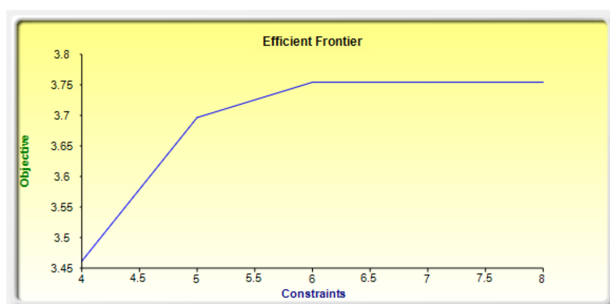


Figura 4.7 - Generare Vincoli Mutevoli in una Frontiera Efficiente

Frontiera efficiente

Problem Parameters:	
Number of variables	12
Number of functions	2
Objective function will be	Maximized



STEP1, D17 <= 5000, J17 <= 4

Functions

Starting Values

No.	Function Name	Status	Type	Initial Value	Lower Bound	Upper Bound
1	G		OBJ	2.45726		
2	G	****	RNGE	8.00000	-1E+10	0

Final Results

No.	Function Name	Initial Value	Final Value	Status	Distance from Nearest Bound	Lagrange Multiplier
1	G	2.45726	3.46137	Objective		
2	G	8.00000	0.00000	UpperBnd	0.0000:U	0.23768

Variables

Starting Values

No.	Variable Name	Status	Initial Value	Lower Bound	Upper Bound
1	Project 1	UL	1.00000	0	1

Final Results

No.	Variable Name	Initial Value	Final Value	Status	Distance from Nearest Bound	Reduced Gradient
1	Project 1	1.00000	1.00000	NonBasic	UpperBnd	0.0511

FigurA 4.8 - Risultati della Frontiera Efficiente

4.5 Ottimizzazione Stocastica

Il prossimo esempio illustra l'applicazione di una ottimizzazione stocastica usando un modello d'esempio con 4 classi di assets, ciascuna con differenti caratteristiche di rischio e rendimento. L'idea qui è di trovare la migliore allocazione del portafoglio in modo da massimizzare il rendimento ("bang-for-the-buck") o il rapporto rendimenti/rischio del portafoglio. In altre parole, l'obiettivo è di allocare il 100% degli investimenti di un individuo fra varie classi differenti di assets (per esempio, tipi diversi di fondi comuni o stili d'investimento: crescita, valore, crescita aggressiva, reddito, globale, indice, "contrarian", "momentum" e così via). Questo modello è diverso dagli altri in quanto ci sono svariate ipotesi di simulazione (valori del rischio e del rendimento per ciascun asset nelle colonne C e D), come mostrato nella Figura 4.9.

Viene eseguita una simulazione, poi viene eseguita una ottimizzazione e l'intero processo è ripetuto multiple volte per ottenere le distribuzioni di ciascuna variabile decisionale. L'intera analisi può essere automatizzata usando l'Ottimizzazione stocastica.

Per poter eseguire una ottimizzazione , bisogna prima identificare varie specifiche fondamentali del modello:

Obiettivo: *Massimizzare il rapporto rendimento/rischio (C12)*

Variabili Decisionali: *Pesi di allocazione (E6:E9)*

Restrizioni sulle variabili decisionali: Minimo e Massimo richiesti (F6:G9)

Vincoli: *Pesi di allocazione totali del portafoglio 100% (E11 è impostato su 100%)*

Ipotesi della Simulazione: *Valori del Rendimento e del Rischio (C6:D9)*

Il modello mostra le varie classi di assets. Ciascuna classe di assets ha il suo set di rendimenti annualizzati e volatilità annualizzate. Queste misure di rendimento e rischio sono valori annualizzati così da poterli confrontare coerentemente con classi di assets differenti. I rendimenti sono calcolati usando la media geometrica dei rendimenti relativi, mentre i rischi sono calcolati usando il metodo logaritmico dei rendimenti relativi di titoli.

I Pesi di allocazione nella colonna E contengono le variabili decisionali. Queste sono le variabili che devono essere "aggiustate" e testate in modo che il peso totale sia limitato a 100% (cella E11). Tipicamente, per avviare l'ottimizzazione s'impostano queste celle su un valore uniforme: in questo caso le celle da E6 ad E9 sono impostate ciascuna al 25%. In aggiunta, ciascuna variabile decisionale può avere specifiche restrizioni nel suo campo consentito. In quest'esempio, le allocazioni inferiori e superiori consentite sono 10% e 40%, come mostrato nelle colonne F e G. Questa impostazione significa che ciascuna classe di assets può avere i suoi limiti di allocazione.

MODELLO DI OTTIMIZZAZIONE - ALLOCAZIONE DEGLI ASSETS

Descrizione della classe degli assets	Rendimenti annualizzati	Rischio di volatilità	Pesi di allocazione	Allocazione minima richiesta	Allocazione massima richiesta	Rapporto rendimenti/ rischi
Asset 1	10.60%	12.41%	25.00%	10.00%	40.00%	0.8544
Asset 2	11.21%	16.16%	25.00%	10.00%	40.00%	0.6937
Asset 3	10.61%	15.93%	25.00%	10.00%	40.00%	0.6660
Asset 4	10.52%	12.40%	25.00%	10.00%	40.00%	0.8480
Totale Portafoglio	*10.7356%*	*7.17%*	*100.00%*			
Rapporto Rendimenti/Rischi	*1.4970*					

Figura 4.9: Modello di Allocazione degli Assets, pronto per una ottimizzazione stocastica

Proseguendo, la colonna H mostra il rapporto rendimenti/rischi, che è semplicemente la percentuale dei rendimenti divisa per la percentuale dei rischi per ciascun asset, dove tanto più alto è questo valore, tanto più alto è il rendimento ("bang-for-the-buck"). Le parti restanti del modello mostrano la classifica delle individuali classi di assets per rendimento, rischio, rapporto rendimenti/rischi e allocazione. In altre parole, questa classifica mostra in un'occhiata quale classe di assets ha il rischio più basso o il rendimento più alto e così via.

Eseguire una Ottimizzazione

Per eseguire questo modello, cliccate semplicemente su **Simulatore di Rischio | Ottimizzazione | Esegui ottimizzazione**. Alternativamente e per esercitarvi, potete impostare il modello usando i seguenti passi.

- Avviate un nuovo profilo (Simulatore di Rischio | Nuovo profilo).
- Per una ottimizzazione stocastica, impostate ipotesi distribuzionali sul rischio e sui rendimenti per ciascuna classe di assets. In altre parole, selezionate la cella C6 e impostate un'ipotesi (Simulatore di Rischio | Imposta ipotesi di input) e fate la vostra ipotesi come richiesto. Ripetete la procedura per le celle da C7 a D9.

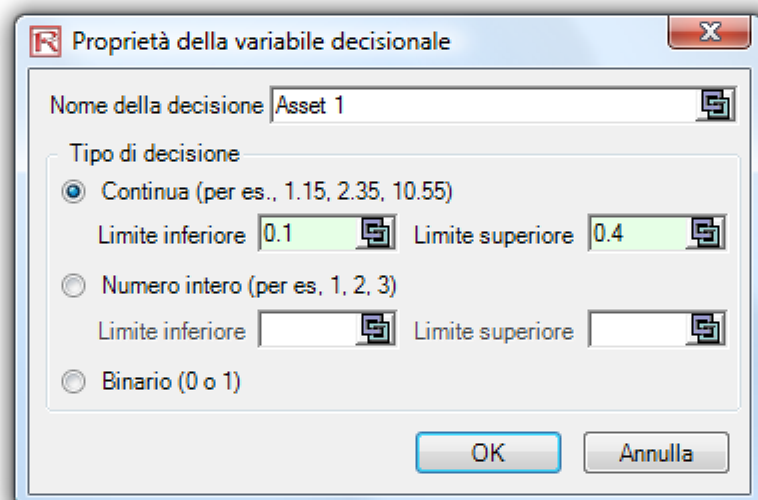
- Selezionate la cella E6 e definite la variabile decisionale (Simulatore di Rischio | Ottimizzazione | Imposta decisione oppure cliccate sull'icona *Imposta decisione D*). Impostate questa come una Variabile continua e poi collegate il nome e il minimo/massimo richiesto della variabile decisionale alle celle pertinenti (B6, F6, G6).

- Dopo di questo, usate la funzione Copia di Simulatore di Rischio sulla cella E6, selezionate le celle da E7 ad E9 e usate la funzione Incolla di Simulatore di Rischio (Simulatore di Rischio | Copia parametro e Simulatore di Rischio | Incolla parametro o usate le icone Copia e Incolla). Ricordatevi di non usare le normali funzioni di Copia e Incolla di Excel.

- Poi, impostate i vincoli dell'ottimizzazione selezionando prima **Simulatore di Rischio | Ottimizzazione | Vincoli** e poi **AGGIUNGI**. Ora selezionate la cella **E11** e impostatela uguale a **100%** (allocazione totale; non scordatevi il simbolo %).

- Selezionate la cella **C12**, l'obiettivo da massimizzare, e impostatela come l'obiettivo: **Simulatore di Rischio | Ottimizzazione | Imposta obiettivo** oppure cliccate sull'icona **O**.

- Eseguite l'ottimizzazione selezionando **Simulatore di Rischio | Ottimizzazione | Esegui ottimizzazione**. Esaminate le varie linguette per assicuravi che tutti gli inputs richiesti nei passi 2 e 3 sono corretti. Selezionate **Ottimizzazione stocastica** e fatela eseguire per 500 prove ripetute 20 volte (la Figura 4.10 illustra questi passi di setup).



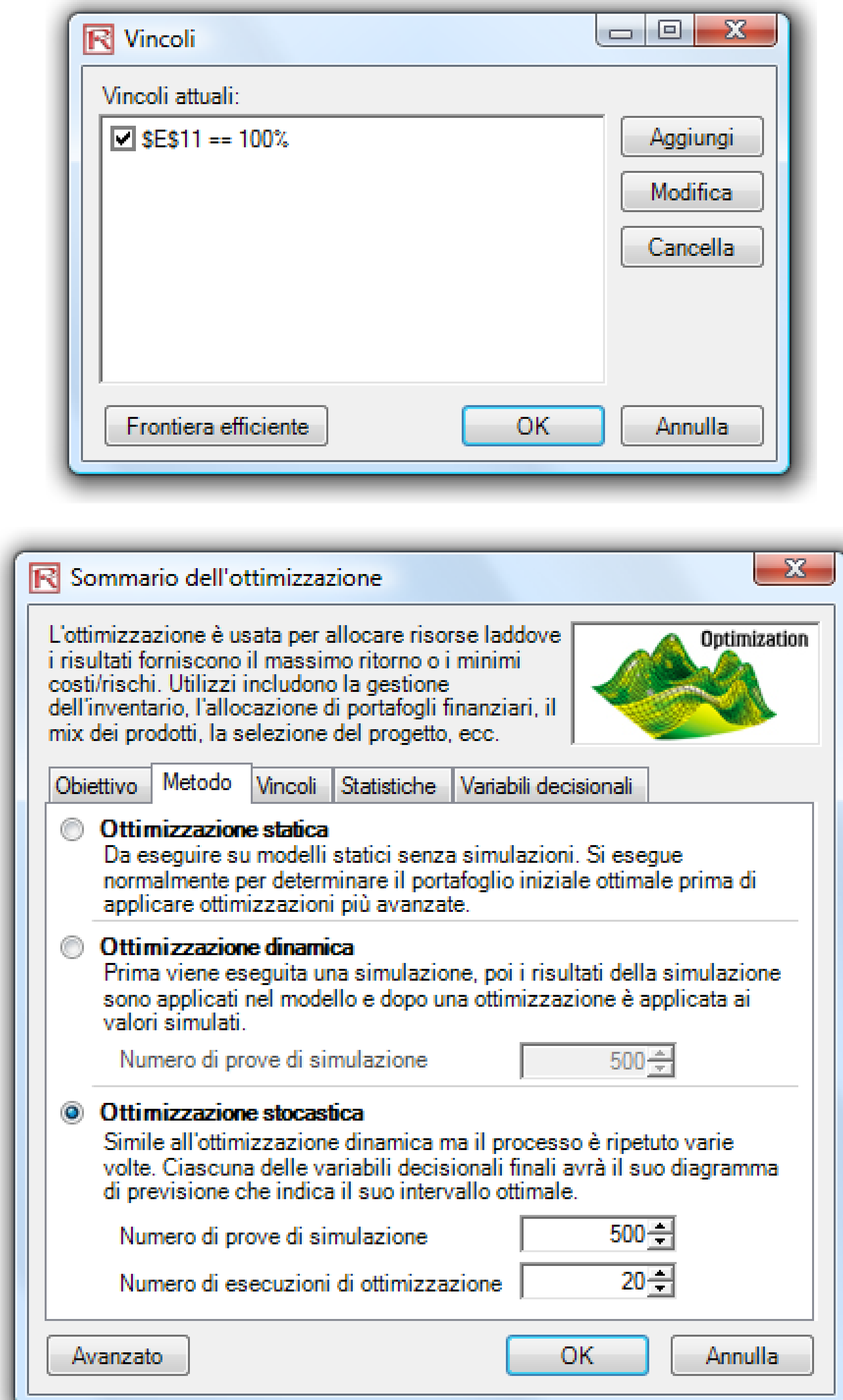


Figura 4.10 - Impostare una ottimizzazione stocastica

Cliccate su **OK** al termine della simulazione e sarà generato un report dettagliato sull'ottimizzazione stocastica insieme ai diagrammi di previsione delle variabili decisionali.

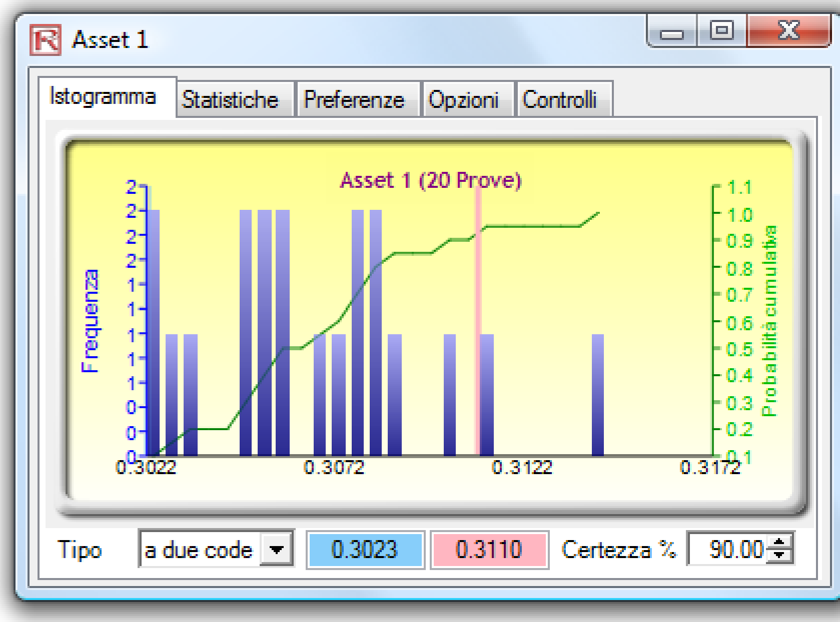
Visualizzare e Interpretare i Risultati della Previsione

L'ottimizzazione stocastica viene eseguita, quando si esegue prima una simulazione e dopo una ottimizzazione. Poi l'intera analisi è ripetuta molte volte. Il risultato è una distribuzione di ciascuna variabile decisionale piuttosto che una stima di un singolo punto (Figura 4.11). Questo significa che invece di dire che si dovrebbe investire il 30.57% nell'Asset 1, la decisione ottimale è di investire tra il 30.10% ed il 30.99% purché il totale del portafoglio sia di 100%. In questa maniera, i risultati forniscono al management o ai decisori un campo di flessibilità per le decisioni ottimali, tenendo al contempo conto dei rischi e delle incertezze degli inputs.

Note:

- **Simulazione Super Veloce con Ottimizzazione**. Potete anche eseguire l'ottimizzazione stocastica con la Simulazione Super Veloce. Per eseguire questa combinazione, per prima cosa resettate l'ottimizzazione e reimpostate tutte le quattro variabili decisionali al 25%. Dopo cliccate su ***Esegui ottimizzazione*** e poi sul bottone ***Avanzato*** (Figura 4.10). Ora selezionate la casella di controllo ***Esegui Simulazione Super Veloce***. Per continuare, nell'interfaccia utente dell'esecuzione dell'ottimizzazione, selezionate ***Ottimizzazione stocastica*** nella linguetta del ***Metodo*** e impostatela per eseguire ***500*** prove e ***20*** esecuzioni dell'ottimizzazione. Adesso cliccate su ***OK***. Questo metodo integra la Simulazione Super Veloce con una ottimizzazione. Notate la maggiore velocità di esecuzione dell'ottimizzazione stocastica. In questo modo potete velocemente rieseguire l'ottimizzazione con un numero più grande di prove di simulazione.
- **Statistiche della Simulazione per Ottimizzazioni Stocastiche e Dinamiche**. Vi preghiamo di notare che se sono state inserite delle ipotesi di simulazione nel modello di ottimizzazione (queste ipotesi d'input sono obbligatorie per poter eseguire le procedure di una ottimizzazione dinamica o stocastica), la linguetta delle ***Statistiche*** nell'interfaccia utente di ***Esegui ottimizzazione*** sarà popolata. Dal menu a tendina potete selezionare le statistiche che v'interessano: media, deviazione standard, coefficiente di variazione, media condizionata, varianza condizionata, un percentile specifico e così via. Questo significa che se eseguite una ottimizzazione stocastica, verrà prima eseguita una simulazione di migliaia di prove, poi la statistica selezionata verrà calcolata e questo valore sarà inserito temporaneamente nella cella dell'ipotesi della simulazione, e dopo verrà eseguita una ottimizzazione basata su questa statistica. Infine, l'intero processo viene ripetuto multiple volte. Questo

metodo è importante e utile per le applicazioni bancarie nel calcolo del Valore a rischio condizionato o VaR condizionato.



Asset 1

Istogramma | Statistiche | Preferenze | Opzioni | Controlli

Statistiche	Risultato
Numero di prove	20
Media	0.3064
Mediana	0.3063
Deviazione standard	0.0031
Varianza	0.0000
Coefficiente di variazione	0.0102
Massimo	0.3142
Minimo	0.3019
Intervallo	0.0123
Assimetria	0.6905
Curtosi	0.6370
25% Percentile	0.3045
75% Percentile	0.3079
Precisione d'errore percentuale a 95% di confi...	0.4459%

Figura 4.11 - Risultati simulati del metodo di ottimizzazione stocastica

5

5. STRUMENTI ANALITICI DI SIMULATORE DI RISCHIO

Questo capitolo tratta gli strumenti analitici di Simulatore di Rischio. Questi strumenti analitici sono discussi mediante l'uso di esempi di applicazioni del software Simulatore di Rischio, comprese le illustrazioni passo passo. Questi strumenti sono preziosi per gli analisti che lavorano nel campo dell'analisi del rischio. L'applicabilità di ciascuno strumento è dettagliatamente discussa in questo capitolo.

5.1 Strumenti Tornado e Sensibilità nella Simulazione

Teoria:

Uno degli strumenti più potenti di simulazione è l'analisi Tornado – questa cattura gli impatti statici di ciascuna variabile sull'esito del modello. In altre parole, lo strumento perturba automaticamente ciascuna variabile nel modello di una quantità predefinita, cattura le fluttuazioni sulla previsione o sui risultati finali del modello ed elenca le risultanti perturbazioni disposte dalla più alla meno significativa. Le Figure da 5.1 a 5.6 illustrano le applicazioni di un'analisi tornado. Per esempio, la Figura 5.1 è un esempio di un modello di cash flow scontato dove sono mostrate le ipotesi d'input. La domanda è: quali sono i critici elementi motore del successo che influenzano maggiormente l'output del modello? In altre parole, cosa "muove" veramente il valore netto attuale di $96.63 o quale variabile d'input ha l'impatto maggiore su questo valore?

Si può accedere allo strumento del diagramma tornado sotto ***Simulatore di Rischio | Strumenti | Analisi Tornado***. Per poter seguire il primo esempio, aprite il file ***Diagrammi Tornado e Sensibilità (Lineari)*** nella cartella degli esempi. La Figura 5.2 mostra questo modello d'esempio, dove la cella G6 che contiene il valore attuale netto è scelta come il risultato dell'obiettivo da analizzare. I precedenti ("precedents") della cella d'obiettivo nel modello sono usati per creare il diagramma tornado. I Precedenti sono tutti gli inputs e tutte le variabili intermedie che influiscono sull'esito del modello. Per esempio, se il modello è composto da A = B + C, dove C = D + E, allora

B, D ed E sono i precedenti per A (C non è un precedente in quanto è solo un valore calcolato intermedio). La Figura 5.2 mostra anche il campo di analisi di ciascuna variabile precedente che è stata usata per stimare il risultato dell'obiettivo. Se le variabili precedenti sono semplici inputs, allora il campo di analisi sarà una semplice perturbazione basata sul campo scelto (per esempio, il default è ±10%). Ciascuna variabile precedente, se necessario, può essere perturbata di percentuali differenti. Un intervallo più ampio è importante, dato che può testare meglio i valori estremi piuttosto perturbazioni più piccole attorno ai valori attesi. In certe circostanze, i valori estremi potrebbero avere un impatto più grande, più piccolo o non equilibrato (per esempio, potrebbero accadere delle non linearità laddove crescenti o decrescenti economie di scala ed economie di diversificazione s'inseriscono lentamente per valori più grandi o più piccoli di una variabile) e solo un intervallo più ampio può catturare questo impatto non lineare.

Modello Cash Flow Scontato

Anno base	*2005*	Somma valore attuale (PV) benefici netti	***$1,896.63***
Tasso di sconto del mercato corretto per il rischio	*15.00%*	Somma valore attuale (PV) investimenti	***$1,800.00***
Tasso di sconto per rischio privato	*5.00%*	Valore attuale netto (NPV)	***$96.63***
Tasso di crescita annualizzato delle vendite	*2.00%*	Tasso di rendimento interno (IRR)	*18.80%*
Tasso di erosione dei prezzi	*5.00%*	Rendimento sull'investimento (ROI)	*5.37%*
Aliquota d'imposta effettiva	*40.00%*		

	2005	*2006*	*2007*	*2008*	*2009*
Prod A Prezzo Medio	$10.00	$9.50	$9.03	$8.57	$8.15
Prod B Prezzo Medio	$12.25	$11.64	$11.06	$10.50	$9.98
Prod C Prezzo Medio	$15.15	$14.39	$13.67	$12.99	$12.34
Prod A Quantità	50.00	51.00	52.02	53.06	54.12
Prod B Quantità	35.00	35.70	36.41	37.14	37.89
Prod C Quantità	20.00	20.40	20.81	21.22	21.65
Ricavi totali	***$1,231.75***	***$1,193.57***	***$1,156.57***	***$1,120.71***	***$1,085.97***
Costo dei beni venduti	$184.76	$179.03	$173.48	$168.11	$162.90
Profitto lordo	***$1,046.99***	***$1,014.53***	***$983.08***	***$952.60***	***$923.07***
Spese operative	$157.50	$160.65	$163.86	$167.14	$170.48
Spese generali, admin. e di vendita	$15.75	$16.07	$16.39	$16.71	$17.05
Utile di esercizio (EBITDA)	***$873.74***	***$837.82***	***$802.83***	***$768.75***	***$735.54***
Deprezzamento	$10.00	$10.00	$10.00	$10.00	$10.00
Ammortamento	$3.00	$3.00	$3.00	$3.00	$3.00
Utile al lordo di tasse e interessi (EBIT)	***$860.74***	***$824.82***	***$789.83***	***$755.75***	***$722.54***
Pagamenti di interessi	$2.00	$2.00	$2.00	$2.00	$2.00
Utile al lordo di tasse (EBT)	***$858.74***	***$822.82***	***$787.83***	***$753.75***	***$720.54***
Tasse	$343.50	$329.13	$315.13	$301.50	$288.22
Utile netto	***$515.24***	***$493.69***	***$472.70***	***$452.25***	***$432.33***
Deprezzamento	$13.00	$13.00	$13.00	$13.00	$13.00
Cambio nel capitale circolante netto	$0.00	$0.00	$0.00	$0.00	$0.00
Spese per capitale	$0.00	$0.00	$0.00	$0.00	$0.00
Cash Flow libero	***$528.24***	***$506.69***	***$485.70***	***$465.25***	***$445.33***
Investimenti	***$1,800.00***				

Figura 5.1 - Esempio di un Modello

Procedura:

- Selezionate la singola cella di output (vale a dire, una cella con una funzione o equazione) in un modello Excel (ad esempio, è stata selezionata la cella G6 nel nostro esempio)
- Selezionate Simulatore di Rischio | Strumenti | Analisi Tornado
- Esaminate i precedenti e rinominateli se necessario (rinominare i precedenti con nomi più corti rende il diagramma tornado e ragno visivamente più attraente) e cliccate su **OK**

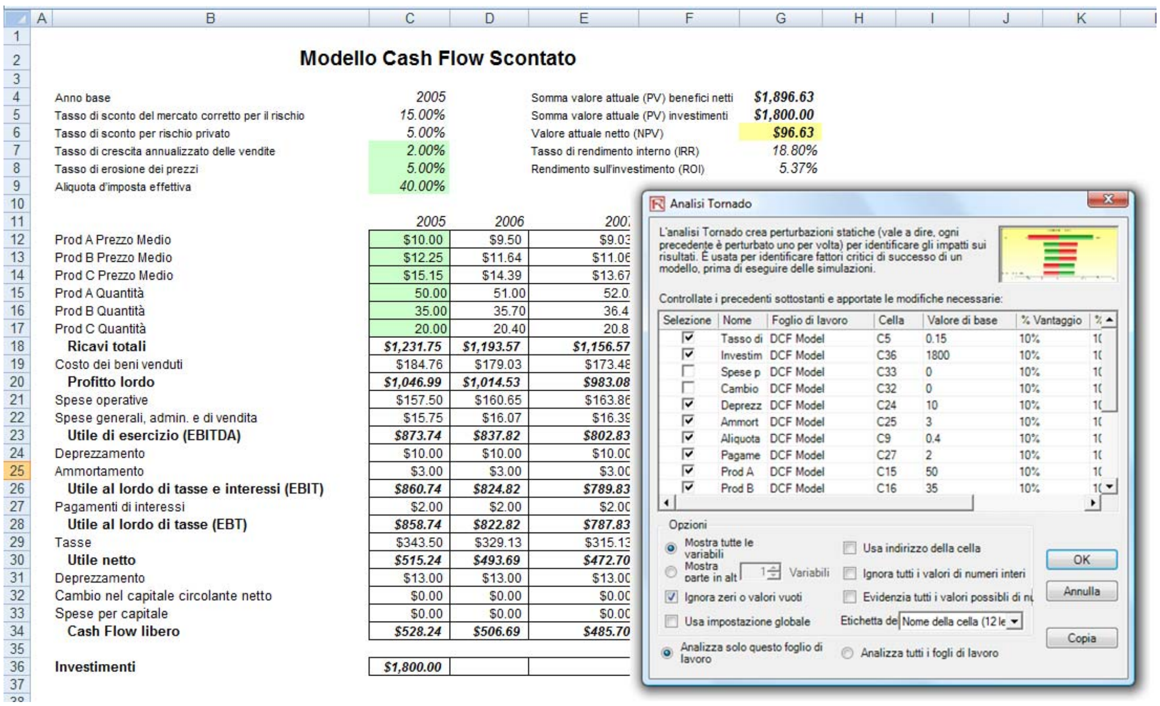


Figura 5.2 – Eseguire una Analisi Tornado

Interpretare i Risultati:

La Figura 5.3 mostra il risultante report dell'analisi tornado, che indica che l'investimento di capitale ha l'impatto maggiore sul valore attuale netto, seguiti dall'aliquota d'imposta, il prezzo medio di vendita e la quantità richiesta delle linee del prodotto e così via. Il report contiene quattro elementi distinti:

- Il Sommario statistico, che elenca le procedure eseguite.
- La Tabella di Sensibilità (Figura 5.4) mostra il valore base iniziale del valore attuale netto (NPV) di 96.63 e come ogni input è modificato (per esempio, Investimento è cambiato da $1800 a $1980 con un'oscillazione a vantaggio del +10% swing, e da $1800 a $1620 con un'oscillazione a

svantaggio del –10%). I risultanti valori di Vantaggio e Svantaggio per il valore attuale netto (NPV) sono –$83.37 e $276.63, con un cambiamento totale di $360. Questo lo rende la variabile con il maggior impatto sul valore attuale netto (NPV). Le variabili precedenti sono classificate dall'impatto maggiore all'impatto minore.

- Il diagramma ragno (Figura 5.5) illustra questi effetti graficamente. L'asse y è il valore di obiettivo del valore attuale netto (NPV), mentre l'asse x rappresenta il cambio percentuale rispetto a ciascuno dei valori precedenti (il punto centrale è il valore del caso base a 96.63 con 0% di cambiamento dal valore base di ciascun precedente). Una linea con inclinazione positiva indica una relazione positiva o un effetto positivo, mentre linee con inclinazione negativa indicano una relazione negativa (per esempio, Investimento è inclinata negativamente; questo significa che tanto più alto è il livello d'investimento, tanto più alto è il valore attuale netto (NPV). Il valore assoluto della pendenza indica la magnitudine dell'effetto (una forte pendenza indica un impatto più alto sull'asse y del valore attuale netto (NPV), dato un cambiamento nell'asse x del precedente).

- Il diagramma tornado illustra questo in un'altra maniera grafica, dove il precedente con l'impatto maggiore è elencato per primo. L'asse x è il valore del valore attuale netto (NPV); il centro del diagramma è la condizione del caso base. Le barre verdi nel diagramma indicano un effetto positivo, mentre le barre rosse indicano un effetto negativo. Per gli investimenti, quindi, le barre rosse sul lato destro indicano un effetto negativo dell'investimento sul valore attuale netto (NPV) – in altre parole, investimento di capitale e valore attuale netto (NPV) sono negativamente correlati. Il contrario è vero per prezzo e quantità dei prodotti da A a C (le loro barre verdi sono sul lato destro del diagramma).

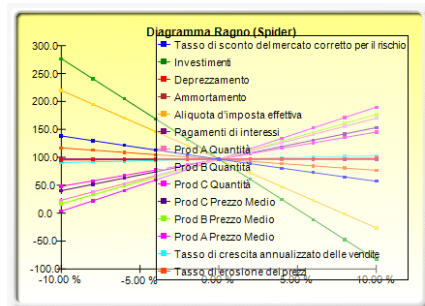
Diagrammi Tornado e Ragno (Spider)

Sommario statistico

Uno degli strumenti più potenti di simulazione è il diagramma Tornado – esso cattura gli impatti statici di ciascuna variabile sull'esito del modello. In altre parole, lo strumento perturba automaticamente ciascuna variabile precedente nel modello di una quantità predefinita impostata dall'utente, cattura le fluttuazioni sulla previsione o sui risultati finali del modello ed elenca le risultanti perturbazioni disposte dalla più alla meno significativa. I Precedenti sono tutti gli inputs e tutte le variabili intermedie che influiscono sull'esito del modello. Per esempio, se il modello è composto da A = B + C, dove C = D + E, allora B, D ed E sono i precedenti per A (C non è un precedente in quanto è solo un valore calcolato intermedio). L'intervallo ed il numero di valori perturbati sono specificati dall'utente e possono essere impostati per testare valori estremi piuttosto che perturbazioni più piccole attorno ai valori attesi. In certe circostanze, i valori estremi potrebbero avere un impatto più grande, più piccolo o non equilibrato (per esempio, potrebbero accadere non linearità laddove si verificano crescenti o decrescenti economie di scala e allontanamento dallo scopo ("scope creep") per valori più grandi o più piccolo di una variabile) e solo un intervallo più ampio può catturare questo impatto non lineare.
Un diagramma Tornado elenca tutti gli inputs che muovono il modello, iniziando dalla variabile di input che ha il maggiore effetto sui risultati. Il diagramma è ottenuto perturbando ciascun input precedente ad un certo intervallo consistente (per es., ±10% dal caso-base) uno alla volta e paragonando i loro risultati al caso-base. Un diagramma Ragno (Spider) assomiglia ad un ragno con un corpo centrale e le sue molte zampe sporgenti. La linea con inclinazione positiva indica una relazione positiva, mentre una linea con inclinazione negativa indica una relazione negativa. Si possono inoltre usare i diagrammi Ragno per visualizzare relazioni lineari e non lineari. I diagrammi Tornado e Ragno aiutano ad identificare i fattori critici di successo di una cella di output in modo da poter identificare gli inputs da simulare. Le variabili critiche identificate come incerte sono quelle che dovrebbero essere simulate. Non perdete tempo a simulare né le variabili che non sono incerte né quelle che hanno poco impatto sui risultati.

Risultato

	Valore di base: 96.6261638553219			Cambiamenti dell'input		
Cella precedente	Svantaggio dell'output	Vantaggio dell'output	Intevallo efficace	Svantaggio dell'input	Vantaggio dell'input	Valore del caso-base
C36: Investimenti	276.62616	-83.373836	360.00	$1,620.00	$1,980.00	$1,800.00
C9: Aliquota d'imposta effettiva	219.72693	-26.474599	246.20	36.00%	44.00%	40.00%
C12: Prod A Prezzo Medio	3.4255424	189.82679	186.40	$9.00	$11.00	$10.00
C13: Prod B Prezzo Medio	16.706631	176.5457	159.84	$11.03	$13.48	$12.25
C15: Prod A Quantità	23.177498	170.07483	146.90	45.00	55.00	50.00
C16: Prod B Quantità	30.533	162.71933	132.19	31.50	38.50	35.00
C14: Prod C Prezzo Medio	40.146587	153.10574	112.96	$13.64	$16.67	$15.15
C17: Prod C Quantità	48.047369	145.20496	97.16	18.00	22.00	20.00
C5: Tasso di sconto del mercato corretto pe	138.23913	57.029841	81.21	13.50%	16.50%	15.00%
C8: Tasso di erosione dei prezzi	116.80381	76.640952	40.16	4.50%	5.50%	5.00%
C7: Tasso di crescita annualizzato delle ven	90.588354	102.68541	12.10	1.80%	2.20%	2.00%
C24: Deprezzamento	95.084173	98.168155	3.08	$9.00	$11.00	$10.00
C25: Ammortamento	96.163566	97.088761	0.93	$2.70	$3.30	$3.00
C27: Pagamenti di interessi	97.088761	96.163566	0.93	$1.80	$2.20	$2.00



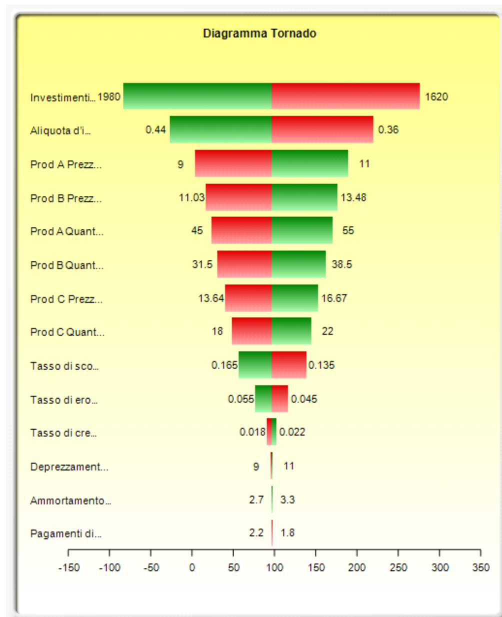


Figura 5.3 – Report dell'Analisi Tornado

Note: Ricordatevi che l'analisi tornado è un'analisi di sensibilità *statica* applicata su ciascuna variabile d'input nel modello – vale a dire, ciascuna variabile è perturbata individualmente e gli effetti risultanti sono tabulati. Questo rende l'analisi tornado una componente chiave da eseguire prima di effettuare una simulazione. Uno dei primi passi nell'analisi del rischio è dove si catturano e si identificano i più importanti elementi motore dell'impatto. Il passo successivo è di identificare quali di questi elementi motore sono incerti. Questi elementi motore incerti sono i critici elementi motore del successo di un progetto, dove i risultati del modello dipendono da questi critici elementi motore del successo. Sono queste le variabili da simulare. Non perdete tempo a simulare né le variabili che non sono incerte né quelle che hanno poco impatto sui risultati. I diagrammi Tornado aiutano ad identificare questi critici elementi motore del successo velocemente e facilmente. Seguendo quest'esempio, si può simulare prezzo e quantità, presumendo che l'investimento richiesto e l'aliquota d'imposta effettiva siano noti in anticipo e non mutevoli.

	Valore di base: 96.6261638553219			Cambiamenti dell'input		
Cella precedente	Svantaggio dell'output	Vantaggio dell'output	Intevallo efficace	Svantaggio dell'input	Vantaggio dell'input	Valore del caso-base
C36: Investimenti	276.62616	-83.373836	360.00	$1,620.00	$1,980.00	$1,800.00
C9: Aliquota d'imposta effettiva	219.72693	-26.474599	246.20	36.00%	44.00%	40.00%
C12: Prod A Prezzo Medio	3.4255424	189.82679	186.40	$9.00	$11.00	$10.00
C13: Prod B Prezzo Medio	16.706631	176.5457	159.84	$11.03	$13.48	$12.25
C15: Prod A Quantità	23.177498	170.07483	146.90	45.00	55.00	50.00
C16: Prod B Quantità	30.533	162.71933	132.19	31.50	38.50	35.00
C14: Prod C Prezzo Medio	40.146587	153.10574	112.96	$13.64	$16.67	$15.15
C17: Prod C Quantità	48.047369	145.20496	97.16	18.00	22.00	20.00
C5: Tasso di sconto del mercato corretto pe	138.23913	57.029841	81.21	13.50%	16.50%	15.00%
C8: Tasso di erosione dei prezzi	116.80381	76.640952	40.16	4.50%	5.50%	5.00%
C7: Tasso di crescita annualizzato delle ven	90.588354	102.68541	12.10	1.80%	2.20%	2.00%
C24: Deprezzamento	95.084173	98.168155	3.08	$9.00	$11.00	$10.00
C25: Ammortamento	96.163566	97.088761	0.93	$2.70	$3.30	$3.00
C27: Pagamenti di interessi	97.088761	96.163566	0.93	$1.80	$2.20	$2.00

Figura 5.4 – Tabella di Sensibilità

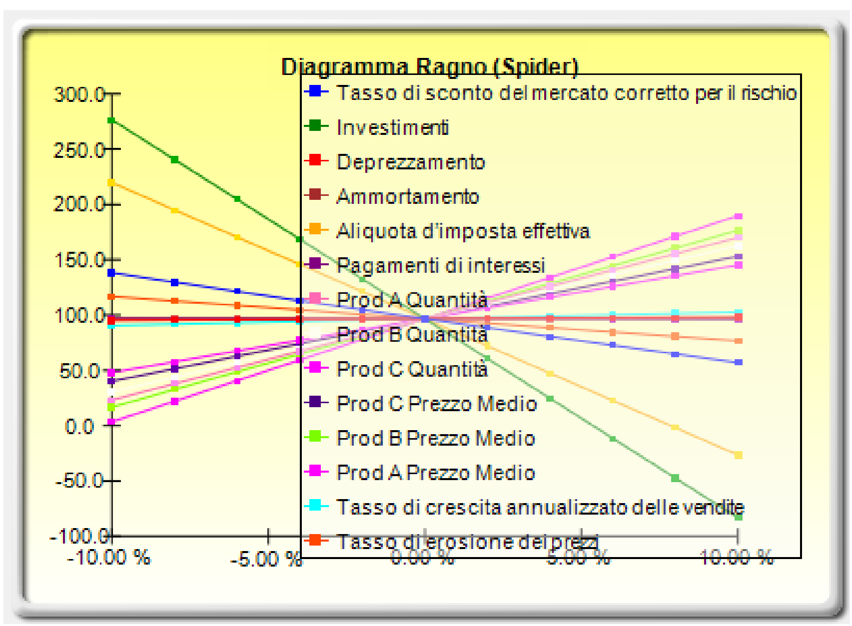


Figura 5.5 – Diagramma Spider

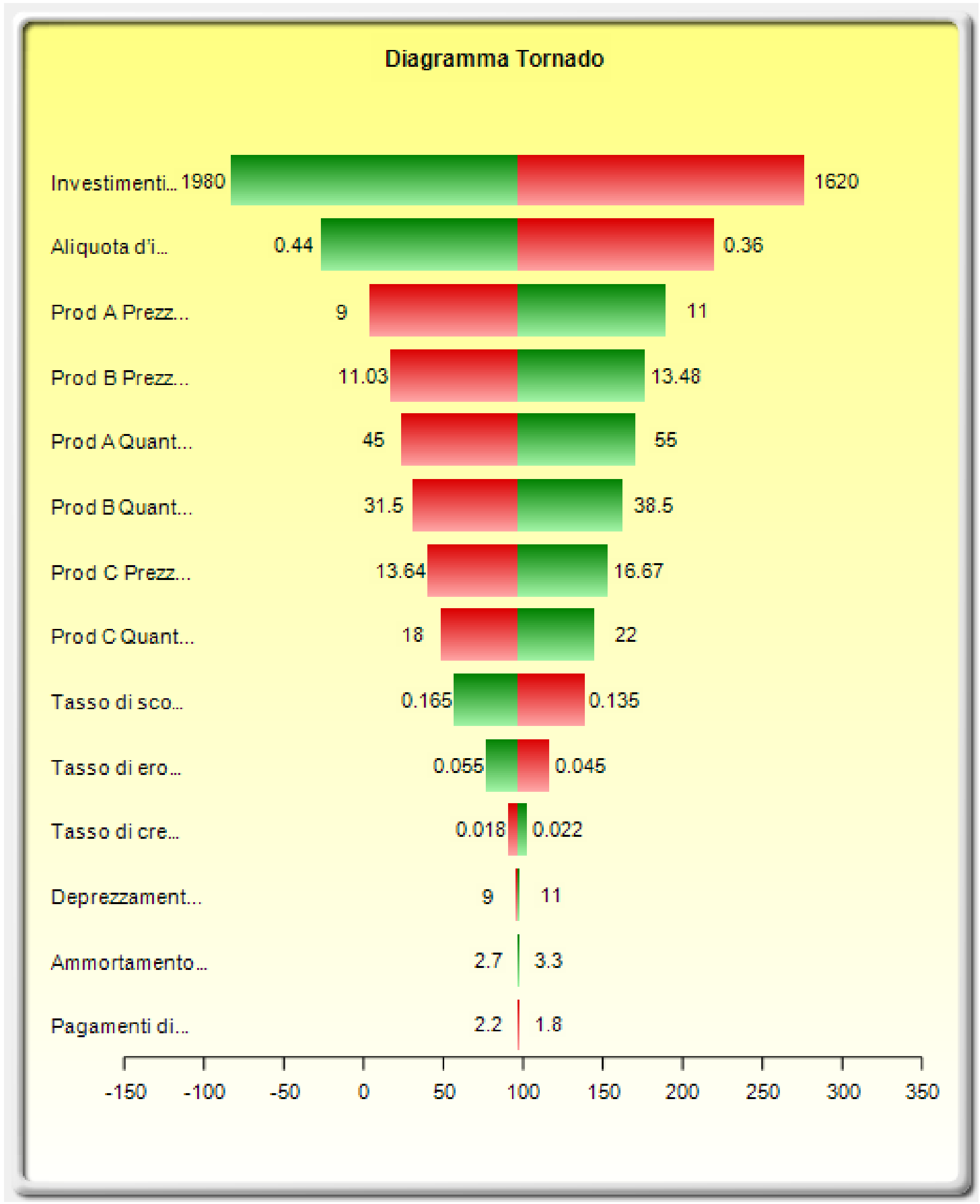


Figura 5.6 – Diagramma Tornado

Anche se il diagramma tornado è più facile da leggere, il diagramma spider è importante per determinare se esistono delle non linearità nel modello. Per esempio, la Figura 5.7 mostra un altro diagramma spider, dove le non linearità sono abbastanza chiare (le linee nel diagramma non sono diritte ma curve). Il modello usato è ***Diagrammi Tornado e di Sensibilità (Non lineari)***, che utilizza il modello per la valutazione delle opzioni di Black-Scholes come esempio. Tali non linearità non possono essere accertate mediante un diagramma tornado e possono costituire informazioni importanti per il modello o fornire concetti importanti sulle dinamiche del modello ai decisori.

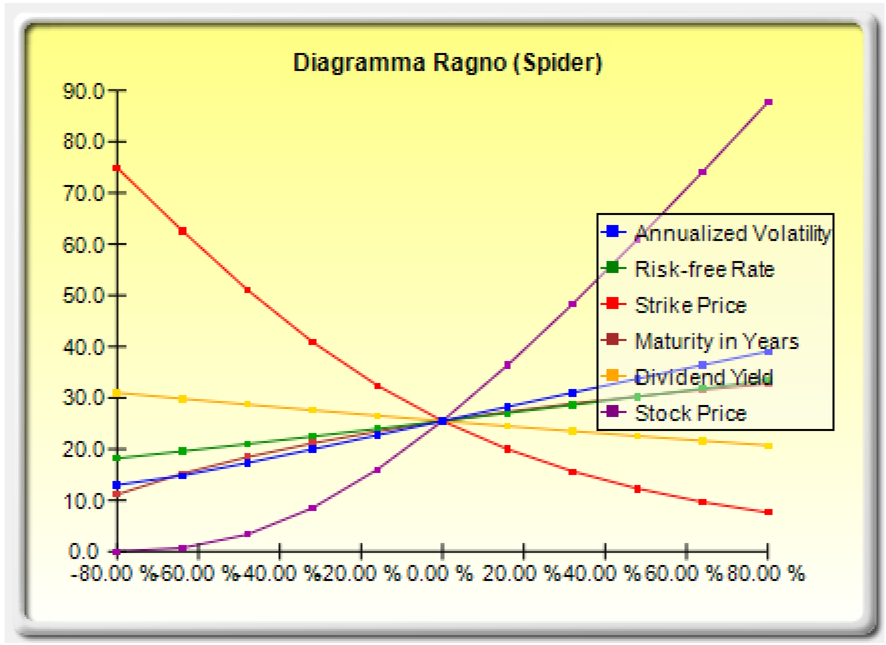


Figura 5.7 – Diagramma Ragno non lineare

Note aggiuntive su Tornado:

La Figura 5.2 mostra l'interfaccia utente dello strumento di analisi Tornado. Prego notare che ci sono alcuni nuovi miglioramenti dalla versione 4 di Simulatore di Rischio in poi. Seguono alcuni suggerimenti su come eseguire l'analisi Tornado e altri dettagli su questi nuovi miglioramenti:

- L'analisi Tornado non dovrebbe mai essere eseguita una sola volta. È intesa come uno strumento diagnostico per un modello, il che significa che dovrebbe idealmente essere eseguito varie volte sullo stesso modello. Per esempio, in un modello grande, l'analisi Tornado può essere eseguita una prima volta usando tutte le impostazioni di default e visualizzando tutti i precedenti (selezionate **Mostra tutte le variabili**). Il risultato potrebbe essere un report grande e dei lunghi (e potenzialmente inestetici) diagrammi Tornado. Tuttavia, quest'analisi fornisce un ottimo punto iniziale per determinare quanti dei precedenti sono da considerare come fattori critici del successo. Per esempio, il diagramma Tornado potrebbe mostrare che le prime 5 variabili hanno un grande impatto sull'output, mentre le rimanenti 200 variabili non hanno un impatto o hanno poco impatto, in qual caso, viene eseguita una seconda analisi Tornado che visualizza meno variabili (per esempio, selezionate **Mostra le prime 10**

variabili se le prime 5 sono quelle critiche; così create un report e diagramma Tornado ben fatto che mostra un contrasto tra i fattori fondamentali e i fattori meno critici. Non dovreste mai visualizzare un diagramma Tornado con le sole variabili fondamentali senza includere alcune variabili meno critiche per mostrare il contrasto dei loro effetti sull'output). Per finire, i punti di test di default possono essere aumentati dal ±10% del parametro ad un altro valore più grande per testare se ci sono delle non linearità (il diagramma Spider mostra linee non lineari e i diagrammi Tornado saranno asimmetrici verso un lato se gli effetti precedenti sono non lineari).

- L'opzione **Usa indirizzo di cella** è sempre una buona idea se il vostro modello è grande. Vi permette di identificare la posizione (nome del foglio di lavoro e indirizzo di cella) di una cella precedente. Se questa opzione non è selezionata, il software applicherà la propria *logica fuzzy* nel tentativo di determinare il nome di ciascuna variabile precedente (i nomi possono qualche volta causare confusione in un modello grande con variabili ripetute o possono essere troppo lunghi, rendendo il diagramma potenzialmente inestetico).
- Le opzioni **Analizza questo foglio di lavoro** e **Analizza tutti i fogli di lavoro** vi permettono di controllare se i precedenti dovrebbero fare parte solo del foglio di lavoro attuale o essere inclusi in tutti i fogli di lavoro della stessa cartella di lavoro. Questa opzione è utile se volete solo analizzare un output basato sui valori nel foglio attuale invece di eseguire una ricerca globale di tutti i precedenti collegati in multipli fogli di lavoro nella stessa cartella di lavoro.
- L'opzione **Usa impostazione globale** è utile se avete un modello grande e volete testare tutti i precedenti a, diciamo, ±50% piuttosto che al 10% di default. Invece di dover cambiare i valori del test di ciascun precedente uno per volta, potete selezionare questa opzione, cambiare una impostazione e *cliccare da un'altra parte* nell'interfaccia utente per modificare la lista completa dei precedenti. Se disattivate questa opzione, potete controllare la modifica dei punti di test un precedente per volta.
- L'opzione **Ignora zero o valori vuoti** è attivata di default. Significa che le celle precedenti che contengono zero o valori vuoti non saranno eseguite nell'analisi Tornado. Questa è l'impostazione tipica.
- L'opzione **Evidenzia possibili valori di numeri interi** identifica velocemente tutte le possibili celle precedenti che al momento contengono

inputs di valori interi. Questa funzione è talvolta importante se il vostro modello usa commutatori (per esempio, funzioni come "SE (IF) una cella è 1, allora accade qualcosa, e SE (IF) una cella ha il valore di 0, accade qualcos'altro", o numeri interi come 1, 2, 3 e così via, in altre parole elementi che non volete testare). Per esempio, ±10% di un valore flag di commutazione di 1 fornirà valori di test di 0.9 e 1.1. Entrambi sono valori d'input non pertinenti e non corretti nel modello, ed Excel potrebbe interpretare la funzione come un errore. Questa opzione, se selezionata, evidenzierà velocemente aree di problemi potenziali per l'analisi Tornado. Così potete determinare quali precedenti sono da attivare/disattivare manualmente o usare l'opzione **Ignora possibili valori di numeri interi** per disattivarli tutti simultaneamente.

5.2 Analisi di Sensibilità

Teoria:

Una funzione associata è l'analisi di sensibilità. Mentre l'analisi tornado (diagrammi tornado e diagrammi ragno) applica perturbazioni statiche *prima* dell'esecuzione di una simulazione, l'analisi di sensibilità applica perturbazioni dinamiche create *dopo* l'esecuzione di una simulazione. I diagrammi tornado e ragno sono i risultati di perturbazioni statiche, il che significa che ciascun precedente o variabile d'ipotesi è perturbata di una quantità predeterminata uno/a alla volta e le fluttuazioni nei risultati sono tabulate. Per contrasto, i diagrammi di sensibilità sono il risultato di perturbazioni dinamiche, nel senso che multiple ipotesi sono perturbate simultaneamente, e le loro interazioni nel modello e le correlazioni tra le variabili sono catturate nelle fluttuazioni dei risultati. I diagrammi Tornado identificano, perciò, quali variabili "muovono" di più i risultati e sono quindi adatti per una simulazione. I diagrammi di sensibilità viceversa identificano l'impatto sui risultati, quando multiple variabili interagenti sono simulate insieme nel modello. Questo effetto è chiaramente illustrato nella Figura 5.8. Prego notare che la classifica dei critici elementi motore del successo è simile a quella del diagramma tornado negli esempi precedenti. Tuttavia, se si aggiungono le correlazioni tra le ipotesi, la Figura 5.9 mostra un quadro molto diverso. Notate, per esempio, che l'erosione del prezzo aveva un impatto piccolo sul valore attuale netto (NPV). Ma quando alcune delle ipotesi d'input vengono correlate, l'interazione che esiste tra queste variabili correlate fa sì che l'erosione del prezzo abbia un impatto maggiore.

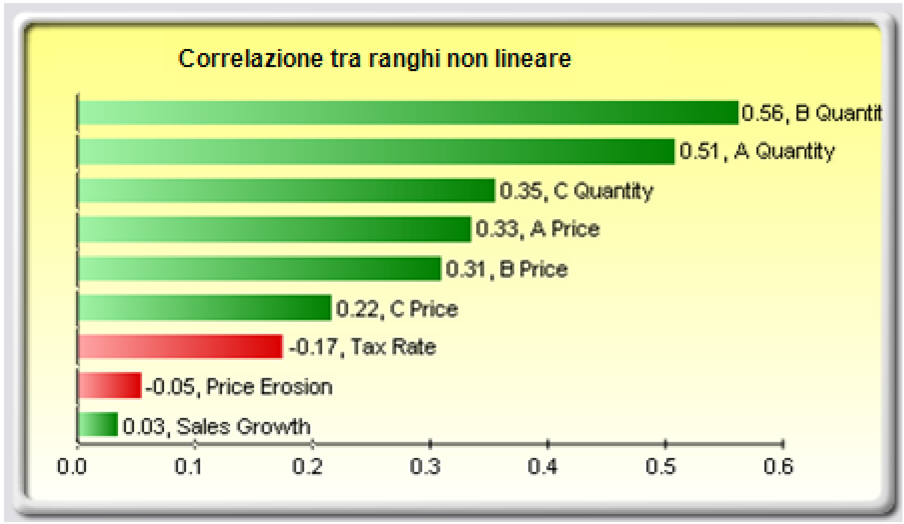


Figura 5.8 – Diagramma di Sensibilità senza Correlazioni

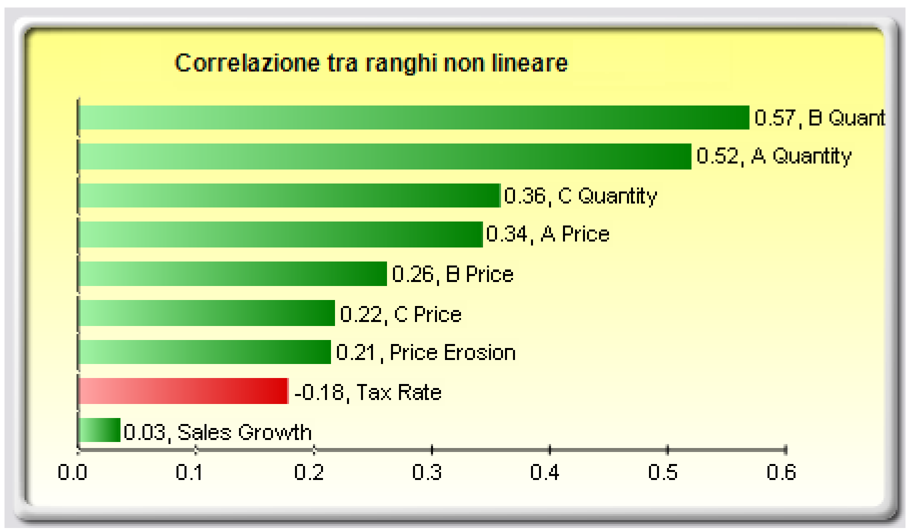


Figura 5.9 – Diagramma di Sensibilità con Correlazioni

Procedura:

- Aprite o create un modello, definite ipotesi e previsioni ed eseguite la simulazione (questo esempio usa il file *Diagrammi Tornado e di Sensibilità (Lineari)*)
- Selezionate Simulatore di Rischio | Strumenti |Analisi di Sensibilità
- Selezionate la previsione che desiderate analizzare e cliccate su **OK** (Figura 5.10)

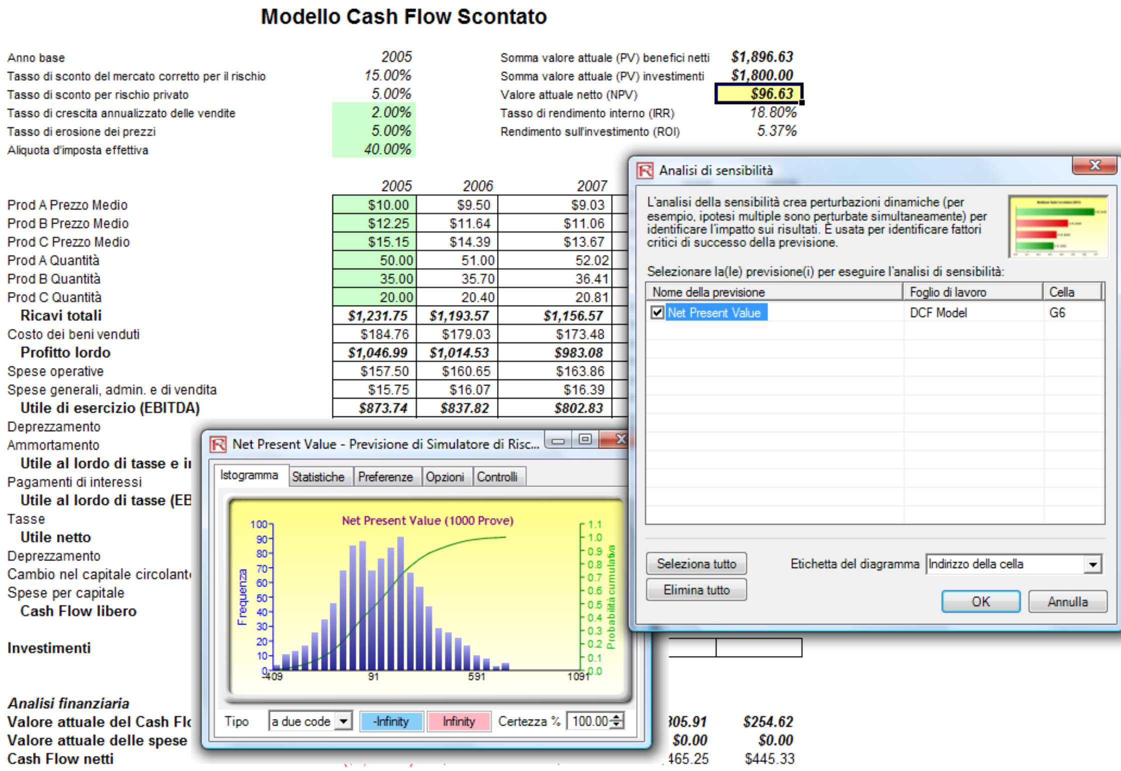


Figura 5.10 – Eseguire una Analisi di Sensibilità

Interpretare i Risultati:

I risultati dell'analisi di sensibilità comprendono un report e due diagrammi fondamentali. Il primo è il diagramma di una correlazione tra ranghi non lineare (Figura 5.11) che classifica le coppie di correlazione ipotesi-previsione dalla più alta alla più bassa. Queste correlazioni sono non lineari e non parametriche, il che le rende libere da ogni requisito distribuzionale (vale a dire, un'ipotesi con una distribuzione Weibull può essere confrontata con un'altra ipotesi con una distribuzione Beta). I risultati di questo diagramma sono abbastanza simili a quelli dell'analisi tornado vista in precedenza (naturalmente senza il valore dell'investimento di capitale, che avevamo deciso essere un valore noto e che, quindi, non è stato simulato), con un'eccezione particolare. L'aliquota d'imposta è stata relegata ad una posizione molto inferiore nel diagramma dell'analisi di sensibilità (Figura 5.11) paragonato alla sua posizione nel diagramma tornado (Figura 5.6). Questo perché da se, l'aliquota d'imposta avrà un impatto significativo, ma una volta che le altre variabili incominciano ad interagire nel modello, sembra che l'aliquota d'imposta abbia un effetto meno dominante (questo sia perché l'aliquota d'imposta ha una distribuzione più piccola, dato che le aliquote d'imposta storiche tendono a non fluttuare troppo, ma anche perché l'aliquota d'imposta è un diretto valore percentuale dei ricavi prima delle tasse, dove altre variabili precedenti esercitano una grande influenza). Questo esempio dimostra che l'esecuzione di un'analisi di sensibilità dopo l'esecuzione di una simulazione è importante per accertare se ci sono eventuali interazioni nel modello e se gli effetti di certe variabili sono ancora validi. Il secondo diagramma (Figura 5.12) illustra la variazione percentuale spiegata. In altre parole, per quanto riguarda le fluttuazioni nella previsione, quanta della variazione può essere spiegata da ciascuna delle ipotesi, dopo aver tenuto conto di tutte le interazioni tra le variabili? Notate che la somma di tutte le variazioni spiegate è normalmente vicina al 100% (talvolta ci sono altri elementi che influiscono sul modello ma che non sono direttamente catturate qui) e che, se ci sono delle correlazioni, la somma può qualche volta superare il 100% (a causa degli effetti di interazione che sono cumulativi).

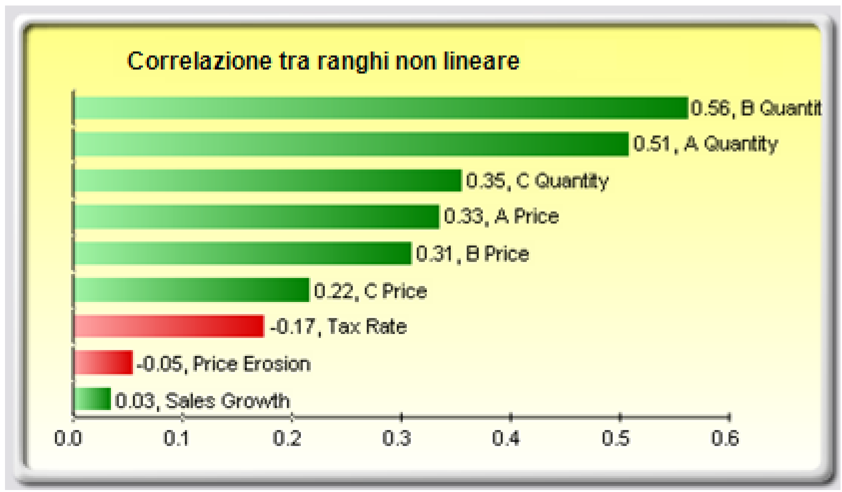


Figura 5.11 – Diagramma di Correlazione tra Ranghi

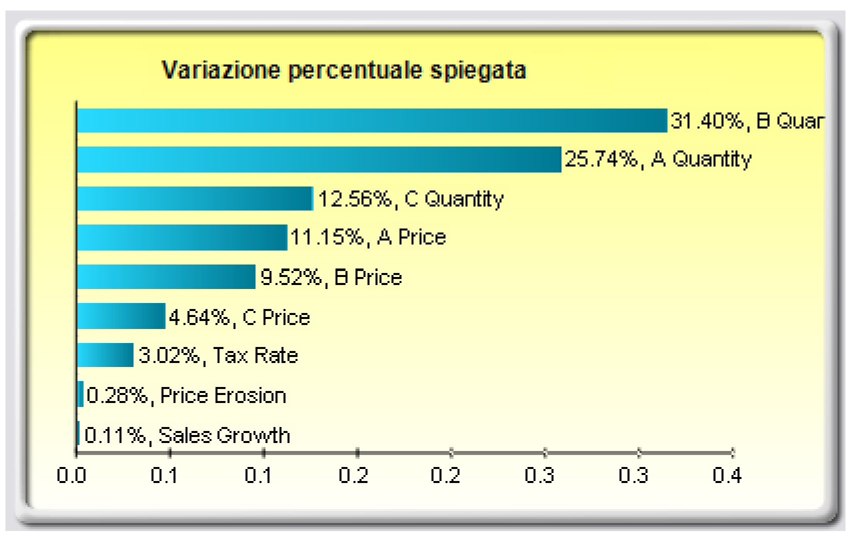


Figura 5.12 – Contributo al Diagramma di Varianza

Nota: L'analisi tornado viene eseguita prima dell'esecuzione di una simulazione, mentre l'analisi di sensibilità viene eseguita dopo l'esecuzione di una simulazione. I diagrammi ragno in un'analisi tornado possono tenere conto di non linearità, mentre i diagrammi di correlazione tra ranghi in un'analisi di sensibilità possono dare conto di condizioni non lineari e non richiedono distribuzioni.

5.3 Adattamento Distribuzionale: Variabile Singola and Variabili Multiple

Teoria: Un ulteriore strumento potente di simulazione è l'adattamento distribuzionale. In altre parole, quale distribuzione deve usare un'analista per una particolare variabile d'input in un modello? Quali sono i parametri distribuzionali pertinenti? Se non esistono dati storici, allora l'analista deve fare delle ipotesi sulle variabili in questione. Un approccio è di usare il metodo Delphi, dove un gruppo di esperti ha il compito di stimare il comportamento di ciascuna variabile. Per esempio, ad un gruppo d'ingegneri meccanici può essere assegnato il compito di valutare le possibilità estreme del diametro di una molla elicoidale mediante una rigorosa sperimentazione o tramite congetture ("guesstimates"). Questi valori possono poi essere usati come i parametri d'input della variabile (per esempio, una distribuzione uniforme con valori estremi tra 0.5 e 1.2). Quando un'analisi non è possibile (per esempio, quota di mercato e tasso di crescita dei ricavi), il management può comunque fare delle valutazioni sugli esiti potenziali e fornire gli scenari del caso migliore, del caso più probabile e del caso peggiore.

D'altra parte, se sono disponibili dei dati storici affidabili, si può eseguire un adattamento distribuzionale. Se si assume che gli schemi storici siano validi che la storia tende a ripetersi, allora si possono usare i dati storici per trovare la distribuzione, e i suoi parametri attinenti, con il miglior adattamento per definire meglio le variabili da simulare. Le Figure da 5.13 a 5.15 illustrano un esempio di adattamento distribuzionale. Questa illustrazione usa il file ***Adattamento dati*** nella cartella degli esempi.

Procedura:

- Aprite un foglio di lavoro con dati esistenti da adattare
- Selezionate i dati che desiderate adattare (i dati devono essere in una singola colonna con multiple righe)
- Selezionate Simulatore di Rischio | Strumenti | Adattamento distribuzionale (variabile singola)

- Selezionate le distribuzioni specifiche alle quali volete adattare o mantenete l'impostazione di default, dove sono selezionate tutte le distribuzioni, e cliccate su **OK** (Figura 5.13)
- Esaminate i risultati dell'adattamento, scegliete la distribuzione attinente che desiderate e cliccate su **OK** (Figura 5.14)

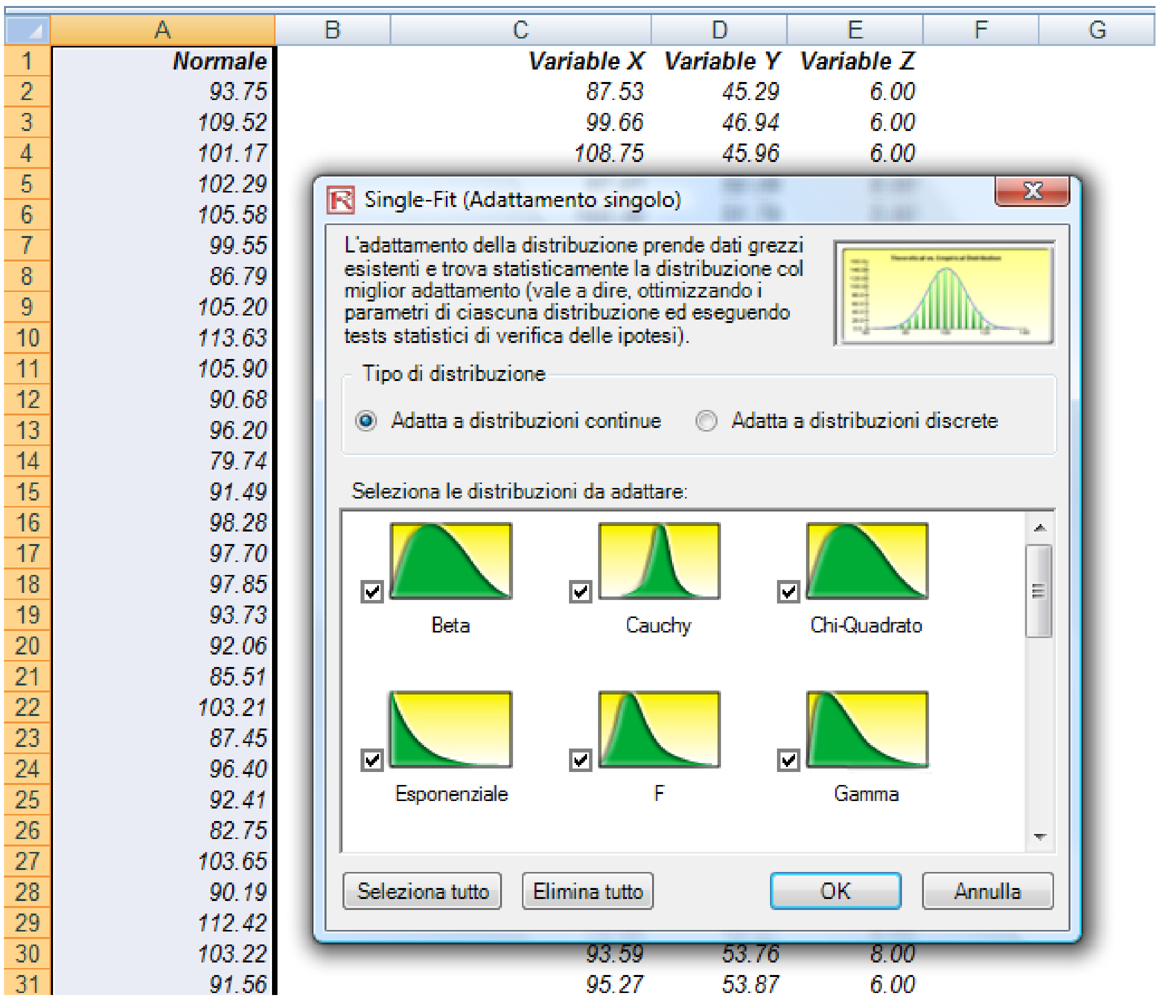
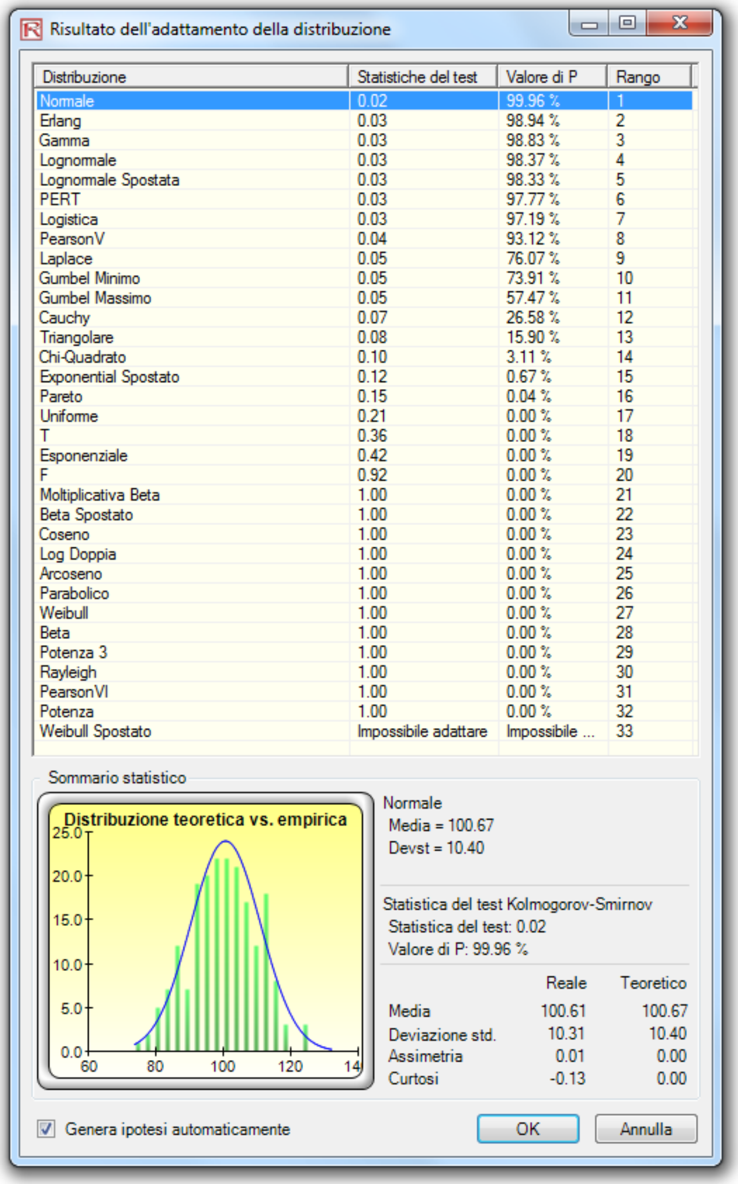


Figura 5.13 – Adattamento Distribuzionale - Variabile Singola

Interpretare i Risultati:

L'ipotesi nulla sotto esame è tale che la distribuzione adattata è la stessa distribuzione come la popolazione da cui provengono i dati campione da adattare. Perciò, se il valore calcolato di p è minore di un livello critico di alfa (tipicamente 0.10 o 0.05), allora questa distribuzione è la distribuzione sbagliata. Viceversa, tanto più alto è il valore di p, tanto meglio la distribuzione si adatta ai dati. Generalmente parlando, potete pensare al valore di p come ad una *percentuale spiegata*. In altre parole, se il valore di p è di

0.9727 (Figura 5.14), allora impostare una distribuzione normale, con una media di 99.28 ed una deviazione standard di 10.17, spiega circa 97.27% della variazione nei dati. Questo indica un adattamento notevolmente buono. Sia i risultati (Figura 5.14) che il report (Figura 5.15) mostrano le statistiche del test, il valore di p, le statistiche teoretiche (basato sulla distribuzione scelta), le statistiche empiriche (basati sui dati grezzi), i dati originali (per mantenere un record dei dati usati) e le ipotesi insieme ai parametri distribuzionali pertinenti (se avete cioè selezionato l'opzione per la generazione automatica delle ipotesi e se esiste già un profilo di simulazione. I risultati classificano anche tutte le distribuzioni selezionate e la bontà di adattamento ai dati.


Risultato dell'adattamento della distribuzione

Distribuzione	Statistiche del test	Valore di P	Rango
Normale	0.02	99.96 %	1
Erlang	0.03	98.94 %	2
Gamma	0.03	98.83 %	3
Lognormale	0.03	98.37 %	4
Lognormale Spostata	0.03	98.33 %	5
PERT	0.03	97.77 %	6
Logistica	0.03	97.19 %	7
PearsonV	0.04	93.12 %	8
Laplace	0.05	76.07 %	9
Gumbel Minimo	0.05	73.91 %	10
Gumbel Massimo	0.05	57.47 %	11
Cauchy	0.07	26.58 %	12
Triangolare	0.08	15.90 %	13
Chi-Quadrato	0.10	3.11 %	14
Exponential Spostato	0.12	0.67 %	15
Pareto	0.15	0.04 %	16
Uniforme	0.21	0.00 %	17
T	0.36	0.00 %	18
Esponenziale	0.42	0.00 %	19
F	0.92	0.00 %	20
Moltiplicativa Beta	1.00	0.00 %	21
Beta Spostato	1.00	0.00 %	22
Coseno	1.00	0.00 %	23
Log Doppia	1.00	0.00 %	24
Arcoseno	1.00	0.00 %	25
Parabolico	1.00	0.00 %	26
Weibull	1.00	0.00 %	27
Beta	1.00	0.00 %	28
Potenza 3	1.00	0.00 %	29
Rayleigh	1.00	0.00 %	30
PearsonVI	1.00	0.00 %	31
Potenza	1.00	0.00 %	32
Weibull Spostato	Impossibile adattare	Impossibile ...	33

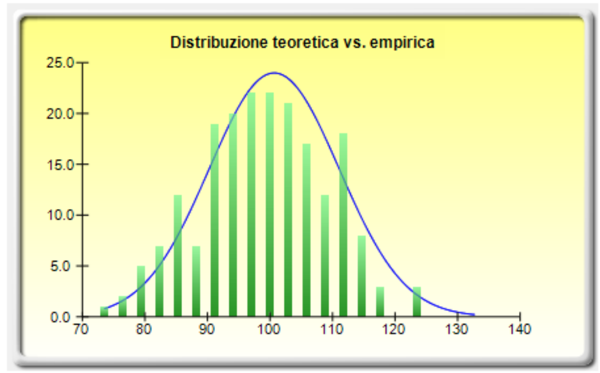
Figura 5.14 - Risultati dell'Adattamento Distribuzionale

Adattamento distribuzionale variabile singola

Sommario statistico

Ipotesi adattata	100.61	
Distribuzione adattata	**Normale**	
Media	100.67	
Sigma	10.40	
Statistica Kolmogorov-Smirnov	0.02	
Valore di P per statistica del test	0.9996	

	Reale	Teoretica
Media	100.61	100.67
Deviazione standard	10.31	10.40
Assimetria	0.01	0.00
Curtosi di eccesso	-0.13	0.00



Dati adattati originali

73.53	78.21	78.52	79.50	79.72	79.74	81.56	82.08	82.68	82.75	83.34	83.64	84.09
84.66	85.00	85.35	85.51	86.04	86.79	86.82	86.91	87.02	87.03	87.45	87.53	87.66
88.05	88.45	88.51	89.95	90.19	90.54	90.68	90.96	91.25	91.49	91.56	91.94	92.06
92.36	92.41	92.45	92.70	92.80	92.84	93.21	93.26	93.48	93.73	93.75	93.77	93.82
94.00	94.15	94.51	94.57	94.64	94.69	94.95	95.57	95.62	95.71	95.78	95.83	95.97
96.20	96.24	96.40	96.43	96.47	96.81	96.88	97.00	97.07	97.21	97.23	97.48	97.70
97.77	97.85	98.15	98.17	98.24	98.28	98.32	98.33	98.35	98.65	99.03	99.27	99.46
99.47	99.55	99.73	99.96	100.08	100.24	100.36	100.42	100.44	100.48	100.49	100.83	101.17
101.28	101.34	101.45	101.46	101.55	101.73	101.74	101.81	102.29	102.55	102.58	102.60	102.70
103.17	103.21	103.22	103.32	103.34	103.45	103.65	103.66	103.72	103.81	103.90	103.99	104.46
104.57	104.76	105.20	105.44	105.50	105.52	105.58	105.66	105.87	105.90	105.90	106.29	106.35
106.59	107.01	107.68	107.70	107.93	108.17	108.20	108.34	108.42	108.43	108.49	108.70	109.15
109.22	109.35	109.52	109.75	110.04	110.16	110.25	110.54	111.05	111.06	111.44	111.76	111.90
111.95	112.07	112.19	112.29	112.32	112.42	112.48	112.85	112.92	113.50	113.59	113.63	113.70
114.13	114.14	114.21	114.91	114.95	115.40	115.58	115.66	116.58	116.98	117.60	118.67	119.24
119.52	124.14	124.16	124.39	132.30								

Figura 5.15 - Report dell'Adattamento Distribuzionale

Per l'adattamento di variabili multiple, il processo è abbastanza simile all'adattamento di variabili individuali. Tuttavia, i dati devono essere disposti in colonne (vale a dire, ciascuna variabile è disposta come una colonna) e tutte le variabili sono adattate una alla volta.

Procedura:

- Aprite un foglio di lavoro con dati esistenti da adattare
- Selezionate i dati che desiderate adattare (i dati devono essere in colonne multiple con multiple righe)
- Selezionate Simulatore di Rischio | Strumenti | Adattamento Distribuzionale (Variabili Multiple)
- Esaminate i dati, scegliete i tipi pertinenti di distribuzione che desiderate e cliccate su **OK**

Nota: Prego notare che I metodi di classificazione statistica usati nelle procedure di adattamento distribuzionale sono il test Chi-Quadrato e il test di Kolmogorov-Smirnov. Il primo è usato per testare distribuzioni discrete e il secondo per testare le distribuzioni continue. Brevemente, viene usato un test di verifica d'ipotesi insieme con una procedura interna di ottimizzazione per trovare i parametri col miglior adattamento per ciascuna distribuzione testata e i risultati sono classificati dal migliore al peggiore adattamento.

5.4 Simulazione Bootstrap

Teoria: La simulazione Bootstrap è una tecnica semplice che valuta l'affidabilità e la precisione delle statistiche della previsione o di altri dati grezzi del campione. Sostanzialmente, la simulazione bootstrap è usata nel test di verifica d'ipotesi . I metodi classici usati nel passato dipendevano da formule matematiche per descrivere la precisione delle statistiche campione. Questi metodi assumono che la distribuzione di una statistica campione si avvicina ad una distribuzione normale, rendendo il calcolo dell'errore standard o dell'intervallo di confidenza della statistica relativamente facile. Tuttavia, quando la distribuzione del campionamento della statistica non è distribuita normalmente o non è facilmente determinabile, questi metodi classici sono difficili da usare o non sono validi. Per contrasto, il bootstrapping analizza le statistiche campione in modo empirico, campionando ripetutamente i dati e creando distribuzioni delle differenti statistiche da ciascun campionamento.

Procedura:

- Eseguite una simulazione
- Selezionate Simulatore di Rischio | Strumenti | Bootstrap non parametrico
- Selezionate solo una previsione sulla quale eseguire il bootstrap, selezionate la/le statistica/statistiche sulla/sulle quale/quali eseguire il bootstrap, inserite il numero di prove di bootstrap e cliccate su **OK** (Figura 5.16)

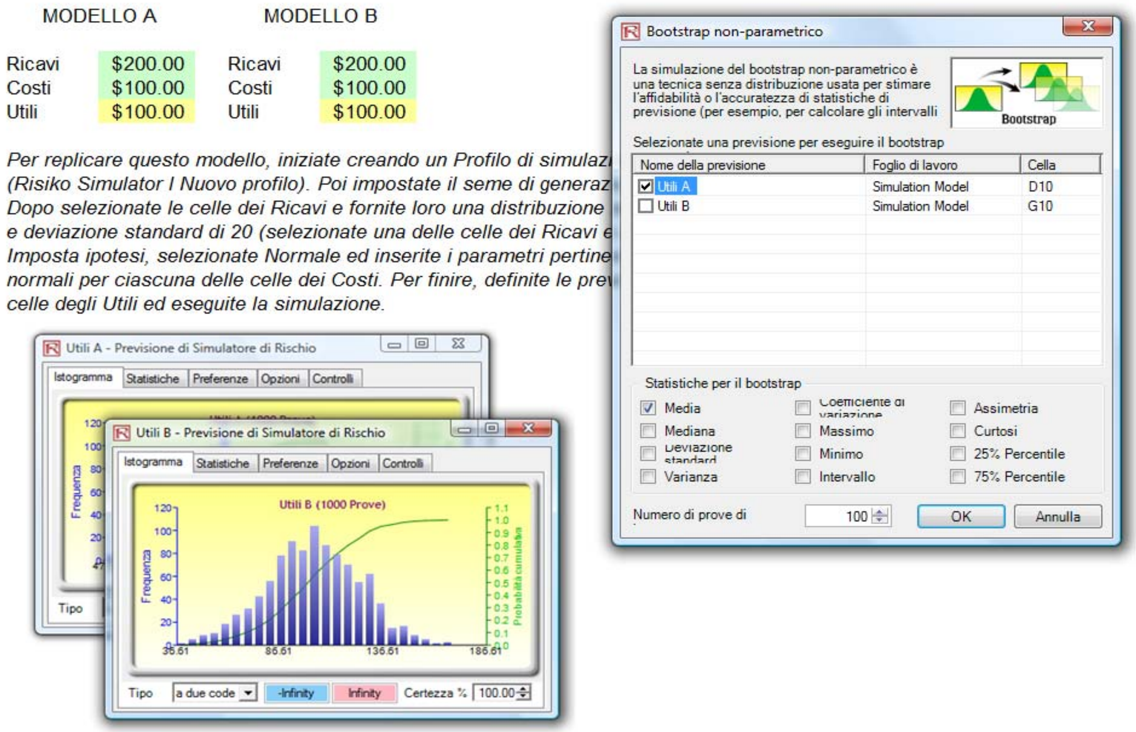


Figura 5.16 –Simulazione Bootstrap Non Parametrica

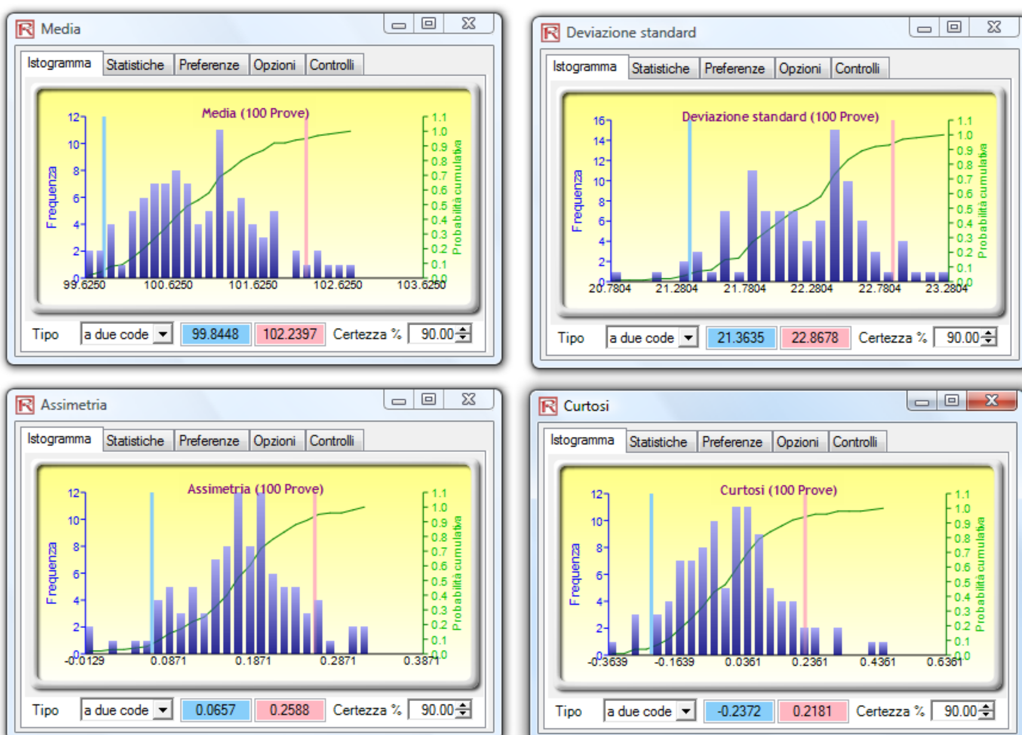


Figura 5.17 – Risultati della Simulazione Bootstrap

Interpretare i Risultati:

Sostanzialmente, la simulazione bootstrap non parametrica può essere denominata come *una simulazione basata su una simulazione*. Perciò, dopo l'esecuzione di una simulazione vengono visualizzate le statistiche risultanti, ma la precisione di tali statistiche e la loro significatività statistica sono talvolta in dubbio. Per esempio, se la statistica dell'asimmetria dell'esecuzione della simulazione è di – 0.10, è questa distribuzione veramente negativamente asimmetrica o è il leggero valore negativo da attribuire alla pura casualità? E se fosse –0.15, –0.20 e così via. In altre parole, quanto lontano è abbastanza lontano in modo che questa distribuzione possa essere considerata come negativamente asimmetrica? Si può applicare la stessa domanda a tutte le altre statistiche. Una distribuzione è veramente statisticamente identica ad un'altra distribuzione, rispetto a delle determinate statistiche calcolate, o sono queste distribuzioni significativamente differenti? La Figura 5.17 illustra alcuni esempi di risultati di un bootstrap. Per esempio, la confidenza del 90 percento per la statistica dell'asimmetria è tra –0.0189 e 0.0952, tale che il valore 0 cade entro questa confidenza. Questo indica che con una confidenza del 90 percento, l'asimmetria di questa previsione non è statisticamente significativamente differente da zero, o che questa distribuzione può essere considerata come simmetrica e non asimmetrica. Viceversa, se il valore 0 cade fuori di questa confidenza, allora è vero l'opposto: la distribuzione è asimmetrica (positivamente asimmetrica se la statistica della previsione è positiva, e negativamente asimmetrica se la statistica della previsione è negativa).

Note:

Il termine *bootstrap* viene dal detto americano, "tirarsi su da terra usando i propri lacci (bootstraps)" ed è applicabile perché questo metodo usa la distribuzione delle statistiche stesse per analizzare la precisione delle statistiche. La simulazione non parametrica è semplicemente l'estrazione con rimpiazzo di palline da golf a caso da una grande cesta, dove ogni pallina da golf si basa su un punto dati storico. Supponiamo che ci siano 365 palline da golf nella cesta (in rappresentanza di 365 punti dati storici). Immaginate, se volete, che il valore di ciascuna pallina da golf estratta a caso sia scritto su una grande lavagna bianca. I risultati delle 365 palline estratte con rimpiazzo sono scritti nella prima colonna della lavagna con 365 righe di numeri. Le statistiche attinenti (per esempio, media, mediana, deviazione standard e così via) sono calcolate su queste 365 righe. Il processo è poi ripetuto, diciamo, cinque mila volte. La lavagna sarà ora riempita con 365 righe e 5000 colonne. Quindi, 5000 insiemi di statistiche (vale a dire, ci saranno 5000 media, 5000 mediane, 5000 deviazioni standard e così via) sono tabulati e le loro distribuzioni visualizzate. Le pertinenti *statistiche delle statistiche* sono poi tabulate e usando questi risultati si può accertare il grado di confidenza delle statistiche simulate. In altre parole, da una semplice simulazione con 10000 prove risulta una media della previsione di, diciamo, \$5.00. Quanto è certo l'analista dei risultati? Il Bootstrapping permette all'utente di determinare l'intervallo di confidenza della statistica calcolata della media, indicando la distribuzione delle statistiche. Per finire, i risultati del bootstrap sono importanti perché, secondo la *Legge dei Grandi Numeri* e il *Teorema del Limite Centrale* nella statistica, la media delle medie dei campioni è uno stimatore imparziale e si avvicina alla vera media della popolazione quando la dimensione del campione aumenta.

5.5 Test di Verifica d'Ipotesi

Teoria: Un test di verifica d'ipotesi viene eseguita quando si analizza le medie e le varianze di due distribuzioni per determinare se sono statisticamente identiche o statisticamente differenti l'una dall'altra. In altre parole, per vedere se le differenze che si verificano tra le medie e le varianze di due diverse previsioni si basano sulla pura casualità o se sono, in effetti, statisticamente significativamente differenti.

Procedura:

- Eseguita una simulazione
- Selezionate Simulatore di Rischio | Strumenti | Test di verifica d'ipotesi
- Selezionate solo *due* previsioni da testare per volta, selezionate il tipo di test di verifica d'ipotesi che desiderate eseguire e cliccate su **OK** (Figura 5.18)

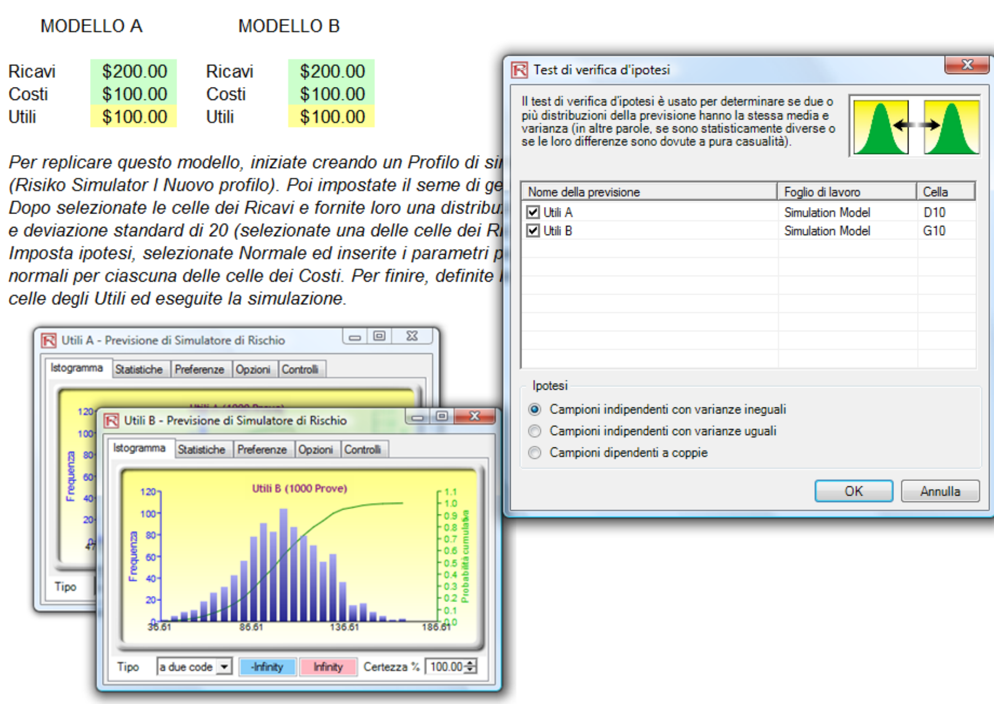


Figura 5.18 –Test di Verifica d'Ipotesi

Interpretare i Risultati:

Un test di verifica d'ipotesi a due code è eseguito sull'ipotesi nulla (H_o), tale che le medie della popolazione delle due variabili sono statisticamente identiche. L'ipotesi alternativa (H_a) è che le medie della popolazione siano statisticamente diverse. Se i valori calcolati di p sono meno di o uguali a 0,01, 0,05 o 0,10, questo significa che l'ipotesi nulla è respinta. Questo implica che le medie della previsione sono statisticamente significativamente diverse ai livelli di significatività del 1%, 5% e 10%. Se l'ipotesi nulla non è respinta quando i valori di p sono alti, allora le medie delle due distribuzioni della previsione sono statisticamente simili. La stessa analisi è eseguita sulle varianze di due previsioni alla volta usando il test F a coppie. Se i valori di p sono bassi, allora le varianze (e le deviazioni standard) sono statisticamente diverse; altrimenti, per valori alti di p, le varianze sono statisticamente identiche.

Test di verifica d'ipotesi sulle medie e le varianze di due previsioni

Sommario statistico

Un test di verifica d'ipotesi viene eseguito quando si esaminano le medie e le varianze di due distribuzioni per determinare se sono statisticamente identici o statisticamente diversi tra loro. In altre parole, per vedere se le differenze che si verificano tra due medie e due varianze sono basate su pura casualità o se sono realmente diversi uno dall'altro. Il t-test a due variabili con varianze ineguali (si prevede cha la varianza popolazione della previsione 1 sia diversa dalla varianza popolazione della previsione 2) è adatto quando le distribuzioni della previsione provengono da popolazioni diverse (per esempio, dati raccolti da due località geografiche diverse, due unità aziendali operative diverse e così via). Il t-test a due variabili con varianze uguali (si prevede cha la varianza popolazione della previsione 1 sia uguale alla varianza popolazione della previsione 2) è adatto quando le distribuzioni della previsione provengono da popolazioni uguali (per esempio, dati raccolti da due progetti di motori diversi con specifiche simili e così via). Il t-test a due variabili dipendenti accoppiate è adatto quando le distribuzioni della previsione provengono da popolazioni simili (per esempio, dati raccolti dallo stesso gruppo di clienti ma in occasioni diverse e così via).

Un test di verifica d'ipotesi a due code è eseguito sull'ipotesi nulla Ho per vedere se le medie della popolazione delle due variabili sono statisticamente identiche. L'ipotesi alternativa è che le medie della popolazione siano statisticamente diverse. Se i valori calcolati di p sono meno di o uguale a 0,01, 0,05 o 0,10, significa che l'ipotesi è respinta. Questo implica che le medie della previsione sono statisticamente significativamente diverse ai livelli di significatività di 1%, 5% e 10%. Se l'ipotesi nulla non è respinta quando i valori di p sono alti, allora le medie delle due distribuzioni della previsione sono statisticamente simili. La stessa analisi è eseguita su varianze di due previsioni alla volta usando il F-test a coppie. Se i valori di p sono bassi, allora le varianze (e le deviazioni standard) sono statisticamente diverse; altrimenti, per valori di p alti, le varianze sono statisticamente identiche.

Risultato

Presupposto del test di verifica di'i	Varianze ineguali:
Statistica di t calcolata:	1.015722
Valore di p per la statistica di t:	0.309885
Statistica di F calcolata:	1.063476
Valore di p per la statistica di F:	0.330914

Figura 5.19 – Risultati del Test di Verifica d'Ipotesi

Note:

Il t-test a due variabili con varianze ineguali (si prevede cha la varianza della popolazione della previsione 1 sia diversa dalla varianza della popolazione della previsione 2) è adatto quando le distribuzioni della previsione provengono da popolazioni diverse (per esempio, dati raccolti da due località geografiche diverse, due unità aziendali operative diverse e così via). Il t-test a due variabili con varianze uguali (si prevede cha la varianza della popolazione della previsione 1 sia uguale alla varianza della popolazione della previsione 2) è adatto quando le distribuzioni della previsione provengono da popolazioni simili (per esempio, dati raccolti da due progetti di motori diversi con specifiche simili e così via). Il t-test a due variabili dipendenti accoppiate è adatto quando le distribuzioni della previsione provengono da popolazioni esattamente uguali (per esempio, dati raccolti dallo stesso gruppo di clienti ma in occasioni diverse e così via).

5.6 Estrarre i Dati e Salvare i Risultati della Simulazione

Si possono facilmente estrarre i dati grezzi di una simulazione usando la procedura di *Estrazione dati* di Simulatore di Rischio. Si possono estrarre sia ipotesi che previsioni, ma prima si deve eseguire una simulazione. I dati estratti possono poi essere usati per una varietà di altre analisi.

Procedura:

- Aprite o create un modello, definite le ipotesi e le previsioni ed eseguite la simulazione
- Selezionate Simulatore di Rischio | Strumenti | Estrazione dati
- Selezionate le ipotesi e/o le previsioni dalle quali volete estrarre i dati e cliccate su **OK**

I dati possono essere estratti in vari formati:

- Come dati grezzi in un nuovo foglio di lavoro, dove si possono poi o salvare i valori simulati (sia ipotesi che previsioni) o analizzarli ulteriormente come richiesto
- Come file piatto di testo, dal quale si possono poi esportare i dati in altri software d'analisi di dati
- Come file di Simulatore di Rischio, dal quale si possono più tardi richiamare i risultati (sia ipotesi che previsioni), selezionando ***Simulatore di Rischio | Strumenti | Apri/Importa dati***.

La terza opzione è la scelta più comune, vale a dire, salvare i risultati simulati come un file *.risksim, dal quale si possono più tardi richiamare i risultati. Così non si deve rieseguire ogni volta una simulazione. La Figura 5.21 mostra la casella di dialogo per estrarre o esportare e per salvare i risultati della simulazione.

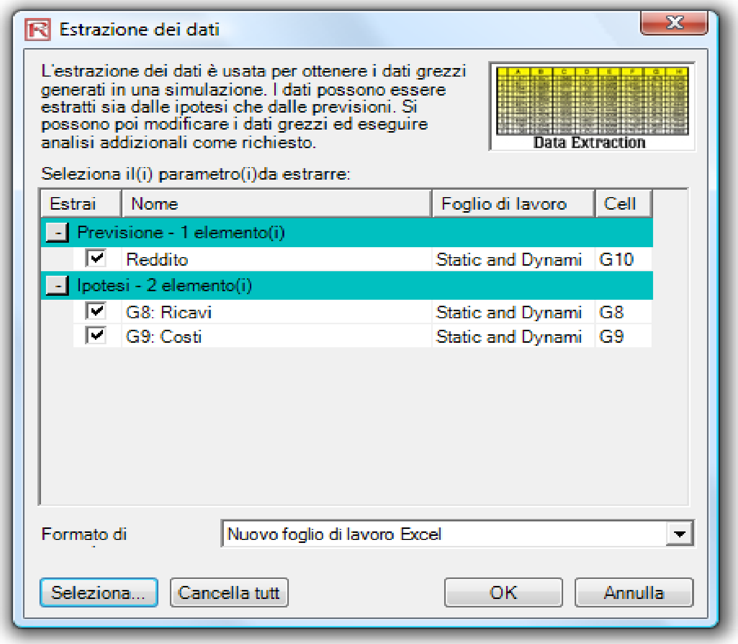


Figura 5.21 – Esempio del Report di una Simulazione

5.7 Creare un Report

Dopo l'esecuzione di una simulazione, potete generare un report delle ipotesi, delle previsioni e dei risultati ottenuti durante l'esecuzione della simulazione.

Procedura:

- Aprite o create un modello, definite le ipotesi e le previsioni ed eseguite la simulazione
- Selezionate Simulatore di Rischio | Crea Report

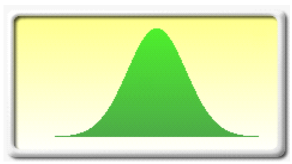
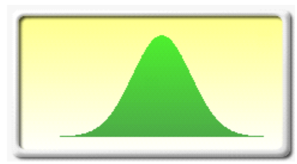
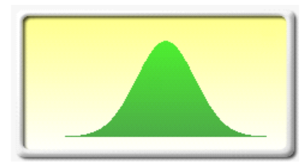
Simulazione - Modello Cash Flow Scontato / ROI

Generale

Numero di prove	1000
Interrompi la simulazione	No
Seme di generazione casu	123456
Attiva le correlazioni	Si

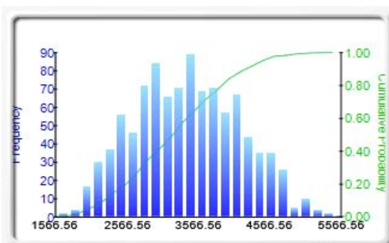
Ipotesi

Nome	**oduct A Avg Price/Unit**	**Nome**	**Product B Avg Price/Unit**	**Nome**	**oduct C Avg Price/Unit**
Attivato	Si	Attivato	Si	Attivato	Si
Cella	C12	Cella	C13	Cella	C14
Simulazione dinamica	No	Simulazione dinamica	No	Simulazione dinamica	No
Intervallo		**Intervallo**		**Intervallo**	
Minimo	-Infinity	Minimo	-Infinity	Minimo	-Infinity
Massimo	Infinity	Massimo	Infinity	Massimo	Infinity
Distribuzione	Normale	**Distribuzione**	Normale	**Distribuzione**	Normale
Media	10	Media	12.25	Media	15.15
Deviazione standard	1	Deviazione standard	1.225	Deviazione standard	1.515

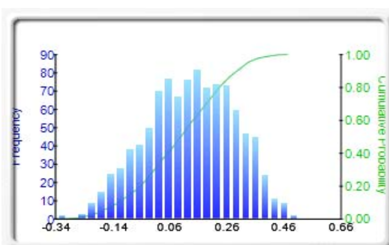




Previsioni

Nome	**Net Present Value**	Numero di punti dati	1000
Attivato	Si	Media	3221.5390
Cella	G6	Mediana	3204.1955
		Deviazione standard	744.0429
Precisione della previsione		Varianza	553599.8476
Livello di precisione	---	Coefficiente di variazione	0.2310
Livello dell'errore	---	Massimo	5443.1341
		Minimo	1405.0321
		Intervallo	4038.1020
		Assimetria	0.1363
		Curtosi	-0.6015
		25% Percentile	2663.3412
		75% Percentile	3764.0708
		Precisione dell'errore a 95	0.0143



Nome	**ıtermediate X Variable**	Numero di punti dati	1000
Attivato	Si	Media	0.0935
Cella	C48	Mediana	0.0997
		Deviazione standard	0.1549
Precisione della previsione		Varianza	0.0240
Livello di precisione	---	Coefficiente di variazione	1.6570
Livello dell'errore	---	Massimo	0.4720
		Minimo	-0.3726
		Intervallo	0.8446
		Assimetria	-0.1733
		Curtosi	-0.5771
		25% Percentile	-0.0179
		75% Percentile	0.2111
		Precisione dell'errore a 95	0.1027



Matrice di correlazione

	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
Variable 1	1.00							
Variable 2	0.00	1.00						
Variable 3	0.00	0.00	1.00					
Variable 4	0.00	0.00	0.00	1.00				
Variable 5	0.00	0.00	0.00	0.00	1.00			
Variable 6	0.00	0.00	0.00	0.00	0.00	1.00		
Variable 7	0.00	0.00	0.00	0.00	0.00	0.00	1.00	
Variable 8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Figura 5.21 – Esempio del Report di una Simulazione

5.8 Strumento Diagnostico per Regressioni e Previsioni

Questo strumento analitico avanzato di Simulatore di Rischio è usato per determinare le proprietà econometriche dei vostri dati. La diagnostica comprende il controllo dei dati per i seguenti elementi: eteroschedasticità, non linearità, valori abnormi, errori delle specifiche, micronumerosità, proprietà stocastiche e di stationarità, normalità e sfericità degli errori e multicollinearità. Ciascun test è descritto più dettagliatamente nel suo rispettivo report nel modello.

Procedura:

- Aprite il modello d'esempio (**Simulatore di Rischio | Esempi | Diagnostica della Regressione**), andate al foglio di lavoro *Dati di serie temporali* e selezionate i dati, incluso i nomi delle variabili (celle **C5:H55**).
- Cliccate su Simulatore di Rischio | Strumenti | Strumento diagnostico.
- Controllate i dati e selezionate la *Variabile dipendente Y* dal menu a tendina. Cliccate su **OK** quando avete terminato (Figura 5.22).

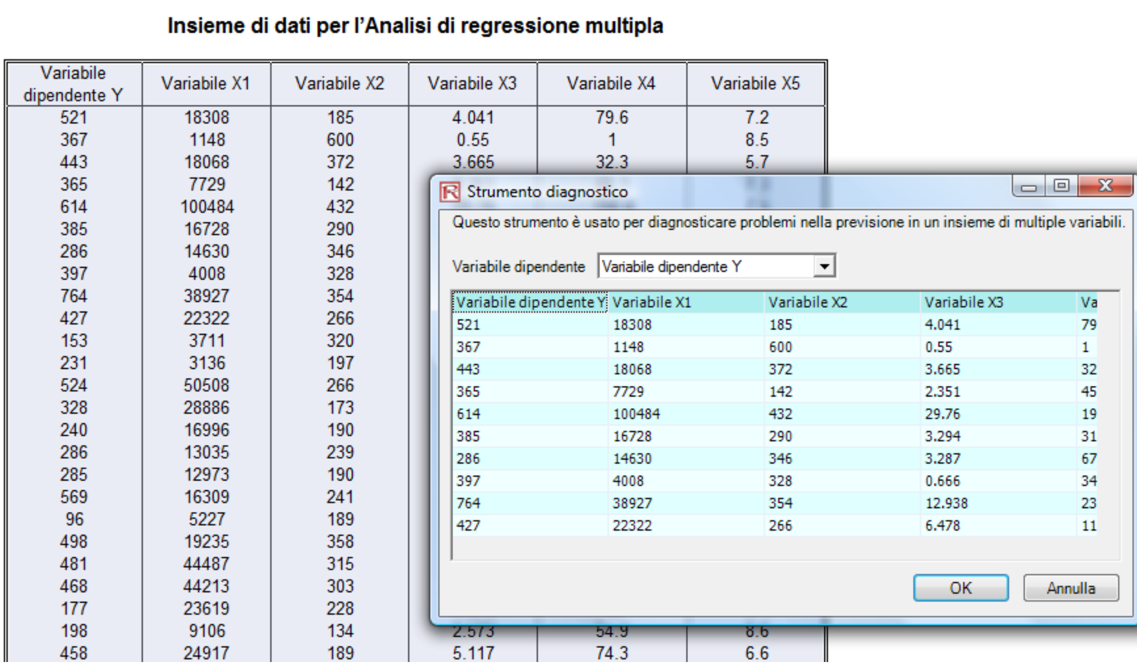


Figura 5.22 – Eseguire lo Strumento Diagnostico dei Dati

Una violazione comune nella previsione e nell'analisi di regressione è l'eteroschedasticità, vale a dire, la varianza degli errori aumenta col passare del tempo (vedere la Figura 5.23 per i risultati del test usando lo strumento diagnostico). Visivamente, la larghezza delle fluttuazioni verticali dei dati aumenta o si allarga a ventaglio nel tempo ed il coefficiente di determinazione (coefficiente R-Quadrato) diminuisce tipicamente in modo significativo quando esiste l'eteroschedasticità. Se la varianza della variabile dipendente non è constante, allora neanche la varianza dell'errore sarà constante. A meno che l'eteroschedasticità della variabile dipendente non sia pronunciata, i suoi effetti non saranno severi: le stime per mezzo dei minimi quadrati non saranno distorte e le stime della pendenza e dell'intercetta saranno distribuite normalmente, se gli errori sono distribuiti normalmente, o per lo meno distribuite normalmente asintoticamente (col ingrandirsi del numero di punti dati), se gli errori non sono distribuiti normalmente. Le stime della varianza della pendenza e della varianza complessiva saranno inesatte, ma è improbabile che l'inesattezza sia sostanziale se i valori delle variabili indipendenti sono simmetrici attorno alla loro media.

Se il numero di punti dati è piccolo (micronumerosità), può essere difficile rilevare le violazioni d'ipotesi. Con campioni piccoli le violazioni d'ipotesi, come la non normalità o l'eteroschedasticità delle varianze, sono difficili da rilevare, anche se presenti. Con un numero piccolo di punti dati, la regressione lineare offre meno protezione contro le violazioni d'ipotesi. Con pochi punti dati può essere difficile determinare quanto adeguatamente la linea adattata corrisponda ai dati o se non sia più adatta una funzione non lineare. Anche se non viene violata nessuna ipotesi del test, una regressione lineare su un numero piccolo di punti dati può non avere abbastanza potenza da rilevare una differenza significativa tra la pendenza e lo zero, anche se la pendenza non è zero. La potenza dipende dall'errore residuo, dalla variazione osservata nella variabile indipendente, dal livello selezionato di significatività alfa del test e dal numero di punti dati. La potenza diminuisce con l'aumento della varianza residua, diminuisce con la diminuzione del livello di significatività (vale a dire, quando il test è reso più rigido), aumenta con l'aumento della variazione nella variabile indipendente osservata e aumenta con l'aumento del numero i punti dati.

I valori possono non essere distribuiti in maniera identica a causa della presenza di valori abnormi. I valori abnormi sono valori anomali nei dati. I valori abnormi possono esercitare una forte influenza sulla pendenza e sull'intercetta adattate, fornendo un adattamento scadente alla maggior parte dei punti dati. I valori abnormi tendono ad aumentare la stima della varianza residua, diminuendo la possibilità di respingere l'ipotesi nulla, creando, in pratica, errori di previsione più alti. Essi possono essere causati da errori di rilevazione, una condizione che può essere corretta, o dati dal fatto che non tutti i valori delle variabili dipendenti sono campionati dalla stessa popolazione. I valori abnormi apparenti possono anche essere causati dal fatto che i valori delle variabili dipendenti provengono dalla stessa popolazione, ma da una che è non normale. Tuttavia, un punto può avere un valore inusuale, sia in una variabile indipendente che in una variabile dipendente, senza essere per forza un valore abnorme nel diagramma di dispersione. In un'analisi di regressione, la line adattata può essere altamente sensibile ai valori abnormi. In altre parole, la regressione dei minimi quadrati non è resistente ai valori abnormi e, perciò, non lo è neanche la stima della pendenza adattata. Un punto verticalmente lontano da altri punti può causare che la linea adattata passi vicino ad esso invece di seguire la generale tendenza lineare del resto dei dati, specialmente se il punto è relativamente lontano orizzontalmente dal centro dei dati.

Si deve però prestare grande attenzione nel decidere se i valori abnormi siano da eliminare. Anche se nella maggioranza dei casi i risultati di una regressione hanno un aspetto migliore se i valori abnormi sono rimossi, deve prima sussistere una giustificazione *a priori*. Per esempio, se si sta eseguendo la regressione della prestazione dei rendimenti azionari di una determinata azienda, i valori abnormi causati da flessioni nel mercato azionario dovrebbero essere inclusi: questi non sono veramente valori abnormi, visto che sono delle inevitabilità del ciclo economico. Se si tralasciano questi valori abnormi e si usa l'equazione della regressione per prevedere il proprio fondo di pensionamento basato sui titoli dell'azienda, si otterranno risultati erronei nel migliore dei casi. Per contrasto, se supponiamo che i valori abnormi siano causati da una singola condizione commerciale non ricorrente (per es. fusione ed acquisizione) e che non sia prevista una ripetizione di tali cambiamenti nella struttura aziendale, allora si devono eliminare questi valori abnormi e pulire i dati prima di eseguire un'analisi di regressione. Questa analisi identifica semplicemente i valori abnormi. La decisione di conservare o escludere questi valori abnormi è lasciata all'utente.

Una relazione non lineare tra le variabili dipendenti e le variabili indipendenti è talvolta più appropriata di una relazione lineare. In questi casi, eseguire una regressione lineare non sarà ottimale. Se il modello lineare non è nella forma corretta, le stime della pendenza e dell'intercetta e i valori adattati dalla regressione lineare saranno distorti e le stime adattate della pendenza e dell'intercetta non saranno significativi. Per un campo limitato di variabili indipendenti e di variabili dipendenti, i modelli non lineari possono essere ben approssimati dai modelli lineari (questo è difatti la base dell'interpolazione lineare), ma per una previsione precisa si deve selezionare un modello adatto ai dati. Si deve prima applicare una trasformazione non lineare ai dati avanti di eseguire una regressione. Un metodo semplice è di ottenere il logaritmo naturale della variabile indipendente (altri metodi includono: ottenere la radice quadrata o elevare la variabile indipendente alla seconda o terza potenza) e di eseguire una regressione o previsione usando i dati trasformati in formato non lineare.

Risultati della Diagnostica

	Eteroschedasticità		Micronumerosità	Valori abnormi (Outliers)			Non linearità	
Variabile	W-Test Valore di p	Risultato del test di verifica d'ipotesi	Risultato della approssimazione	Limite inferiore naturale	Limite superiore naturale	Numero dei valori abnormi potenziali	Test non lineare Valore di p	Risultato del test di verifica d'ipotesi
Y			nessuni problemi	-7.86	671.70	2		
Variable X1	0.2543	Homoskedastic	nessuni problemi	-21377.95	64713.03	3	0.2458	linéaire
Variable X2	0.3371	Homoskedastic	nessuni problemi	77.47	445.93	2	**0.0335**	**non linéaire**
Variable X3	0.3649	Homoskedastic	nessuni problemi	-5.77	15.69	3	**0.0305**	**non linéaire**
Variable X4	0.3066	Homoskedastic	nessuni problemi	-295.96	628.21	4	0.9298	linéaire
Variable X5	0.2495	Homoskedastic	nessuni problemi	3.35	9.38	3	0.2727	linéaire

Figura 5.23 – Risultati dai Tests per Valori Abnormi, Eteroschedasticità, Micronumerosità e Nonlinearità

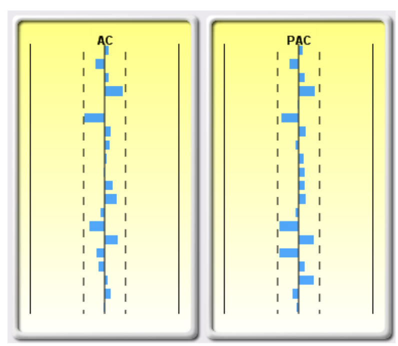
Un'altra questione tipica che emerge quando si eseguono previsioni di dati di serie temporali, è se i valori delle variabili indipendenti sono veramente indipendenti gli uni dagli altri o se sono piuttosto dipendenti. I valori di variabili dipendenti raccolti nel corso di una serie temporale possono essere autocorrelati. Per valori di variabili dipendenti correlati serialmente, le stime della pendenza e dell'intercetta saranno non distorte, ma le stime della loro previsione e delle loro varianze non saranno affidabili e, quindi, la validità di certi tests della bontà di adattamento statistica sarà viziata. Per esempio, i tassi d'interesse, i tassi d'inflazione, le vendite, i ricavi e molti altri dati di serie temporali sono tipicamente autocorrelati, dove il valore del periodo attuale è in relazione con il valore del periodo precedente e così via (il tasso d'inflazione di marzo è chiaramente collegato al livello di febbraio, che è, a sua volta, in relazione con il livello di gennaio e così via). Ignorare tali ovvie relazioni fornirà delle previsioni distorte e meno precise. In tali circostanze può essere più adatto un modello autocorrelato di regressione o un modello ARIMA (**Simulatore di Rischio** l **Previsione** l **ARIMA**). Per finire, le funzioni di autocorrelazione di una serie che è non stazionaria tendono a decadere lentamente (vedere il report Non stazionario nel modello).

Se l'autocorrelazione AC(*1*) è non-zero, significa che la serie è correlata serialmente di primo ordine. Se AC(*k*) si esaurisce più o meno geometricamente con uno sfasamento crescente, questo implica che la serie segue un processo autoregressivo di basso ordine. Se AC(*k*) scende a zero dopo un numero limitato di sfasamenti, questo implica che la serie segue un processo a media mobile di basso ordine. La correlazione parziale PAC(*k*) misura la correlazione dei valori che sono separati da *k* periodi, dopo la rimozione della correlazione dagli sfasamenti intercorsi. Se lo schema di autocorrelazione può essere catturato da una autoregressione di ordine minore di *k*, allora l'autocorrelazione parziale allo sfasamento k sarà vicino a zero. Le statistiche di Q di Ljung-Boxdi Ljung-Box ed i loro valori di p allo sfasamento *k* hanno l'ipotesi nulla che non esiste nessuna autocorrelazione fino all'ordine k. Le linee punteggiate nei tracciati delle autocorrelazioni sono i due approssimativi limiti standard d'errore. Se l'autocorrelazione è entro questi limiti, essa non è significativamente diversa da zero al livello di significatività del 5%.

L'autocorrelazione misura la relazione con il passato della variabile dipendente Y con se stessa. GLI sfasamenti distributivi, per contrasto, sono relazioni di sfasamento temporale tra la variabile dipendente Y e differenti variabili indipendenti X. Per esempio, il movimento e la direzione dei tassi ipotecari tendono a seguire il Tasso dei Fondi Federali, ma con uno sfasamento temporale (tipicamente da 1 a 3 mesi). Talvolta gli sfasamenti temporali seguono cicli e stagionalità (per esempio, le vendite di gelato tendono a culminare durante i mesi estivi e sono, quindi, in relazione con le vendite dell'estate passata, 12 mesi nel passato). L'analisi di sfasamento distributivo (Figura 5.24) mostra come la variabile dipendente è collegata a ciascuna delle variabili indipendenti a vari sfasamenti temporali, quando tutti gli sfasamenti sono persi in considerazione simultaneamente, per determinare quali sfasamenti temporali sono statisticamente significativi e da tenere in considerazione.

Autocorrelazione

Sfasamento temporale	AC	PAC	Limite inferiore	Limite superiore	Statistica di Q	Prob
1	0.0580	0.0580	-0.2828	0.2828	0.1786	0.6726
2	-0.1213	-0.1251	-0.2828	0.2828	0.9754	0.6140
3	0.0590	0.0756	-0.2828	0.2828	1.1679	0.7607
4	0.2423	0.2232	-0.2828	0.2828	4.4865	0.3442
5	0.0067	-0.0078	-0.2828	0.2828	4.4890	0.4814
6	-0.2654	-0.2345	-0.2828	0.2828	8.6516	0.1941
7	0.0814	0.0939	-0.2828	0.2828	9.0524	0.2489
8	0.0634	-0.0442	-0.2828	0.2828	9.3012	0.3175
9	0.0204	0.0673	-0.2828	0.2828	9.3276	0.4076
10	-0.0190	0.0865	-0.2828	0.2828	9.3512	0.4991
11	0.1035	0.0790	-0.2828	0.2828	10.0648	0.5246
12	0.1658	0.0978	-0.2828	0.2828	11.9466	0.4500
13	-0.0524	-0.0430	-0.2828	0.2828	12.1394	0.5162
14	-0.2050	-0.2523	-0.2828	0.2828	15.1738	0.3664
15	0.1782	0.2089	-0.2828	0.2828	17.5315	0.2881
16	-0.1022	-0.2591	-0.2828	0.2828	18.3296	0.3050
17	-0.0861	0.0808	-0.2828	0.2828	18.9141	0.3335
18	0.0418	0.1987	-0.2828	0.2828	19.0559	0.3884
19	0.0869	-0.0821	-0.2828	0.2828	19.6894	0.4135
20	-0.0091	-0.0269	-0.2828	0.2828	19.6966	0.4770



Distributive Lags

P-Values of Distributive Lag Periods of Each Independent Variable

Variable	1	2	3	4	5	6	7	8	9	10	11	12
X1	0.8467	0.2045	0.3336	0.9105	0.9757	0.1020	0.9205	0.1267	0.5431	0.9110	0.7495	0.4016
X2	0.6077	0.9900	0.8422	0.2851	**0.0638**	**0.0032**	0.8007	0.1551	0.4823	0.1126	**0.0519**	0.4383
X3	0.7394	0.2396	0.2741	0.8372	0.9808	**0.0464**	0.8355	**0.0545**	0.6828	0.7354	0.5093	0.3500
X4	**0.0061**	0.6739	0.7932	0.7719	0.6748	0.8627	0.5586	0.9046	0.5726	0.6304	0.4812	0.5707
X5	0.1591	0.2032	0.4123	0.5599	0.6416	0.3447	0.9190	0.9740	0.5185	0.2856	0.1489	0.7794

Figura 5.24 – Risultati dell'Autocorrelazione e dello Sfasamento Distributivo

Un altro requisito per l'esecuzione di un modello di regressione è l'ipotesi di normalità e sfericità del termine d'errore. Se l'ipotesi di normalità è violata o se sono presenti valori abnormi, allora il test di bontà dell'adattamento della regressione lineare può non essere il test più potente o informativo disponibile, e questo può significare la differenza tra il rilevamento di un adattamento lineare o meno. Se gli errori non sono indipendenti e non sono distribuiti normalmente, questo può indicare che i dati sono autocorrelati o che soffrono di non linearità o di altri errori più distruttivi. L'indipendenza degli errori può anche essere rilevata con i tests di eteroschedasticità (Figura 5.25).

Il test di normalità degli errori che è stato eseguito è un test non parametrico, che non fa delle ipotesi sulla forma specifica della popolazione da cui il campione è stato tratto, permettendo così l'analisi di campioni più piccoli di insiemi di dati. Questo test valuta l'ipotesi nulla che gli errori del campione sono tratti da una popolazione distribuita normalmente, contro un'ipotesi alternativa, che il campione dei dati non è distribuito normalmente. Se la Statistica di D calcolata è maggiore dei o uguale ai valori D-Critici a vari valori di significatività, allora rifiutate l'ipotesi nulla e accettate l'ipotesi alternativa (gli errori non sono distribuiti normalmente). Altrimenti, se la Statistica di D è minore del valore D-Critico, non rifiutate l'ipotesi nulla (gli sono distribuiti normalmente). Questo test dipende da due frequenze cumulative: una derivata dal campione dell'insieme di dati, la seconda da una distribuzione teorica basata sulla media e la deviazione standard del campione dei dati.

Test di Normalità e Sfericità degli Errori

Un altro requisito per l'esecuzione di un modello di regressione è l'ipotesi di normalità e sfericità del termine d'errore. Se l'ipotesi di normalità è violata o se sono presenti valori abnormi, allora il test di bontà dell'adattamento della regressione lineare può non essere il test più potente o informativo disponibile, e questo potrebbe significare la differenza tra il rilevamento di un adattamento lineare o meno. Se gli errori non sono indipendenti e non sono distribuiti normalmente, questo può indicare che i dati sono autocorrelati o che soffrono di non linearità o di altri errori più distruttivi. L'indipendenza degli errori può anche essere rilevata nei tests di eteroschedasticità (vedere il report Diagnostica).

Il test di normalità degli errori che è stato eseguito è un test non parametrico, che non fa delle ipotesi sulla forma specifica della popolazione da cui i campioni sono stati tratti, permettendo così l'analisi di campioni più piccoli di insiemi di dati. Questo test valuta l'ipotesi nulla, che gli errori campioni sono tratti da una popolazione distribuita normalmente, contro una ipotesi alternativa, che il campione dei dati non è distribuito normalmente. Se la Statistica di D calcolata è maggiore dei o uguale ai valori D-Critici a vari valori di significatività, allora rifiutate l'ipotesi nulla e accettate l'ipotesi alternativa (gli errori non sono distribuiti normalmente). Altrimenti se la Statistica di D è minore del valore D-Critico, non rifiutate l'ipotesi nulla (gli sono distribuiti normalmente). Questo test dipende da due frequenze cumulative: una derivata dall'insieme campione di dati, la seconda da una distribuzione teoretica basata sulla media e la deviazione standard dei dati campioni.

Risultato del Test

Media dell'errore di regressione	*0.00*
Deviazione standard degli errori	*141.83*
Statistica di D	*0.1036*
D-Critico a 1%	*0.1138*
D-Critico a 5%	*0.1225*
D-Critico a 10%	*0.1458*

Ipotesi nulla: gli errori sono distribuiti normalmente.

Errori	*Frequenza relativa*	*Osservato*	*Atteso*	*O-A*
-219.04	0.02	0.02	0.0612	-0.0412
-202.53	0.02	0.04	0.0766	-0.0366
-186.04	0.02	0.06	0.0948	-0.0348
-174.17	0.02	0.08	0.1097	-0.0297
-162.13	0.02	0.10	0.1265	-0.0265
-161.62	0.02	0.12	0.1272	-0.0072
-160.39	0.02	0.14	0.1291	0.0109
-145.40	0.02	0.16	0.1526	0.0074
-138.92	0.02	0.18	0.1637	0.0163
-133.81	0.02	0.20	0.1727	0.0273
-120.76	0.02	0.22	0.1973	0.0227
-120.12	0.02	0.24	0.1985	0.0415
-113.25	0.02	0.26	0.2123	0.0477

Figura 5.25 – Test per la Normalità degli Errori

Talvolta alcuni tipi di dati di serie temporali possono essere modellati solo usando il metodo del processo stocastico, dato che gli eventi sottostanti sono di natura stocastica. Per esempio, non si possono adeguatamente modellare e prevedere i prezzi di titoli, i tassi d'interesse, il prezzo del petrolio ed i prezzi di altre commodities usando un semplice modello di regressione, perché queste variabili sono molto incerte e volatili e non seguono una predefinita regola statica di comportamento. In altre parole, il processo è non stazionario. La stazionarità è controllata qui usando il Test dei Runs, mentre un altro indizio visivo si trova nel report di Autocorrelazione (la funzione di autocorrelazione (ACF) tende a decadere lentamente). Un processo stocastico è una sequenza di eventi o percorsi generati da leggi probabilistiche. In altre parole, eventi casuali possono verificarsi nel tempo, ma essi sono governati da specifiche regole statistiche e probabilistiche. I processi stocastici principali comprendono la Passeggiata casuale o Moto Browniano, il Ritorno alla media e la Diffusione a salti. Questi processi possono essere usati per prevedere una molteplicità di variabili che sembrano seguire tendenze casuali, ma che sono tuttavia limitate da leggi probabilistiche. L'equazione che genera il processo è nota in anticipo, ma i risultati reali generati non sono noti (Figura 5.26).

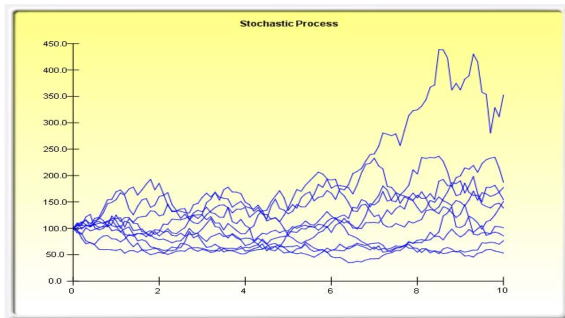
Il processo di Passeggiata casuale di Moto Browniano può essere usato per prevedere i prezzi di titoli, i prezzi di commodities ed altri dati stocastici di serie temporali, dato una deriva o tasso di crescita e una volatilità intorno al percorso della deriva. Il processo di Ritorno alla media può essere usato per ridurre le fluttuazioni del processo di Passeggiata casuale, permettendo al percorso di stabilire come obiettivo un valore a lungo termine. Questo lo rende utile nella previsione di variabili di serie temporali che hanno un tasso a lungo termine, come i tassi d'interesse e dell'inflazione (questi sono tassi di obiettivo a lungo termine delle autorità di regolamentazione o del mercato). Il processo di Diffusione a salti è utile nella previsione di dati di serie temporali quando la variabile può occasionalmente manifestare salti casuali, come nel caso dei prezzi del petrolio o dell'elettricità (gli shock di un evento esogeno discreto possono fare balzare i prezzi in alto o in basso). In conclusione, questi tre processi stocastici possono essere combinati e composti come richiesto.

Analisi di Non Stazionarità della Variabile Dipendente

Sommario statistico

Talvolta alcuni dati di serie temporali possono essere modellati solo usando il metodo del processo stocastico, dato che gli eventi sottostanti sono di natura stocastica. Per esempio, non si possono adeguatamente modellare e prevedere i prezzi di titoli, i tassi d'interesse, il prezzo del petrolio ed i prezzi di altre commodities usando una semplice modello di regressione, perché queste variabili sono molto incerte e volatili e non seguono una regola statica di comportamento. In altre parole, il processo è non stazionario. La stazionarità è controllata qui usando il Test dei Runs, mentre un altro indizio visivo si trova nel report Autocorrelazione (la funzione di autocorrelazione ACF tende a decadere lentamente). Un processo stocastico è una sequenza di eventi o percorsi generati da leggi probabilistiche. In altre parole, eventi casuali possono verificarsi nel tempo, ma essi sono governati da specifiche regole statistiche e probabilistiche. I principali processi stocastici includono Passeggiata casuale o Moto Browniano, Ritorno alla media e Diffusione a salti. Questi processi possono essere usati per prevedere una molteplicità di variabili che sembrano seguire tendenze casuali, ma che sono tuttavia limitate da leggi probabilistiche. L'equazione che genera il processo è nota in anticipo, ma i risultati reali generati non sono noti.

Il processo di Passeggiata casuale di Moto Browniano può essere usato per prevedere i prezzi di titoli, i prezzi di commodities ed altri dati di serie temporali, dato una deriva o tasso di crescita e una volatilità intorno al percorso della deriva. Il processo di Ritorno alla media può essere usato per ridurre le fluttuazioni del processo di Passeggiata casuale permettendo al percorso di stabilire come obiettivo un valore a lungo termine. Questo lo rende utile nella previsione di variabili di serie temporali che hanno un tasso a lungo termine, come i tassi di interesse e dell'inflazione (questi sono tassi di obiettivo a lungo termine delle autorità di regolamentazione o del mercato). Il processo di Diffusione a salti è utile nella previsione di dati di serie temporali quando la variabile può occasionalmente manifestare salti casuali, come nel caso dei prezzi del petrolio o dell'elettricità (gli shock di un evento esogeno discreto possono fare balzare i prezzi in alto o in basso). In conclusione, questi tre processi stocastici possono essere combinati e composti come richiesto.



Sommario statistico

I seguenti sono i parametri stimati di un processo stocastico, considerato i dati forniti. Sta a voi determinare se la probabilità di adattamento (simile al calcolo della bontà di adattamento) è sufficiente da giustificare l'uso di una previsione con processo stocastico e, se sì, se debba essere un modello con passeggiata casuale, con ritorno alla media o con diffusione a salti, o combinazioni di questi. Nella scelta del corretto modello con processo stocastico, dovete dipendere dalle esperienze passate e dalle aspettative economiche e finanziarie a priori, su cosa possa meglio rappresentare l'insieme dei dati sottostanti. Questi parametri possono essere inseriti in un processo stocastico (**Simulatore di Rischio I Previsione I Processi stocastici**).

Periodico					
Tasso della deriva	*-1.48%*	*Tasso di ritorno*	*283.89%*	*Tasso del salto*	*20.41%*
Volatilità	*88.84%*	*Valore a lungo termine*	*327.72*	*Dimensione del salto*	*237.89*

Probabilità di adattamento al modello stocastico:
Un adattamento alto significa che un modello stocastico è meglio di modelli convenzionali.

Esecuzioni	*20*	*Standard Normale*	*-1.7321*
Positive	*25*	*Valore di P (1 coda)*	*0.0416*
Negative	*25*	*Valore di P (2 code)*	*0.0833*
Esecuzione prevista	*26*		

Un valore di P basso (sotto 0.10, 0.05, 0.01) significa che la sequenza non è casuale e, quindi, soffre di problemi di stazionarità e che un modello ARIMA potrebbe essere più adatto. Per contro, valori di P più alti indicano una casualità e processi stocastici potrebbero essere appropriati.

Figura 5.26 – Stima dei Parametri del Processo Stocastico

A questo punto è d'uopo una nota di cautela. La calibratura dei parametri stocastici mostra tutti i parametri per tutti i processi e non distingue quale processo è migliore e qual è peggiore o quale processo è più appropriato da usare. Sta all'utente fare questa determinazione. Per esempio, se vediamo un tasso di reversione del 283%, è probabile che un processo con ritorno alla media non sia appropriato, o un tasso molto alto di salti, diciamo del 100%, significa molto probabilmente che un processo con diffusione a salti non è appropriato, e così via. L'analisi, inoltre, non può determinare cosa sia la variabile e quale sia la sorgente dei dati. Per esempio, i dati grezzi provengono da prezzi storici di titoli, o da prezzi storici dell'elettricità, o da tassi d'inflazione, o dal moto molecolare di particelle subatomiche, e così via? Solo l'utente può saperlo e, usando la teoria e la conoscenza a priori, essere in grado, quindi, di scegliere il processo corretto da usare (per esempio, i prezzi dei titoli tendono a seguire una passeggiata casuale di Moto Browniano, mentre i tassi d'inflazione seguono un processo con ritorno alla media; dall'altra parte, se state prevedendo il prezzo dell'elettricità, è più appropriato un processo con diffusione a salti).

La multicollinearità esiste se c'è una relazione lineare tra le variabili indipendenti. Quando questo accade, l'equazione di regressione non può essere stimata in nessun modo. In situazioni di quasi collinearità, l'equazione di regressione stimata sarà distorta e fornirà risultati inesatti. Questa situazione è specialmente vera se si usa il metodo della regressione a passi successivi, dove le variabili indipendenti statisticamente significative saranno eliminate dal mix della regressione prima del previsto, dando come risultato un'equazione di regressione che non è né efficiente né precisa. Un test veloce per la presenza della multicollinearità in un'equazione di regressione multipla è che il valore di R-quadrato è relativamente alto, mentre le statistiche di t sono relativamente basse.

Un altro test veloce è di creare una matrice di correlazione tra le variabili indipendenti. Un'alta correlazione incrociata indica un potenziale per autocorrelazione. La regola empirica è che una correlazione con un valore assoluto maggiore di 0.75 è indicativa di una severa multicollinearità. Un altro test della multicollinearità è di usare il Fattore d'inflazione della varianza (VIF), ottenuto regredendo ciascuna variabile indipendente a tutte le altre variabili indipendenti, ottenendo poi il valore di R-Quadrato e calcolando infine il Fattore d'inflazione della varianza (VIF). Un Fattore d'inflazione della varianza (VIF) superiore a 2.0 può essere considerato come una multicollinearità severa. Un Fattore d'inflazione della varianza (VIF) superiore a 10.0 indica una multicollinearità distruttiva (Figura 5.27).

Matrice di Correlazione

CORRELAZIONE	X2	X3	X4	X5
X1	0.333	0.959	0.242	0.237
X2	1.000	0.349	0.319	0.120
X3		1.000	0.196	0.227
X4			1.000	0.290
X5				1.000

Fattore d'inflazione della varianza (VIF)

VIF	X2	X3	X4	X5
X1	1.12	12.46	1.06	1.06
X2	N/A	1.14	1.11	1.01
X3		N/A	1.04	1.05
X4			N/A	1.09
X5				N/A

Figura 5.27 –Errori di Multicollinearità

La Matrice di correlazione elenca le Correlazioni Prodotto Momento di Pearson (indicate comunemente come R di Pearson) tra coppie di variabili. Il coefficiente di correlazione oscilla tra -1.0 and + 1.0 incluso. Il segno indica la direzione dell'associazione tra le variabili, mentre il coefficiente indica la magnitudine o forza d'associazione. La R di Pearson misura solamente una relazione lineare ed è meno efficace nella misurazione di relazioni non lineari.

Per testare se le correlazioni sono significative, si esegue un test di verifica d'ipotesi a due code ed i risultanti valori di p sono elencati di sopra. Valori di p minori di 0.10, 0.05 e 0.01 sono evidenziati in blu per indicare una significatività statistica. In altre parole, un valore di p per una copia di correlazioni che è minore di un dato livello di significatività è statisticamente significativamente diverso da zero, indicando che esiste una significativa relazione lineare tra le due variabili.

Il Coefficiente di Correlazione Prodotto Momento di Pearson (*R*) tra due variabili (*x* e *y*) è collegato alla misura di covarianza (*cov*), dove $R_{x,y} = \frac{COV_{x,y}}{s_x s_y}$. Il beneficio nel dividere la covarianza per il prodotto della deviazione standard (*s*) delle due variabili è che il risultante coefficiente di correlazione è confinato tra -1.0 e +1.0 incluso. Questo rende la correlazione una buona misura relativa per fare paragoni tra variabili diverse (particolarmente con unità e magnitudini differenti). Anche la correlazione non parametrica per ranghi di Spearman è inclusa sotto. La R di Spearman è collegata alla R di Pearson nel senso che i dati sono prima classificati per rango e poi correlati. Le correlazioni per ranghi forniscono una stima migliore della relazione tra due variabili, quando una o entrambe sono non lineari.

Bisogna sottolineare che una correlazione significativa non implica una causazione. Associazioni tra variabili non implicano in nessun modo che il cambiamento di una variabile causa il cambiamento di un'altra variabile. Due variabili che si muovono indipendentemente, ma in un percorso imparentato, possono essere correlate, ma la loro relazione può essere falsa (per esempio, una correlazione tra macchie solari ed il mercato azionario può essere forte, ma si può presumere che non esiste una causalità e che questa relazione è completamente falsa).

5.9 Strumento di Analisi Statistica

Un altro strumento molto potente in Simulatore di Rischio è lo Strumento di Analisi Statistica, che determina le proprietà statistiche dei dati. La diagnostica eseguita include il controllo dei dati per varie proprietà statistiche, dalle statistiche descrittive di base all'analisi e la calibratura delle proprietà stocastiche dei dati.

Procedura:

- Aprite il modello d'esempio (**Simulatore di Rischio | Esempi | Analisi statistica**), andate al foglio di lavoro *Dati* e selezionate i dati, incluso i nomi delle variabili (celle **C5:E55**).
- Cliccate su Simulatore di Rischio | Strumenti | Analisi statistica (Figura 5.28).
- Controllate il *tipo dei dati*, se provengono da una singola variabile o da multiple variabili disposte in righe. Nel nostro esempio assumiamo che le aree selezionate dei dati provengono da multiple variabili. Cliccate su **OK** quando avete terminato.
- *Scegliete i tests statistici* che desiderate eseguire. Il suggerimento (e di default) è di scegliere tutti i tests. Cliccate su **OK** quando avete terminato (Figura 5.29).

Prendete del tempo per esaminare i reports generati per comprendere meglio i tests statistici che sono stati eseguiti (esempi di reports sono mostrati nelle Figure 5.30-5.33).

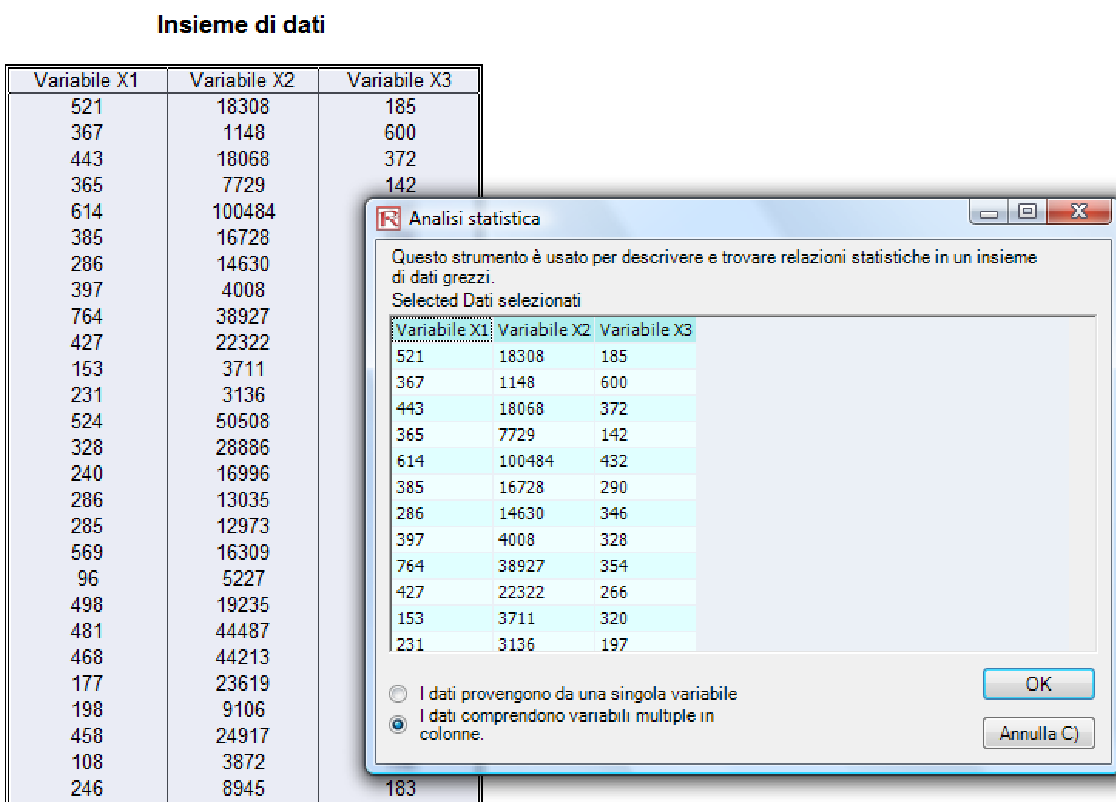


Figura 5.28 – Eseguire lo Strumento di Analisi Statistica

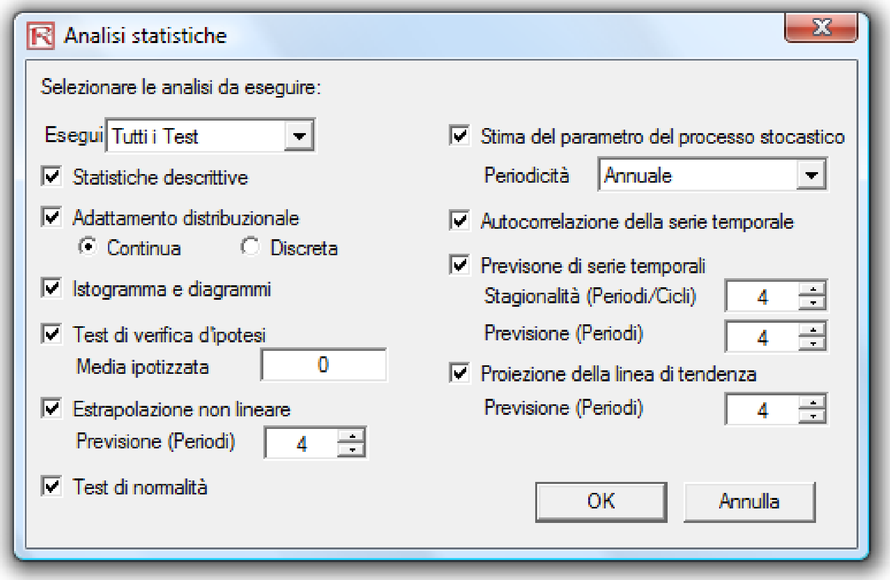


Figura 5.29 –Tests Statistici

Statistica Descrittiva

Analisi della statistica

Quasi tutte le distribuzioni possono essere descritte da 4 momenti (alcune distribuzioni richiedono un momento, mentre altre richiedono due momenti e così via). La statistica descrittiva cattura questi momenti quantitativamente. Il primo momento descrive la posizione di una distribuzione (vale a dire, media, mediana e moda) ed è interpretata come il valore atteso, i rendimenti attesi o il valore medio degli eventi.

La Media aritmetica calcola la media di tutti gli eventi sommando tutti i punti dati e dividendoli per il numero di punti. La Media geometrica è calcolata ottenendo la radice potenziata dei prodotti di tutti i punti dati; è necessario che siano tutti positivi. La Media geometrica è più precisa per percentuali o tassi che fluttuano in modo significativo. Per esempio, potete usare la media geometrica per calcolare il tasso medio di crescita dato un interesse composto con tassi variabili. La Media troncata calcola la Media aritmetica dell'insieme di dati, dopo il taglio dei valori abnormi estremi. Dato che le medie sono soggette a distorsioni significative se ci sono valori abnormi, la Media troncata riduce queste distorsioni in distribuzioni asimmetriche.

L'Errore standard della media calcola l'errore che circonda la media del campione. Tanto più grande è la dimensione del campione, tanto più piccolo è l'errore, tale che per un campione infinitamente grande l'errore si avvicina a zero, indicando che il parametro della popolazione è stato stimato. A causa di errori di campionamento, viene fornito l'Intervallo di confidenza per la media del 95%. Basandosi su un'analisi dei punti dati del campione, la media reale della popolazione dovrebbe cadere tra questi Intervalli inferiori e superiori per la media.

La Mediana è il punto dati dove il 50% di tutti i punti dati si trovano sopra questo valore e il 50% sotto questo valore. Tra i primi tre momenti statistici, la mediana è la meno sensibile ai valori estremi. Una distribuzione simmetrica ha la Mediana uguale alla Media aritmetica. Si ha una distribuzione asimmetrica, quando la Mediana è molto lontana dalla Media. La Moda misura il punto dati che accade più spesso.

Minimo è il valore più piccolo nell'insieme di dati, mentre Massimo è il valore più grande. L'Intervallo è la differenza tra il valore Massimo e quello Minimo.

Il secondo momento misura la dispersione o l'ampiezza di una distribuzione ed è spesso descritta usando misure come le Deviazioni standard, le Varianze, i Quartili e gli Intervalli interquartili. La Deviazione standard indica la deviazione media di tutti i punti dati dalle loro medie. È una misura diffusa dato che è associata con il rischio (deviazioni standard più grandi significano una distribuzione più ampia, un rischio più alto o una dispersione più larga attorno alla media) e le sue unità sono identiche a quelle dell'insieme di dati originale. La Deviazione standard del campione si distingue dalla Deviazione standard della popolazione, in quanto la prima usa una correzione di gradi di libertà per rendere conto di dimensioni piccole del campione. Vengono anche forniti gli Intervalli di confidenza inferiori e superiori per la Deviazione standard e la deviazione standard reale della popolazione cade all'interno di questo intervallo. Se il vostro insieme di dati copre tutti gli elementi della popolazione, usate piuttosto la Deviazione standard della popolazione. Le due misure della Varianza sono semplicemente i valori al quadrato delle deviazioni standard.

Il Coefficiente di variabilità è la deviazione standard del campione diviso per la media del campione, fornendo una misura della dispersione libera da unità che può essere confrontata tra distribuzioni diverse (potete ora confrontare una distribuzione con valori denominati in milioni di dollari con una in miliardi di dollari, o metri e chilogrammi, ecc.). Il 1° Quartile misura il 25° percentile dei punti dati, quando sono disposti dal valore più piccolo a quello più grande. Il 3° Quartile è il valore del punto dati al 75° percentile. I quartili sono talvolta usati come gli intervalli superiori ed inferiori di una distribuzione, dato che questo tronca l'insieme di dati per ignorare i valori abnormi. L'Intervallo interquartile è la differenza tra il 1° ed il 3° quartile ed è spesso usato per misurare la larghezza del centro di una distribuzione.

L'asimmetria è il terzo momento in una distribuzione. L'asimmetria caratterizza il grado di asimmetria di una distribuzione attorno alla sua media. L'asimmetria positiva indica una distribuzione con una coda asimmetrica che si estende più verso valori positivi. L'asimmetria negativa indica una distribuzione con una coda asimmetrica che si estende più verso valori negativi.

La Curtosi caratterizza la relativa pinguedine ("peakedness") o piattezza di una distribuzione paragonata ad una distribuzione normale. È il quarto momento in una distribuzione. Un valore positivo di Curtosi indica una distribuzione relativamente pingue ("peaked"). Una Curtosi negativa indica una distribuzione relativamente piatta. La Curtosi misurata qui è stata centrata allo zero (certe altre misure di Curtosi sono centrate attorno a 3,0). Anche se entrambe sono ugualmente valide, la centratura attraverso lo zero rende l'interpretazione più facile. Un'alta Curtosi positiva indica una distribuzione pingue intorno al suo centro e code leptocurtiche o "grasse". Questo indica una probabilità più alta di eventi estremi (per esempio, eventi catastrofici, attacchi terroristici, crolli della Borsa) che quella prevista da una distribuzione normale.

Sommario delle Statistiche

Statistiche	*Variable X1*		
Osservazioni	50.0000	*Deviazione standard (campione)*	172.9140
Media aritmetica	331.9200	*Deviazione standard (popolazione)*	171.1761
Media geometrica	281.3247	*Intervallo di confidenza inferiore per la deviazione standard*	148.6090
Media troncata	325.1739	*Intervallo di confidenza superiore per la deviazione standard*	207.7947
Errore standard della media aritmetica	24.4537	*Varianza (campione)*	29899.2588
Intervallo di confidenza inferiore per la media	283.0125	*Varianza (popolazione)*	29301.2736
Intervallo di confidenza superiore per la media	380.8275	*Coefficiente di variabilità*	0.5210
Mediana	307.0000	*1° Quartile (Q1)*	204.0000
Minimo	47.0000	*3° Quartile (Q3)*	441.0000
Massimo	764.0000	*Intervallo interquartile*	237.0000
Intervallo	717.0000	*Asimmetria*	0.4838
		Curtosi	-0.0952

Figura 5.30 – Esempio di un Report dello Strumento di Analisi Statistica

Test di verifica d'ipotesi (t-Test sulla media della popolazione di una variabile)

Sommario statistico

Statistiche dall'insieme di dati:		***Statistiche calcolate:***	
Osservazioni	*50*	*Statistica di t*	*13.5734*
Media del campione	*331.92*	*Valore di P (a coda destra)*	***0.0000***
Deviazione standard del campione	*172.91*	*Valore di P (a coda sinistra)*	***1.0000***
		Valore di P (a due code)	***0.0000***
Statistiche fornite dall'utente:			
		Ipotesi nulla (Ho):	*μ = Media ipotizzata*
Media ipotizzata	*0.00*	*Ipotesi alternativa (Ha):*	*μ = Media ipotizzata*

Note: "<>" denota "maggiore di" per test a coda destra, "minore di" per test a coda sinistra, o "non uguale a" per tests a due code

Sommario dei Tests di verifica d'ipotesi

Il t-test con una variabile è appropriato quando la deviazione standard della popolazione non è nota, ma si presume che la distribuzione della campionatura sia approssimativamente normale (il t-test è usato quando la dimensione del campione è minore di 30, ma è anche appropriato e fornisce, in fatti, risultati più conservativi con insiemi di dati più grandi). Questo t-test può essere applicato a tre tipi di tests di verifica d'ipotesi: un test a due code, un test a coda destra e un test a coda sinistra. I tre tests e i loro rispettivi risultati sono elencati sotto per vostro riferimento.

Test di verifica d'ipotesi a due code

Un'ipotesi a due code verifica l'ipotesi nulla Ho che la media della popolazione è statisticamente identica alla media ipotizzata. L'ipotesi alternativa è che la media reale della popolazione è statisticamente differente dalla media ipotizzata, quando testato usando l'insieme campione di dati. Usando il t-test, se il valore di P calcolato è minore di una quantità specificata di significatività (tipicamente 0.10, 0.05 o 0.01), questo significa che la media della popolazione è statisticamente significativamente differente dalla media ipotizzata al valore di significatività del 10%, 5% e 1% (o alla confidenza statistica del 90%, 95% e 99%). Per contro, se il valore di P è maggiore di 0.10, 0.05 o 0.01, la media della popolazione è statisticamente identica alla media ipotizzata e qualsiasi differenza è causata da pura casualità.

Test di verifica d'ipotesi a coda destra

Un'ipotesi a coda destra verifica l'ipotesi nulla Ho che la media della popolazione è statisticamente minore della o uguale alla media ipotizzata. L'ipotesi alternativa è che la media reale della popolazione è statisticamente maggiore della media ipotizzata, quando testato usando l'insieme campione di dati. Usando il t-test, se il valore di P calcolato è minore di una quantità specificata di significatività (tipicamente 0.10, 0.05 o 0.01), questo significa che la media della popolazione è statisticamente significativamente maggiore della media ipotizzata al valore di significatività del 10%, 5% e 1% (o alla confidenza statistica del 90%, 95% e 99%). Per contro, se il valore di P è maggiore di 0.10, 0.05 o 0.01, la media della popolazione è statisticamente simile alla o minore della media ipotizzata.

Test di verifica d'ipotesi a coda sinistra

Un'ipotesi a coda sinistra verifica l'ipotesi nulla Ho che la media della popolazione è statisticamente maggiore della o uguale alla media ipotizzata. L'ipotesi alternativa è che la media reale della popolazione è statisticamente minore della media ipotizzata, quando testato usando l'insieme campione di dati. Usando il t-test, se il valore di P calcolato è minore di una quantità specificata di significatività (tipicamente 0.10, 0.05 o 0.01), questo significa che la media della popolazione è statisticamente significativamente minore della media ipotizzata al valore di significatività del 10%, 5% e 1% (o alla confidenza statistica del 90%, 95% e 99%). Per contro, se il valore di P è maggiore di 0.10, 0.05 o 0.01, la media della popolazione è statisticamente simile alla o maggiore della media ipotizzata e qualsiasi differenza è causata da pura casualità.

Dato che il t-test è più conservativo e non richiede una deviazione standard nota della popolazione, come lo Z-test, usiamo solo questo t-test.

Figura 5.31 – Esempio di un Report dello Strumento di Analisi Statistica (Test di Verifica d'ipotesi di Una Variabile)

Test di Normalità

Il test di normalità è una forma di test non parametrico, che non fa delle ipotesi sulla forma specifica della popolazione da cui i campioni sono stati tratti, permettendo così l'analisi di campioni più piccoli di insiemi di dati. Questo test valuta l'ipotesi nulla, che il campione di dati è stato tratto da una popolazione distribuita normalmente, contro una ipotesi alternativa, che il campione dei dati non è distribuito normalmente. Se il valore di P calcolato è maggiore del o uguale al valore di significatività alfa, allora rifiutate l'ipotesi nulla e accettate l'ipotesi alternativa. Altrimenti se il valore di P è maggiore del valore di significatività alfa, non rifiutate l'ipotesi nulla. Questo test dipende da due frequenze cumulative: una derivata dall'insieme campione di dati, la seconda da una distribuzione teoretica basata sulla media e la deviazione standard dei dati campioni. Un'alternativa a questo test è il test di normalità Chi-quadrato. Il test Chi-quadrato richiede più punti dati paragonato al test di normalità usato qui.

Risultato del test

Media dei dati	*331.92*
Deviazione standard	*172.91*
Statistica di D	*0.0859*
D-Critico a 1%	*0.1150*
D-Critico a 5%	*0.1237*
D-Critico a 10%	*0.1473*

Ipotesi nulla: i dati sono distribuiti normalmente.

Dati	Frequenza relativa	Osservato	Atteso	O-A
47.00	0.02	0.02	0.0497	-0.0297
68.00	0.02	0.04	0.0635	-0.0235
87.00	0.02	0.06	0.0783	-0.0183
96.00	0.02	0.08	0.0862	-0.0062
102.00	0.02	0.10	0.0918	0.0082
108.00	0.02	0.12	0.0977	0.0223
114.00	0.02	0.14	0.1038	0.0362
127.00	0.02	0.16	0.1180	0.0420
153.00	0.02	0.18	0.1504	0.0296
177.00	0.02	0.20	0.1851	0.0149
186.00	0.02	0.22	0.1994	0.0206
188.00	0.02	0.24	0.2026	0.0374
198.00	0.02	0.26	0.2193	0.0407

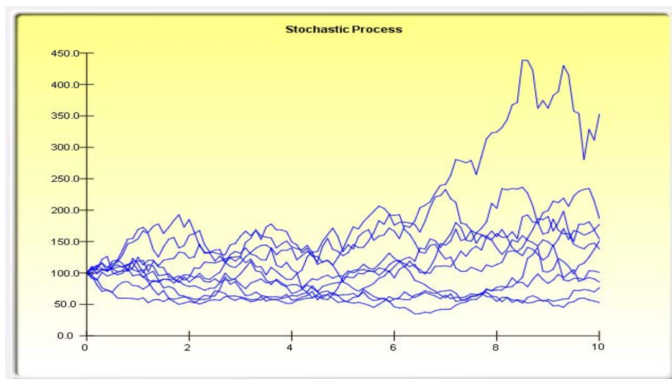
Figura 5.32 – Esempio di un Report dello Strumento di Analisi Statistica (Test di Normalità)

Processo stocastico – Stime dei Parametri

Sommario statistico

Un processo stocastico è una sequenza di eventi o percorsi generati da leggi probabilistiche. In altre parole, eventi casuali possono verificarsi nel tempo, ma essi sono governati da specifiche regole statistiche e probabilistiche. I principali processi stocastici includono Passeggiata casuale o Moto Browniano, Ritorno alla media e Diffusione a salti. Questi processi possono essere usati per prevedere una molteplicità di variabili che sembrano seguire tendenze casuali, ma che sono tuttavia limitate da leggi probabilistiche. L'equazione che genera il processo è nota in anticipo, ma i risultati reali generati non sono noti.

Il processo di Passeggiata casuale di Moto Browniano può essere usato per prevedere i prezzi di titoli, i prezzi di commodities ed altri dati di serie temporali, dato una deriva o tasso di crescita e una volatilità intorno al percorso della deriva. Il processo di Ritorno alla media può essere usato per ridurre le fluttuazioni del processo di Passeggiata casuale permettendo al percorso di stabilire come obiettivo un valore a lungo termine. Questo lo rende utile nella previsione di variabili di serie temporali che hanno un tasso a lungo termine, come i tassi di interesse e dell'inflazione (questi sono tassi di obiettivo a lungo termine delle autorità di regolamentazione o del mercato). Il processo di Diffusione a salti è utile nella previsione di dati di serie temporali quando la variabile può occasionalmente manifestare salti casuali, come nel caso dei prezzi del petrolio o dell'elettricità (gli shock di un evento esogeno discreto possono fare balzare i prezzi in alto o in basso). In conclusione, questi tre processi stocastici possono essere combinati e composti come richiesto.



Sommario statistico

I seguenti sono i parametri stimati di un processo stocastico, considerato i dati forniti. Sta a voi determinare se la probabilità di adattamento (simile al calcolo della bontà di adattamento) è sufficiente da giustificare l'uso di una previsione con processo stocastico e, se sì, se debba essere un modello con passeggiata casuale, con ritorno alla media o con diffusione a salti, o combinazioni di questi. Nella scelta del corretto modello con processo stocastico, dovete dipendere dalle esperienze passate e dalle aspettative economiche e finanziarie a priori, su cosa possa meglio rappresentare l'insieme dei dati sottostanti. Questi parametri possono essere inseriti in un processo stocastico (**Simulatore di Rischio I Previsione I Processi stocastici**).

(Annualizzati)					
*Tasso della deriva**	*-1.48%*	*Tasso di ritorno***	*283.89%*	*Tasso del salto***	*20.41%*
*Volatilità**	*88.84%*	*Valore a lungo termine***	*327.72*	*Dimensione del salto***	*237.89*

Probabilità di adattamento al modello stocastico: *46.48%*

**I valori sono annualizzati*

*** I valori sono periodici*

Figura 5.33 – Esempio di un Report dello Strumento di Analisi Statistica (Stima dei Parametri Stocastici)

5.10 Strumento di Analisi Distribuzionale

Lo strumento di Analisi Distribuzionale è uno strumento di probabilità statistica in Simulatore di Rischio che è piuttosto utile in una varietà di campi. Può essere usato per calcolare la funzione di densità della probabilità (PDF), denominata anche la funzione di massa della probabilità (PMF) per distribuzioni discrete (questi termini sono intercambiabili), dove, dato una determinata distribuzione e i suoi parametri, possiamo determinare la probabilità di accadimento, dato un determinato esito *x*. In aggiunta, si può calcolare la funzione di distribuzione cumulativa (CDF), che è la somma dei valori della PDF fino a questo valore *x*. Per finire, la funzione di distribuzione cumulativa inversa (ICDF) è usata per calcolare il valore di *x*, data la probabilità cumulativa di accadimento.

Questo strumento è accessibile tramite **Simulatore di Rischio** | **Strumenti** | **Analisi distribuzionale**. Come esempio del suo utilizzo, la Figura 5.34 mostra il calcolo di una distribuzione binomiale (vale a dire una distribuzione con due esiti, come il lancio di una moneta, dove l'esito è Testa o Coda, con una determinata probabilità prescritta di teste e code). Supponiamo di lanciare la moneta per due volte e di impostare l'esito di una Testa come un successo. Usiamo la distribuzione binomiale con Prove = 2 (due lanci della moneta) e Probabilità = 0.50 (la probabilità del successo di ottenere una Testa). Se selezioniamo PDF ed impostiamo l'intervallo dei valori di *x* da 0 a 2 con una dimensione del passo di 1 (questo significa che cerchiamo i valori di 0, 1 e 2 per *x*), otteniamo le probabilità risultanti sia in una tabella che in un formato grafico come anche i quattro momenti teorici della distribuzione. Dato che gli esiti di un lancio della moneta possono essere Testa-Testa, Coda-Coda, Testa-Coda e Coda-Testa, la probabilità di non ottenere una Testa è del 25%, di ottenere una Testa è del 50% e di ottenere due Teste è del 25%.

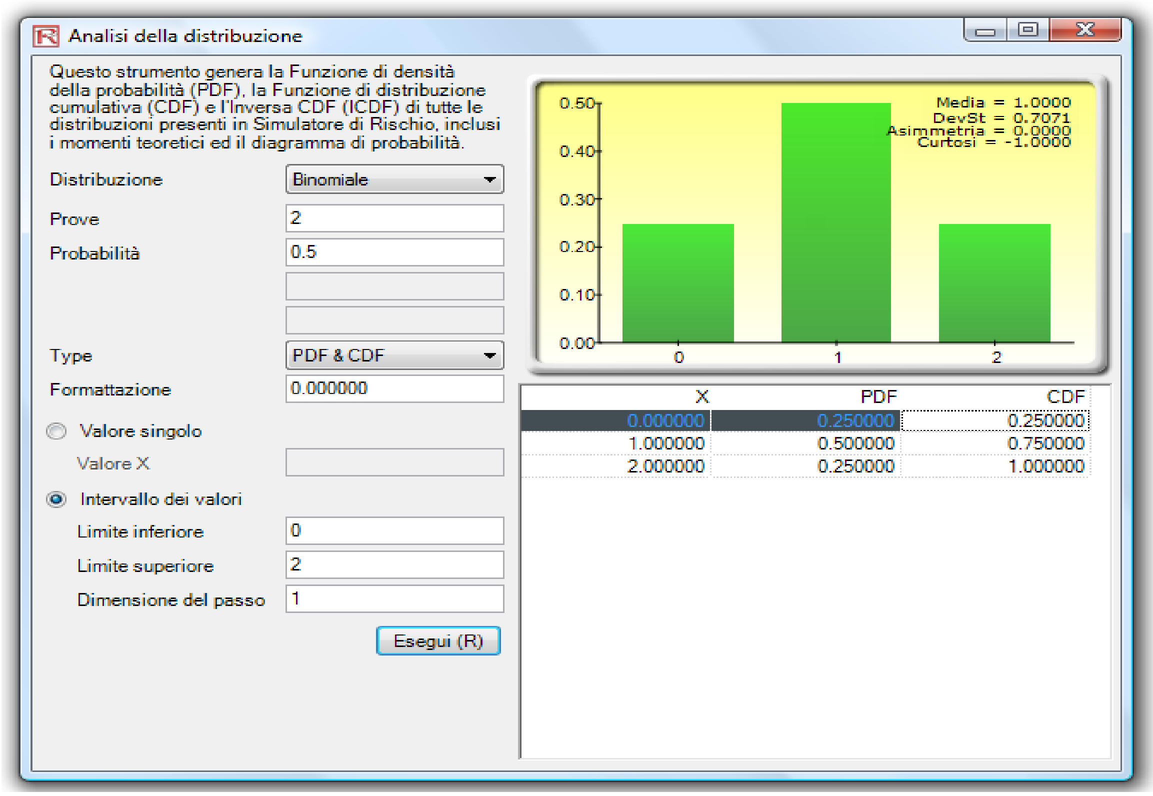


Figura 5.34 – Strumento di Analisi Distribuzionale (Distribuzione Binomiale con 2 Prove)

In modo simile possiamo ottenere le probabilità esatte, per esempio, di 20 lanci, come visto nella Figura 5.35. I risultati sono presentati sia come tabella che in formati grafici.

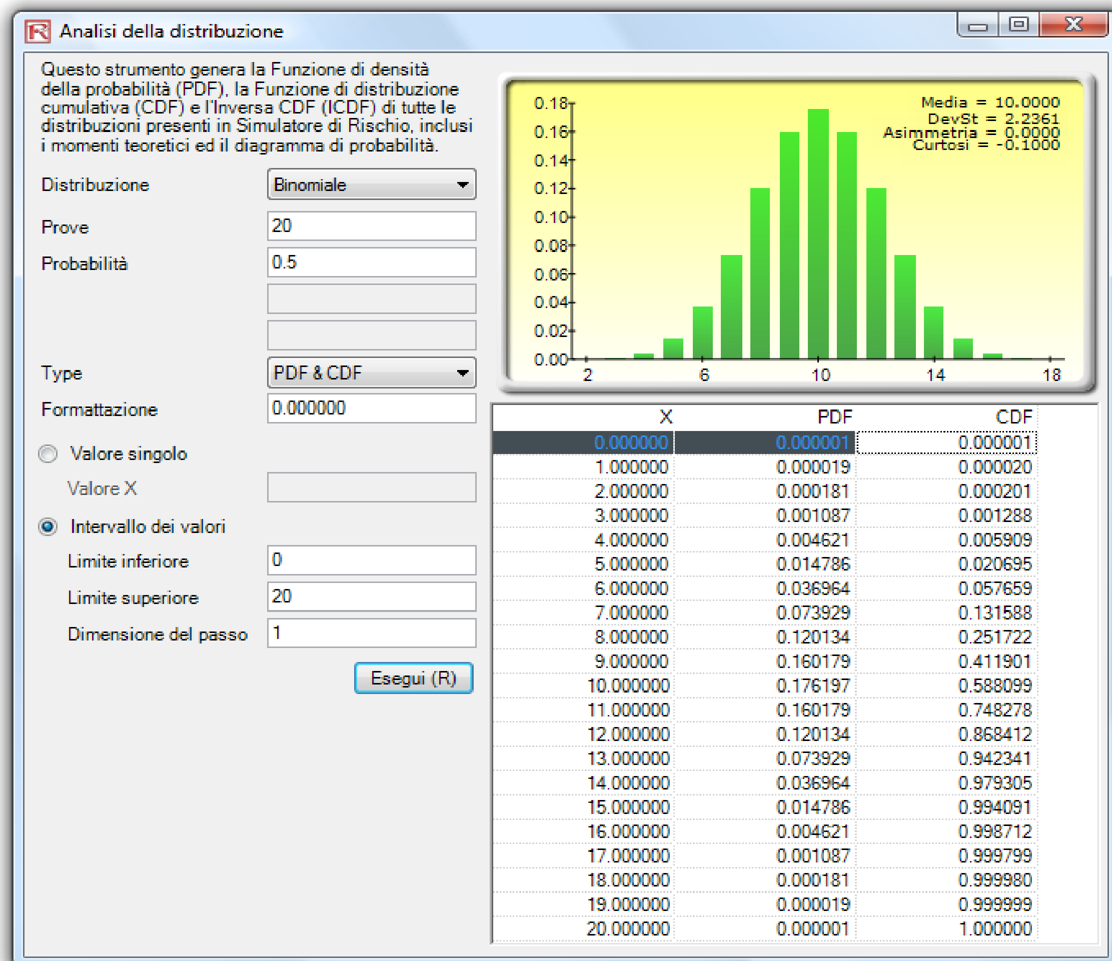


Figura 5.35 – Strumento di Analisi Distribuzionale (Distribuzione Binomiale con 20 Prove)

La Figura 5.36 mostra la stessa distribuzione binomiale, ma ora viene calcolata la CDF. La CDF è semplicemente la somma dei valori della PDF fino al punto *x*. Per esempio, nella Figura 5.35, vediamo che le probabilità di 0, 1 e 2 sono 0.000001, 0.000019 e 0.000181. La somma di queste probabilità è 0.000201, che è il valore della CDF a *x* = 2 nella Figura 5.36. Mentre la PDF calcola le probabilità di ottenere esattamente 2 teste, la CDF calcola le probabilità di ottenere non più di 2 teste o fino a 2 teste (ovvero le probabilità di 0, 1 e 2 teste). Prendere il complemento (vale a dire, 1 – 0.00021) dà come risultato 0.999799 o 99.9799%, che è la probabilità di ottenere almeno 3 teste o di più.

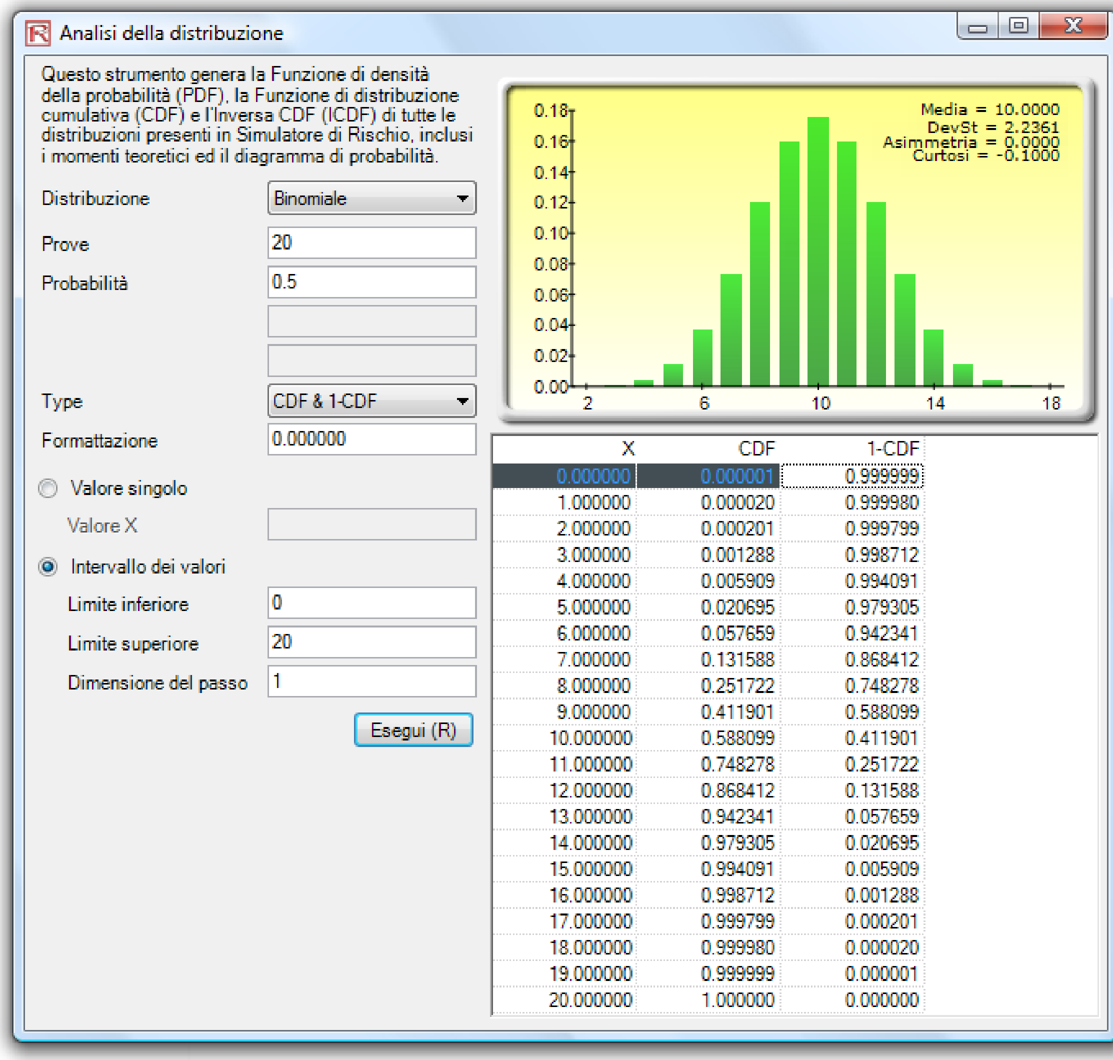


Figura 5.36 – Strumento di Analisi Distribuzionale (CDF della Distribuzione Binomiale con 20 Prove)

L'uso di questo strumento di Analisi Distribuzionale in Simulatore di Rischio permette l'analisi di distribuzioni ancora più avanzate, come la gamma, la beta, la binomiale negativa e molte altre. Come ulteriore esempio dell'uso dello strumento in una distribuzione continua e delle funzionalità della ICDF, la Figura 5.37 mostra la distribuzione normale standard (una distribuzione normale con la media dello zero e la deviazione standard di uno) sulla quale applichiamo la ICDF per trovare il valore di *x* che corrisponde alla probabilità cumulativa del 97.50% (CDF). In altre parole, una CDF ad una coda del 97.50% è equivalente ad un intervallo di confidenza a due code del 95% (c'è una probabilità del 2.50% nella coda destra e del 2.50% nella coda sinistra, lasciando un 95% nel centro ovvero nell'area dell'intervallo di confidenza;

questo è equivalente ad un'area del 97.50% per una coda). Il risultato è il noto Punteggio Z di 1.96. Quindi, l'uso di questo strumento di Analisi Distribuzionale permette di ottenere velocemente e facilmente sia i punteggi standardizzati che le probabilità esatte e cumulative per altre distribuzioni.

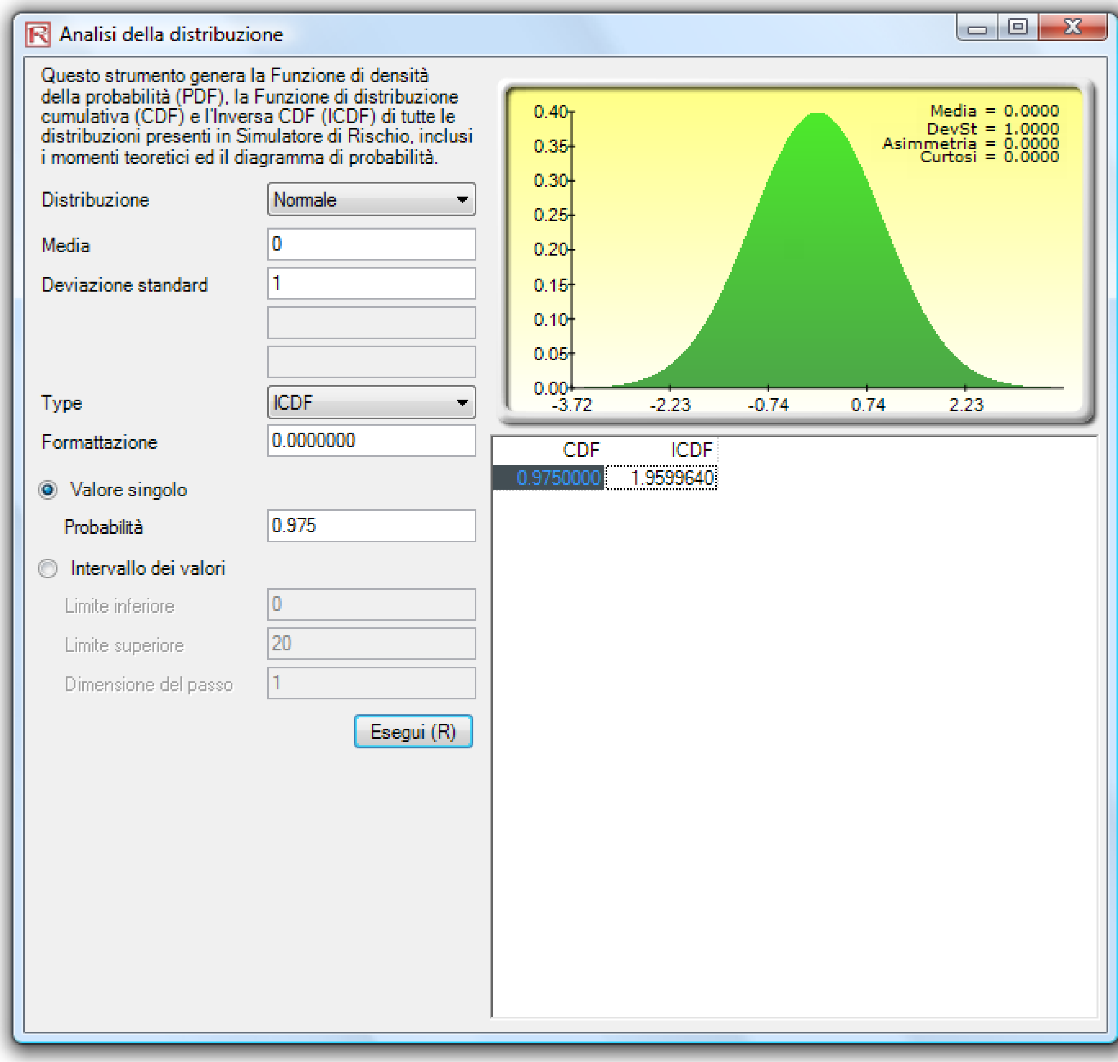


Figura 5.37 – Strumento di Analisi Distribuzionale (ICDF and Punteggio Z della Distribuzione Normale)

5.11 Strumenti di Analisi degli Scenari

Lo strumento di analisi degli scenari in Simulatore di Rischio vi permette di eseguire multipli scenari velocemente e facilmente, modificando uno o due parametri d'input per determinare l'output di una variabile. La Figura 5.38 illustra il funzionamento di questo strumento sul modello d'esempio del cash flow scontato (Modello 7 nella cartella dei Modelli d'esempio in Simulatore di Rischio). In quest'esempio, la cella G6 (valore attuale netto) è selezionata come l'output che c'interessa, mentre le celle C9 (aliquota effettiva d'imposta) e C12 (prezzo del prodotto) sono selezionate come gli inputs da perturbare. Potete impostare sia i valori iniziali e finali da testare che la dimensione del passo o il numero di passi da eseguire tra questi valori iniziali e finali. Il risultato e una tabella dell'analisi dello scenario (Figura 5.39), dove le intestazioni della riga e della colonna sono le due variabili d'input e il corpo della tabella mostra i valori attuali netti.

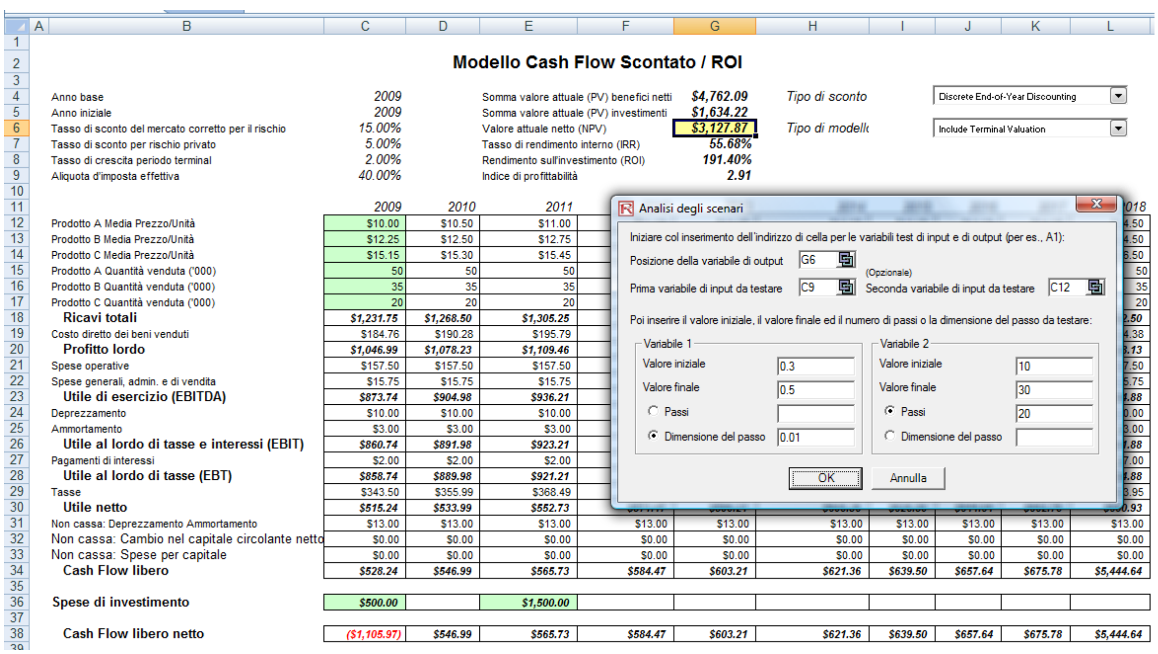


Figura 5.38 – Strumento di Analisi degli Scenari

TABELLA DELL'ANALISI DEGLI SCENARI

Variabile di output:	G6	**Valore del caso-base in**			$3,128				
Variabile della colonna	C12	**Min:**	10	**Max:**	30	**Passi:**	20	**Valore del caso-base in**	$10.00
Variabile della riga:	C9	**Min:**	0.3	**Max:**	0.5	**Passi:**	---	**Valore del caso-base in**	40.00%

	$10.00	**$11.00**	**$12.00**	**$13.00**	**$14.00**	**$15.00**	**$16.00**	**$17.00**	**$18.00**	**$19.00**
30.00%	$3,905	$4,134	$4,364	$4,594	$4,823	$5,053	$5,282	$5,512	$5,742	$5,971
31.00%	$3,827	$4,053	$4,280	$4,506	$4,732	$4,959	$5,185	$5,411	$5,638	$5,864
32.00%	$3,749	$3,972	$4,196	$4,419	$4,642	$4,865	$5,088	$5,311	$5,534	$5,757
33.00%	$3,672	$3,892	$4,111	$4,331	$4,551	$4,771	$4,990	$5,210	$5,430	$5,650
34.00%	$3,594	$3,811	$4,027	$4,243	$4,460	$4,676	$4,893	$5,109	$5,326	$5,542
35.00%	$3,516	$3,730	$3,943	$4,156	$4,369	$4,582	$4,796	$5,009	$5,222	$5,435
36.00%	$3,439	$3,649	$3,858	$4,068	$4,278	$4,488	$4,698	$4,908	$5,118	$5,328
37.00%	$3,361	$3,568	$3,774	$3,981	$4,188	$4,394	$4,601	$4,807	$5,014	$5,221
38.00%	$3,283	$3,487	$3,690	$3,893	$4,097	$4,300	$4,503	$4,707	$4,910	$5,114
39.00%	$3,206	$3,406	$3,606	$3,806	$4,006	$4,206	$4,406	$4,606	$4,806	$5,006
40.00%	$3,128	$3,325	$3,521	$3,718	$3,915	$4,112	$4,309	$4,505	$4,702	$4,899
41.00%	$3,050	$3,244	$3,437	$3,631	$3,824	$4,018	$4,211	$4,405	$4,598	$4,792
42.00%	$2,972	$3,163	$3,353	$3,543	$3,733	$3,924	$4,114	$4,304	$4,494	$4,685
43.00%	$2,895	$3,082	$3,269	$3,456	$3,643	$3,830	$4,017	$4,204	$4,390	$4,577
44.00%	$2,817	$3,001	$3,184	$3,368	$3,552	$3,735	$3,919	$4,103	$4,287	$4,470
45.00%	$2,739	$2,920	$3,100	$3,281	$3,461	$3,641	$3,822	$4,002	$4,183	$4,363
46.00%	$2,662	$2,839	$3,016	$3,193	$3,370	$3,547	$3,724	$3,902	$4,079	$4,256
47.00%	$2,584	$2,758	$2,932	$3,106	$3,279	$3,453	$3,627	$3,801	$3,975	$4,149
48.00%	$2,506	$2,677	$2,847	$3,018	$3,189	$3,359	$3,530	$3,700	$3,871	$4,041
49.00%	$2,429	$2,596	$2,763	$2,930	$3,098	$3,265	$3,432	$3,600	$3,767	$3,934
50.00%	$2,351	$2,515	$2,679	$2,843	$3,007	$3,171	$3,335	$3,499	$3,663	$3,827

	$20.00	**$21.00**	**$22.00**	**$23.00**	**$24.00**	**$25.00**	**$26.00**	**$27.00**	**$28.00**	**$29.00**	**$30.00**
30.00%	$6,201	$6,430	$6,660	$6,890	$7,119	$7,349	$7,578	$7,808	$8,038	$8,267	$8,497
31.00%	$6,090	$6,317	$6,543	$6,769	$6,996	$7,222	$7,448	$7,675	$7,901	$8,127	$8,354
32.00%	$5,980	$6,203	$6,426	$6,649	$6,872	$7,095	$7,318	$7,541	$7,764	$7,987	$8,210
33.00%	$5,869	$6,089	$6,309	$6,529	$6,748	$6,968	$7,188	$7,408	$7,627	$7,847	$8,067
34.00%	$5,759	$5,975	$6,192	$6,408	$6,625	$6,841	$7,058	$7,274	$7,491	$7,707	$7,924
35.00%	$5,648	$5,862	$6,075	$6,288	$6,501	$6,714	$6,928	$7,141	$7,354	$7,567	$7,780
36.00%	$5,538	$5,748	$5,958	$6,168	$6,378	$6,587	$6,797	$7,007	$7,217	$7,427	$7,637
37.00%	$5,427	$5,634	$5,841	$6,047	$6,254	$6,461	$6,667	$6,874	$7,081	$7,287	$7,494
38.00%	$5,317	$5,520	$5,724	$5,927	$6,130	$6,334	$6,537	$6,740	$6,944	$7,147	$7,350
39.00%	$5,206	$5,406	$5,607	$5,807	$6,007	$6,207	$6,407	$6,607	$6,807	$7,007	$7,207
40.00%	$5,096	$5,293	$5,489	$5,686	$5,883	$6,080	$6,277	$6,473	$6,670	$6,867	$7,064
41.00%	$4,985	$5,179	$5,372	$5,566	$5,759	$5,953	$6,147	$6,340	$6,534	$6,727	$6,921
42.00%	$4,875	$5,065	$5,255	$5,446	$5,636	$5,826	$6,016	$6,207	$6,397	$6,587	$6,777
43.00%	$4,764	$4,951	$5,138	$5,325	$5,512	$5,699	$5,886	$6,073	$6,260	$6,447	$6,634
44.00%	$4,654	$4,838	$5,021	$5,205	$5,389	$5,572	$5,756	$5,940	$6,123	$6,307	$6,491
45.00%	$4,543	$4,724	$4,904	$5,085	$5,265	$5,445	$5,626	$5,806	$5,987	$6,167	$6,347
46.00%	$4,433	$4,610	$4,787	$4,964	$5,141	$5,319	$5,496	$5,673	$5,850	$6,027	$6,204
47.00%	$4,322	$4,496	$4,670	$4,844	$5,018	$5,192	$5,365	$5,539	$5,713	$5,887	$6,061
48.00%	$4,212	$4,382	$4,553	$4,724	$4,894	$5,065	$5,235	$5,406	$5,576	$5,747	$5,918
49.00%	$4,101	$4,269	$4,436	$4,603	$4,771	$4,938	$5,105	$5,272	$5,440	$5,607	$5,774
50.00%	$3,991	$4,155	$4,319	$4,483	$4,647	$4,811	$4,975	$5,139	$5,303	$5,467	$5,631

Figura 5.39 – Tabella dell'Analisi degli Scenari

5.12 Strumento di Segmentazione Clustering

Un'ultima tecnica analitica interessante è quella della segmentazione clustering (raggruppamento). La Figura 6.25 illustra un esempio di un insieme di dati. Potete selezionare i dati ed eseguire lo strumento sotto **Simulatore di Rischio | Strumenti | Segmentazione Clustering**. La Figura 5.40 mostra un esempio della segmentazione di due gruppi. In altre parole, prendiamo l'insieme di dati originale ed eseguiamo alcuni algoritmi interni (una combinazione o un clustering gerarchico "k-means" e altri metodi di momenti per trovare i gruppi col miglior adattamento o i clusters statistici naturali) per dividere o segmentare statisticamente l'insieme di dati originale in due gruppi. Potete vedere le appartenenze dei due gruppi nella Figura 5.40. Ovviamente potete segmentare questo insieme di dati nel numero desiderato di gruppi. Questa tecnica è molto utile in una varietà di campi, compreso il marketing (segmentazione di mercato dei clienti in vari gruppi di management delle relazioni con i clienti e così via), la fisica, l'ingegneria ed altro.

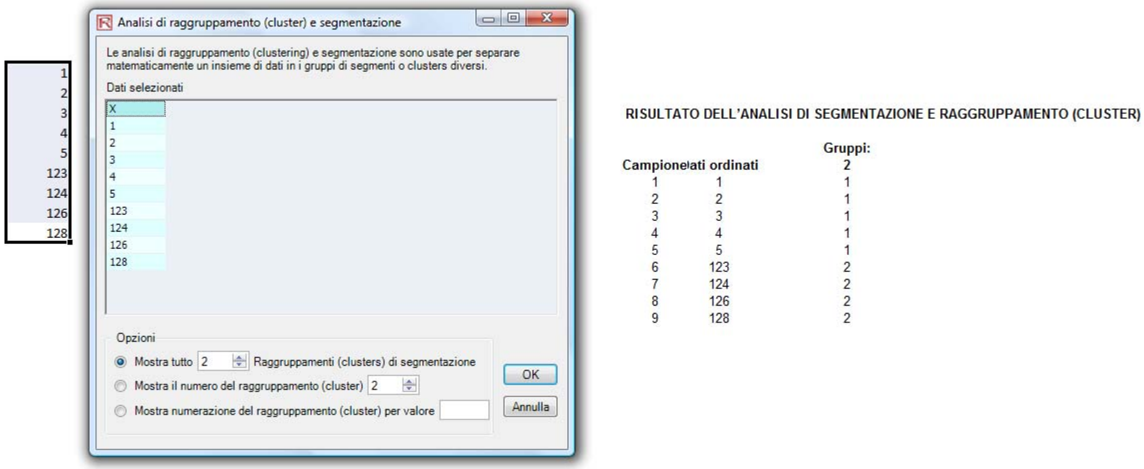


Figura 5.40 – Strumento di Segmentazione Clustering e i Risultati

5.13 Simulatore di Rischio 2011/2012 Nuovi Strumenti

5.14 Generazione di numeri casuali; Monte Carlo in confronto a Ipercubo latino; Metodi di copula di correlazione

A partire dalla versione del 2011/2012 sono presenti 6 Generatori di numeri casuali, 3 Copule di correlazione e 2 Metodi di campionamento della simulazione tra cui scegliere (Figura 5.41). Queste preferenze si impostano dalla posizione ***Simulatore di Rischio*** | ***Opzioni***.

Il Generatore di numeri casuali (Random Number Generator, RNG) è il cuore di tutti i software di simulazione. Basato sul numero casuale generato, è possibile costruire differenti distribuzioni matematiche. Il metodo di default è la metodologia proprietaria di ROV Simulatore di Rischio, che fornisce i numeri casuali migliori e più robusti. Sono supportati 6 generatori di numeri casuali; in generale, i due approcci consigliati sono il metodo di default di ROV Simulatore di Rischio e il metodo Mescolata casuale sottrattiva avanzata (Advanced Subtractive Random Shuffle). Non applicate gli altri metodi a meno che il loro utilizzo non sia esplicitamente richiesto dal vostro modello o dalla vostra analisi e anche allora vi consigliamo di testare i risultati contro questi due approcci consigliati. Tanto più in basso si scende nella lista dei Generatori di numeri casuali, tanto più semplice diventa l'algoritmo e tanto più velocemente viene eseguito. Per converso, tanto più in alto si sale nella lista dei Generatori di numeri casuali, tanto più robusti sono i risultati.

Nella sezione delle correlazioni sono supportati tre metodi: la Copula normale, la Copula a T e la Copula quasi normale. Questi metodi si basano su tecniche di integrazione matematica e, quando sussiste un dubbio, la copula normale fornisce i risultati più sicuri e prudenti. La copula a T fornisce valori estremi nelle code delle distribuzioni simulate, mentre la copula quasi normale fornisce risultati che si trovano fra questi valori.

Nella sezione sui metodi di simulazione sono supportati i metodi della Simulazione Monte Carlo (Monte Carlo Simulation, MCS) e del Campionamento Ipercubo Latino (Latin Hypercube Sampling, LHS). Si prega di notare che le copule e le altre funzioni multivariate ***non*** sono compatibili con il Campionamento Ipercubo Latino. La ragione è che il Campionamento Ipercubo Latino può essere applicato ad una singola variabile casuale ma non ad una distribuzione congiunta. In realtà, il Campionamento Ipercubo Latino ha un impatto molto limitato sull’accuratezza dell'output del modello, tanto più numerose sono le distribuzioni nel modello visto che il Campionamento Ipercubo Latino viene applicata alle distribuzioni solo in maniera individuale. Il beneficio del Campionamento Ipercubo Latino viene anche ridotto se non si completa il numero di campionamenti definito all'inizio, per esempio, se la simulazione è fermata a metà. Il Campionamento Ipercubo Latino applica anche un notevole peso sui modelli di simulazione con molti input, dato che ha bisogno di generare e organizzare i campioni da ciascuna delle distribuzioni prima di poter eseguire il primo campione da una distribuzione. Questo può determinare un lungo tempo di attesa se si esegue un modello di grandi dimensioni, ma fornisce poca accuratezza addizionale. Per finire, è meglio applicare il Campionamento Ipercubo Latino quando le distribuzioni esibiscono un buon comportamento, sono simmetriche e non hanno nessuna correlazione. Ciononostante, il Campionamento Ipercubo Latino è un metodo potente che fornisce una distribuzione con un campionamento uniforme, mentre la Simulazione Monte Carlo può talvolta generare delle distribuzioni “bitorzolute” (i dati campionati possono essere talvolta maggiormente concentrati in un’area della distribuzione) paragonato ad una distribuzione con un campionamento più uniforme (ogni parte della distribuzione viene campionata) quando si applica il Campionamento Ipercubo Latino.

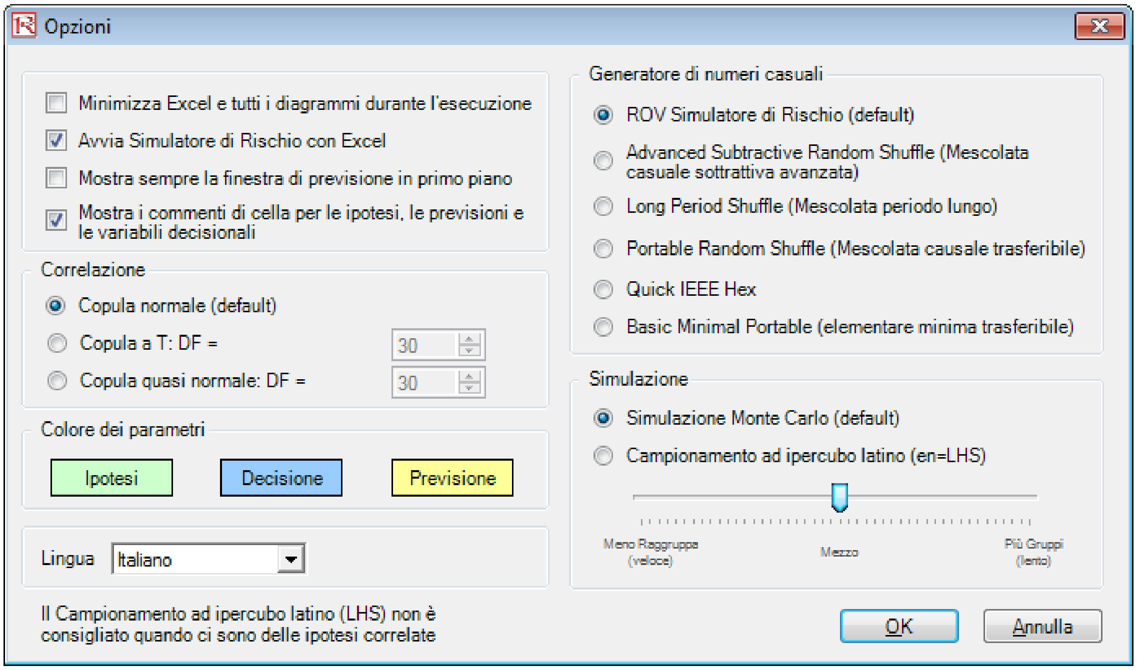


Figura 5.41 –Opzioni di Simulatore di Rischio

5.15 Destagionalizzazione dei dati, Detrending (rimozione del trend) dei dati e Test di stagionalità

Lo strumento per la destagionalizzazione dei dati ed il detrending di Simulatore di Rischio vi permette di eliminare tutti gli elementi stagionali o di trend nei vostri dati. Questo processo vi permette di visualizzare solamente i cambi assoluti dei valori da periodo a periodo, rendendo così possibile l'identificazione di schemi ciclici nei vostri dati di serie temporali. La destagionalizzazione ed il detrending eliminano tutte le derive generali, la tendenza, i twist, le deviazioni (bends) ed altri cicli stagionali che potrebbero influenzare i vostri dati di serie temporali, fornendo come risultato il vero comportamento strutturale dei dati nel tempo.

I periodi di stagionalità rappresentano quanti periodi devono passare prima che il ciclo si ripeti (p.e., 24 ore in un giorno, 12 mesi in un anno, 4 trimestri in un anno, 60 minuti in un ora, e così via). Tuttavia esistono talvolta altri periodi stagionali che non sono completamente evidenti semplicemente guardando i dati. Usando questa stagionalità, potete ora correggere per gli effetti stagionali usando lo Strumento Destagionalizzare come mostrato sopra per determinare la periodicità stagionale col miglior adattamento per i dati o usando lo Strumento Analisi di serie temporali per ottenere una previsione migliore.

Procedura di destagionalizzazione e di detrending:

- Selezionate i dati che desiderate analizzare (p.e., B9:B28) e cliccate su Simulatore di rischio | Strumenti | Dati Destagionalizzazione e Eliminazione tendenza dai dati (Detrend)
- ***Selezionate Destagionalizzare dati*** e/o ***Eliminare la tendenza dei dati***, selezionate il modello di detrending che volete eseguire, inserite gli ordini pertinenti (p.e., ordine polinomiale, ordine media mobile, ordine differenza, e ordine tasso) e cliccate su OK
- Esaminate i due report generati per altri dettagli sulla metodologia, sull'applicazione, sui diagrammi risultanti e sui dati destagionalizzati/corretti per tendenza.

Procedura per il Test di stagionalità:

- Selezionate i dati che desiderate analizzare (p.e., B9:B28) e cliccate su Simulatore di rischio | Strumenti | Dati Test di stagionalità
- Inserite il periodo massimo di stagionalità da testare. In altre parole, se inserite 6, il Simulatore di rischio testerà i seguenti periodi di stagionalità: 1, 2, 3, 4, 5, 6. Il Periodo 1 implica, naturalmente, che non esiste stagionalità nei dati.
- Esaminate il report generato per maggiori dettagli sulla metodologia, sull'applicazione, sui diagrammi risultanti e sui risultati del test di stagionalità. La periodicità stagionale migliore è elencata per prima (classificata secondo la misura d'errore RMSE più bassa) e tutte le pertinenti misure d'errore sono incluse per un confronto: radice dell'errore quadratico medio (RMSE), errore quadratico medio (MSE), deviazione assoluta media (MAD) ed errore percentuale assoluto medio (MAPE)

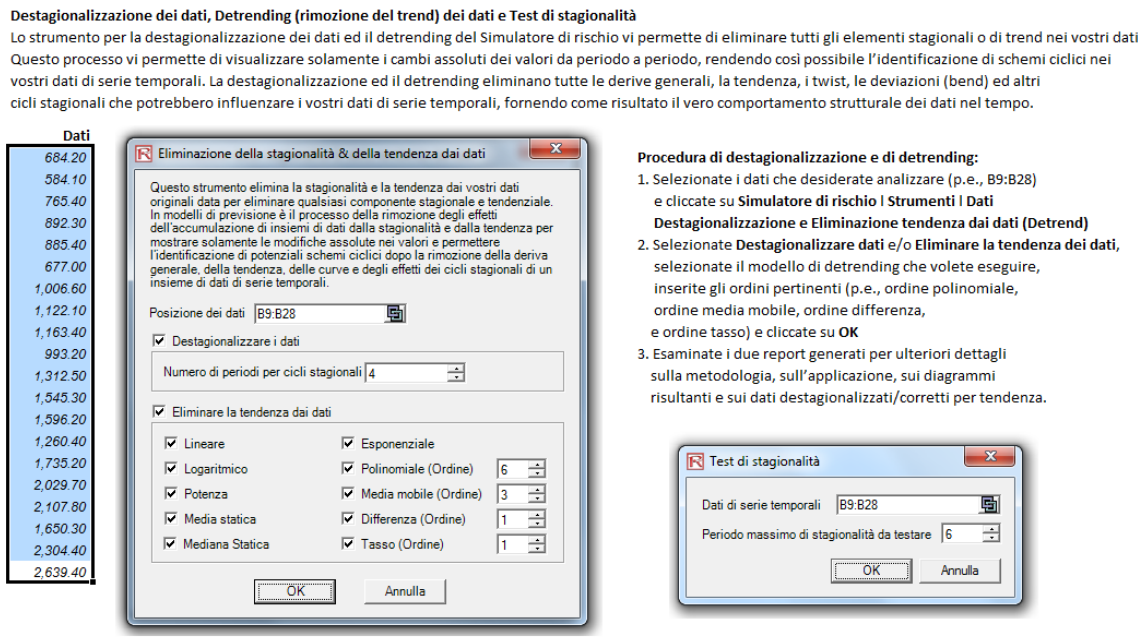


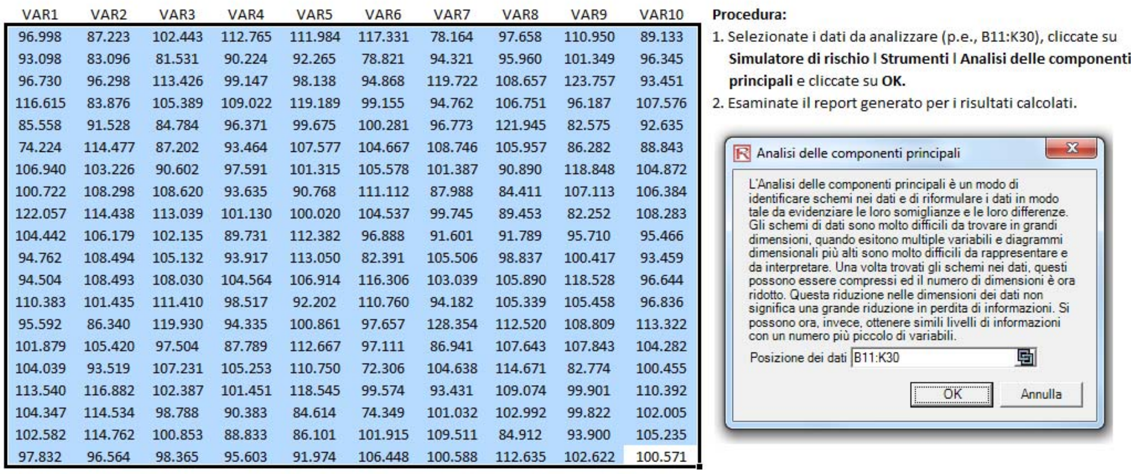
Figura 5.42 – Destagionalizzazione e Detrending dei dati

5.16 Analisi delle componenti principali

L'Analisi delle componenti principali è un modo di identificare schemi nei dati e di riformulare i dati in modo tale da evidenziare le loro somiglianze e le loro differenze. Gli schemi di dati sono molto difficili da trovare in grandi dimensioni, quando esistono multiple variabili e diagrammi dimensionali più alti sono molto difficili da rappresentare e da interpretare. Una volta trovati gli schemi nei dati, questi possono essere compressi ed il numero di dimensioni è ora ridotto. Questa riduzione nelle dimensioni dei dati non significa una grande riduzione nella perdita di informazioni. Si possono ora, invece, ottenere simili livelli di informazioni con un numero più piccolo di variabili.

Procedura:

- Selezionate i dati da analizzare (p.e., B11:K30), cliccate su Simulatore di rischio | Strumenti | Analisi delle componenti principali e cliccate su OK.
- Esaminate il report generato per i risultati calcolati.



VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8	VAR9	VAR10
96.998	87.223	102.443	112.765	111.984	117.331	78.164	97.658	110.950	89.133
93.098	83.096	81.531	90.224	92.265	78.821	94.321	95.960	101.349	96.345
96.730	96.298	113.426	99.147	98.138	94.868	119.722	108.657	123.757	93.451
116.615	83.876	105.389	109.022	119.189	99.155	94.762	106.751	96.187	107.576
85.558	91.528	84.784	96.371	99.675	100.281	96.773	121.945	82.575	92.635
74.224	114.477	87.202	93.464	107.577	104.667	108.746	105.957	86.282	88.843
106.940	103.226	90.602	97.591	101.315	105.578	101.387	90.890	118.848	104.872
100.722	108.298	108.620	93.635	90.768	111.112	87.988	84.411	107.113	106.384
122.057	114.438	113.039	101.130	100.020	104.537	99.745	89.453	82.252	108.283
104.442	106.179	102.135	89.731	112.382	96.888	91.601	91.789	95.710	95.466
94.762	108.494	105.132	93.917	113.050	82.391	105.506	98.837	100.417	93.459
94.504	108.493	108.030	104.564	106.914	116.306	103.039	105.890	118.528	96.644
110.383	101.435	111.410	98.517	92.202	110.760	94.182	105.339	105.458	96.836
95.592	86.340	119.930	94.335	100.861	97.657	128.354	112.520	108.809	113.322
101.879	105.420	97.504	87.789	112.667	97.111	86.941	107.643	107.843	104.282
104.039	93.519	107.231	105.253	110.750	72.306	104.638	114.671	82.774	100.455
113.540	116.882	102.387	101.451	118.545	99.574	93.431	109.074	99.901	110.392
104.347	114.534	98.788	90.383	84.614	74.349	101.032	102.992	99.822	102.005
102.582	114.762	100.853	88.833	86.101	101.915	109.511	84.912	93.900	105.235
97.832	96.564	98.365	95.603	91.974	106.448	100.588	112.635	102.622	100.571

Figure 5.43 –Analisi delle componenti principali

5.17 Rottura Strutturale

Una rottura strutturale analizza se i coefficienti di insiemi di dati differenti sono uguali. Questo test è più comunemente usato nell'analisi di serie temporali per verificare la presenza di una rottura strutturale. Un insieme di dati di serie temporali può essere diviso in due sottoinsiemi e ciascun sottoinsieme è testato l'uno contro l'altro e contro l'insieme completo di dati per determinare statisticamente se c'è difatti una rottura che inizia in un periodo temporale particolare. Il test di rottura strutturale è spesso usato per determinare se le variabili indipendenti hanno un impatto diverso su differenti sottoinsiemi della popolazione, come il testare se una nuova campagna di marketing, un'attività, un evento importante, un'acquisizione, un disinvestimento, e così via, hanno un impatto sui dati di serie temporali. Supponiamo che l'insieme di dati ha 100 punti dati di serie temporali. Potete impostare vari punti di interruzione da testare, per esempio, i punti dati 10, 30 e 51 (questo significa che saranno eseguiti tre tests di rottura strutturale sui seguenti insiemi di dati, i punti dati 1-9 rispetto ai punti dati 10-100; i punti dati 1-29 rispetto ai punti dati 30-100; i punti dati 1-50 rispetto ai punti dati 51-100, per vedere se c'è difatti una rottura nella struttura sottostante all'inizio dei punti dati 10, 30 e 51).

Procedura:

- Selezionate i dati da analizzare (p.e., B15:D34), cliccate su ***Simulatore di rischio | Strumenti | Test di rottura strutturale***, inserite i punti di test pertinenti che desiderate applicare ai dati (p.e., 6, 10, 12) e cliccate su OK.
- Esaminate il report per determinare quali di questi punti di test indicano un punto di rottura statisticamente significativo nei vostri dati e quali punti non lo fanno.

Rottura Strutturale

Una rottura strutturale analizza se i coefficienti di insiemi di dati differenti sono uguali. Questo test è più comunemente usato nell'analisi di serie temporali per verificare la presenza di una rottura strutturale. Un insieme di dati di serie temporali può essere diviso in due sottoinsiemi e ciascun sottoinsieme è testato l'uno contro l'altro e contro l'insieme completo di dati per determinare statisticamente se c'è difatti una rottura che inizia in un periodo temporale particolare. Il test di rottura strutturale è spesso usato per determinare se le variabili indipendenti hanno un impatto diverso su differenti sottoinsiemi della popolazione, come il testare se una nuova campagna di marketing, un'attività, un evento importante, un'acquisizione, un disinvestimento, e così via, hanno un impatto sui dati di serie temporali. Supponiamo che l'insieme di dati ha 100 punti dati di serie temporali. Potete impostare vari punti di interruzione da testare, per esempio, i punti dati 10, 30 e 51 (questo significa che saranno eseguiti tre tests di rottura strutturale sui seguenti insiemi di dati, i punti dati 1-9 rispetto ai punti dati 10-100; i punti dati 1-29 rispetto ai punti dati 30-100; i punti dati 1-50 rispetto ai punti dati 51-100, per vedere se c'è difatti una rottura nella struttura sottostante all'inizio dei punti dati 10, 30 e 51).

Y	X1	X2
521	18308	185
367	1148	600
443	18068	372
365	7729	142
614	100484	432
385	16728	290
286	14630	346
397	4008	328
764	38927	354
427	22322	266
153	3711	320
231	3136	197
524	50508	266
328	28886	173
240	16996	190
286	13035	239
285	12973	190
569	16309	241
96	5227	189
498	19235	358

Procedura:

1. Selezionate i dati da analizzare (p.e., B15:D34), cliccate su **Simulatore di rischio I Strumenti I Test di rottura strutturale**, inserite i punti di test pertinenti che desiderate applicare ai dati (p.e., 6, 10, 12) e cliccate su **OK**.
2. Esaminate il report per determinare quali di questi punti di test indicano un punto di rottura statisticamente significativo nei vostri dati e quali punti non lo fanno.

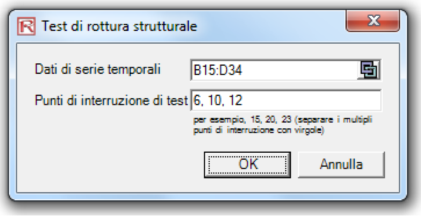


Figura 5.44 –Analisi di rottura strutturale

5.18 Previsioni linea di tendenza

Le linee di tendenza (trendlines) possono essere usate per determinare se un insieme di dati di serie temporali segue una qualsiasi apprezzabile tendenza (Figura 5.45). Le tendenze possono essere lineari o non lineari (come per esempio esponenziali, logaritmici, media mobile, potenza, polinomiali o potenza).

Procedura:

- Selezionate i dati da analizzare e cliccate su ***Simulatore di Rischio | Previsione | Linea di tendenza***, selezionate le attinenti linee di tendenza da applicare ai dati (p. es., selezionate tutti i metodi come default), inserite il numero di periodi da prevedere (p. es., 6 periodi) e cliccate su *OK*.
- Esaminate il report per determinare quali di questi test delle linee di tendenza forniscono il miglior adattamento e la migliore previsione per i vostri dati.

Ricavi storici da vendite

Anno	Trimestre	Periodo	Vendite
2006	1	1	$684.20
2006	2	2	$584.10
2006	3	3	$765.40
2006	4	4	$892.30
2007	1	5	$885.40
2007	2	6	$677.00
2007	3	7	$1,006.60
2007	4	8	$1,122.10
2008	1	9	$1,163.40
2008	2	10	$993.20
2008	3	11	$1,312.50
2008	4	12	$1,545.30
2009	1	13	$1,596.20
2009	2	14	$1,260.40
2009	3	15	$1,735.20
2009	4	16	$2,029.70
2010	1	17	$2,107.80
2010	2	18	$1,650.30
2010	3	19	$2,304.40
2010	4	20	$2,639.40

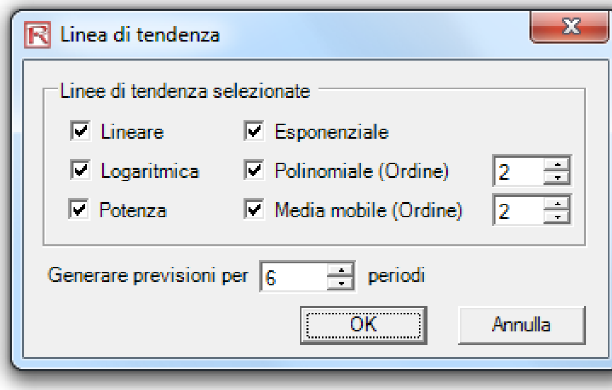


Figura 5.45 – Previsioni Linee di tendenza

5.19 Strumento di verifica del modello

Dopo aver creato un modello e aver impostato delle ipotesi e delle previsioni, potete eseguire la simulazione come al solito o eseguire lo Strumento di verifica del modello (Figura 5.46) per verificare se il modello è stato impostato correttamente. Alternativamente, se il modello non viene eseguito e sospettate che alcune impostazioni potrebbero essere sbagliate, potete eseguire questo strumento da ***Risk Simulatore di Rischio | Strumenti | Verifica modello*** per identificare i punti dove potrebbero esservi dei problemi con il vostro modello. Si prega di notare che questo strumento esegue una verifica solo per i problemi più comuni nei modelli come anche per i problemi nelle ipotesi e nelle simulazioni di Simulatore di Rischio e che questa verifica non è assolutamente abbastanza completa da poter testare per tutti i tipi di problemi…rimane sempre il compito dello sviluppatore del modello accertarsi che il modello funzioni correttamente.

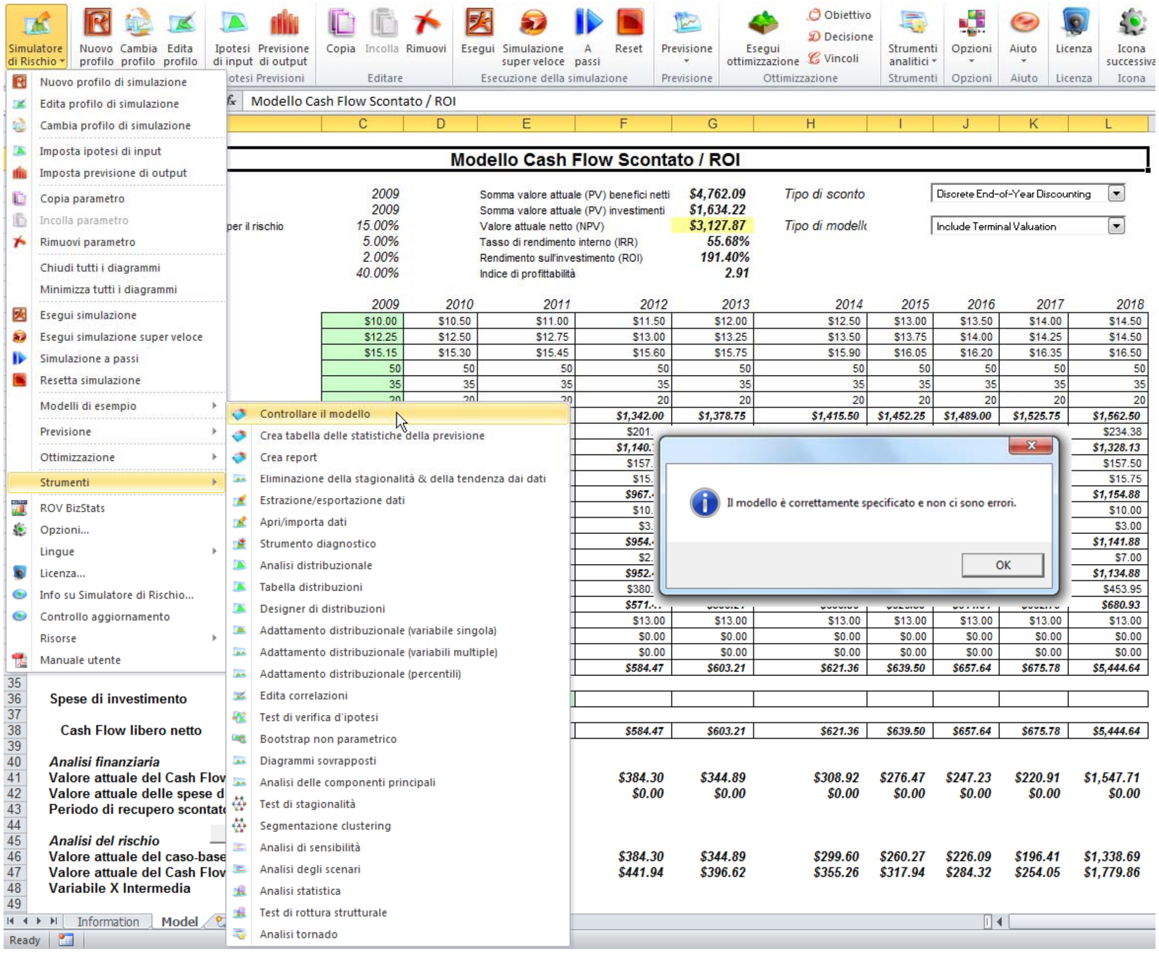


Figura 5.46 – Strumento di verifica del modello

5.20 Strumento di adattamento distribuzionale percentile

Lo Strumento di adattamento distribuzionale percentile (Figure 5.47) è un altro metodo alternativo di adattamento delle distribuzioni di probabilità. Ci sono altri strumenti collegati e ciascuno ha i suoi utilizzi e i suoi vantaggi:

- Adattamento distribuzionale (percentili) — usando il metodo alternativo di inserimento (percentili e combinazioni di primo/secondo momento), potete identificare i parametri col miglior adattamento di una distribuzione specificata senza la necessità di avere dei dati grezzi. L'uso di questo metodo è appropriato se non ci sono dati a sufficienza, se ci sono solo percentili e momenti a disposizione, o come mezzo per recuperare l'intera distribuzione con solo due o tre punti dati e si richiede di poter ipotizzare o conoscere il tipo di distribuzione.

- Adattamento distribuzionale (singola variabile) — vengono usati metodi statistici per adattare i vostri dati grezzi a tutte e 42 le distribuzioni per trovare la distribuzione col miglior adattamento e i suoi parametri di input. Sono necessari multipli punti dati per avere un buon adattamento; il tipo di distribuzione può essere noto o ignoto in anticipo.
- Adattamento distribuzionale (multiple variabili) — vengono usati metodi statistici per adattare i vostri dati grezzi a multiple variabili contemporaneamente, usando gli stessi algoritmi come per l'adattamento della singola variabile ma incorporando una matrice di correlazione a coppie tra le variabili. Sono necessari multipli punti dati per avere un buon adattamento; il tipo di distribuzione può essere noto o ignoto in anticipo.
- Distribuzione personalizzata (imposta ipotesi) — vengono usate tecniche non parametriche di ricampionamento per generare una distribuzione personalizzata con i dati grezzi esistenti e per simulare la distribuzione basato su questa distribuzione empirica. Sono necessari meno punti dati; il tipo di distribuzione non è noto in anticipo.

Procedura:

- Cliccate su ***Simulatore di Rischio | Strumenti | Adattamento distribuzionale (percentili)***, scegliete la distribuzione di probabilità e i tipi di input che desiderate usare, inserite i parametri e cliccate su ***Esegui*** per ottenere i risultati. Esaminate i risultati del R-quadrato adattato e confrontate i risultati dell'adattamento empirico con quelli dell'adattamento teoretico per determinare se la vostra distribuzione ha un buon adattamento.

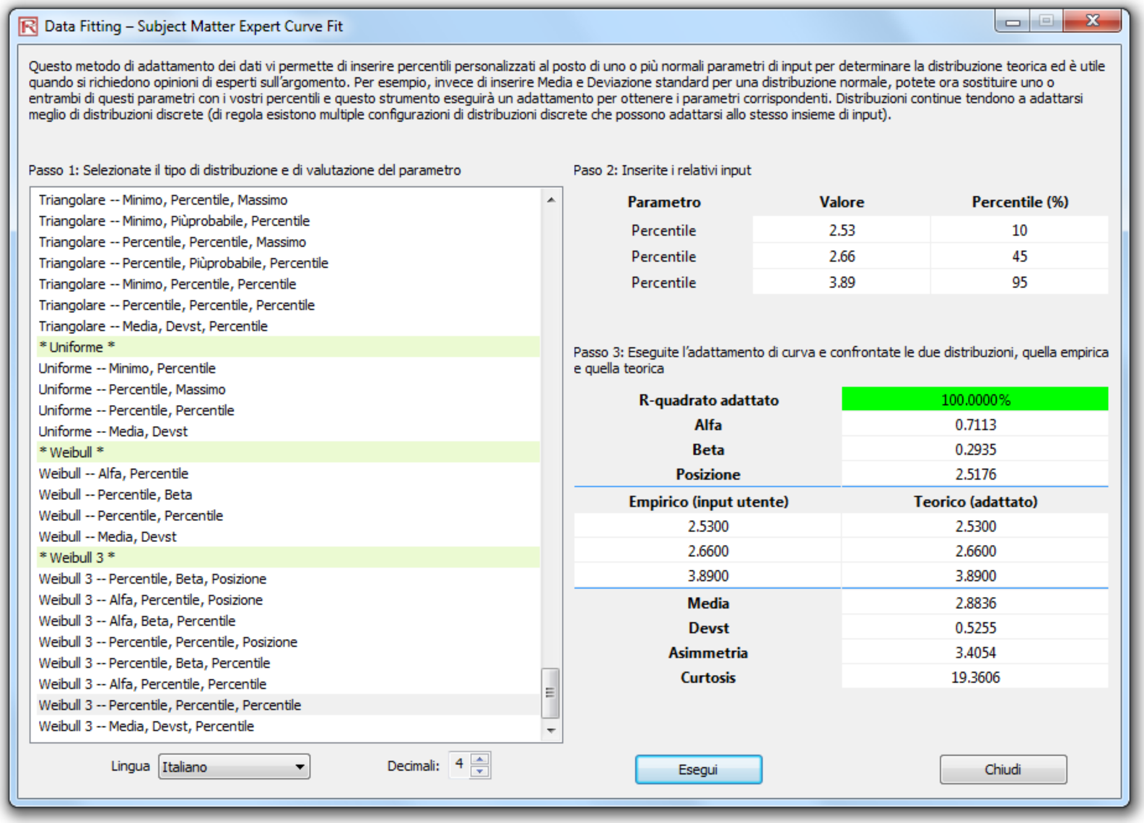


Figura 5.47 – Strumento di adattamento distribuzionale percentile

5.21 Diagrammi e tabelle di distribuzioni: Strumento di distribuzione di probabilità

Questo nuovo strumento di distribuzione della probabilità è un modulo molto potente e veloce usato per generare i diagrammi e le tabelle di distribuzioni (Figure 5.48-5.51). Si prega di notare che esistono tre strumenti simili in Simulatore di Rischio, ma che ciascuno di essi esegue operazioni molte diverse:

- Analisi distribuzionale — usata per calcolare velocemente PDF, CDF e ICDF delle 42 distribuzioni di probabilità disponibili in Simulatore di Rischio e per fornire una *tabella di probabilità* di questi valori.

- Diagrammi e tabelle distribuzionali — questo è lo strumento di distribuzioni di probabilità descritto qui. Viene usato per paragonare *differenti parametri della stessa distribuzione* (p. es., le forme e i valori PDF, CDF e ICDF di una distribuzione Weibull con Alfa e Beta di [2, 2], [3, 5] e [3.5, 8] e li sovrappone uno sull'altro).

- Diagrammi di sovrapposizione — usati per paragonare *distribuzioni differenti* (ipotesi di input teoretici e previsioni di output simulate empiricamente) e li sovrappone uno sull'altro per un confronto visuale.

Procedura:

- Eseguite ROV BizStats da ***Simulatore di Rischio | Diagrammi e tabelle distribuzionali***, cliccate sul bottone ***Applicare input globali*** per caricare un insieme campione di parametri di input o inserite i vostri input e cliccate su ***Esegui*** per calcolare i risultati. I risultanti quattro momenti e i valori CDF, ICDF, PDF sono calcolati per ciascuna delle 45 distribuzioni di probabilità (Figura 5.48).

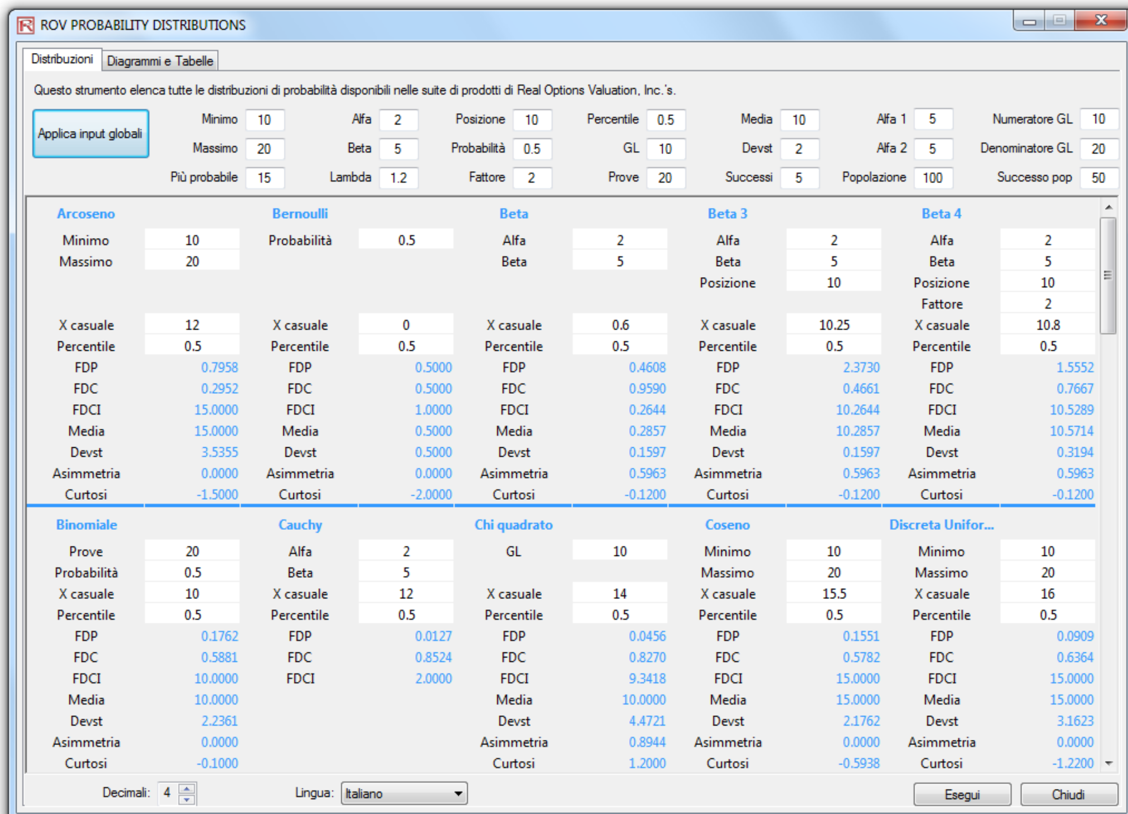


Figura 5.48 – Strumento di distribuzione di probabilità (45 distribuzioni di probabilità)

- Cliccate sulla linguetta ***Diagrammi e tabelle*** (Figura 5.49), selezionate una distribuzione [**A**] (p. es., Arcoseno), scegliete se desiderate eseguire CDF, ICDF o PDF [**B**], inserite gli input attinenti e cliccate su ***Esegui diagramma*** o ***Esegui tabella*** [**C**]. Potete alternare tra le linguette

Diagramma e Tabella per visionare i risultati e anche provare alcune delle icone del diagramma [**E**] per vedere gli effetti sul diagramma.

- Potete anche modificare i due parametri [**H**] per generare multipli diagrammi e multiple tabelle di distribuzione inserendo l'input ***Da/A/Passo*** o usando gli input ***Personalizzati*** e premendo su ***Esegui***. Per esempio, come illustrato nella Figura 5.50, eseguite la distribuzione Beta e selezionate PDF [**G**], selezionate Alfa e Beta da modificare [**H**] usando gli input Personalizzati [**I**], inserite i relativi parametri di input (2;5;5 per Alfa e 5;3;5 per Beta) [**J**] e cliccate su ***Esegui diagramma***. Questo genererà tre distribuzioni Beta [**K**]: Beta (2,5), Beta (5,3) e Beta (5,5) [**L**]. Esaminate i vari tipi di diagrammi, le linee della griglia, le impostazioni della lingua e dei decimali [**M**] e provate a rieseguire la distribuzione usando i valori teoretici in confronto ai valori simulati empiricamente [**N**].

- La Figura 5.51 illustra le tabelle di probabilità generate per una distribuzione binomiale dove la probabilità di successo e il numero di prove di successo (variabile casuale X) sono stati selezionati come elementi da modificare [**O**] usando l'opzione ***Da/A/Passo***. Provate a riprodurre il calcolo come mostrato e cliccate sulla linguetta tabella [**P**] per visionare i risultati della funzione di densità di probabilità che è stata creata. In questo esempio, la distribuzione binomiale ha un insieme di input iniziale di Prove = 20, Probabilità di successo = 0.5 e Numero di prove di successo X = 10, dove alla Probabilità di successo è permesso di cambiare da 0., 0,25, …, 0,50 ed è mostrata come la variabile di riga, e al Numero di prove di successo e altresì permesso di cambiare da 0, 1, 2, …, 8, ed è mostrata come la variabile di colonna. PDF è scelta e quindi i risultati nella tabella mostrano la probabilità che gli eventi specificati accadano. Per esempio, la probabilità di avere esattamente 2 successi quando vengono eseguite 20 prove e dove ciascuna prova ha una possibilità di successo del 25% è di 0.0669 o una probabilità del 6.69%.

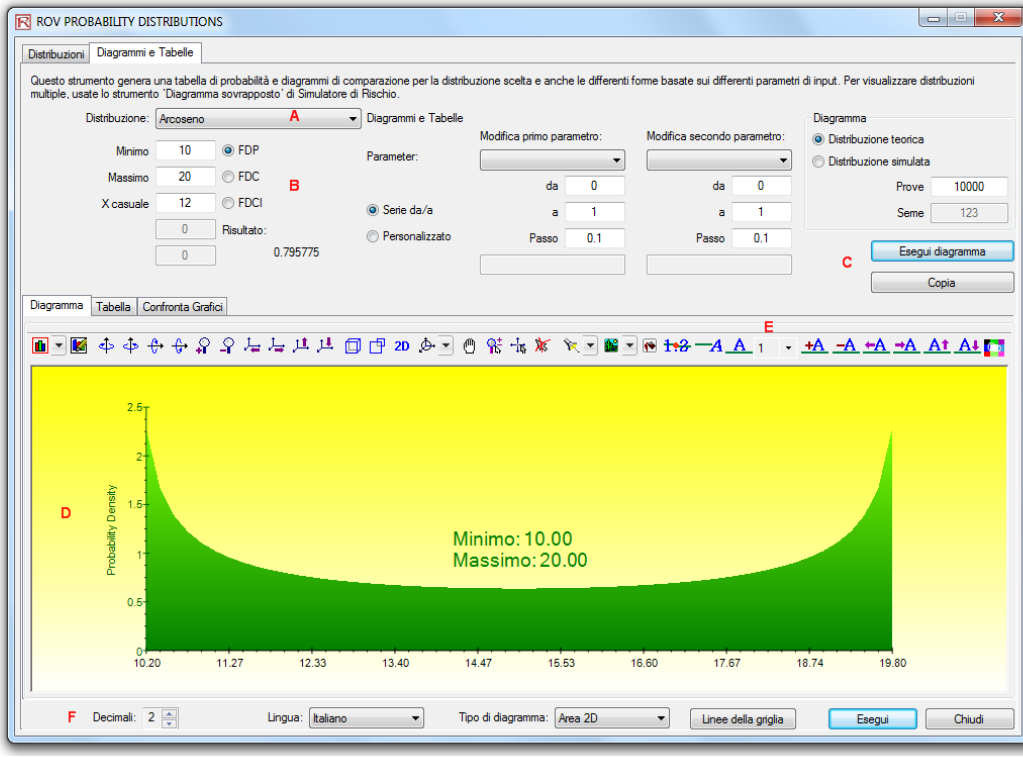


Figura 5.49 – ROV Distribuzione di probabilità (Diagrammi PDF e CDF)

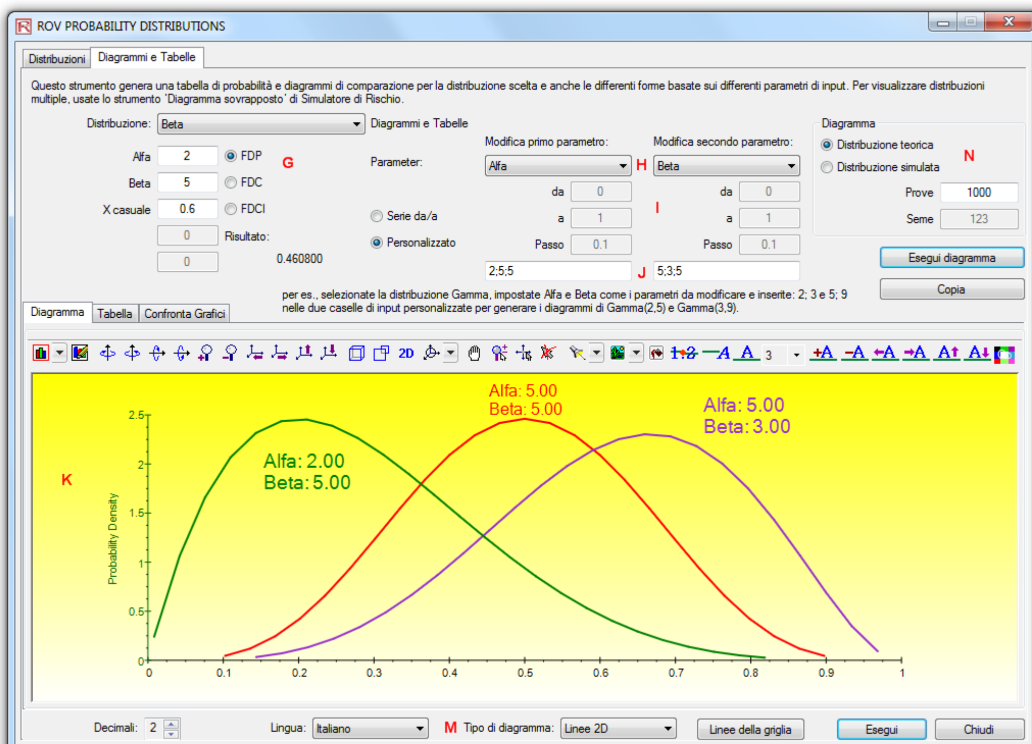


Figura 5.50 – ROV Distribuzione di probabilità (Multipli diagrammi sovrapposti)

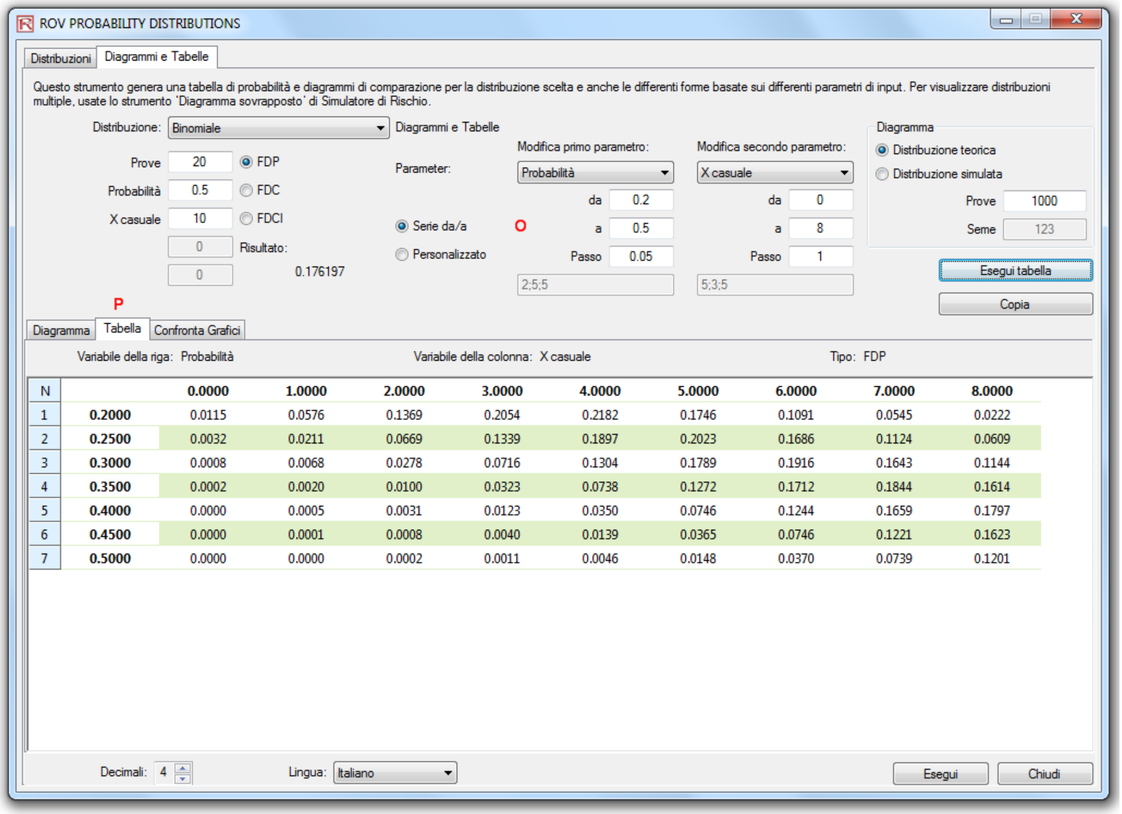


Figura 5.51 – ROV Distribuzione di probabilità (Tabelle di distribuzione)

5.22 ROV BizStats

Questo nuovo strumento ROV BizStats è un modulo molto potente e veloce in Simulatore di Rischio che viene usato per eseguire statistiche di business e modelli analitici sui vostri dati data, e che comprende più di 130 statistiche di business e modelli analitici (Figure 5.52-5.55). Di seguito sono forniti alcuni veloci suggerimenti su come iniziare ad eseguire il modulo e dettagli su ciascuno degli elementi nel software.

Procedura:

- Eseguite ROV BizStats da ***Risk Simulator | ROV BizStats***, cliccate su ***Esempio*** per caricare un esempio di profilo di dati e di modello [**A**] o inserite i vostri dati o copiate/incollate nella griglia dei dati [**D**] (Figura 5.52). Potete aggiungere le vostre note o i nomi delle variabili nella prima riga per le Note [**C**].
- Selezionate il relativo modello [**F**] da eseguire nel Passo 2 e, usando le impostazioni d'esempio per l'input dei dati [**G**], inserite le relative variabili [**H**]. Separate le variabili per lo stesso parametro usando i punti e virgole e usate una nuova linea (premete Enter per creare una nuova linea) per parametri differenti.
- Cliccate su ***Esegui*** [**I**] per calcolare i risultati [**J**]. Potete visionare tutti i relativi risultati analitici, i diagrammi o le statistiche tramite le varie linguette nel Passo 3.
- Se necessario, potete dare un nome al modello da salvare nel profilo nel Passo 4 [**L**]. Multipli modelli possono essere salvati nello stesso profilo. Modelli esistenti possono essere modificati o cancellati [**M**], l'ordine di comparsa può essere riassegnato [**N**] e tutte le modifiche possono essere salvate [**O**] in un singolo profilo in un nome di file con l'estensione *.bizstats.

Note:

- La dimensione della griglia dei dati può essere impostata nel menu, dove la griglia può contenere fina a 1000 colonne di variabili con 1 milione di righe di dati per variabile. Il menu vi permette anche di cambiare le impostazioni della lingua e le impostazioni dei decimali per i vostri dati.
- Per iniziare, è sempre una buona idea caricare il file d'esempio [**A**] che viene fornito completo di alcuni dati e modelli creati in anticipo [**S**]. Potete fare un doppio click su qualsiasi di questi modelli per eseguirli e i risultati vengono mostrati nell'area del report [**J**], che può talvolta essere un diagramma o delle statistiche del modello [**T/U**]. Usando questo file

d'esempio, potete ora vedere come i parametri di [input **H**] sono inseriti basati sulla descrizione del modello [**G**] e potete ora creare i vostri modelli personalizzati.

- Cliccate sull'intestazione delle variabili [**D**] per selezionare una o più variabili alla volta, e poi cliccate col tasto destro per aggiungere, cancellare, copiare, incollare o visualizzare [**P**] le variabili selezionate.
- I modelli possono anche essere inseriti usando una Consolle di commando [**V/W/X**]. Per vedere come questa funziona, fate un doppio clic per eseguire un modello [**S**] e andate alla consolle di commando [**V**]. Potete replicare il modello o creare un vostro modello e cliccare su Esegui commando [**X**] quando siete pronti. Ciascuna riga nella Consolle rappresenta un modello e i suoi relativi parametri.
- L'intero profilo **.bizstats* (dove dati e multipli modelli sono creati e salvati) può essere modificato direttamente in XML [**Z**] aprendo l'editore XML dal menu File. Modifiche al profilo possono essere fate qui in maniera programmatica e saranno rese effettive una volta che il profilo è stato salvato.

Suggerimenti

- Potete cliccare sull'intestazione (sulle intestazioni) della colonna della griglia dei dati per selezionare l'intera colonna o variabile (le intere colonne o variabili). Una volta fatta la selezione, potete cliccare sull'intestazione per *Auto adattare* la colonna, *Tagliare*, *Copiare*, *Cancellare* o *Incollare* i dati. Potete anche cliccare su e selezionare multiple intestazioni di colonne per selezionare multiple variabili e dopo cliccare col tasto destro e selezionare *Visualizzare* per diagrammare i dati.
- Se una cella contiene un valore grande che non è completamente visualizzato, cliccate su e tenete il puntatore del mouse posizionato sopra la cella; vedrete apparire un commento popup che indica l'intero valore. In alternativa, potete ridimensionare la colonna della variabile (trascinate la colonna per allargarla, fate doppio clic sul bordo della colonna per auto-adattare la colonna o fate doppio clic sull'intestazione e selezionate auto-adattare).
- Usate i tasti su, giù, destra, sinistra per muovervi all'interno della griglia o usate i tasti *Home* e *Fine* sulla tastiera per andare all'estremo sinistro e all'estremo destro di una riga. Potete anche usare una combinazione di tasti come *Ctrl+Home* per saltare alla cella in alto a sinistra, *Ctrl+Fine* per saltare alla cella in basso a destra, *Maiusc+Su/Giù* per selezionare una specifica area e così via.

- Potete inserire delle brevi note nella riga Note. Ricordatevi di mantenere le vostre note brevi e semplici.
- Testate le varie icone dei diagrammi nella linguetta Visualizza per cambiare l'aspetto dei diagrammi (p. es. ruotare, spostare, zoomare, cambiare colori, aggiungere una legenda e così via).
- Il pulsante *Copia* è usato per copiare le linguette *Risultati*, *Diagrammi* e *Statistiche* nel Passo 3 dopo l'esecuzione di un modello. Se non viene eseguito un modello, la funzione Copia copierà solo una pagina vuota.
- Il pulsante *Report* funzionerà solo se sono salvati dei modelli nel *Passo 4* o se ci sono dei dati nella griglia, altrimenti il report generato sarà vuoto. Microsoft Excel deve essere installato per eseguire i report dell'estrazione dei dati e dei risultati; Microsoft PowerPoint deve essere disponibile per eseguire i report dei diagrammi.
- Se avete dei dubbi su come eseguire un determinato modello o metodo statistico, avviate il profilo *Esempio* ed esaminate come si impostano i dati nel *Passo 1* o come si inseriscono i parametri di input nel *Passo 2*. Potete usare questi esempi come guide per iniziare e come modelli per i vostri dati e modelli.
- La lingua può essere cambiata nel menu *Lingue*. Si prega di notare che al momento sono disponibili 10 lingue nel software; altre saranno aggiunte. Talvolta, tuttavia, alcuni risultati limitati saranno visualizzati in inglese.
- Potete modificare la visualizzazione della lista dei vostri modelli nel *Passo 2* mediante la modifica della lista a tendina *Visualizza*. Potete elencare i modelli in ordine alfabetico, per categoria e secondo i requisiti per l'input dei dati — si prega di notare che in taluni linguaggi Unicode (p. es. cinese, giapponese e coreano), non esistono disposizioni alfabetiche e, quindi, la prima opzione non sarà disponibile.
- Il software è capace di gestire differenti regionali impostazioni decimali e numeriche (p. es. la cifra di mille dollari e cinquanta centesimi può essere scritta come 1,000.50 o 1.000,50 o 1'000,50 e così via). Le impostazioni decimali possono essere impostate nel menu ***Dati | Impostazioni decimali*** di ROV BizStats**.** Tuttavia, se avete un dubbio, si prega di cambiare le impostazioni regionali del computer a English USA e di mantenere il North America di default 1,000.50 in ROV BizStats (il funzionamento di questa impostazione è garantito sia con ROV BizStats che con gli esempi di default).

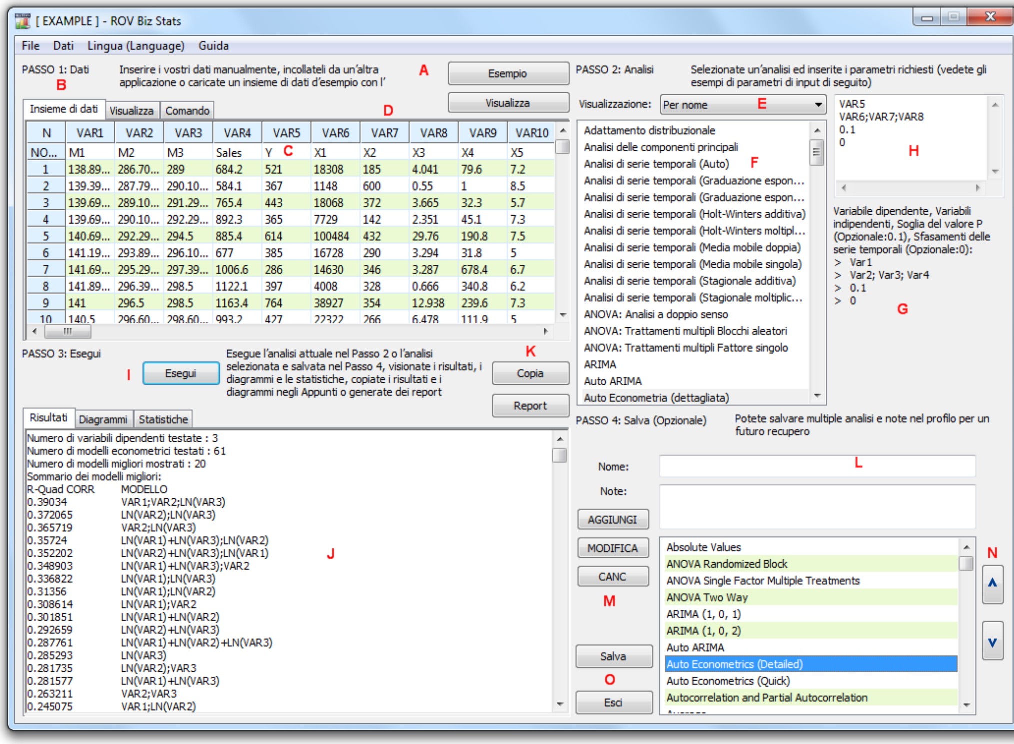


Figura 5.52 – ROV BizStats (Analisi statistica)

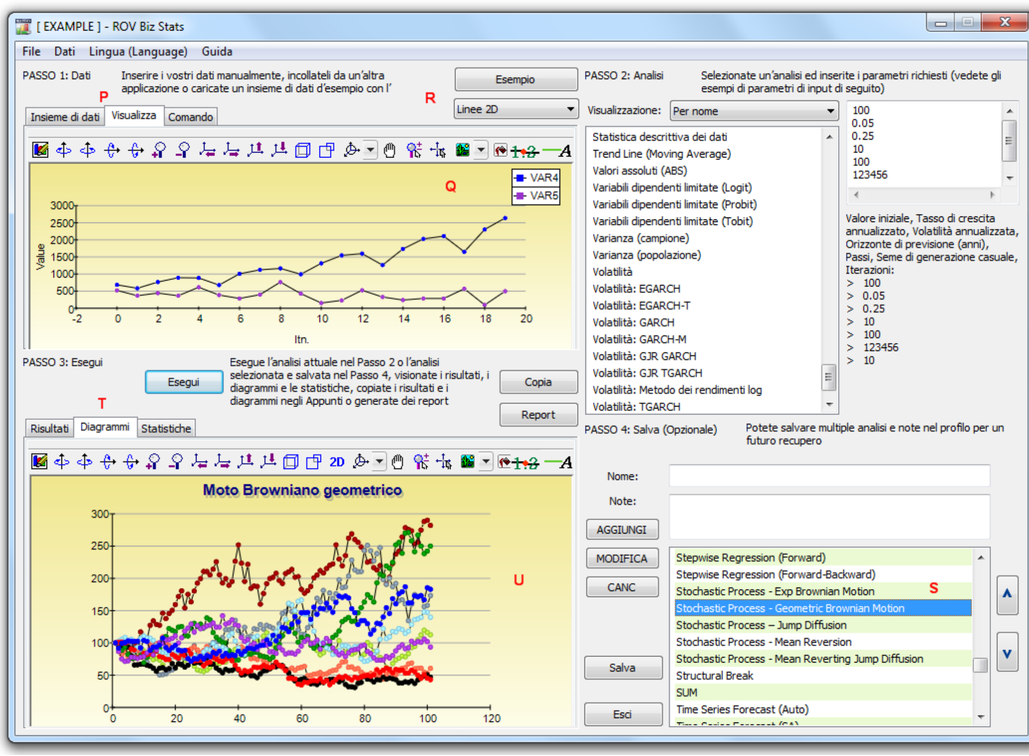


Figura 5.53 – ROV BizStats (Visualizzazione dei dati e Diagrammi dei risultati)

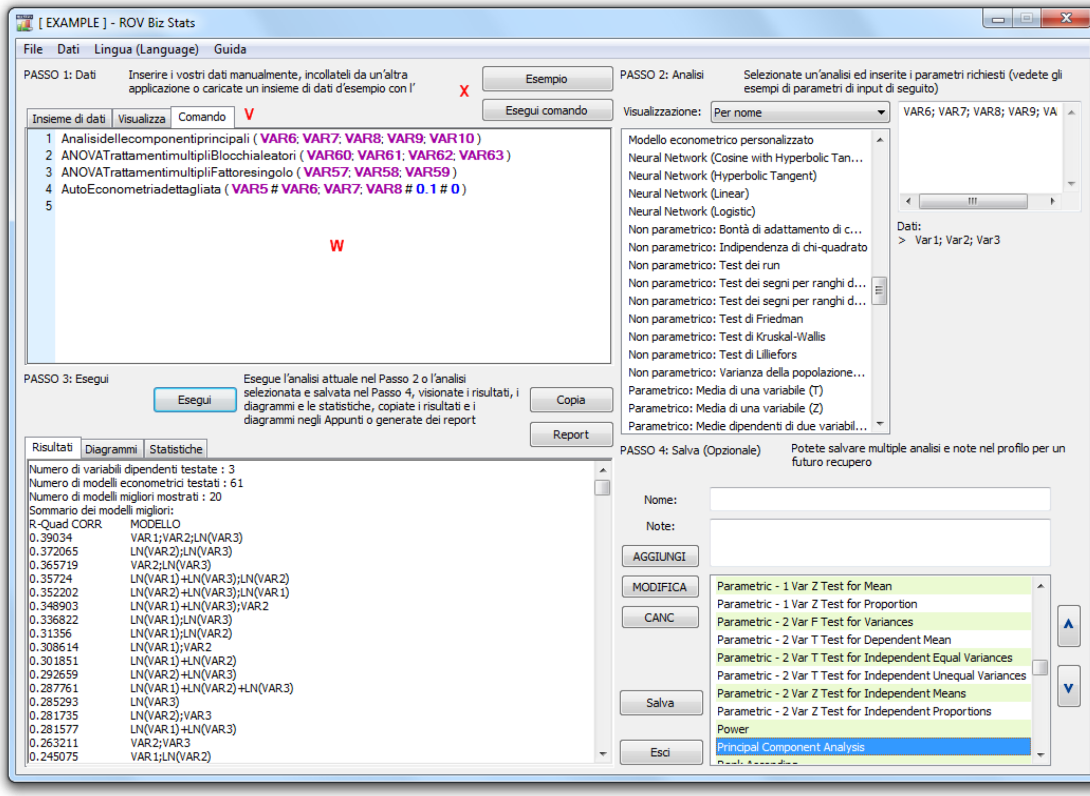


Figura 5.54 – ROV BizStats (Consolle di commando)

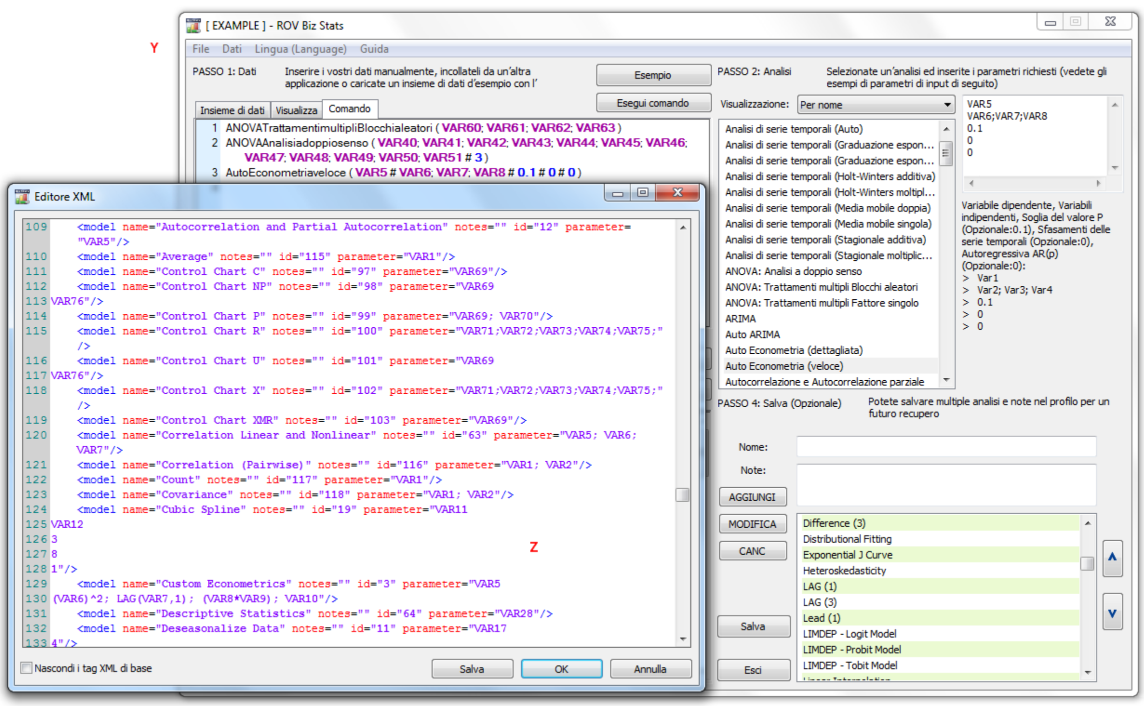


Figura 5.55 – ROV BizStats (Editore XML)

5.23 Metodologie di previsione - Rete neurale e Logica fuzzy combinatoria

Il termine Rete neurale è spesso usato per indicare una rete o un circuito of neuroni biologici, mentre l'uso moderno del termine si riferisce a reti neurali artificiali che consistono di neuroni o nodi artificiali ricreati in un ambiente software. Questa metodologia tenta di imitare il cervello o i neuroni umani per quanto riguarda i modi di pesare e di identificare gli schemi, e nel nostro caso, di identificare gli schemi con lo scopo di eseguire previsioni di dati di serie temporali. Questa metodologia si trova all'interno del modulo **ROV BizStats** in Simulatore di Rischio, sotto **Simulatore di Rischio | ROV BizStats | Rete neurale** e anche sotto **Simulatore di Rischio | Previsione | Rete neurale.** La Figura 5.56 mostra la metodologia di previsione di tipo Rete neurale.

Procedura

- Cliccate su Simulatore di Rischio | Previsione | Rete neurale.
- Avviate l'operazione o inserendo manualmente dei dati o incollando dei dati dagli Appunti (p. es. selezionate e copiate dei dati da Excel, avviate questo strumento ed incollate i dati cliccando sul bottone *Incolla*).
- Selezionate se desiderate eseguire un modello di Rete neurale *lineare* o *non lineare*, inserito il numero di *Periodi di previsione* che desiderate (p. es. 5), il numero di *Livelli* nascosti nella Rete neurale (p. es. 3) e il numero di *Periodi di test* (p. es. 5).
- Cliccate su *Esegui* per eseguire l'analisi ed esaminate i risultati e i diagrammi che sono stati calcolati. Potete anche *Copiare* i risultati e il diagramma negli Appunti e incollarli in un'altra applicazione software.

Si prega di notare che il numero di livelli nascosti nella Rete neurale è un parametro di input e dovrà essere tarato ai vostri dati. In generale, tanto più complicato è lo schema dei dati, tanto più alto è il numero di livelli nascosti e tanto più lungo è il tempo richiesto per il calcolo. Consigliamo di iniziare con tre livelli. I periodi di test sono semplicemente il numero di punti dati usati nella taratura finale del modello di Rete neurale e consigliamo di usare almeno lo stesso numero per i periodi di test di quello dei periodi che desiderate prevedere.

Per contrasto, il termine logica fuzzy deriva dalla teoria degli insiemi fuzzy per affrontare ragionamenti che sono approssimativi anziché accurati — contrapposta alla logica "crisp" (incisiva), dove insiemi binari hanno logica binaria, le variabili della logica fuzzy possono avere un valore di verità che varia tra 0 e 1 e non è vincolato ai due valori di verità della classica logica preposizionale. Questo schema di ponderazione fuzzy è usato insieme con un metodo combinatorio per fornire risultati di previsioni di serie temporali in Simulatore di Rischio come illustrato nella Figura 5.57 ed è applicabile al meglio, quando si applica a dati di serie temporali che hanno stagionalità e tendenza. Questa metodologia si trova all'interno del modulo ROV BizStats in Simulatore di Rischio, sotto **Simulatore di Rischio | ROV BizStats | Logica fuzzy combinatoria** e anche sotto **Simulatore di Rischio | Previsione | Logica fuzzy combinatoria.** La Figura 5.57 mostra la metodologia di previsione di tipo Rete neurale.

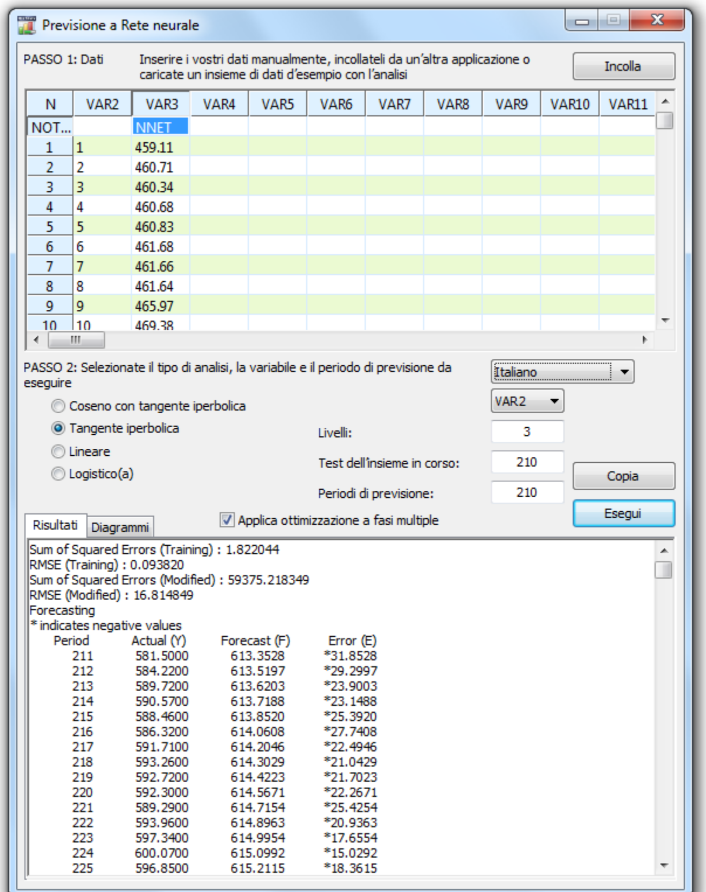


Figura 5.56 – Previsione Rete neurale

Procedura

- Cliccate su Simulatore di Rischio | Previsione | Logica fuzzy combinatoria.
- Avviate l'operazione o inserendo manualmente dei dati o incollando dei dati dagli Appunti (p. es. selezionate e copiate dei dati da Excel, avviate questo strumento ed incollate i dati cliccando sul bottone *Incolla*).
- Selezionate la variabile sulla quale desiderate eseguire l'analisi dalla lista a tendina, inserite il periodo di stagionalità (p. es. 4 per dati trimestrali, 12

per dati mensili e così via) e il numero desiderato di *Periodi di previsione* (p. es. 5).

- Cliccate su *Esegui* per eseguire l'analisi ed esaminate i risultati e i diagrammi che sono stati calcolati. Potete anche *Copiare* i risultati e il diagramma negli Appunti e incollarli in un'altra applicazione software.

Si prega di notare che né le reti neurali né le tecniche di logica fuzzy sono ancora state accettate come metodi validi ed affidabili nel campo della previsione economico-finanziaria, sia sul livello strategico, che tattico che operativo. Una grande mole di ricerca continuativa è ancora necessaria in questo campo delle previsioni avanzate; tuttavia, Simulatore di Rischio fornisce le basi di queste due tecniche con lo scopo di eseguire previsioni di serie temporali. Consigliamo di non usare qualsiasi di queste tecniche in maniera isolata, ma in combinazione con altre metodologie di previsione di Simulatore di Rischio per costruire modelli più robusti.

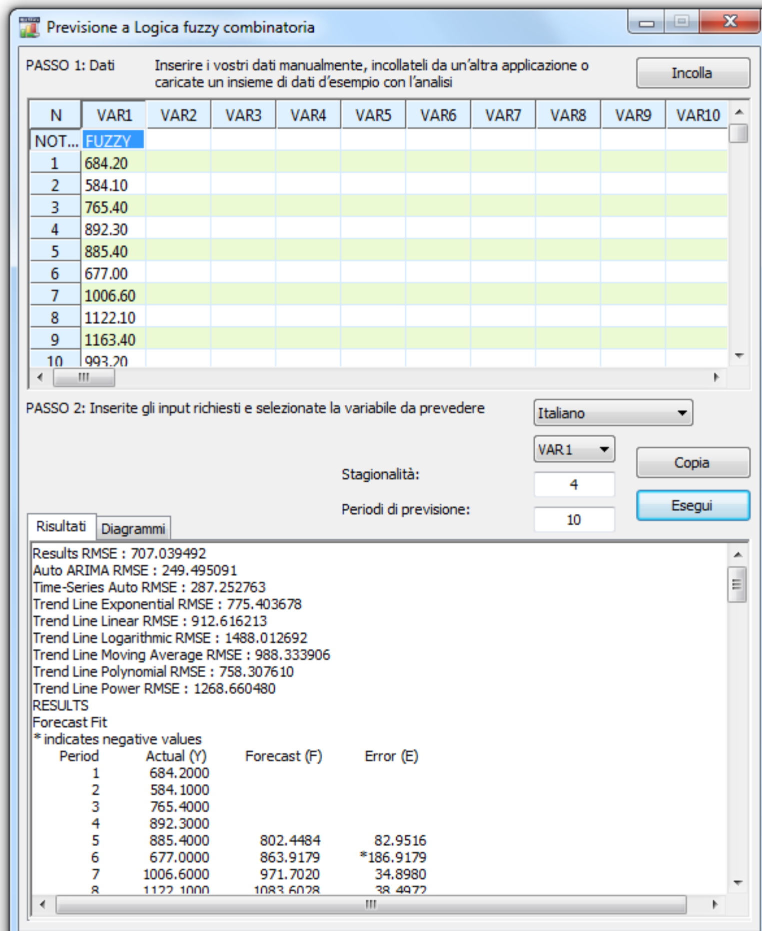


Figura 5.57 – Previsione di serie temporali Logica fuzzy

5.24 Ottimizzatore Ricerca obiettivo

Lo strumento Ricerca obiettivo è un algoritmo di ricerca applicato per trovare la soluzione di una singola variabile all'interno di un modello. Se conoscete il risultato che desiderate da una formula o da un modello, ma non siete sicuri quale valore di input la formula richiede per ottenere quel risultato, usate la funzione **Simulatore di Rischio | Strumenti | Ricerca obiettivo**. Si prega di notare che Ricerca obiettivo funziona solo con un valore di una variabile di input. Se volete accettare più di un valore di input, usate le routine avanzate di Ottimizzazione di Simulatore di Rischio. La Figura 5.58 mostra un semplice modello e come si applica la Ricerca obiettivo.

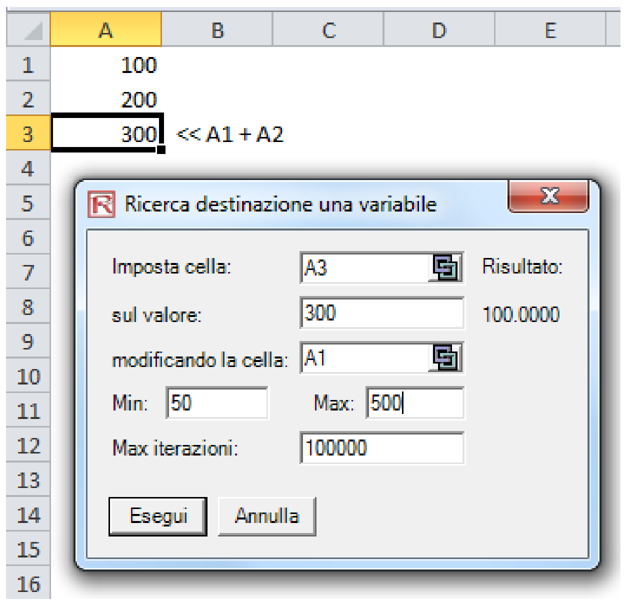


Figura 5.58 – Ricerca obiettivo

5.25 Ottimizzatore singola variabile

Lo strumento Ottimizzatore singola variabile è un algoritmo di ricerca usato per trovare la soluzione di una singola variabile all'interno di un modello, esattamente come la routine Ricerca obiettivo discussa in precedenza. Se desiderate il massimo o minimo risultato possibile da un modello, ma non siete sicuri quale valore di input la formula richiede per ottenere quel risultato, usate la funzione **Simulatore di Rischio | Strumenti | Ottimizzatore singola variabile** (Figura 5.59). Si prega di notare che questo Ottimizzatore singola variabile viene eseguito molto velocemente ma che è solo applicabile per trovare una variabile di input. Se volete accettare più di un valore di input, usate le routine avanzate di Ottimizzazione di Simulatore di Rischio. Si prega di notare che strumento è incluso in Simulatore di Rischio perché talvolta potreste avere la necessità di un veloce calcolo di ottimizzazione per una singola variabile decisionale e questo strumento fornisce questa potenzialità senza il bisogno di impostare un modello di ottimizzazione con profili, ipotesi di simulazione, variabili decisionali, obiettivi e vincoli.

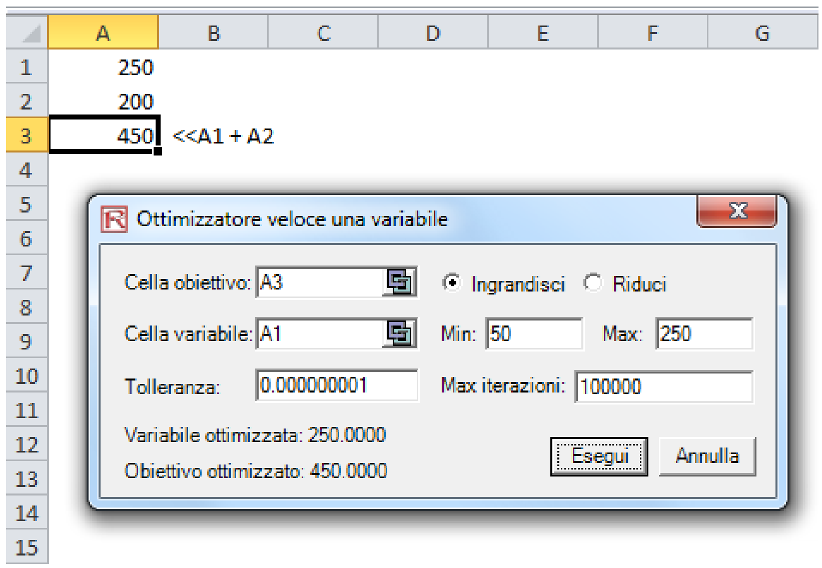


Figura 5.59 – Ottimizzatore singola variabile

5.26 Ottimizzazione algoritmo genetico

L'algoritmo genetico è un'euristica di ricerca che imita il processo dell'evoluzione naturale. Questa euristica è comunemente usata per generare soluzioni utili per problemi di ottimizzazione e ricerca. Gli algoritmi genetici appartengono alla classe più ampia degli algoritmi evoluzionali, i quali generano soluzioni per problemi di ottimizzazione usando tecniche inspirate alla evoluzione naturale, come ereditarietà, mutazione, selezione e crossover.

L'algoritmo genetico è disponibile sotto **Simulatore di Rischio | Strumenti | Algoritmo genetico** (Figura 5.60). Si deve prestare attenzione nella taratura degli input del modello dato che i risultati sono piuttosto sensibili agli input (gli input di default sono forniti come guida generale ai livelli di input più comuni) e consigliamo di scegliere il Test di ricerca gradiente per un insieme di risultati più robusti (potete deselezionare questa opzione per iniziare, poi selezionare questa funzionalità, rieseguire l'analisi e confrontare i risultati).

Nota: In molti problemi, gli algoritmi genetici possono tendere a convergere verso gli optima locali o dei addirittura punti arbitrari piuttosto che verso l'optimum globale del problema. Questo significa che non ha il "know-how" per sacrificare l'idoneità a breve termine per ottenere l'idoneità a lungo termine. Per specifici problemi di ottimizzazione e problemi di istanze, altri algoritmi di ottimizzazione potrebbero trovare soluzioni migliori che gli algoritmi genetici (dato lo stesso tempo di calcolo). Consigliamo, perciò, di eseguire prima l'algoritmo genetico e poi di eseguirlo di nuovo selezionando la casella *Applica test di ricerca gradiente Search* (Figura 5.60) per verificare la robustezza del modello. Questo test di ricerca gradiente tenterà di eseguire combinazioni di tecniche tradizionali di ottimizzazione e di metodi di algoritmi genetici e fornire la migliore soluzione possibile. Per finire, a meno che non ci sia uno specifico bisogno teoretico di usare l'algoritmo genetico, consigliamo di usare il modulo Ottimizzazione di Simulatore di Rischio per risultati più robusti, il che vi permette di eseguire routine avanzate di ottimizzazione dinamiche e stocastiche basate sul rischio.

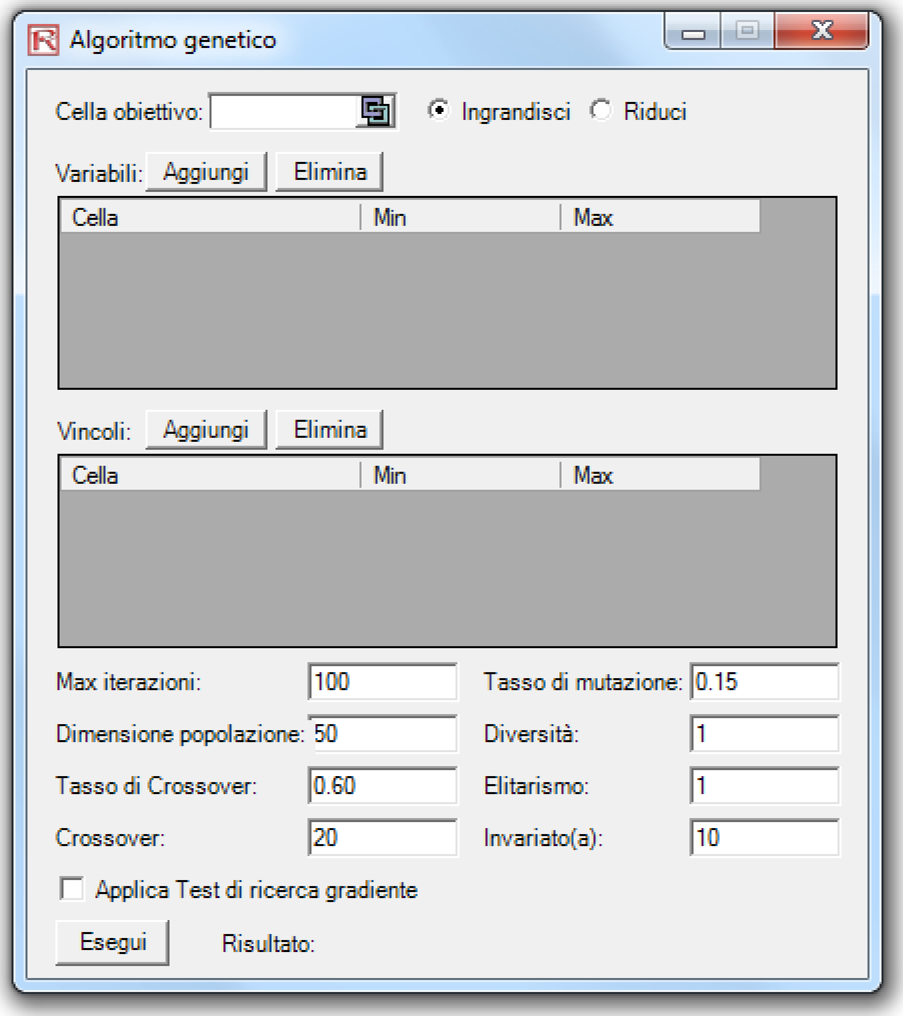


Figura 5.60 – Algoritmo genetico

5.27 Modulo ROV Albero decisionale

ROV Albero decisionale (Figura 5.61) è usato per creare e valutare modelli di alberi decisionali. Sono anche incluse addizionali metodologie e analitiche avanzate.

- Modelli di alberi decisionali
- Simulazione di rischio Monte Carlo
- Analisi di sensibilità
- Analisi di scenario
- Analisi Bayesiana (Aggiornamento probabilità congiunte e posteriori)
- Valore atteso dell'informazione
- MINIMAX
- MAXIMIN
- Prifili di rischio

- Ci sono 11 lingue locali disponibili in questo modulo e la lingua attuale può essere modificata attraverso il menu Lingua
- *Inserire nodo opzione* o *Inserire nodo terminale* selezionando prima un qualsiasi nodo esistente e cliccando poi sull'icona nodo opzione (quadrato) o sull'icona nodo terminale (triangolo) o usando le funzioni nel menu *Inserisci*.
- Modificate singole proprietà del *Nodo opzione* o del *Nodo terminale* eseguendo un doppio click su un nodo. Talvolta, quando cliccate su un nodo, vengono selezionati anche tutti i successivi nodi figli (questo vi permette di spostare l'intero albero iniziando da quel nodo selezionato). Se desiderate selezionare solo quel nodo, potrebbe essere necessario cliccare sullo sfondo vuoto e poi cliccare di nuovo su quel nodo per selezionarlo singolarmente. Inoltre, potete spostare i singoli nodi o l'intero albero iniziando dal nodo selezionato secondo l'impostazione attuale (clic tasto destro o dal menu *Editare* e poi selezionare *Sposta nodi singolarmente* o *Sposta nodi insieme*).
- Seguono alcune brevi descrizioni degli elementi che possono essere personalizzati e configurati nell'interfaccia dell'utente delle proprietà del

nodo. La cosa più semplice è di provare diverse impostazioni per ciascuno dei seguenti elementi per vedere i loro effetti nell'albero delle strategie:

  - *Nome*. Nome visualizzato sopra il nodo.
  - *Valore*. Valore visualizzato sotto il nodo.
  - *Collegamento Excel*. Collega il valore di una cella di un foglio di calcolo Excel.
  - *Note*. Note possono essere inserite sopra o sotto un nodo.
  - *Mostra nel modello*. Mostra qualsiasi combinazione di Nome, Valore e Note.
  - *Colore locale* in confronto a *Colore globale*. I colori dei nodi possono essere modificati localmente per un nodo o globalmente.
  - *Etichetta dentro la forma*. È possibile posizionare testo dentro il nodo (è possibile che sia necessario allargare il nodo per contenere il testo più lungo).
  - *Nome evento diramazione*. È possibile posizionare testo sulla diramazione che porta al nodo per indicare l'evento che porta a questo nodo.
  - *Selezionare le opzioni reali*. È possibile assegnare un tipo specifico di opzione reale al nodo attuale. Assegnare opzioni reali ai nodi permette allo strumento di generare una lista di variabili di input obbligatorie.

- *Elementi globali* sono tutti personalizzabili, incluso gli elementi *Sfondo, Linee di collegamento, Nodi opzioni, Nodi terminali* e *Caselle di testo* dell'albero delle strategie. Per esempio, le seguenti impostazioni possono essere modificati per ciascuno degli elementi:
  - Le impostazioni del tipo di *Carattere* per Nome, Valore, Note, Etichetta, Nomi eventi.
  - *Dimensione del nodo* (altezza e larghezza minima e massima).
  - *Bordo* (stili della linea, larghezza e colore).
  - *Ombreggiatura* (colori e se applicare una ombreggiatura o no).

- Colore globale.
- Forma globale.

- Il commando *Editare* del menu *Visualizza finestra requisiti dati* apre una finestra ancorata sulla destra dell'albero delle strategie in modo che quando un nodo opzione o un nodo terminale viene selezionato , le proprietà di quel nodo saranno visualizzate e potranno essere aggiornate direttamente. Questa funzione fornisce un'alternativa al doppio clic su un nodo ogni volta.
- *Esempi di file* sono disponibili nel menu *File* per aiutarvi ad iniziare a costruire alberi delle strategie.
- *Proteggi file* nel menu *File* permette di crittografare l'albero delle strategie con una crittografia password fino a 256-bit. Fate attenzione se crittografate un file: se la password viene persa, il non non potrà più essere aperto.
- La funzione *Cattura schermata* o stampa il modello esistente può essere eseguito dal menu *File*. La schermata catturata può successivamente essere incollata in altre applicazioni software.
- Le funzioni *Aggiungi, Duplica, Rinomina* e *Elimina un albero delle strategie* possono essere eseguite mediante un doppio clic sulla scheda Albero delle strategie o nel menu *Edit*.
- Potete anche eseguire le operazioni *Inserisci collegamento file* e *Inserisci commento* per qualsiasi nodo opzione o nodo terminale o *Inserisci testo* o *Inserisci immagine* ovunque sullo sfondo o nell'area di disegno.
- Potete inoltre eseguire le operazioni *Cambia stili esistenti* o *Gestisci e crea stili personalizzati* sul vostro albero delle strategie (questo comprende dimensione, forma, schemi del colore e le specifiche di dimensione/colore del carattere del completo albero delle strategie).
- *Inserisci nodo decisionale*, *Inserisci nodo incertezza* o *Inserisci nodo terminale* selezionando qualsiasi nodo esistente e cliccando poi sull'icona nodo decisionale (quadrato), sull'icona nodo incertezza (cerchio) o sull'icona nodo terminale (triangolo), o usando le funzionalità nel menu *Inserisci*.
- Modificate le proprietà di singoli nodi decisionali, incertezza o terminali eseguendo un doppio clic su un nodo. Seguono alcuni aggiuntivi elementi

unici nel modulo Albero decisionale che possono essere personalizzati e configurati nell'interfaccia dell'utente delle proprietà del nodo.

  - Nodi decisionali: *Sostituzione personalizzata* o *Calcolo automatico* del valore di un nodo. L'opzione Calcolo automatico è impostata di default; quando cliccate su *ESEGUI* per un modello completato di Albero decisionale, i nodi decisionali saranno aggiornati con i risultati.
  - Nodi incertezza: *Nomi eventi*, *Probabilità* e *Imposta ipotesi di simulazione*. Potete aggiungere nomi degli eventi di probabilità, probabilità e ipotesi di simulazione solo dopo che siano stati creati le diramazioni incertezza.
  - Nodi terminali: *Inserimento manuale*, *Collegamento Excel* e *Imposta ipotesi di simulazione*. I payoff di eventi terminali possono essere inseriti manualmente o collegati ad una cella Excel (p.es. se avete un modello grande di Excel che calcola i payoff, potete collegare il modello alla cella di output di questo modello Excel); alternativamente potete impostare ipotesi distribuzionali di probabilità per eseguire simulazioni.

- *Visualizza finestra proprietà nodo* è disponibile nel menu *Edita* e le proprietà del nodo saranno aggiornate quando si seleziona un nodo.
- Il modulo albero decisionale comprende anche le seguenti analitiche avanzate:
  - Modellazione di Simulazioni Monte Carlo su alberi decisionali
  - Analisi Bayesiana per ottenere le probabilità posteriori
  - Valore atteso della perfetta informazione, Analisi MINIMAX e MAXIMIN, Profili di rischio e Valore dell'imperfetta informazione
  - Analisi di sensibilità
  - Analisi di scenario
  - Analisi della Funzione di utilità

5.27.1 Modellazione della simulazione

Questo strumento esegue una Simulazione di rischio Monte Carlo sull'albero decisionale (Figura 5.62). Questo strumento vi permette di impostare le distribuzioni di probabilità come ipotesi di input per l'esecuzione di simulazioni. Potete impostare una ipotesi per il nodo selezionato oppure impostare una nuova ipotesi e usare questa nuova ipotesi (o anche ipotesi create in precedenza) in una equazione o formula numerica. Per esempio, potete impostare una nuova ipotesi chiamata Normale (p. es., una distribuzione normale con un valore medio di 100 e una deviazione standard di 10) ed eseguire una simulazione nell'albero decisionale oppure usare questa ipotesi in una equazione del tipo (100*Normal+15.25). Create il vostro modello nella casella di espressione numerica. Potete usare calcoli di base o aggiungere variabili esistenti nella vostra equazione eseguendo un doppio clic sulla lista delle variabili esistenti. Potete aggiungere nuove variabili alla lista come richiesto, sia come espressione numerica sia come ipotesi.

5.27.2 Analisi bayesiana

Questo strumento di analisi bayesiana (Figura 5.63) può essere eseguito su qualsiasi due eventi di incertezza purché collegati lungo un percorso. Per esempio nell'esempio a destra, le incertezze A e B sono collegate, di cui l'evento A si verifica per primo nella linea del tempo e l'evento B si verifica come secondo. Il primo evento A è Ricerca di mercato con due esiti (favorevole o sfavorevole). Il secondo evento B è Condizioni di mercato, anch'esso con due esiti (forte e debole). Questo strumento è utilizzato per calcolare le probabilità congiunte, le probabilità marginali e le probabilità bayesiane aggiornate a posteriori mediante l'inserimento delle probabilità precedenti e delle probabilità condizionali di affidabilità. Oppure si possono calcolare le probabilità di affidabilità quando si hanno le probabilità condizionali aggiornate a posteriori. Selezionate l'analisi pertinente desiderata sotto e cliccate su Carica esempio per vedere sia gli input campione corrispondenti all'analisi selezionata e i risultati mostrati nella griglia a destra sia quali risultati sono usati come input nell'albero decisionale nella figura.

Quick Procedures

- PASSO 1: Inserite il nome per il primo e il nome per il secondo evento d'incertezza e scegliete la quantità di eventi di probabilità (stati della natura o esiti) per ogni evento.
- PASSO 2: Inserite i nomi di ciascun evento di probabilità o esito.
- PASSO 3: Inserite le probabilità precedenti del secondo evento e le probabilità condizionali per ciascun evento o esito. Le probabilità devono sommarsi a 100%.

5.27.3 Valore atteso della perfetta informazione, Analisi Minimax e Maximin, Profili di rischio e Valore dell'informazione imperfetta

Questo strumento calcola il Valore atteso della perfetta informazione (Expected Value of Perfect Information - EVPI), l'Analisi Minimax e Maximin, il Profilo di rischio e il Valore dell'informazione imperfetta (Figura 5.64). Per iniziare, inserite il numero di diramazioni decisionali o strategie in esame (p. es., costruire un impianto grande, medio, piccolo) e il numero di eventi incerti o di esiti di stati di natura (p. es., mercato favorevole, mercato sfavorevole) e inserite i payoff attesi sotto ciascun scenario.

Il Valore atteso della perfetta informazione (EVPI). In altre parole, supponendo che aveste una perfetta preveggenza e che sapeste esattamente il da farsi (attraverso ricerche di mercato o mediante qualche altro mezzo per meglio discernere gli esiti probabilistici), l'EVPI calcola se esiste un valore aggiunto in tale informazione (cioè, se le ricerche di mercato aggiungeranno del valore) in confronto a stime più semplicistiche degli stati di natura probabilistici. Per iniziare, inserite il numero di diramazioni decisionali o strategie in esame (p. es., costruire un impianto grande, medio, piccolo) e il numero di eventi incerti o di esiti di stati di natura (p. es., mercato favorevole, mercato sfavorevole) e inserite i payoff attesi sotto ciascun scenario.

Minimax (minimizzare la massima perdita) e Maximin (massimizzare il minimo payoff) sono due apporci alternativi per trovare il percorso decisionale ottimale. Questi due metodi non sono usati spesso, ma offrono tuttavia informazioni aggiuntive riguardo al processo decisionale. Inserite il numero di diramazioni decisionali o percorsi che esistono (p. es., costruire un impianto grande, medio, piccolo) e anche il numero di eventi d'incertezza o stati di natura sotto ciascun percorso (p. es., economia forte contro economia debole). Ora completate la tabella dei payoff per i vari scenari e calcolate i risultati Minimax e Maximin. Potete anche cliccare su Carica esempio per visionare un calcolo d'esempio.

5.27.4 Sensibilità

L'analisi di sensibilità sulle probabilità (Figura 5.65) di input viene eseguita per determinare i suoi effetti sui valori dei percorsi decisionali. Per primo, selezionate sotto un Nodo decisionale da analizzare e dopo selezionate dalla lista un evento di probabilità da testare Se ci sono multiple eventi di incertezza con identiche probabilità, questi possono essere analizzati sia indipendentemente sia contemporaneamente.

I diagrammi di sensibilità mostrano i valori dei percorsi decisionali in condizioni di livelli variabili di probabilità. I valori numerici sono indicati nella tabella dei risultati. La posizione delle eventuali linee di crossover rappresentano a quali eventi probabilistici un certo percorso decisionale diventa dominante sopra un altro.

5.27.5 Tabelle degli scenari

Le tabelle degli scenari (Figura 5.66) possono essere generate per determinare i valori di output considerate alcune modifiche al input. Potete scegliere uno o più percorsi decisionali da analizzare (i risultati di ciascun percorso scelto sarà rappresentato come una tabella separata e un diagramma separato) e uno o più nodi di incertezza o nodi terminali come variabili di input per la tabella degli scenari.

Quick Procedures

- Selezionate dalla lista sottostante uno o più percorsi decisionali da analizzare.
- Selezionate uno o due eventi di incertezza o payoff terminali da modellare.
- Decidete se desiderate modificare la probabilità dell'evento da solo o modificare tutti gli eventi di probabilità identici contemporaneamente.
- Inserisci il campo di variazione dello scenario di input.

5.27.6 Generazione della funzione di utilità

Le funzioni di utilità (Figura 5.67), o U(x), vengono a volte usate al posto dei valori attesi dei payoff terminali in un albero decisionale. U(x) possono essere sviluppate in due modi: usando la sperimentazione noiosa e dettagliata di ciascun esito possibile o mediate un metodo di estrapolazione esponenziale (usato qui). Possono essere modellate per un decisore che è avverso al rischio (gli svantaggi sono più disastrosi o dolorosi di un uguale vantaggio potenziale), neutrale al rischio (vantaggi e svantaggi sono ugualmente interessanti) o amante del rischio (un vantaggio potenziale è più interessante). Inserite il valore atteso minimo e massimo dei vostri payoff terminali e il numero di punti dati in mezzo a loro per calcolare la curva e la tabella di utilità.

Se aveste una scommessa 50:50 con la quale potreste vincere $X o perdere -$X/2 paragonato a non scommettere e ricevere un payoff di $0, cosa sarebbe questo $X? Per esempio, se fare una scommessa dove potete vincere $100 o perdere -$50 con uguale probabilità o non scommettere affatto vi è indifferente, allora il vostro X è $100. Inserite X nella casella sottostante Ricavi positivi. Notare che tanto più grande è questo X, tanto mento siete avverso al rischio, mentre un X più piccolo indica che siete maggiormente avverso al rischio.

Inserite gli input obbligatori, selezionate il tipo di U(x) e cliccate su Calcola utilità per ottenere i risultati. Potete anche applicare i valori calcolati di U(x) all'albero decisionale per eseguirlo di nuovo o riportare l'albero all'uso dei valori attesi dei payoff.

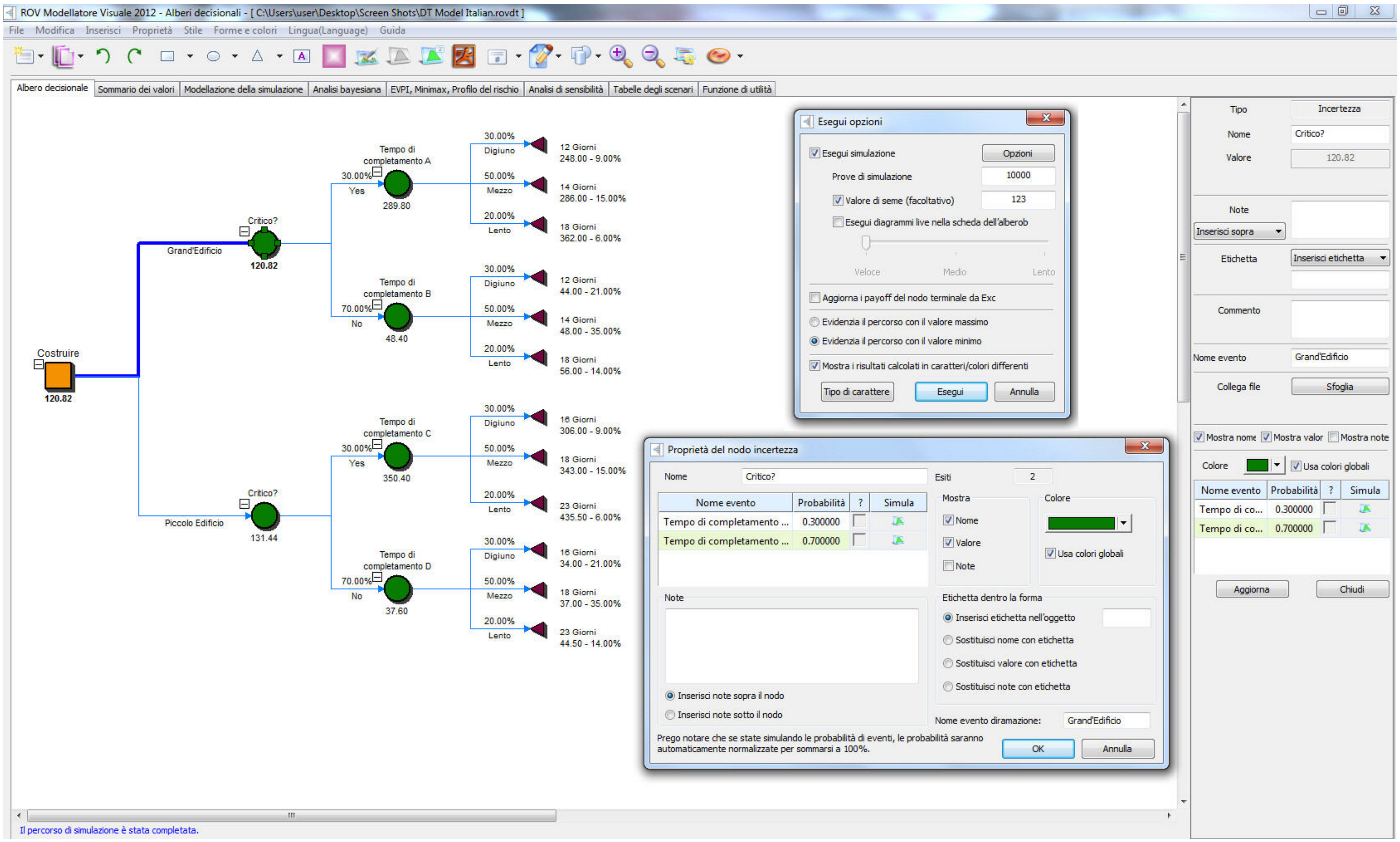


Figura 5.61 – ROV Albero decisionale (Albero decisionale)

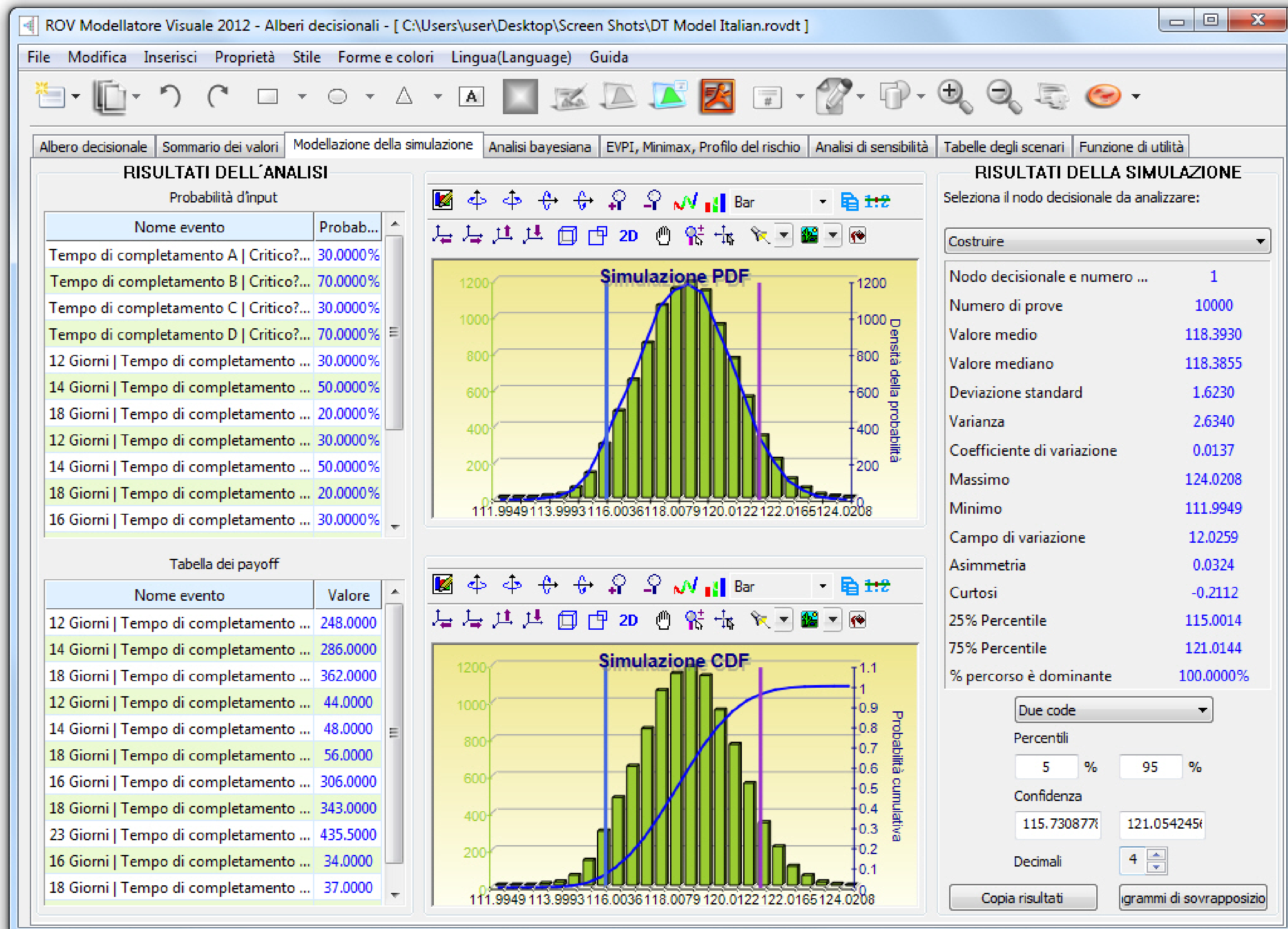


Figura 5.62 – ROV Albero decisionale (Risultati della simulazione)

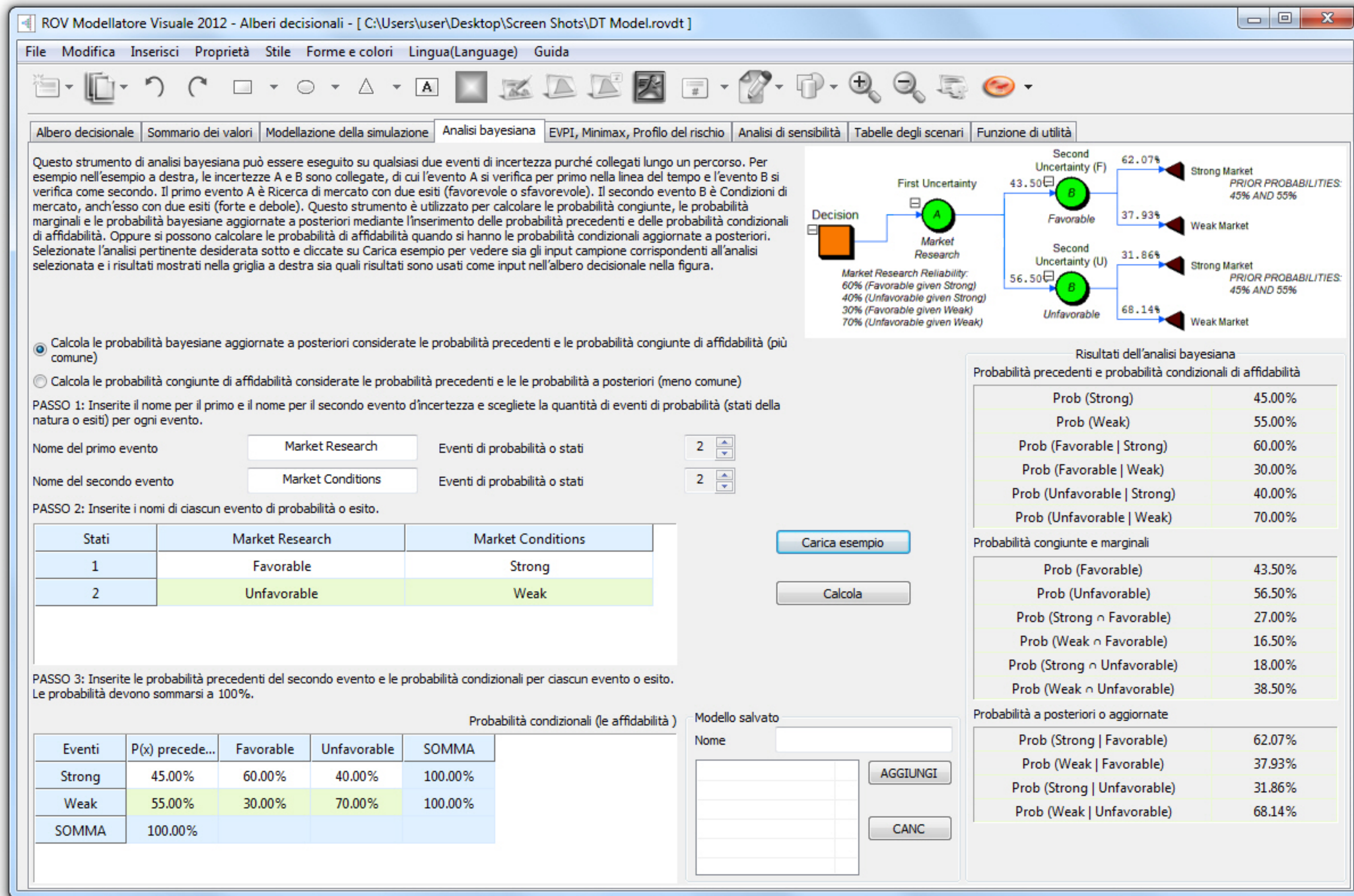


Figura 5.63 – ROV Albero decisionale (Analisi Bayesiana)

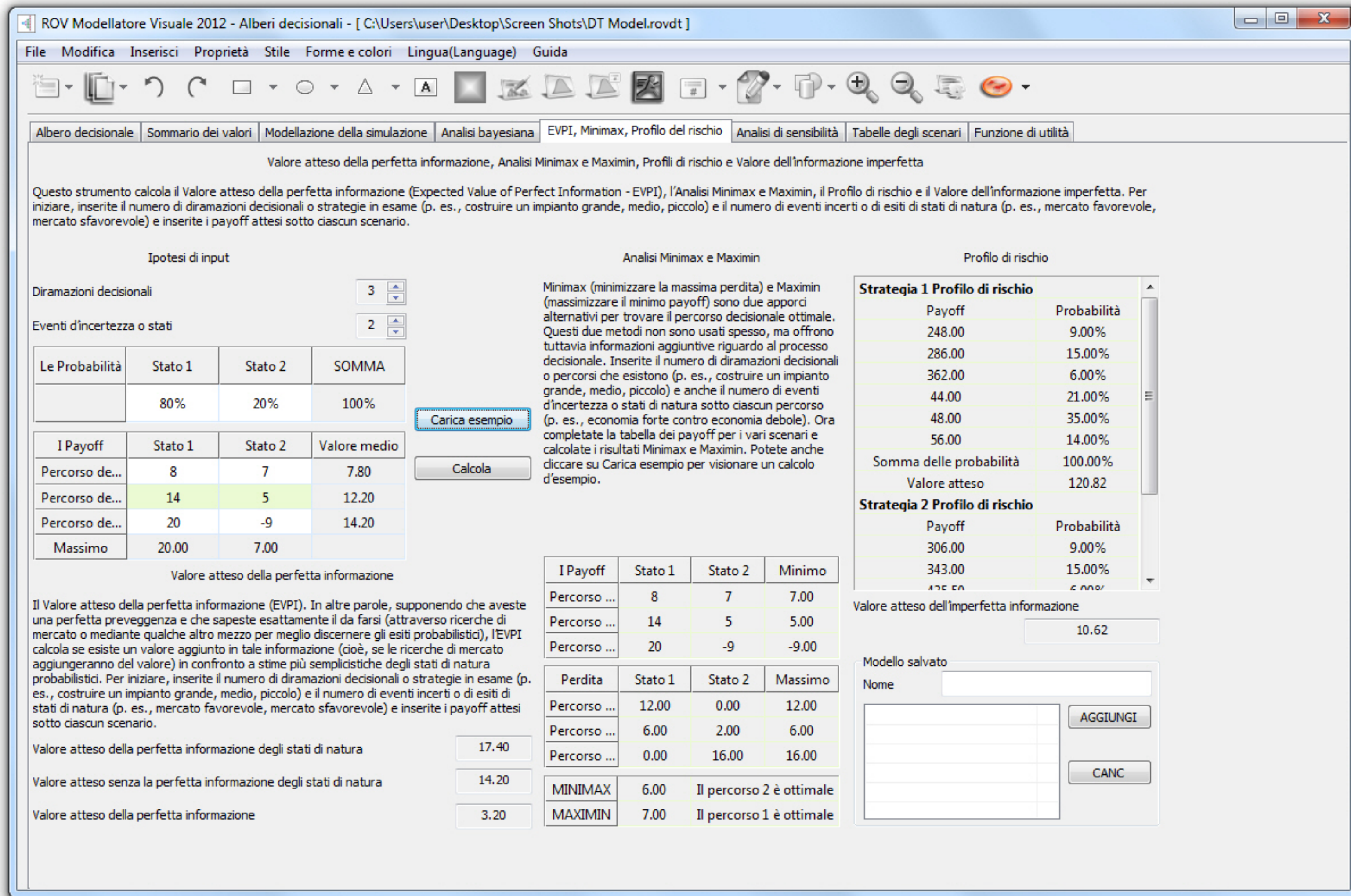


Figura 5.64 – ROV Albero decisionale (VAPI, MINIMAX, Profilo di rischio)

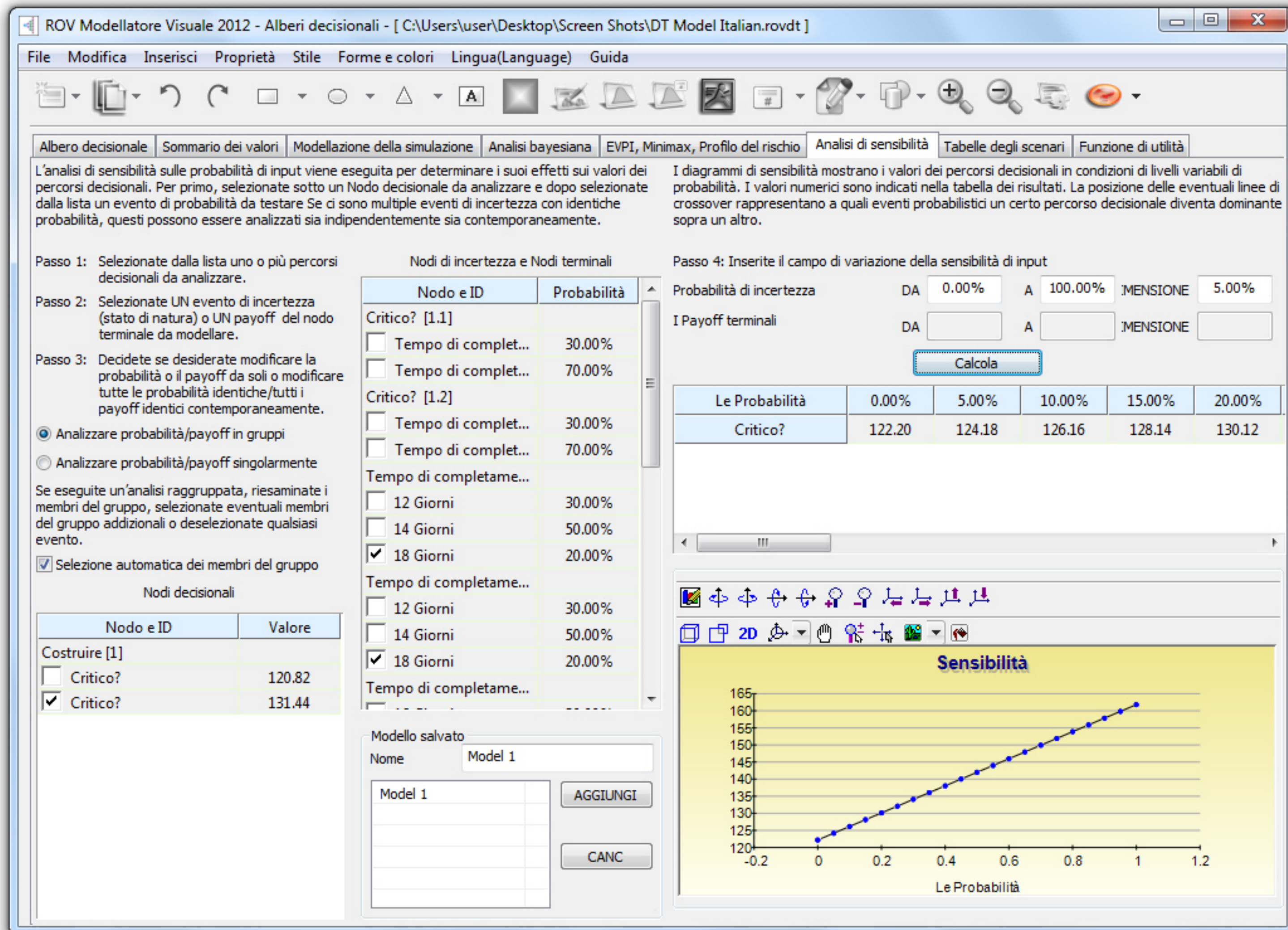


Figura 5.65 – ROV Albero decisionale (Analisi di sensibilità)

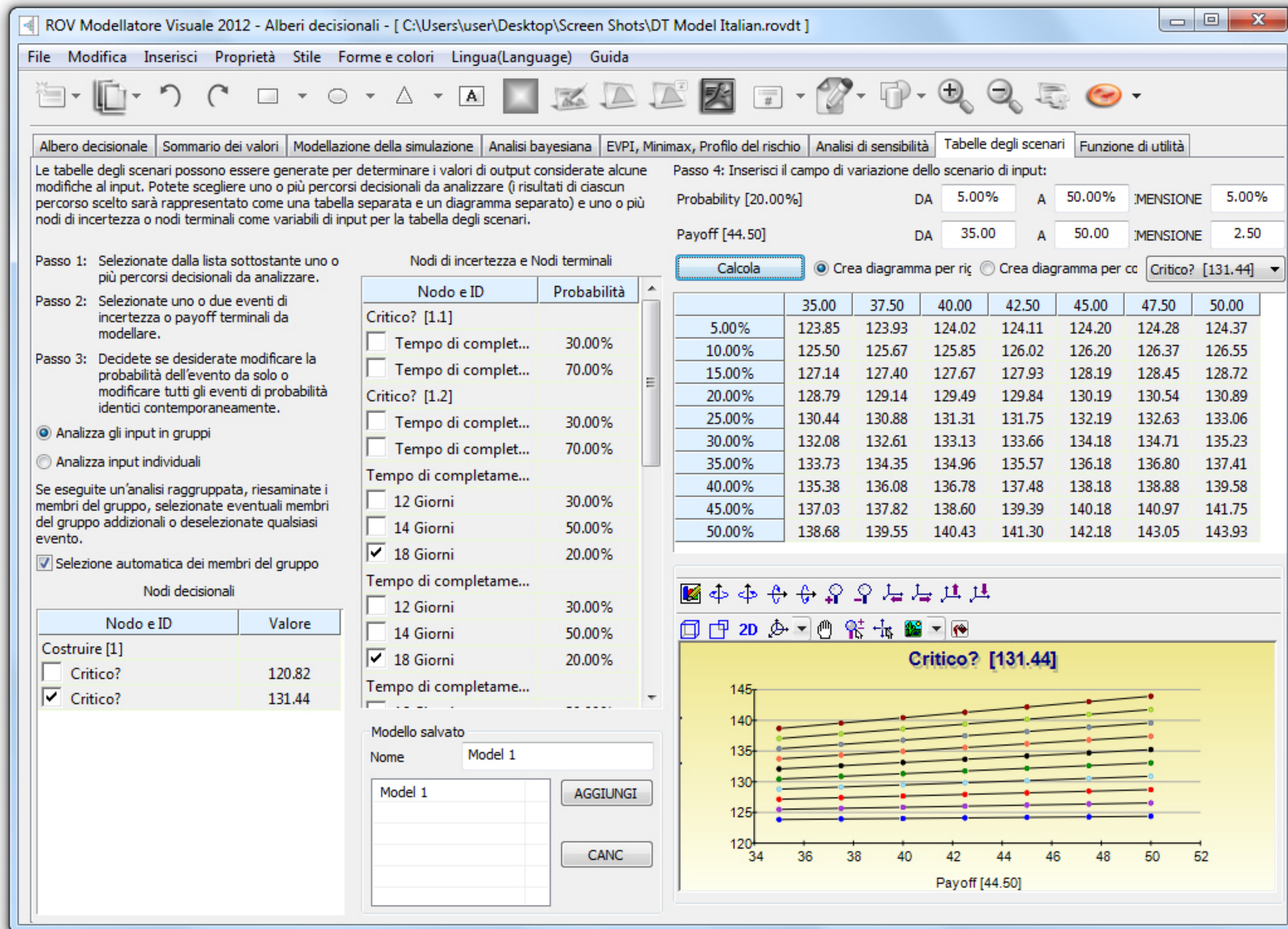


Figura 5.66 – ROV Albero decisionale (Tabelle degli scenari)

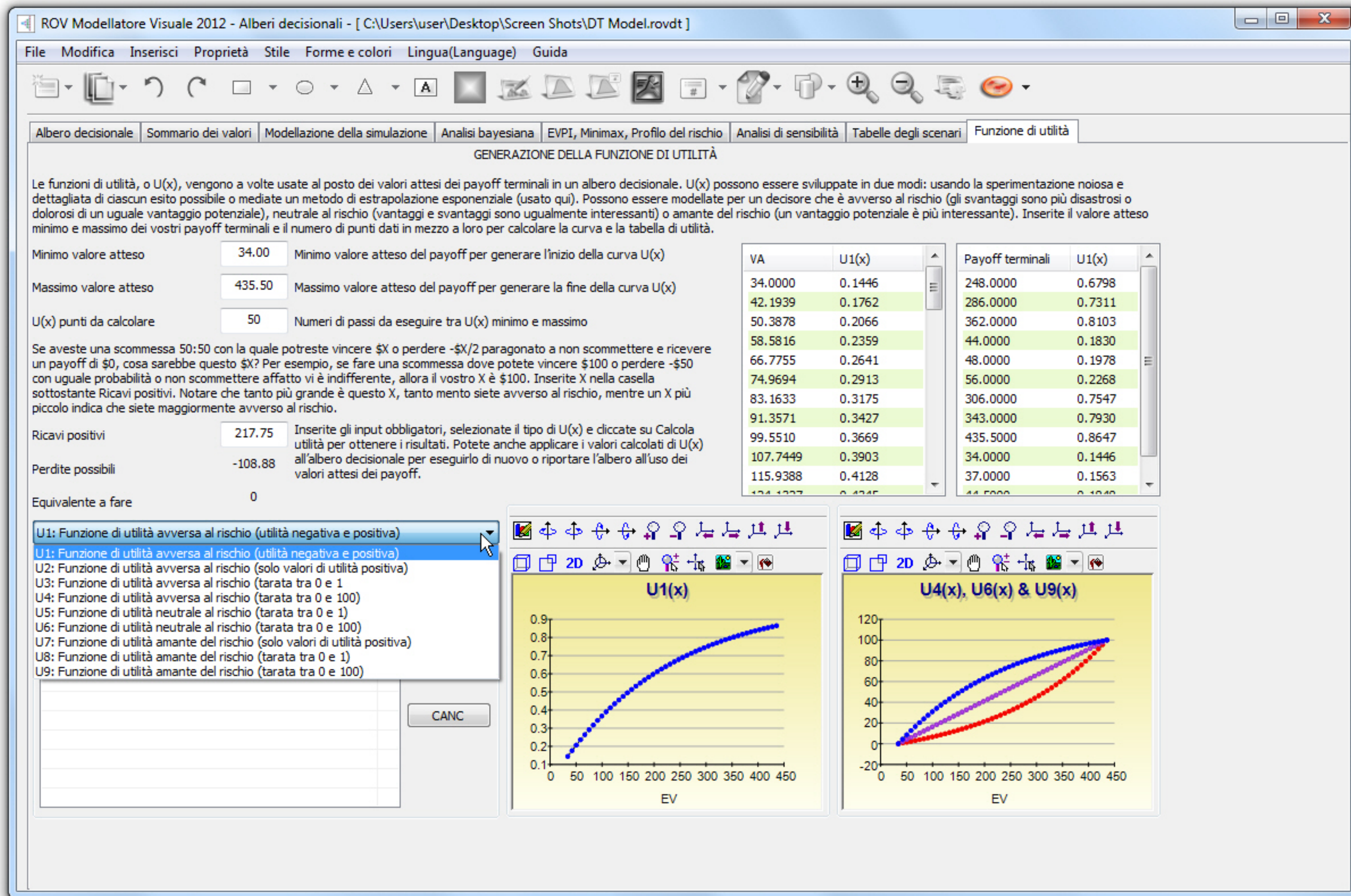


Figura 5.67 – ROV Albero decisionale (Funzioni di utilità)

6

6. SUGGERIMENTI E TECNICHE UTILI

Di seguito forniamo alcuni suggerimenti utili e alcune tecniche di scelta rapida per utenti avanzati di Simulatore di Rischio. Per dettagli sull'uso degli specifici strumenti, siete pregati di consultare le relative sezioni nel manuale dell'utente.

SUGGERIMENTI: Ipotesi (Interfaccia dell'utente per l'impostazione dell'ipotesi di input)

- Salto veloce — selezionate qualsiasi distribuzione ed inserite qualsiasi lettera e "salterete" alla prima distribuzione che inizia con quella lettera (p. es., cliccate su Normale e digitate W e sarete portati alla distribuzione Weibull).
- Visualizzazioni clic destro — selezionate qualsiasi distribuzione, cliccate con tasto destro e selezionate-le diverse visualizzazioni delle distribuzioni (icone grandi, icone piccole, lista).
- Tab per aggiornare i diagrammi — dopo aver inserito dei nuovi parametri di input (p. es., digitate un nuovo valore per la media o per la deviazione standard), premete Tab sulla tastiera o cliccate ovunque sull'interfaccia utente al di fuori della casella di input per vedere che il diagramma distribuzionale si aggiorna automaticamente.
- Inserire delle correlazioni — potete inserire direttamente qui delle correlazioni a coppie (le colonne possono essere ridimensionate come richiesto), usare lo strumento di adattamento distribuzionale multiplo per calcolare automaticamente ed inserire tutte le correlazioni a coppie, o, dopo aver impostato delle ipotesi, usare lo strumento di editazione delle correlazioni per inserire la vostra matrice di correlazione.
- Equazioni in una cella d'ipotesi — solo celle vuote o celle con valori statici possono essere impostate come ipotesi; tuttavia, ci possono essere

delle occasioni quando una funzione o una equazione è richiesta in una cella d'ipotesi e questo può essere eseguito inserendo per prima cosa l'ipotesi di input nella cella e digitando poi l'equazione o la funzione (quando la simulazione è in corso di esecuzione, i valori simulati sostituiranno la funzione e, dopo che la simulazione sarà completata, la funzione o equazione appariranno di nuovo).

SUGGERIMENTI: Copia e incolla

- Copia e incolla usando *Escape* — quando selezionate una cella e usate la funzione Copia di Simulatore di Rischio, questa operazione copia tutto il contenuto della cella negli Appunti di Windows, incluso il valore, l'equazione, la funzione, il colore, il font e la dimensione, così come le ipotesi, le previsioni o le variabili decisionali di Simulatore di Rischio. In seguito, quando applicate la funzione Incolla di Simulatore di Rischio, avrete due opzioni. La prima opzione è di applicare direttamente la funzione Incolla di Simulatore di Rischio e tutto il contenuto della cella, valori, colore, font, equazione, funzioni e parametri, sarà incollato nella nuova cella. La seconda opzione è di premere prima il tasto Escape sulla tastiera e di applicare poi la funzione Incolla di Simulatore di Rischio. Escape comunica al Simulatore di Rischio che desiderate incollare solo le ipotesi, le previsioni o le variabili decisionali di Simulatore di Rischio e non i valori della cella, color, equazione, funzione, font e così via. Premere Escape prima di incollare vi permette di mantenere i valori e i calcoli della cella di destinazione e incolla solo i parametri di target Simulatore di Rischio.
- Copia e incolla con celle multiple — potete selezionare multiple celle per il copia e incolla (con ipotesi contigue e non contigue).

SUGGERIMENTI: Correlazioni

- Imposta ipotesi — imposta correlazioni a coppie usando il dialogo Imposta ipotesi di input (ideale per inserire solo alcune correlazioni).
- Edita correlazioni — imposta una matrice di correlazione inserendola manualmente o incollandola dagli Appunti di Windows (ideale per matrici di correlazioni di grandi dimensioni e per multiple correlazioni).
- Adattamento distribuzionale multiplo — calcola automaticamente ed inserisce correlazioni a coppie (ideale quando si eseguono multipli adattamenti di variabili e si fanno calcolare automaticamente le

correlazioni per decidere quale costituisce una correlazione statisticamente significativa).

SUGGERIMENTI: Diagnostica dei dati e Analisi statistica

- Stima parametro stocastico — nei report dell'Analisi statistica e della Diagnostica dei dati si trova una linguetta sulle stime dei parametri stocastici che stima la volatilità, la deriva, il tasso di ritorno alla media e i tassi della diffusione a salti basati su tassi storici. Siate consapevoli che questi risultati dei parametri sono basati solamente sui dati storici utilizzati, che i parametri possono variare nel tempo e che dipendono sulla quantità di dati storici adattati. Inoltre, i risultati dell'analisi mostrano tutti i parametri, ma non suggeriscono quale modello di processo stocastico (p. es., Moto Browniano, Ritorno alla media, Diffusione a salti o processi combinati) è il miglior adattamento. È compito dell'utente fare questa determinazione in funzione alla variabile di serie temporali da prevedere. L'analisi non può determinare quale sia il processo migliore, solo l'utente può fare questo (p. es., il processo con Moto Browniano è il migliore per modellare prezzi di titoli, ma l'analisi non può determinare se i dati storici analizzati provengono da un titolo o da qualche altra variabile, solo l'utente può sapere questo). Per finire, un buon indizio è che se un determinato parametro si trova fuori del suo normale campo, allora il the processo che richiede questo parametro di input probabilmente non è il processo giusto (p. es., se il tasso di ritorno alla media è del 110%, è probabile che il processo con ritorno alla media non sia il processo giusto e così via).

SUGGERIMENTI: Analisi distribuzionale, Diagrammi e tabelle di probabilità

- Analisi distribuzionale — usata per calcolare velocemente PDF, CDF e ICDF delle 42 distribuzioni di probabilità disponibili in Simulatore di Rischio e fornire una *tabella* di questi valori.
- Diagrammi e tabelle distribuzionali — usati per confrontare *differenti parametri della stessa distribuzione* (p. es., le forme ed i valori PDF, CDF, ICDF di una distribuzione Weibull con Alfa e Beta di [2, 2], [3, 5], e [3.5, 8], sovrapponendoli uno sopra l'altro).
- Diagrammi sovrapposti — usati per paragonare *distribuzioni differenti* (ipotesi di input teoretici e previsioni di output simulate empiricamente) e li sovrappone uno sull'altro per un confronto visivo.

SUGGERIMENTI: Frontiera efficiente

- Variabili di Frontiera efficiente — per accedere alle variabili di frontiera, impostate prima i Vincoli del modello prima di impostare le variabili di frontiera efficiente.

SUGGERIMENTI: Celle di previsione

- Celle di previsione senza equazioni — potete impostare previsioni di output in celle senza equazioni o valori (ignorate semplicemente il messaggio d'avviso) ma siate consapevoli che il risultante diagramma di previsione sarà vuoto. Previsioni di output sono tipicamente impostate in celle vuote quando ci sono dei macros che sono calcolati, per cui la cella sarà continuamente aggiornata.

SUGGERIMENTI: Diagrammi di previsione

- *Tab* in confronto a *Barra spaziatrice* — premete Tab sulla tastiera per aggiornare il diagramma di previsione e per ottenere i valori di percentile e di confidenza dopo aver inserito alcuni input, e premete la Barra spaziatrice per alternare tra le varie linguette nel diagramma di previsione.
- Vista normale in confronto a Vista globale — cliccate su queste viste per alternare tra una interfaccia a linguetta ed una interfaccia globale dove tutti gli elementi dei diagrammi di previsione sono visibili insieme.
- Copia — questa funzione copierà il diagramma di previsione o l'intera vista globale a seconda se siete nella vista normale o quella globale.

SUGGERIMENTI: Previsione

- Indirizzo di collegamento cella — se selezionate prima i dati nel foglio di lavoro e poi eseguite uno strumento di previsione, l'indirizzo della cella dei dati selezionati sarà automaticamente inserito nell'interfaccia utente; altrimenti dovrete inserire manualmente l'indirizzo della cella o usare la icona di collegamento per collegare alla relativa posizione dei dati.
- Previsione RMSE — usate questa misura di errore universale su multipli modelli di previsione per confronti diretti riguardo la accuratezza di ciascun modello.

SUGGERIMENTI: Previsione: ARIMA

- Periodi di previsione — Il numero di righe di dati esogeni deve superare le righe dei dati di serie temporali di almeno i periodi di previsione desiderati

(p. es., se desiderate prevedere 5 periodi nel futuro e avete 100 punti dati di serie temporali, saranno necessari almeno 105 o più punti dati sulla variabile esogena); altrimenti eseguite semplicemente ARIMA senza la variabile esogena per prevedere tutti i periodi che desiderate senza limitazioni.

SUGGERIMENTI: Previsione: Econometria di base

- Separazione delle variabili con punti e virgole — separate le variabili indipendenti con un punto e virgola.

SUGGERIMENTI: Previsione: Logit, Probit e Tobit

- Requisiti dei dati — le variabili dipendenti per eseguire i modelli logit e probit devono essere solamente binarie (0 e 1), mentre il modello tobit può assumere valori binari e altri valori decimali numerici. Le variabili indipendenti per tutti e 3 i modelli possono assumere qualsiasi valore numerico.

SUGGERIMENTI: Previsione: Processi stocastici

- Input d'esempio di default — se sussiste un dubbio, usate gli input di default come punto d'inizio per sviluppare il vostro modello.
- Strumento di analisi statistica per la stima del parametro —usate questo strumento per tarare i parametri di input nei modelli con the processi stocastici stimandoli dai vostri dati grezzi.
- Modello con processo stocastico — qualche volta, se l'interfaccia utente del processo stocastico rimane ferma per molto tempo, è probabile che i vostri input siano sbagliati e che il modello non sia correttamente specificato (p. es., se il tasso di ritorno alla media è del 110%, è probabile che il processo di ritorno alla media non sia il processo giusto e così via). Si prega di riprovare con input diversi o di usare un modello differente.

SUGGERIMENTI: Previsione: Linee di tendenza

- Risultati della previsione — scendete verso la fine del report per vedere i valori previsti.

SUGGERIMENTI: Chiamate di funzioni

- Funzioni RS — Qui potete impostare le ipotesi di input impostate e ottenere funzioni statistiche della previsione che potete usare all'interno del vostro foglio di lavoro di Excel. Per usare queste funzioni dovete prima installare le Funzioni RS (Start, Programmi, Real Options Valuation, Simulatore di Rischio, Strumenti e Installa funzioni) e poi eseguire una simulazione prima di impostare le funzioni RS all'interno di Excel. Consultate il modello d'esempio 24 per esempi su come usare queste funzioni.

SUGGERIMENTI: Esercizi su come iniziare e Video su come iniziare

- Esercizi su come iniziare — ci sono multipli esempi pratici passo per passo ed esercizi su come interpretare i risultati disponibili nella posizione di scelta rapida Start, Programmi, Real Options Valuation, Simulatore di Rischio. Questi esercizi sono intesi per aiutarvi ad iniziare velocemente ad usare il software.
- Video su come iniziare — tutti questi sono disponibili senza costo dal nostro sito web

  www.realoptionsvaluation.com/download.html o da www.rovdownloads.com/download.html.

SUGGERIMENTI: ID Hardware (HWID)

- Copia HWID con clic destro — nell'interfaccia utente di Installa licenza, selezionate o fate un doppio clic sull'HWID per selezionare il suo valore, cliccate col tasto destro per copiare o cliccate sul collegamento E-mail HWID per generare una e-mail con il HWID.
- Strumento di risoluzione dei problemi (Troubleshooter) — eseguite il Troubleshooter dalla cartella Start, Programmi, Real Options Valuation, Risk Simulatore di Rischio ed eseguite lo strumento Ottieni HWID per ottenere il HWID del vostro computer.

SUGGERIMENTI: Campionamento Ipercubo Latino (Latin Hypercube Sampling, LHS) in confronto a Simulazione Monte Carlo (Monte Carlo Simulation, MCS)

- Correlazioni — quando impostate correlazioni a coppie tra le ipotesi di input, consigliamo di usare l'impostazione Monte Carlo in Simulatore di Rischio, menu Opzioni. Il Campionamento Ipercubo Latino non è compatibile con il metodo della copula correlata per la simulazione.
- LHS Bin — un numero più alo di bin rallenterà la simulazione, ma fornirà un insieme più uniforme di risultati della simulazione.

- Casualità — tutte le tecniche di simulazione casuale nel menu Opzioni sono state testate e sono buoni simulatori e si avvicinano agli stessi livelli di casualità quando si eseguono un grande numero di prove.

SUGGERIMENTI: Risorse on-line

- Libri, Video su come iniziare, Modelli, White Paper — disponibili senza costo sul nostro sito web www.realoptionsvaluation.com/download.html o www.rovdownloads.com/download.html.

SUGGERIMENTI: Ottimizzazione

- Risultati irrealizzabili — se l'esecuzione dell'ottimizzazione fornisce risultati irrealizzabili, potete provare a cambiare i vincoli da Uguale (=) a Non uguale (>= o <=) e riprovare. Questo è applicabile anche quando eseguite un'analisi di frontiera efficiente.

SUGGERIMENTI: Profili

- Multipli profili — create e alternate tra multipli profili in un singolo modello. Questo vu permette di eseguire scenari sulla simulazione mediante la possibilità di cambiare i parametri di input o i tipi di distribuzione nel vostro modello per vedere gli effetti sui risultati.
- Profilo richiesto — ipotesi, previsioni o variabili decisionali non possono essere creati se non esiste un profilo attivo. Tuttavia, una volta che esiste un profilo, non ci sarà bisogno che creiate un nuovo profilo ogni volta. Infatti, se desiderate eseguire un modello di simulazione aggiungendo altre ipotesi o previsioni, dovreste mantenere lo stesso profilo.
- Profilo attivo — l'ultimo profilo usato quando salvate Excel sarà automaticamente aperta la prossima volta che aprirete il file Excel.
- Multipli file Excel — quando alternate tra vari modelli aperti in Excel, il profilo attivo sarà dal modello attuale ed attivo di Excel.
- Profili cartella di lavoro incrociata — fate attenzione se avete multipli file Excel aperti e se solo uno di questi file Excel ha un profilo attivo e se voi accidentalmente cambiate ad un altro file Excel ed impostate ipotesi e previsioni su questo file, le ipotesi e previsioni non potranno essere eseguite e saranno non valide.
- Cancellare profili — potete clonare profili esistenti e cancellare profili esistenti ma prendete nota che almeno un profilo deve esistere nel file Excel se cancellate i profili.

- Posizione del profilo — i profili che create (contenenti le ipotesi, le previsioni, le variabili decisionali, gli obiettivi, i vincoli, ecc.) sono salvati come un foglio di lavoro nascosto crittografato. Questa è la ragione per la quale, quando salvate il file del foglio di lavoro di Excel, viene salvato automaticamente anche il profilo.

SUGGERIMENTI: Scelta rapida clic destro e altri tasti di scelta rapida

- Clic destro — potete aprire il menu di scelta rapida di Simulatore di Rischio cliccando col tasto destro su una cella qualsiasi in Excel.

SUGGERIMENTI: Salvare

- Salvare il file Excel — questo salva le impostazioni, le ipotesi, le previsioni e le variabili decisionali del profilo e anche il vostro modello Excel (incluso tutti i report, i diagrammi e i dati estratti di Simulatore di Rischio).
- Salvare le impostazioni del diagramma — questo salva le impostazioni del diagramma di previsione in modo che le stesse impostazioni possano essere recuperate e applicate a digrammi di previsione futuri (usate le icone salve e apri nei digrammi di previsione).
- Salvare ed estrarre i dati simulati in Excel — questo estrae le ipotesi e le previsioni dell'esecuzione di una simulazione, ma anche il file Excel stesso dovrà essere salvato per poter salvare i dati per un futuro recupero.
- Salvare i dati e i diagrammi simulati in Simulatore di Rischio — usando la funzione Simulatore di Rischio, Estrai dati e salvando in un file **.RiskSim*, vi permetterà di riaprire il diagramma di previsione dinamico e attivo con gli stessi dati in un futuro momento senza dover rieseguire la simulazione.
- Salvare e generare report — i report della simulazione e altri report analitici sono estratti come separati fogli di lavoro nella vostra cartella di lavoro, ma anche l'intero file Excel dovrà essere salvato per poter salvare i dati per un futuro recupero.

SUGGERIMENTI: Tecniche di campionamento e simulazione

- Generatore di numeri casuali — sono supportati 6 generatori di numeri casuali (consultare il manuale dell'utente per dettagli) e, in generale, i due approcci che consigliamo di usare sono il metodo di default *ROV Simulatore di Rischio* e il metodo *Mescolata casuale sottrattiva avanzata*. Non applicate gli altri metodi a meno che il loro utilizzo non sia esplicitamente richiesto dal vostro modello o dalla vostra analisi e anche allora vi consigliamo di testare i risultati contro questi due approcci consigliati.

SUGGERIMENTI: Software Development Kit (SDK) e Librerie DLL

- SDK, DLL e OEM — tutte le analitiche in Simulatore di Rischio possono essere chiamate fuori da questo software e integrate in qualsiasi software proprietario dell'utente. Si prega di contattare admin@realoptionsvaluation.com per dettagli sull'uso del nostro Software Development Kit per accedere ai file analitici DLL (Dynamic Link Library).

SUGGERIMENTI: Avviare Simulatore di Rischio con Excel

- ROV Strumento di risoluzione dei problemi (Troubleshooter) — eseguite questo Troubleshooter per ottenere il HWID del vostro computer per scopi di concessione licenza, per visualizzare le impostazioni e i prerequisiti del vostro computer e per riattivare Simulatore di Rischio se fosse stato accidentalmente disattivato.
- Avvio di Simulatore di Rischio quando si avvia Excel — potete far avviare Simulatore di Rischio automaticamente ogni volta che viene avviato Excel o avviarlo manualmente dalla posizione di scelta rapida Start, Programmi, Real Options Valuation, Simulatore di Rischio. Questa preferenza può essere impostata dal menu Opzioni di Simulatore di Rischio.

SUGGERIMENTI: Simulazione super veloce

- Sviluppo del modello — se desiderate eseguire il vostro modello ad altissima velocità, può avere senso eseguire alcune esecuzioni di prova di simulazioni super veloci durante la costruzione del modello per assicurarvi che il prodotto finale eseguirà la simulazione super veloce. Non aspettate che il modello finale sia completo prima di testare a super velocità e dover poi riesaminare all'indietro per identificare dove si trovano le posizioni di collegamenti rotti o funzioni incompatibili.
- Velocità normale — se sussiste un dubbio, la simulazione a velocità normale funziona sempre.

SUGGERIMENTI: Analisi Tornado

- Analisi Tornado — l'analisi Tornado non dovrebbe mai essere eseguita una sola volta. è intesa come uno strumento diagnostico del modello, il che significa che idealmente dovrebbe essere eseguito più volte sullo stesso modello Per esempio, in un modello di grandi dimensioni, Tornado può essere eseguito per la prima volta usando tutte le impostazioni di default e tutti i precedenti dovrebbero essere mostrati (selezionate Mostra tutte le variabili). Questa singola analisi può risultare in un report di grandi dimensioni e in diagrammi Tornado lunghi (e potenzialmente brutti). Tuttavia, essa fornisce un ottimo punto d'inizio per determinare quanti di questi precedents sono considerati come critici fattori di successo (p. es., il diagramma Tornado può mostrare che le prime 5 variabili hanno un forte impatto sull'output, mentre le rimanenti 200 variabili hanno un minimo o nessun impatto), in qual caso, una seconda analisi tornado viene eseguita mostrando meno variabili (p. es., selezionate the Mostra le migliori 10 variabili se le prime 5 sono critiche, creando così un bel report ed un diagramma Tornado che mostra un contrasto tra i fattori chiave e i fattori meno critici. In altre parole, non dovreste mai visualizzare un diagramma Tornado con solo le variabili chiave senza mostrare qualche variabile meno critica come contrasto ai loro effetti sull'output).
- Valori di default — i punti di test di default possono essere aumentati da ±10% ad un valore più alto per testare per le non linearità (il diagramma Spider mostrerà linee non lineari e i diagrammi Tornado saranno asimmetrici verso un lato se gli effetti precedenti sono non lineari).
- Valori di zero e numeri interi — input con zero o con solo valori di numeri interi dovrebbero essere deselezionati in una analisi Tornado prima di eseguirla. Altrimenti la perturbazione percentuale potrebbe invalidare il vostro modello (p. es., se il vostro modello usa una tabella di ricerca dove Gen = 1, Feb = 2, Mar = 3 e così via, perturbare il valore 1 ad a +/- 10% fornisce 0,9 e 1,1 che non ha senso per il modello).
- Opzioni diagrammi — provate le varie opzioni dei diagrammi per trovare le migliori opzioni da attivare o disattivare nel vostro modello.

SUGGERIMENTI: Strumento di risoluzione dei problemi (Troubleshooter)

- ROV Strumento di risoluzione dei problemi (Troubleshooter) — eseguite questo Troubleshooter per ottenere il HWID del vostro computer per scopi di concessione licenza, per visualizzare le impostazioni e i prerequisiti del vostro computer e per riattivare Simulatore di Rischio se fosse stato accidentalmente disattivato.

INDICE