

Appendix A: Solution Key

This appendix provides brief solutions to some of the problems in the book. Often only the final answer or value is provided without any explanation or justification. Given the sometimes subjective nature of regression analysis, other equally valid answers can also exist. As well, although the author has strived to verify that the answers are correct, it is inevitable that one or two stray errors may appear. This should be borne in mind when comparing answers. If you believe that there is indeed an error in the solution key, please let the author know so that appropriate corrections can be made.

Chapter 1

- (1) F; (2) F; (3) T; (4) F; (5) T; (6) F; (7) T; (8) F; (9) T; (10) T; (11) F; (12) T; (13) F; (14) F; (15) T; (16) T; (17) F; (18) T; (19) T; (20) T
(21) (a) $\mu = 4.3$, mode = 5, median = 5; (b) $\sigma^2 = 5.12$, MAD = 1.84, range = 7;
(c) Q1 = 2, Q2 = 5, Q3 = 5
(24) Maybe!

Chapter 2

- (1) T; (2) F; (3) T; (4) T; (5) T; (6) F; (7) T; (8) T; (9) F; (10) F; (11) T; (12) T; (13) F; (14) T; (15) T; (16) T; (17) F
(18) (a) $\mathbb{S} = \{Q\heartsuit, Q\spadesuit, Q\clubsuit\}$; $\mathbb{F} = \{\{\}, \{Q\heartsuit\}, \{Q\spadesuit\}, \{Q\clubsuit\}, \{Q\heartsuit, Q\spadesuit\}, \{Q\heartsuit, Q\clubsuit\}, \{Q\spadesuit, Q\clubsuit\}, \{Q\heartsuit, Q\spadesuit, Q\clubsuit\}\}$; $P(\{\}) = 0$; $P(\{Q\heartsuit\}) = \frac{1}{3}$, $P(\{Q\spadesuit\}) = \frac{1}{3}$; $P(\{Q\clubsuit\}) = \frac{1}{3}$; $P(\{Q\heartsuit, Q\spadesuit\}) = \frac{2}{3}$; $P(\{Q\heartsuit, Q\clubsuit\}) = \frac{2}{3}$; $P(\{Q\spadesuit, Q\clubsuit\}) = \frac{2}{3}$; $P(\{Q\heartsuit, Q\spadesuit, Q\clubsuit\}) = 1$;
(b) $\frac{1}{3}$; (c) $\mu = 0$; $\sigma^2 = 2$.

Table A.1 Answers for question 27 in Chap. 2

| | $\hat{\mu}_1$ | $\hat{\mu}_2$ | $\hat{\mu}_3$ | $\hat{\mu}_4$ | $\hat{\mu}_5$ |
|------------|---------------|---------------|--------------------------|------------------------------------|----------------|
| Bias | 0 | 0 | 0.5μ | $\frac{1}{3}\mu$ | 0 |
| σ^2 | σ^2 | $0.5\sigma^2$ | $1.25\sigma^2$ | $0.2\bar{\sigma}^2$ | σ^2 / N |
| MSE | σ^2 | $0.5\sigma^2$ | $1.25\sigma^2 + 0.25\mu$ | $0.2\bar{\sigma}^2 + 0.1\bar{\mu}$ | σ^2 / N |

(20) (a) $\mathbb{S} = \{HH, HT, TH, TT\}$; $\mathbb{F} = \{\{\}, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\}, \{HH, TT\}, \{HT, TH\}, \{HT, TT\}, \{TH, TT\}, \{HH, HT, TH\}, \{HH, HT, TT\}, \{HH, TH, TT\}, \{HT, TH, TT\}, \{HH, HT, TH, TT\}\}$; $P(\{\}) = 0$, $P(\{HH\}) = 0.3$, $P(\{HT\}) = 0.3$, $P(\{TH\}) = 0.2$, $P(\{TT\}) = 0.2$, $P(\{HH, HT\}) = 0.6$, $P(\{HH, TH\}) = P(\{HH, TT\}) = P(\{HT, TH\}) = P(\{HT, TT\}) = 0.5$, $P(\{TH, TT\}) = 0.4$, $P(\{HH, HT, TH\}) = P(\{HH, HT, TT\}) = 0.8$, $P(\{HH, TH, TT\}) = P(\{HT, TH, TT\}) = 0.7$, $P(\{HH, HT, TH, TT\}) = 1$.

(21) (a) No; (b) No; (c) Yes, $\mu = 2.5$, $\sigma^2 = 25/12$, $E(|x|) = 2.5$.

(23) $E(2X - 4Y) = -6$; $E(3XY) = 6$; $E(X^2) = 3$.

(27) See Table A.1.

(29) (a) Yes; (b) No; (c) Yes; (d) No.

(34) (b) $\mu = 0.5\pi$; (c) $4.84 \leq \psi \leq 6.41$, yes; (d) sampled mean is not equal to the true value; (e) $P(0.25 < X < 0.5) = 0.046$.

Chapter 3

(1) F; (2) T; (3) F; (4) T; (5) F; (6) T; (7) T; (8) F; (9) F; (10) T; (11) F; (12) F; (13) T; (14) T; (15) F; (16) T; (17) F; (18) T; (19) T; (20) T

(23) The solution is given as

$$\vec{y} = \langle y_3, y_4, \dots, y_{100} \rangle^T$$

$$\mathcal{A} = \begin{bmatrix} -y_2 & -y_1 & u_2 & u_1 \\ -y_3 & -y_2 & u_3 & u_2 \\ \vdots & \vdots & \vdots & \vdots \\ -y_{99} & -y_{98} & u_{99} & u_{98} \end{bmatrix}$$

$$\vec{\beta} = \langle \alpha_1, \alpha_2, \beta_1, \beta_2 \rangle^T$$

(24) (c) $\hat{R} = 2.15 \Omega$, $\hat{\sigma} = 0.85$; (d) $2.1 \pm 0.2 \Omega$; (e) Yes; (f) Fit is not good, as the errors are increasing with current.

(25) (b) $\hat{R} = 2.044 \Omega$; (c) An examination of the residuals (especially as a function of the current, I) and how close to normality they are; (d) Instrument error is often proportional to the magnitude of the measured value.

(27) (a + b) $\beta_0 = 276 \pm 6.8$; $\beta_1 = -1.98 \pm 1.0$; $\beta_2 = 0.012 \pm 0.041$; $\beta_3 = 0.00018 \pm 0.00045$; (c) 229.5 ± 3.3 , 229.5 ± 8.5 ; (d) Yes; (e) $R^2 = 0.963$, F -score = 87.23, No.

(29) (b+c) $\ln \hat{K} = 5.3 \pm 0.45$, $\hat{b} = 0.54 \pm 0.28$; (d) $128 \leq K \leq 317$; \hat{b} stays the same; (e) $\hat{\sigma} = 0.103$, $R^2 = 0.8789$; (g) Yes; (h) $R = 0.467 \text{ kg} \cdot \text{m}^{-0.5} \cdot \text{s}^{-1}$; (i) $\hat{K} = 181 \pm 67$, $\hat{b} = 0.46 \pm 0.25$; $R^2 = 0.8905$, $\hat{\sigma} = 7.397$; (j) Yes, $R = 0.518 \text{ kg} \cdot \text{m}^{-0.5} \cdot \text{s}^{-1}$.

(31) (a) No, plot the errors for the two runs using different symbols; (c) Yes.

Chapter 4

(1) T; (2) T; (3) F; (4) F; (5) F; (6) F; (7) F; (8) F; (9) T; (10) T; (11) T; (12) T; (13) T; (14) F; (15) T; (16) T; (17) F; (18) F; (19) F; (20) T; (21) (a) No; (b) Yes.
 (23) (a) $I = ACD = CDF = AF$; (b) (two examples) $A = F = CD = ACDF$, $B = ABF = ABCD = ABCDF$; (c) III; (d) (one of many) $I = ABCE = ACDF = BDEF$.

(25) (a brief outline of the solution) A fractional factorial design with centre point replicates. Blocking and randomisation should also be considered.

(26) (a) $I = x_1x_2^2x_3 = x_1^2x_2x_3^2$; (b) $x_1 = x_1^2x_2^2x_3 = x_2x_3^2$, $x_2 = x_1x_3 = x_1^2x_2^2x_3^2$, $x_3 = x_1x_2^2x_3^2 = x_1^2x_2$, $x_1^2 = x_2^2x_3 = x_1x_2x_3^2$, $x_2^2 = x_1x_2x_3 = x_1^2x_3^2$, $x_3^2 = x_1x_2^2 = x_1^2x_2x_3$, $x_1x_2 = x_1^2x_3 = x_2^2x_3^2$, $x_2x_3 = x_1x_3^2 = x_1^2x_2^2$, $x_1x_2^2x_3 = x_1^2x_2x_3^2 = I$; (c) $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{33}x_3^2 + \beta_{12}x_1x_2 + \beta_{23}x_2x_3$

(29) (b) $E = -ABCD$; (c) V; (d) BD, AE, DE, BE, and E.

(31) (a) $y = 100 - 4x_1^2 - 12x_2^2 - 9x_3^2$; (b) 100.

Chapter 5

(1) T; (2) F; (3) T; (4) T; (5) F; (6) F; (7) F; (8) T; (9) T; (10) F; (11) T; (12) T; (13) T; (14) F; (15) T; (16) F; (17) T; (18) F; (19) T; (20) F; (24) AR(1).
 (25) seasonal MA(1) with $s = 3$ and a normal AR(4).

Chapter 6

(1) F; (2) T; (3) T; (4) F; (5) F; (6) t; (7) T; (8) T; (9) T; (10) T; (11) F; (12) F; (13) T; (14) F; (15) T; (16) F; (17) T; (18) F; (19) F; (20) T.

(24) left graph: 10, right graph: 5. (25) Model is not adequate. (26) Model is adequate.

References

- Ashley R (1988) On the relative worth of recent macroeconomic forecasts. *Int J Forecast* 4:363–376
- Barnett V, Lewis T (1994) *Outliers in statistical data*, 3rd edn. Wiley, Chichester
- Bloomfield P (2000) *Fourier analysis of time series: an introduction*, 2nd edn. Wiley, New York
- Box GE, Jenkins GM (1970) *Time series analysis, forecasting, and control*. Holden-Day, Oakland
- Box GE, Hunter WG, Hunter JS (1978) *Statistics for experimenters: an introduction to design, data analysis, and model building*. Wiley, New York
- Brillo J (2007) *Excel for scientists and engineers: numerical methods*. Wiley, Hoboken
- Chambers JM, Cleveland WS, Kleiner B, Tukey P (1983) *Graphical methods for data analysis*. Wadsworth International Group, Belmont
- Daniel C, Wood FS (1980) *Fitting equations to data*, 2nd edn. Wiley, New York
- Davies L, Gather U (1993) The identification of multiple outliers. *J Am Stat Assoc* 88 (423):782–792
- Dean JA (1999) *Lange's handbook of chemistry*, 15th edn. McGraw-Hill, Inc., New York
- Elliott JA, Prickett RC, Elmoazzen HY, Porter KR, McGann LE (2007) A multisolute osmotic virial equation for solutions of interest in biology. *J Phys Chem B* 111:1775–1785
- Franke J, Härdle WK, Hafner CM (2011) *Statistics of financial markets: an introduction*, 3rd edn. Springer, Heidelberg. doi:10.1007/978-3-642-16521-4
- Gerhart PM, Gross RJ, Hochstein JI (1992) *Fundamentals of fluid mechanics*. Addison-Wesley Publication Co., Reading
- Hald A (2003) *A history of probability and statistics and their application before 1750*. Wiley, Hoboken
- Hannan EJ, Deistler M (2012) *The statistical theory of linear systems*. Society of Industrial and Applied Mathematics, Philadelphia (Originally published: Wiley, New York, 1988)
- Hare LB (1988) In the soup: a case study to identify contributors to filling variability. *J Qual Technol* 20:36–43
- Harvey AC (1991) *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge
- Harvey G (2013) *Excel® 2013 All-in-One for dummies*. Wiley, Hoboken
- Hassler U (1994) The sample autocorrelation function of I(1) processes. *Statistical Papers—Statistische Hefte* 35:1–16
- Hawkins DM (1980) *Identification of outliers*. Chapman and Hall, London
- Hinkelmann K, Kempthorne O (2007) *Design and analysis of experiments*, vol I, II, III. Wiley, Hoboken

- Hodge VJ, Austin J (2004) A survey of outlier detection methodologies. *Artif Intell Rev* 22:85–126
- Huang B, Kadali R (2008) *Dynamic modeling, predictive control, and performance monitoring*. Springer, London
- Hunt BR, Lipsman RL, Rosenberg J (2014) *A guide to MATLAB: for beginners and experienced users: updated for MATLAB 8 and Simulink 8, 3rd edn*. Cambridge University Press, Cambridge, UK
- Hyndman RJ, Fan Y (1996) Sample quantiles in statistical packages. *Am Stat* 50(4):361–365
- Jochem M, Körber C (1987) Extended phase diagrams for the ternary solutions $\text{H}_2\text{O} - \text{NaCl} - \text{glycerol}$ and $\text{H}_2\text{O} - \text{NaCl} - \text{hydroxyethylstarch}$ (HES) determined by DSC. *Cryobiology* 24:513–536
- Johnson RA, Wichern DW (2007) *Applied multivariate statistical analysis, 6th edn*. Prentice Hall, Upper Saddle River
- Kallenberg O (2002) *Foundations of modern probability*. Springer, New York
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *Trans ASME J Basic Eng* 82:35–45
- Kalman RE, Bucy RS (1961) New results in filtering and prediction theory. *Trans ASME J Basic Eng* 83:95–108
- Khintchine A (1934) Korrelationstheorie der stationären stochastischen prozesse (correlation theory of stocastic processes). *Math Ann* 109(1):604–615. doi:[10.1007/BF01449156](https://doi.org/10.1007/BF01449156)
- Lin B, Recke B, Knudsen JK, Jørgensen SB (2007) A systematic approach for soft sensor development. *Comput Chem Eng* 31:419–425
- Ljung L (1999) *System identification theory for the user*. Prentice Hall, Inc., Upper Saddle River
- Montgomery DC (1991) *Design and analysis of experiments, 3rd edn*. Wiley, New York
- Montgomery DC, Peck EA (1982) *Introduction to linear regression analysis, 1st edn*. Wiley, New York
- Montgomery DC, Runger GC (2007) *Applied statistics and probability for engineers, 4th edn*. Wiley, Hoboken
- Montgomery DC, Jennings CL, Kulahci M (2008) *Introduction to time series analysis and forecasting*. Wiley, Hoboken
- Myers RH (1971) *Response surface methodology*. Allyn and Bacon, Inc., Boston
- Nelson CR (1972) The prediction performance of the FRB-MIT-PENN model of the U.S. Economy. *Am Econ Rev* 62(5):902–917
- Ogunnaik BA (2010) *Random phenomena: fundamentals of probability and statistics for engineers*. CRC Press, Boca Raton
- Prickett RC, Elliott JA, McGann LE (2010) Application of the osmotic virial equation in cryobiology. *Cryobiology* 2010:30–42
- Prickett RC, Elliott JA, McGann LE (2011) Application of the multisolute osmotic virial equation to solutions containing electrolytes. *J Phys Chem B* 115:14531–14543
- Priestley MB (1981) *Spectral analysis and time series: Vol. 1: univariate series and vol 2: multivariate series, prediction, and control*. Academic, New York
- Ross-Rodriguez LU (2009) *Cellular osmotic properties and cellular responses to cooling*. University of Alberta, Edmonton
- Schlittgen R (2012) *Einführung in die statistik analyse und modellierung von daten, 12th edn*. Oldenbourg Wissenschaftsverlag GmbH, Munich
- Schmuller J (2013) *Statistical analysis with Excel® for dummies, 3rd edn*. Wiley, Hoboken
- Seber GA, Wild CJ (1989) *Nonlinear regression*. Wiley, New York
- Shardt Y (2012a) *Data quality assessment for closed-loop system identification and forecasting with application to soft sensors*. Doctoral thesis, University of Alberta, Department of Chemical and Materials Engineering, Edmonton, Alberta, Canada. doi: <http://hdl.handle.net/10402/era.29018>

- Shardt YA (2012b) Data quality assessment for closed-loop system identification and forecasting with application to soft sensors. University of Alberta, Edmonton. doi: <http://hdl.handle.net/10402/era.29018>
- Shardt YA, Huang B (2011) Closed-loop identification with routine operating data: effect of time delay and sampling time. *J Process Control* 21:997–1010
- Shardt YA, Huang B (2014) Minimal required excitation for closed-loop identification: implications for PID control loops. In: ADCONIP conference proceedings, Hiroshima, Japan, pp 296–301. doi:http://www.nt.ntnu.no/users/skoge/prost/proceedings/adconip-2014/pdf/SUBS61TO80/0069/0069_FI.pdf
- Sheynin O (2004) History of the theory of probability to the beginning of the 20th century. NG Verlag, Berlin
- Shumway RH, Stoffer DS (2011) Time series analysis and its applications with R examples, 3rd edn. Springer, New York. doi:10.1007/978-1-4419-7865-3
- Sizemore J, Mueller JP (2015) MATLAB for dummies. Wiley, Hoboken
- Söderström T, Gustavsson I, Ljung L (1975) Identifiability conditions for linear systems operating in closed loop. *Int J Control* 21(2):243–255
- Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. W. H. Freeman and Company, New York
- Steel RG, Torrie JH (1980) Principles and procedures of statistics: a biometrical approach, 2nd edn. McGraw-Hill Kogakusha, Ltd., Tokyo
- Stoica P, Moses R (2005) Spectral analysis of signals. Prentice Hall, Upper Saddle River
- Tufte ER (1997) Visual and statistical thinking: displays of evidence for making decisions. Graphics Press LLC., Cheshire
- Tufte ER (2001) The visual display of quantitative information. Graphics Press LLC., Cheshire
- Varberg DE (1963) The development of modern statistics. *The Mathematics Teacher* 56(4): 252–257
- Welch PD (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified Periodograms. *IEEE Trans Audio Electroacoustics* AU-15(2):70–73
- Wichern DW (1973) The behaviour of the sample autocorrelation function for an integrated moving average process. *Biometrika* 60(2):235–239. doi:<http://www.jstor.org/stable/2334535?origin=JSTOR-pdf>
- Zhu Y (2001) Multivariable system identification for process control. Elsevier Science Ltd., Oxford
- Zorn ME, Gibbons RD, Sonzogni WC (1997) Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. *Anal Chem* 69(15):3069–3075

Subject Index

An index of Excel and Matlab topics is listed as a separate index located after this one on p. 413.

A

ACF. *See* Autocorrelation function (ACF)
Akaike's information criterion (AIC), 251, 297
Aliased. *See* Confounded
 α -error, 59, 60, 99, 255
Alternative hypothesis, 58, 59, 68, 73, 76, 113
Analysis of variance (ANOVA), 2, 99–102, 110, 146, 170
AR, *See* Autoregressive model.
ARIMA. *See* Autoregressive, integrating, moving average model (ARIMA)
ARMA. *See* Autoregressive moving average model (ARMA)
Autocorrelation function (ACF), 213, 214, 225, 230, 237, 238, 240, 244, 245, 250, 296
Autocovariance, 37, 213–215, 223, 224, 226–229, 231, 237, 241, 243, 246, 248, 260, 261, 276, 299
Autoregressive exogenous model (ARX), 287, 347
Autoregressive, integrating, moving average model (ARIMA), 235–237, 276
Autoregressive model (AR), 221, 223, 240–244, 247, 275, 287, 347
Autoregressive moving average exogenous model (ARMAX), 286, 287, 309, 347
Autoregressive moving average model (ARMA), 222, 235–237, 240, 245–250, 254, 271, 275–276

B

Backshift operator, 220, 286, 293
Bandwidth, 299, 300, 312, 313, 322
Bar chart, 9–10
Bayesian information criterion (BIC), 251
Bernoulli distribution, 49
 β -error, 59, 60
Bias, 51, 53, 57, 82, 111
 unbiased, 51
Binomial distribution, 48–50, 68, 77
Black-box model, 145, 147, 202, 283
Blocking, 141, 143, 144, 176–178, 202
Box-and-whisker plot. *See* Boxplot
Box-Jenkins model, 286, 315, 322, 347, 361
Boxplot, 12, 338, 378
 Tukey, 12

C

Causal, 222, 223, 227, 229, 234–236, 267, 275, 276
cdf. *See* Cumulative distribution function (cdf)
Central composite design, 199–200, 203
Central limit theorem, 58, 64, 84
Centre point design, 193, 195
Chi-squared distribution (χ^2), 2, 46–47, 67, 80, 83, 251, 341
Complete defining relationship, 161–168, 170–178, 203–205

Confidence interval, 58–78, 96, 146, 215, 296, 343, 394
 converting, 62–64, 124
 cover, 61
 Confounded, 142–144, 158–159, 161–164, 166, 167, 171, 193, 203
 Confounding, 141–143, 158–166, 170, 177, 202, 204, 205, 344
 pattern, 158, 159, 161–166, 170, 204, 205, 344
 Consistency, 51–52, 194
 Correlation, 20, 39, 89, 108, 138, 139, 141–143, 200, 213–215, 232, 273, 294, 308, 316, 338, 339, 346, 348, 349, 352
 Covariance, 37, 39, 42, 45, 79, 95, 96, 213–215, 217, 228, 249, 266–268, 270, 294, 343, 344
 matrix, 39, 42, 45, 95, 268, 270, 343, 344
 Cramér–Rao lower bound, 52, 55, 299. *See also* Maximum likelihood method; Method of moments; Regression
 Cross-correlation plot, 20, 217, 296, 300–302, 315–319, 322, 323, 338, 352–353
 Cross-covariance, 213–215, 217, 228, 249
 Cumulative distribution function (cdf), 15, 32–34, 44, 46–48

D

Data-driven modelling, 283, 325. *See also* Grey-box model
 Data scaling, 113
 Data sets, 2–8, 10, 12, 13, 15–17, 19, 21–24, 26–30, 52, 54, 57, 70, 71, 79, 80, 84, 87, 111–112, 119, 124, 127, 131, 133, 134, 154, 171, 195, 202, 206, 212, 216, 218, 219, 239, 240, 246, 249, 251, 254, 255, 257, 259, 271, 274, 277–282, 291, 296–298, 303, 313, 315, 316, 321, 324–336, 342, 349, 352, 356, 358, 361, 362, 374, 375, 383, 392, 395, 398
 friction factor, 21, 23
 Data visualisation, 1–30
 Defining relationship, 160–168, 170, 171, 177, 178, 203–205
 Differencing, 216, 220, 266, 272, 302, 348
 periodic, 216
 true, 216
 Direct identification, 306–312. *See also* System identification, closed-loop
 Distribution, 2, 8, 13, 15–18, 43–50, 53, 55, 57, 58, 64, 67, 68, 77, 80, 83, 89, 92, 96, 156, 246, 250, 251, 341, 342, 370.

See also Bernoulli distribution; Binomial distribution; Chi-squared distribution (χ^2); *F*-distribution; Normal distribution; Poisson distribution; Student's *t*-distribution
 Durbin–Levinson Algorithm, 242, 257–259

E

Edmonton temperature, 249, 264–266, 276, 324
 Effects, 16, 17, 70, 134, 142–145, 147, 148, 152–154, 156, 158, 159, 169, 176, 197, 202, 209, 214, 216, 260, 304, 321, 377.
See also Factorial design, influential parameters
 Error, 8, 45, 89, 113, 141, 221, 286, 344, 373.
See also Residual
 Error structure, 89–92, 112, 113, 131, 132, 136, 137, 306
 Expectation operator, 37, 96
 Experimental design, 126, 142, 147, 203, 205, 207
 Externally excited data, 303, 312. *See also* System identification, closed-loop
 Extrapolation, 126, 127, 131

F

Factor, 6, 49, 88, 111, 141, 147, 291, 348
 Factorial design, 147–160, 166, 169–200, 202–204, 207–209, 386–388, 395–398, 405. *See also* Orthogonal basis
 generalised, 178–192
 influential parameters, 152
 projection, 152
 False negative rate. *See* β -error
 False positive rate. *See* α -error
F-distribution, 17, 47–48, 83, 341
 Final prediction error criterion (FPE), 297, 298
 First-principle model, 145, 283
 Forecasting, 145, 253–259, 267, 273, 274, 301.
See also Prediction
 Forward shift operator, 219, 220
 Fourier transform, 212, 259–262, 265, 274, 309, 383
 Fractional factorial design, 157–177, 202.
See also Factorial design

G

Generator, 160, 161, 163, 165–167, 170, 171, 177, 203–206
 Grey-box model, 145, 283, 322

H

- Histogram, 6, 9–11, 19, 20, 28, 29, 338, 339
- Hypothesis testing, 2, 58–78, 94
 - mean, 65
 - one-sided, 59, 73, 74
 - ratio, 69
 - two samples, 69, 77
 - two-sided, 59
 - variance, 2, 94

I

- Identifiability, 298, 304, 321
- Impulse response model, 222
- Independent, 37, 40, 42, 50, 54, 55, 69, 79, 81, 82, 89, 94, 96, 99, 108, 147, 160, 166, 167, 194, 204, 213, 220, 223, 224, 228, 250, 266, 294, 296
- Indirect identification, 309. *See also* System identification, closed-loop
- Innovation, 247, 268, 270. *See also* Maximum likelihood method, for ARMA model
- Integrated/integrating model, 19, 20, 221–222, 274
- Interpolation, 7, 126
- Invertible, 94, 141, 204, 222, 223, 227, 235, 236, 275–276, 298

J

- Joint input-output identification, 308–313. *See also* System identification, closed-loop

K

- Kalman filter, 211, 267–270, 274–276. *See also* Time series, state-space approach
- Kolmogorov's axioms, 31

L

- Least-squares regression, 93–99, 102–104, 116–120, 127, 137–139. *See also* Regression
 - ordinary, 93–99, 102–104, 116–120, 127
 - weighted, 116–120
- Left probability, 63, 66, 366
- Level, 30, 59, 66, 75, 80, 84, 114, 147–151, 158, 159, 168, 169, 178, 180–187, 189–191, 197, 202–204, 207, 208, 283, 299, 310–320, 322, 324–336, 361, 367

- Line chart, 8, 10–12
 - time series, 11

- Ljung–Box–Pierce Q-Statistic, 250–252, 349

M

- MA. *See* Moving average model (MA)
- Main effects, 147, 148, 158
- Maximum likelihood method, 52, 54–57, 241, 248, 270. *See also* Kalman filter; Time series, state-space approach
 - for ARMA model, 222, 240, 276
 - log-likelihood function, 56, 246–248, 270
 - state-space approach, 211, 212, 266
- Mean, 3, 4, 23, 33, 36, 44, 46–50, 57, 64, 66, 69–76, 88, 130–131, 147, 212, 285, 338, 370
 - average, 364
 - conditional, 39
 - response, 96–99, 103, 106, 117, 123, 127, 130–131, 138, 140, 147–150, 161, 269
- Mean square error (MSE), 51, 82
- Measures of central tendency, 3–4, 6. *See also* Mean; Median; Mode
- Measures of dispersion, 4–6. *See also* Range; Skew; Standard deviation; Variance
- Median, 3–6, 8, 12, 23–26, 28, 29, 338, 342, 378, 383. *See also* Quantile
- Median absolute deviation (MAD), 5, 6
- Method of moments, 52–54, 57, 82, 241, 275
- Mode, 3, 4, 18, 29, 285
- Model validation, 88, 107–116, 118, 123, 124, 240, 250–252, 271, 296–298, 308, 314, 323, 324, 346
 - model adequacy, 110–112, 114, 135, 297, 298
 - multicollinearity, 113, 133
 - residual testing, 107, 108
- Moment, 34, 42, 43, 52–54, 56, 57, 82, 213, 241, 275
 - centred, 34
 - uncentred, 34
- Moving average model (MA), 221–223, 240, 249

N

- Normal distribution, 6, 8, 13, 15–18, 43–46, 49, 50, 53–58, 64, 81, 83, 84, 89, 92, 96, 108, 156, 246, 250, 342, 366
 - Gaussian distribution, 2, 43, 220, 266
 - standard normal distribution, 17, 18, 44, 49, 83, 366

Normalisation, 36, 44, 114
 Null hypothesis, 58–60, 65, 66, 68–70, 72,
 74–78, 252

O

Optimal design, 201, 203
 Order of a model, 148, 220
 Order of an interaction, 148
 Orthogonal basis, 101, 146, 151, 159, 168, 170,
 178–184, 193–196, 205
 Orthogonality, 142, 179
 Orthonormal, 146, 147, 151, 169, 170,
 179–182, 184–186, 190, 191, 196
 Outlier, 6, 8, 12, 16, 17, 27, 28, 108, 110, 111,
 114, 116, 132, 152, 266, 356, 390–392
 Output-error model (OE), 287, 322, 347

P

PACF. *See* Partial autocorrelation function
 (PACF)
 Parameter, 16, 43, 50–58, 87, 88, 141, 211,
 240, 245–250, 283, 292–296, 343, 370
 Parameter estimation, 50–57, 240, 270,
 292–296
 Cramér–Rao lower bound, 52, 55, 295
 efficiency, 52
 Partial autocorrelation function (PACF), 214,
 217, 218, 233–241, 244–245, 249, 257,
 259, 275–277, 349
 Partial correlation, 214, 232, 240
 pdf. *See* Probability density function (pdf)
 Pearson's coefficient of regression, 110,
 118, 251
 PEM. *See* Prediction error model (PEM)
 Periodogram, 262–266, 271, 274, 275, 277,
 346, 349, 350, 375, 383–385
 Persistency of excitation, 298, 299, 305, 306,
 322
 Pie chart, 10, 11, 28, 339
 Poisson distribution, 43, 50
 Power spectrum, 259
 Prediction, 96, 110, 139, 140, 201, 203,
 253–259, 267–270, 272, 274, 286, 287,
 289–300, 302–304, 306, 307, 309, 321,
 322, 343, 344, 350. *See also* Forecasting
 Prediction error model (PEM), 287, 292, 298,
 299, 302–304, 306, 322, 344, 346
 Probability, 1, 31, 108, 146, 246, 339, 370
 conditional, 38, 40, 41, 246, 247
 event, 32, 33, 48
 function, 31, 32, 49, 50, 342

 independent, 54, 79
 marginal, 38
 outcome, 32, 33, 38, 81
 Probability density function (pdf), 33–36,
 38–40, 43, 44, 46–48, 54, 55, 58, 59,
 79–81, 84
 Probability function, 31, 32, 49, 50, 342
 Probability plot, 13–18, 29, 44, 108, 109, 114,
 115, 125, 126, 128–130, 135, 136, 146,
 152–154, 156, 157, 170–173, 175, 176,
 191, 192, 197, 198, 250, 252, 272, 357,
 360, 377, 378, 386, 388, 394
 normplot, 339, 356, 359, 362, 377
 P–P plots, 13
 Q–Q plot, 13

Q

Quantile, 6–7, 13–15, 27, 342
 percentile, 12
 quartile, 7, 27

R

Random binary signal, 299, 311, 312, 322, 348
 Randomisation, 141, 145, 202
 Random variable, 32, 37, 43, 51, 213, 220
 Range, 2, 5, 10, 12, 19, 23, 24, 26, 28, 29, 40,
 122, 126, 145, 188, 233, 262, 266, 267,
 283, 299, 348, 363–365, 369, 375, 378,
 386, 394
 Regression, 2, 31, 87–140, 211, 284, 342, 373
 ANOVA, 110, 114
 deterministic component, 89, 90
 Fisher information matrix, 295, 296
 F-statistic, 101, 104, 107, 110, 128, 129,
 152, 169, 191, 197
 matrix, 93, 148, 150, 153, 183, 199, 200,
 204, 205, 207, 344
 mean response, 96, 97, 117, 123, 150
 model, 88, 90, 99–102, 110, 116, 120, 284,
 343, 344
 nonlinear, 90–93, 120–126, 131, 132, 342,
 343, 354, 358–360, 362, 386–388,
 391–395, 398
 output, 89, 147, 358
 predicted response, 96
 R², 101, 118
 regressor, 89, 112, 113, 132, 142, 152, 169,
 358–360
 SSE, 100–104, 106, 118, 123, 152, 174,
 190, 197
 SSR, 100, 103, 118, 123

- stochastic component, 89, 90, 284
 - TSS, 99–101, 104, 118, 123, 152, 169, 174
 - weighted, 116–120, 133, 342, 344
- Regressor selection, 112–113
 - backward selection, 112
 - forward selection, 112
- Replicates, 23, 110, 113, 118–119, 132, 145–147, 149, 150, 153, 154, 169–171, 173, 175, 176, 178, 187, 192–198, 202, 207–209
- Residual, 24, 90, 108–110, 146, 240, 296, 343, 390. *See also* Error
- Resolution, 158, 160, 162, 163, 166–168, 171, 178, 203–206. *See also* Fractional factorial design
- Response surface methodology, 202, 203
- Right probability, 62, 63
- Routine operating data, 303, 308. *See also* System identification, closed-loop

- S**
- Sample, 3, 31, 76–77, 108, 113, 144, 212, 301, 338, 370
- Sample space, 31, 32, 59, 66, 338
- Sampling time, 299, 312, 313, 321, 322, 345, 346
- SARIMA, 220, 276
- Scatter plot, 8, 13, 14, 20, 24, 25, 27, 108, 109, 251, 339, 378
- Screening experiment, 203
- Seasonal model, 346
- Sensitivity, 6, 141, 142, 171, 295, 306
- Set theory, 31, 84–85
- Signal-to-noise ratio, 142
- Skew, 4–7, 16, 20, 28
 - negative skewness, 6
 - positive skewness, 6
- Sparkplot (profile plot), 19, 28
- Spectral analysis, 274
- Spectral density, 259, 261, 262, 276
- Standard deviation, 5, 6, 8, 15, 18, 23, 24, 26, 28, 33, 36, 43, 45, 46, 53, 57, 61, 64–66, 70, 71, 82, 83, 96, 102, 104–106, 114, 117, 119, 122, 138, 140, 151, 168, 209, 212, 215, 249, 250, 271, 315, 338, 341, 342, 344, 348, 366, 377
- Stationary, 212–214, 216, 222, 233, 240, 275, 302
 - jointly, 214
- Step test, 300, 301, 311, 322
- Student's *t*-distribution, 16, 45–46, 64, 83
- System identification, 2, 44, 274, 283–336, 344–348, 354, 361–362
 - basis function, 180, 181, 184, 186, 189, 191, 194–196, 309, 310, 318, 322, 347
 - closed-loop, 274, 285, 306
 - linearity tests, 302–303
 - nonlinear, 309–311
 - open-loop, 285, 298, 300
 - proportional test, 303
 - step-up and step-down, 302–303
 - time invariance, 300, 303

- T**
- Table, 18, 19, 28, 62, 155, 166
- Test statistic, 59–61, 64, 67, 68, 70–72, 74, 76–78, 80, 250, 349
- Time delay, 286, 289, 293, 294, 297, 300–302, 304, 306, 307, 309, 312, 314, 319, 321–323, 345, 347
 - frequency approach, 212
- Time series, 11, 108, 173, 211–282, 339, 390
 - state-space approach, 211, 212, 266, 275
 - transfer function approach, 211, 212, 266
- Treatment, 19, 147, 149, 166, 178–185, 194, 209, 373
- Tukey boxplot, 12. *See also* Boxplot
- Type I error. *See* α -error
- Type II error. *See* β -error

- U**
- Unbiased, 51, 53, 54, 57, 80–82, 95, 97, 98, 132, 159, 215, 275, 293, 295, 296, 307, 322

- V**
- Variance, 5, 33, 36, 51, 57, 67–68, 72, 74, 92, 99–102, 146, 213, 285, 344

- W**
- Water level, 310–320, 324–336, 362
- White-box model, 145, 283
- White noise, 172, 223–225, 231, 240, 246, 250–252, 255, 261, 262, 264, 266, 271, 276, 284, 285, 290, 293–296, 299, 301, 307, 314, 322
- Wiener-Hammerstein, 310
- Word, 160–164, 166, 167

- Y**
- Yule-Walker Equations, 241–244, 248, 275

Index of Excel and MATLAB Topics

This index only considers Excel and Matlab functions that are used in this book. The main index is found on p. 407.

E

Excel

- absolute, 364
- array function, 365
- macros, 366–368, 385, 386
- range, 363–365
- relative, 364
- security, 366–368

Excel 2003/7

- binomdist, 49
- chiinv, 47
- finv, 48
- normdist, 44
- norminv, 44
- poissondist, 50
- tinv, 46

Excel Add-ins

- Data Analysis Add-in, 374–376
- Solver Add-in, 368

Excel Function

- average, 363
- binom.dist, 49
- binom.inv, 49
- chisq.inv, 47
- chisq.pdf, 47
- count, 366
- f.dist, 48
- f.inv, 48, 63, 366
- mdeterm, 365
- minverse, 365
- mmult, 365

- norm.dist, 44
- norm.inv, 44, 63, 366
- norm.s.inv, 366
- poisson.dist, 50
- quartile, 27
- quartile.inc, 26, 27, 379
- rank, 366
- stdev, 344, 366
- sum, 366
- t.dist, 46
- t.inv, 46, 63, 366
- transpose, 365

Excel Template

- boxplot2007.xlt, 378
- boxplot.xltx, 378
- factorialdesigntemplate.xltx, 388
- linearregression.xltx, 385
- nonlinearregression.xltx, 386
- normplot.xltx, 377
- periodogram.xltx, 383

M

MATLAB Function

- ar, 347
- arma, 348
- armax, 347
- arx, 347
- autocorr, 349
- bar, 338

MATLAB Function (*cont.*)

barh, 338
 binoinv, 49
 binopdf, 49
 binornd, 49
 bj, 347
 boxplot, 338
 chi2pdf, 47, 341
 chi2rnd, 47
 colorbar, 338
 colormap, 338
 compare, 348
 cordexch, 344
 cra, 347
 crosscorr, 349
 delayest, 347
 detrend, 346
 estimate, 349
 ff2n, 344
 finv, 48, 63, 341
 fpdf, 48, 341
 fracfact, 344
 frnd, 48
 hist, 338
 iddata, 346
 idinput, 348
 idplot, 346
 idpoly, 346
 imagesc, 338
 infer, 349
 lbqtest, 349
 legend, 339
 lscov, 344
 mad, 342
 mean, 338
 median, 338
 nlarx, 347
 nlinfit, 343
 nlparci, 343
 nlpredci, 343
 norminv, 44, 63, 341
 normpdf, 44
 normplot, 339
 normrnd, 44
 oe, 347
 parcorr, 349
 periodogram, 349
 pie, 339
 plot, 339
 plot3, 339
 plotmatrix, 339
 poissinv, 50
 poisspdf, 50
 poissrnd, 50
 present, 348
 quantile, 342
 rand, 342
 randn, 342
 randperm, 342
 regress, 342, 343
 resid, 348
 scatter, 339
 set, 340
 std, 338
 surf, 340
 tinva, 341
 title, 340
 tpdf, 341
 xcorr, 349
 xlabel, 340
 ylabel, 44
 zlabel, 340
 zscore, 342

MATLAB Recipe

autocorrelation.m, 351
 corrplot1.m, 352
 crosscorrelation.m, 352
 periodogram2.m, 350