

# **Modern Industrial Statistics**

with applications in R, MINITAB and JMP

**Second Edition**

**RON S. KENETT**

Chairman and CEO, the KPA Group, Israel  
Research Professor, University of Turin, Italy, and  
International Professor, NYU, Center for Risk Engineering, New York, USA

**SHELEMYAHU ZACKS**

Distinguished Professor,  
Binghamton University, Binghamton, USA

With contributions from

**DANIELE AMBERTI**

Turin, Italy

**WILEY**

# Contents

<i>Preface to Second Edition</i>	<i>xv</i>
<i>Preface to First Edition</i>	<i>xvii</i>
<i>Abbreviations</i>	<i>xix</i>
<b>PART I PRINCIPLES OF STATISTICAL THINKING AND ANALYSIS</b>	<b>1</b>
<b>1 The Role of Statistical Methods in Modern Industry and Services</b>	<b>3</b>
1.1 The different functional areas in industry and services	3
1.2 The quality-productivity dilemma	5
1.3 Fire-fighting	6
1.4 Inspection of products	7
1.5 Process control	7
1.6 Quality by design	8
1.7 Information quality and practical statistical efficiency	9
1.8 Chapter highlights	11
1.9 Exercises	12
<b>2 Analyzing Variability: Descriptive Statistics</b>	<b>13</b>
2.1 Random phenomena and the structure of observations	13
2.2 Accuracy and precision of measurements	17
2.3 The population and the sample	18
2.4 Descriptive analysis of sample values	19
2.4.1 Frequency distributions of discrete random variables	19
2.4.2 Frequency distributions of continuous random variables	23
2.4.3 Statistics of the ordered sample	26
2.4.4 Statistics of location and dispersion	28
2.5 Prediction intervals	32
2.6 Additional techniques of exploratory data analysis	32
2.6.1 Box and whiskers plot	33
2.6.2 Quantile plots	34
2.6.3 Stem-and-leaf diagrams	34
2.6.4 Robust statistics for location and dispersion	36
2.7 Chapter highlights	38
2.8 Exercises	38
<b>3 Probability Models and Distribution Functions</b>	<b>41</b>
3.1 Basic probability	41
3.1.1 Events and sample spaces: Formal presentation of random measurements	41
3.1.2 Basic rules of operations with events: Unions, intersections	42
3.1.3 Probabilities of events	44

3.1.4	Probability functions for random sampling	46
3.1.5	Conditional probabilities and independence of events	47
3.1.6	Bayes formula and its application	49
3.2	Random variables and their distributions	51
3.2.1	Discrete and continuous distributions	51
3.2.2	Expected values and moments of distributions	55
3.2.3	The standard deviation, quantiles, measures of skewness and kurtosis	57
3.2.4	Moment generating functions	59
3.3	Families of discrete distribution	60
3.3.1	The binomial distribution	60
3.3.2	The hypergeometric distribution	62
3.3.3	The Poisson distribution	65
3.3.4	The geometric and negative binomial distributions	67
3.4	Continuous distributions	69
3.4.1	The uniform distribution on the interval $(a, b), a < b$	69
3.4.2	The normal and log-normal distributions	70
3.4.3	The exponential distribution	75
3.4.4	The gamma and Weibull distributions	77
3.4.5	The Beta distributions	80
3.5	Joint, marginal and conditional distributions	82
3.5.1	Joint and marginal distributions	82
3.5.2	Covariance and correlation	84
3.5.3	Conditional distributions	86
3.6	Some multivariate distributions	88
3.6.1	The multinomial distribution	88
3.6.2	The multi-hypergeometric distribution	89
3.6.3	The bivariate normal distribution	90
3.7	Distribution of order statistics	92
3.8	Linear combinations of random variables	94
3.9	Large sample approximations	98
3.9.1	The law of large numbers	98
3.9.2	The Central Limit Theorem	99
3.9.3	Some normal approximations	99
3.10	Additional distributions of statistics of normal samples	101
3.10.1	Distribution of the sample variance	101
3.10.2	The "Student" $t$ -statistic	102
3.10.3	Distribution of the variance ratio	102
3.11	Chapter highlights	104
3.12	Exercises	105
<b>4</b>	<b>Statistical Inference and Bootstrapping</b>	<b>113</b>
4.1	Sampling characteristics of estimators	113
4.2	Some methods of point estimation	114
4.2.1	Moment equation estimators	115
4.2.2	The method of least squares	116
4.2.3	Maximum likelihood estimators	118
4.3	Comparison of sample estimates	120
4.3.1	Basic concepts	120
4.3.2	Some common one-sample tests of hypotheses	122
4.4	Confidence intervals	128

4.4.1	Confidence intervals for $n \setminus a$ known	129
4.4.2	Confidence intervals for $n \setminus a$ unknown	130
4.4.3	Confidence intervals for $a^2$	130
4.4.4	Confidence intervals for $p$	130
4.5	Tolerance intervals	132
4.5.1	Tolerance intervals for the normal distributions	132
4.6	Testing for normality with probability plots	134
4.7	Tests of goodness of fit	137
4.7.1	The chi-square test (large samples)	137
4.7.2	The Kolmogorov-Smirnov test	139
4.8	Bayesian decision procedures	140
4.8.1	Prior and posterior distributions	141
4.8.2	Bayesian testing and estimation	144
4.8.3	Credibility intervals for real parameters	147
4.9	Random sampling from reference distributions	148
4.10	Bootstrap sampling	150
4.10.1	The bootstrap method	150
4.10.2	Examining the bootstrap method	151
4.10.3	Harnessing the bootstrap method	152
4.11	Bootstrap testing of hypotheses	152
4.11.1	Bootstrap testing and confidence intervals for the mean	153
4.11.2	Studentized test for the mean	153
4.11.3	Studentized test for the difference of two means	155
4.11.4	Bootstrap tests and confidence intervals for the variance	157
4.11.5	Comparing statistics of several samples	158
4.12	Bootstrap tolerance intervals	161
4.12.1	Bootstrap tolerance intervals for Bernoulli samples	161
4.12.2	Tolerance interval for continuous variables	163
4.12.3	Distribution-free tolerance intervals	164
4.13	Non-parametric tests	165
4.13.1	The sign test	165
4.13.2	The randomization test	166
4.13.3	The Wilcoxon Signed Rank test	168
4.14	Description of MINITAB macros (available for download from Appendix VI of the book website)	170
4.15	Chapter highlights	170
4.16	Exercises	171
<b>5</b>	<b>Variability in Several Dimensions and Regression Models</b>	<b>177</b>
5.1	Graphical display and analysis	177
5.1.1	Scatterplots	177
5.1.2	Multiple boxplots	179
5.2	Frequency distributions in several dimensions	181
5.2.1	Bivariate joint frequency distributions	182
5.2.2	Conditional distributions	185
5.3	Correlation and regression analysis	185
5.3.1	Covariances and correlations	185
5.3.2	Fitting simple regression lines to data	187
5.4	Multiple regression	192
5.4.1	Regression on two variables	194
5.5	Partial regression and correlation	198

5.6	Multiple linear regression	200
5.7	Partial F-tests and the sequential SS	204
5.8	Model construction: Step-wise regression	206
5.9	Regression diagnostics	209
5.10	Quantal response analysis: Logistic regression	211
5.11	The analysis of variance: The comparison of means	213
5.11.1	The statistical model	213
5.11.2	The one-way analysis of variance (ANOVA)	214
5.12	Simultaneous confidence intervals: Multiple comparisons	216
5.13	Contingency tables	220
5.13.1	The structure of contingency tables	220
5.13.2	Indices of association for contingency tables	223
5.14	Categorical data analysis	227
5.14.1	Comparison of binomial experiments	227
5.15	Chapter highlights	229
5.16	Exercises	230
<b>PART II</b>	<b>ACCEPTANCE SAMPLING</b>	<b>235</b>
<b>6</b>	<b>Sampling for Estimation of Finite Population Quantities</b>	<b>237</b>
6.1	Sampling and the estimation problem	237
6.1.1	Basic definitions	237
6.1.2	Drawing a random sample from a finite population	238
6.1.3	Sample estimates of population quantities and their sampling distribution	239
6.2	Estimation with simple random samples	241
6.2.1	Properties of $X_n$ and $V$ ; under RSWR	242
6.2.2	Properties of $X_n$ and $S_n^2$ under RSWOR	244
6.3	Estimating the mean with stratified RSWOR	248
6.4	Proportional and optimal allocation	249
6.5	Prediction models with known covariates	252
6.6	Chapter highlights	255
6.7	Exercises	256
<b>7</b>	<b>Sampling Plans for Product Inspection</b>	<b>258</b>
7.1	General discussion	258
7.2	Single-stage sampling plans for attributes	259
7.3	Approximate determination of the sampling plan	262
7.4	Double-sampling plans for attributes	264
7.5	Sequential sampling	267
7.6	Acceptance sampling plans for variables	270
7.7	Rectifying inspection of lots	272
7.8	National and international standards	274
7.9	Skip-lot sampling plans for attributes	276
7.9.1	The ISO 2859 skip-lot sampling procedures	276
7.10	The Deming inspection criterion	278
7.11	Published tables for acceptance sampling	279
7.12	Chapter highlights	280
7.13	Exercises	281

<b>PART III STATISTICAL PROCESS CONTROL</b>	<b>283</b>
<b>8 Basic Tools and Principles of Process Control</b>	<b>285</b>
8.1 Basic concepts of statistical process control	285
8.2 Driving a process with control charts	294
8.3 Setting up a control chart: Process capability studies	298
8.4 Process capability indices	300
8.5 Seven tools for process control and process improvement	302
8.6 Statistical analysis of Pareto charts	305
8.7 The Shewhart control charts	308
8.7.1 Control charts for attributes	309
8.7.2 Control charts for variables	311
8.8 Chapter highlights	316
8.9 Exercises	316
<b>? Advanced Methods of Statistical Process Control</b>	<b>319</b>
9.1 Tests of randomness	319
9.1.1 Testing the number of runs	319
9.1.2 Runs above and below a specified level	321
9.1.3 Runs up and down	323
9.1.4 Testing the length of runs up and down	324
9.2 Modified Shewhart control charts for $\bar{X}$	325
9.3 The size and frequency of sampling for Shewhart control charts	328
9.3.1 The economic design for $\bar{X}$ -charts	328
9.3.2 Increasing the sensitivity of p-charts	328
9.4 Cumulative sum control charts	330
9.4.1 Upper Page's scheme	330
9.4.2 Some theoretical background	333
9.4.3 Lower and two-sided Page's scheme	335
9.4.4 Average run length, probability of false alarm and conditional expected delay	339
9.5 Bayesian detection	342
9.6 Process tracking	346
9.6.1 The EWMA procedure	347
9.6.2 The BECM procedure	348
9.6.3 The Kalman filter	350
9.6.4 Hoadley's QMP	351
9.7 Automatic process control	354
9.8 Chapter highlights	356
9.9 Exercises	357
<b>10 Multivariate Statistical Process Control</b>	<b>361</b>
10.1 Introduction	361
10.2 A review of multivariate data analysis	365
10.3 Multivariate process capability indices	367
10.4 Advanced applications of multivariate control charts	370
10.4.1 Multivariate control charts scenarios	370
10.4.2 Internally derived targets	370
10.4.3 Using an external reference sample	371
10.4.4 Externally assigned targets	372

10.4.5	Measurement units considered as batches	373
10.4.6	Variable decomposition and monitoring indices	373
10.5	Multivariate tolerance specifications	374
10.6	Chapter highlights	376
10.7	Exercises	377
<b>PART IV DESIGN AND ANALYSIS OF EXPERIMENTS</b>		<b>379</b>
<b>11</b>	<b>Classical Design and Analysis of Experiments</b>	<b>381</b>
11.1	Basic steps and guiding principles	381
11.2	Blocking and randomization	385
11.3	Additive and non-additive linear models	385
11.4	The analysis of randomized complete block designs	387
11.4.1	Several blocks, two treatments per block: Paired comparison	387
11.4.2	Several blocks, $t$ treatments per block	391
11.5	Balanced Incomplete block designs	394
11.6	Latin square design	397
11.7	Full factorial experiments	402
11.7.1	The structure of factorial experiments	402
11.7.2	The ANOVA for full factorial designs	402
11.7.3	Estimating main effects and interactions	408
11.7.4	$2^m$ factorial designs	409
11.7.5	$3^m$ factorial designs	417
11.8	Blocking and fractional replications of $2^m$ factorial designs	425
11.9	Exploration of response surfaces	430
11.9.1	Second order designs	431
11.9.2	Some specific second order designs	433
11.9.3	Approaching the region of the optimal yield	438
11.9.4	Canonical representation	440
11.10	Chapter highlights	441
11.11	Exercises	442
<b>12</b>	<b>Quality by Design</b>	<b>446</b>
12.1	Off-line quality control, parameter design and the Taguchi method	447
12.1.1	Product and process optimization using loss functions	447
12.1.2	Major stages in product and process design	448
12.1.3	Design parameters and noise factors	449
12.1.4	Parameter design experiments	450
12.1.5	Performance statistics	452
12.2	The effects of non-linearity	452
12.3	Taguchi's designs	456
12.4	Quality by design in the pharmaceutical industry	458
12.4.1	Introduction to quality by design	458
12.4.2	A quality by design case study – the full factorial design	459
12.4.3	A quality by design case study – the profiler and desirability function	462
12.4.4	A quality by design case study – the design space	462
12.5	Tolerance designs	462
12.6	More case studies	467
12.6.1	The Quinlan experiment at Flex Products, Inc.	467
12.6.2	Computer response time optimization	469

12.7	Chapter highlights	474
12.8	Exercises	474
<b>13</b>	<b>Computer Experiments</b>	<b>477</b>
13.1	Introduction to computer experiments	477
13.2	Designing computer experiments	481
13.3	Analyzing computer experiments	483
13.4	Stochastic emulators	488
13.5	Integrating physical and computer experiments	491
13.6	Chapter highlights	492
13.7	Exercises	492
<b>PART V</b>	<b>RELIABILITY AND SURVIVAL ANALYSIS</b>	<b>495</b>
<b>J4</b>	<b>Reliability Analysis</b>	<b>497</b>
14.1	Basic notions	498
14.1.1	Time categories	498
14.1.2	Reliability and related functions	499
14.2	System reliability	500
14.3	Availability of repairable systems	503
14.4	Types of observations on TTF	509
14.5	Graphical analysis of life data	510
14.6	Non-parametric estimation of reliability	513
14.7	Estimation of life characteristics	514
14.7.1	Maximum likelihood estimators for exponential TTF distribution	514
14.7.2	Maximum likelihood estimation of the Weibull parameters	518
14.8	Reliability demonstration	520
14.8.1	Binomial testing	520
14.8.2	Exponential distributions	521
14.9	Accelerated life testing	528
14.9.1	The Arrhenius temperature model	528
14.9.2	Other models	529
14.10	Burn-in procedures	529
14.11	Chapter highlights	530
14.12	Exercises	531
<b>15</b>	<b>Bayesian Reliability Estimation and Prediction</b>	<b>534</b>
15.1	Prior and posterior distributions	534
15.2	Loss functions and Bayes estimators	537
15.2.1	Distribution-free Bayes estimator of reliability	538
15.2.2	Bayes estimator of reliability for exponential life distributions	538
15.3	Bayesian credibility and prediction intervals	539
15.3.1	Distribution-free reliability estimation	539
15.3.2	Exponential reliability estimation	540
15.3.3	Prediction intervals	540
15.4	Credibility intervals for the asymptotic availability of repairable systems: The exponential case	542
15.5	Empirical Bayes method	543
15.6	Chapter highlights	545
15.7	Exercises	545



<i>List of R Packages</i>	547
<i>References and Further Reading</i>	549
<i>Author Index</i>	555
<i>Subject Index</i>	557

Also available on book's website: [www.wiley.com/go/modern\\_industrial\\_statistics](http://www.wiley.com/go/modern_industrial_statistics)

**Appendix I:** An Introduction to R by Stefano Iacus

**Appendix II:** Basic MINITAB Commands and a Review of Matrix Algebra for Statistics

**Appendix III:** mistat Manual (mistat.pdf) and List of R Scripts, by Chapter (R\_scripts.zip)

**Appendix IV:** Source Version of mistat Package (mistat\_1.0.tar.gz), also available on the Comprehensive R Archive Network (CRAN) Website.

**Appendix V:** Data Sets as csv Files

**Appendix VI:** MINITAB Macros

**Appendix VII:** JMP Scripts by Ian Cox

**Appendix VIII:** Solution Manual