



Springer

of Engineering Statistics

Hoang Pham (Ed.)

With CD-ROM, 377 Figures and 204 Tables



Springer

Contents

List of Tables.....	XXXI
List of Abbreviations	XLI

Part A Fundamental Statistics and Its Applications

1 Basic Statistical Concepts	
<i>Hoang Pham</i>	3
1.1 Basic Probability Measures.....	3
1.2 Common Probability Distribution Functions	7
1.3 Statistical Inference and Estimation	17
1.4 Stochastic Processes	32
1.5 Further Reading	42
References.....	42
1.A Appendix: Distribution Tables	43
1.B Appendix: Laplace Transform	47
2 Statistical Reliability with Applications	
<i>Paul Kvam, Jye-Chyi Lu</i>	49
2.1 Introduction and Literature Review	49
2.2 Lifetime Distributions in Reliability	50
2.3 Analysis of Reliability Data	54
2.4 System Reliability	56
References.....	60
3 Weibull Distributions and Their Applications	
<i>Chin-Diew Lai, D.N. Pra Murthy, Min Xie</i>	63
3.1 Three-Parameter Weibull Distribution	64
3.2 Properties	64
3.3 Modeling Failure Data	67
3.4 Weibull-Derived Models	70
3.5 Empirical Modeling of Data	73
3.6 Applications.....	74
References.....	76
4 Characterizations of Probability Distributions	
<i>H.N. Nagaraja</i>	79
4.1 Characterizing Functions.....	80
4.2 Data Types and Characterizing Conditions.....	81
4.3 A Classification of Characterizations.....	83
4.4 Exponential Distribution	84
4.5 Normal Distribution	85
4.6 Other Continuous Distributions	87

4.7	Poisson Distribution and Process	88
4.8	Other Discrete Distributions	90
4.9	Multivariate Distributions and Conditional Specification	90
4.10	Stability of Characterizations	92
4.11	Applications	92
4.12	General Resources	93
	References	94
5	Two-Dimensional Failure Modeling	
	<i>D.N. Pra Murthy, Jaiwook Baik, Richard J. Wilson, Michael Bulmer</i>	97
5.1	Modeling Failures	98
5.2	Black-Box Modeling Process	98
5.3	One-Dimensional Black-Box Failure Modeling	99
5.4	Two-Dimensional Black-Box Failure Modeling	103
5.5	A New Approach to Two-Dimensional Modeling	107
5.6	Conclusions	110
	References	110
6	Prediction Intervals for Reliability Growth Models with Small Sample Sizes	
	<i>John Quigley, Lesley Walls</i>	113
6.1	Modified IBM Model – A Brief History	114
6.2	Derivation of Prediction Intervals for the Time to Detection of Next Fault	115
6.3	Evaluation of Prediction Intervals for the Time to Detect Next Fault ..	117
6.4	Illustrative Example	119
6.5	Conclusions and Reflections	122
	References	122
7	Promotional Warranty Policies: Analysis and Perspectives	
	<i>Jun Bai, Hoang Pham</i>	125
7.1	Classification of Warranty Policies	126
7.2	Evaluation of Warranty Policies	129
7.3	Concluding Remarks	134
	References	134
8	Stationary Marked Point Processes	
	<i>Karl Sigman</i>	137
8.1	Basic Notation and Terminology	138
8.2	Inversion Formulas	144
8.3	Campbell’s Theorem for Stationary MPPs	145
8.4	The Palm Distribution: Conditioning in a Point at the Origin	146
8.5	The Theorems of Khintchine, Korolyuk, and Dobrushin	146
8.6	An MPP Jointly with a Stochastic Process	147
8.7	The Conditional Intensity Approach	148
8.8	The Non-Ergodic Case	150
8.9	MPPs in \mathbb{R}^d	150
	References	152

9 Modeling and Analyzing Yield, Burn-In and Reliability for Semiconductor Manufacturing: Overview	
<i>Way Kuo, Kyungmee O. Kim, Taeho Kim</i>	153
9.1 Semiconductor Yield	154
9.2 Semiconductor Reliability	159
9.3 Burn-In	160
9.4 Relationships Between Yield, Burn-In and Reliability	163
9.5 Conclusions and Future Research	166
References	166

Part B Process Monitoring and Improvement

10 Statistical Methods for Quality and Productivity Improvement	
<i>Wei Jiang, Terrence E. Murphy, Kwok-Leung Tsui</i>	173
10.1 Statistical Process Control for Single Characteristics	174
10.2 Robust Design for Single Responses	181
10.3 Robust Design for Multiple Responses	185
10.4 Dynamic Robust Design	186
10.5 Applications of Robust Design	187
References	188
11 Statistical Methods for Product and Process Improvement	
<i>Kailash C. Kapur, Qianmei Feng</i>	193
11.1 Six Sigma Methodology and the (D)MAIC(T) Process	195
11.2 Product Specification Optimization	196
11.3 Process Optimization	204
11.4 Summary	211
References	212
12 Robust Optimization in Quality Engineering	
<i>Susan L. Albin, Di Xu</i>	213
12.1 An Introduction to Response Surface Methodology	216
12.2 Minimax Deviation Method to Derive Robust Optimal Solution	218
12.3 Weighted Robust Optimization	222
12.4 The Application of Robust Optimization in Parameter Design	224
References	227
13 Uniform Design and Its Industrial Applications	
<i>Kai-Tai Fang, Ling-Yau Chan</i>	229
13.1 Performing Industrial Experiments with a UD	231
13.2 Application of UD in Accelerated Stress Testing	233
13.3 Application of UD in Computer Experiments	234
13.4 Uniform Designs and Discrepancies	236
13.5 Construction of Uniform Designs in the Cube	237
13.6 Construction of UD for Experiments with Mixtures	240

13.7	Relationships Between Uniform Design and Other Designs	243
13.8	Conclusion	245
	References.....	245
14	Cuscore Statistics: Directed Process Monitoring for Early Problem Detection	
	<i>Harriet B. Nembhard</i>	249
14.1	Background and Evolution of the Cuscore in Control Chart Monitoring.....	250
14.2	Theoretical Development of the Cuscore Chart.....	251
14.3	Cuscores to Monitor for Signals in White Noise	252
14.4	Cuscores to Monitor for Signals in Autocorrelated Data	254
14.5	Cuscores to Monitor for Signals in a Seasonal Process	255
14.6	Cuscores in Process Monitoring and Control	256
14.7	Discussion and Future Work	258
	References.....	260
15	Chain Sampling	
	<i>Raj K. Govindaraju</i>	263
15.1	ChSP-1 Chain Sampling Plan	264
15.2	Extended Chain Sampling Plans	265
15.3	Two-Stage Chain Sampling	266
15.4	Modified ChSP-1 Plan.....	268
15.5	Chain Sampling and Deferred Sentencing	269
15.6	Comparison of Chain Sampling with Switching Sampling Systems	272
15.7	Chain Sampling for Variables Inspection	273
15.8	Chain Sampling and CUSUM.....	274
15.9	Other Interesting Extensions	276
15.10	Concluding Remarks	276
	References.....	276
16	Some Statistical Models for the Monitoring of High-Quality Processes	
	<i>Min Xie, Thong N. Goh</i>	281
16.1	Use of Exact Probability Limits	282
16.2	Control Charts Based on Cumulative Count of Conforming Items.....	283
16.3	Generalization of the c -Chart	284
16.4	Control Charts for the Monitoring of Time-Between-Events	286
16.5	Discussion	288
	References.....	289
17	Monitoring Process Variability Using EWMA	
	<i>Philippe Castagliola, Giovanni Celano, Sergio Fichera</i>	291
17.1	Definition and Properties of EWMA Sequences	292
17.2	EWMA Control Charts for Process Position	295
17.3	EWMA Control Charts for Process Dispersion.....	298

17.4	Variable Sampling Interval EWMA Control Charts for Process Dispersion	310
17.5	Conclusions	323
	References	324
18	Multivariate Statistical Process Control Schemes for Controlling a Mean	
	<i>Richard A. Johnson, Ruojia Li</i>	327
18.1	Univariate Quality Monitoring Schemes	328
18.2	Multivariate Quality Monitoring Schemes	331
18.3	An Application of the Multivariate Procedures	336
18.4	Comparison of Multivariate Quality Monitoring Methods	337
18.5	Control Charts Based on Principal Components	338
18.6	Difficulties of Time Dependence in the Sequence of Observations	341
	References	344
 Part C Reliability Models and Survival Analysis		
19	Statistical Survival Analysis with Applications	
	<i>Chengjie Xiong, Kejun Zhu, Kai Yu</i>	347
19.1	Sample Size Determination to Compare Mean or Percentile of Two Lifetime Distributions	349
19.2	Analysis of Survival Data from Special Cases of Step-Stress Life Tests	355
	References	365
20	Failure Rates in Heterogeneous Populations	
	<i>Maxim Finkelstein, Veronica Esaulova</i>	369
20.1	Mixture Failure Rates and Mixing Distributions	371
20.2	Modeling the Impact of the Environment	377
20.3	Asymptotic Behaviors of Mixture Failure Rates	380
	References	385
21	Proportional Hazards Regression Models	
	<i>Wei Wang, Chengcheng Hu</i>	387
21.1	Estimating the Regression Coefficients β	388
21.2	Estimating the Hazard and Survival Functions	389
21.3	Hypothesis Testing	390
21.4	Stratified Cox Model	390
21.5	Time-Dependent Covariates	390
21.6	Goodness-of-Fit and Model Checking	391
21.7	Extension of the Cox Model	393
21.8	Example	394
	References	395

22 Accelerated Life Test Models and Data Analysis	
<i>Francis Pascual, William Q. Meeker, Jr., Luis A. Escobar</i>	397
22.1 Accelerated Tests	398
22.2 Life Distributions	400
22.3 Acceleration Models	400
22.4 Analysis of Accelerated Life Test Data	407
22.5 Further Examples	412
22.6 Practical Considerations for Interpreting the Analysis of ALT Data	421
22.7 Other Kinds of ATs	421
22.8 Some Pitfalls of Accelerated Testing	423
22.9 Computer Software for Analyzing ALT Data	424
References.....	425
23 Statistical Approaches to Planning of Accelerated Reliability Testing	
<i>Loon C. Tang</i>	427
23.1 Planning Constant-Stress Accelerated Life Tests	428
23.2 Planning Step-Stress ALT (SSALT)	432
23.3 Planning Accelerated Degradation Tests (ADT)	436
23.4 Conclusions.....	439
References.....	440
24 End-to-End (E2E) Testing and Evaluation of High-Assurance Systems	
<i>Raymond A. Paul, Wei-Tek Tsai, Yinong Chen, Chun Fan, Zhibin Cao, Hai Huang</i>	443
24.1 History and Evolution of E2E Testing and Evaluation.....	444
24.2 Overview of the Third and Fourth Generations of the E2E T&E	449
24.3 Static Analyses	451
24.4 E2E Distributed Simulation Framework	453
24.5 Policy-Based System Development.....	459
24.6 Dynamic Reliability Evaluation.....	465
24.7 The Fourth Generation of E2E T&E on Service-Oriented Architecture	470
24.8 Conclusion and Summary.....	473
References.....	474
25 Statistical Models in Software Reliability and Operations Research	
<i>P.K. Kapur, Amit K. Bardhan</i>	477
25.1 Interdisciplinary Software Reliability Modeling	479
25.2 Release Time of Software	486
25.3 Control Problem	489
25.4 Allocation of Resources in Modular Software	491
References.....	495

26 An Experimental Study of Human Factors in Software Reliability Based on a Quality Engineering Approach	
<i>Shigeru Yamada</i>	497
26.1 Design Review and Human Factors	498
26.2 Design-Review Experiment	499
26.3 Analysis of Experimental Results	500
26.4 Investigation of the Analysis Results	501
26.5 Confirmation of Experimental Results	502
26.6 Data Analysis with Classification of Detected Faults	504
References	506
27 Statistical Models for Predicting Reliability of Software Systems in Random Environments	
<i>Hoang Pham, Xiaolin Teng</i>	507
27.1 A Generalized NHPP Software Reliability Model	509
27.2 Generalized Random Field Environment (RFE) Model	510
27.3 RFE Software Reliability Models	511
27.4 Parameter Estimation	513
References	519
 Part D Regression Methods and Data Mining	
28 Measures of Influence and Sensitivity in Linear Regression	
<i>Daniel Peña</i>	523
28.1 The Leverage and Residuals in the Regression Model	524
28.2 Diagnosis for a Single Outlier	525
28.3 Diagnosis for Groups of Outliers	528
28.4 A Statistic for Sensitivity for Large Data Sets	532
28.5 An Example: The Boston Housing Data	533
28.6 Final Remarks	535
References	535
29 Logistic Regression Tree Analysis	
<i>Wei-Yin Loh</i>	537
29.1 Approaches to Model Fitting	538
29.2 Logistic Regression Trees	540
29.3 LOTUS Algorithm	542
29.4 Example with Missing Values	543
29.5 Conclusion	549
References	549
30 Tree-Based Methods and Their Applications	
<i>Nan Lin, Douglas Noe, Xuming He</i>	551
30.1 Overview	552
30.2 Classification and Regression Tree (CART)	555
30.3 Other Single-Tree-Based Methods	561

30.4	Ensemble Trees	565
30.5	Conclusion	568
	References.....	569
31	Image Registration and Unknown Coordinate Systems	
	<i>Ted Chang</i>	571
31.1	Unknown Coordinate Systems and Their Estimation	572
31.2	Least Squares Estimation	575
31.3	Geometry of $\mathcal{O}(p)$ and $\mathcal{SO}(p)$	578
31.4	Statistical Properties of M -Estimates	580
31.5	Diagnostics	587
	References.....	590
32	Statistical Genetics for Genomic Data Analysis	
	<i>Jae K. Lee</i>	591
32.1	False Discovery Rate	592
32.2	Statistical Tests for Genomic Data	593
32.3	Statistical Modeling for Genomic Data	596
32.4	Unsupervised Learning: Clustering	598
32.5	Supervised Learning: Classification	599
	References.....	603
33	Statistical Methodologies for Analyzing Genomic Data	
	<i>Fenghai Duan, Heping Zhang</i>	607
33.1	Second-Level Analysis of Microarray Data	609
33.2	Third-Level Analysis of Microarray Data	611
33.3	Fourth-Level Analysis of Microarray Data	618
33.4	Final Remarks	618
	References.....	619
34	Statistical Methods in Proteomics	
	<i>Weichuan Yu, Baolin Wu, Tao Huang, Xiaoye Li, Kenneth Williams, Hongyu Zhao</i>	623
34.1	Overview.....	623
34.2	MS Data Preprocessing.....	625
34.3	Feature Selection	628
34.4	Sample Classification	630
34.5	Random Forest: Joint Modelling of Feature Selection and Classification	630
34.6	Protein/Peptide Identification.....	633
34.7	Conclusion and Perspective.....	635
	References.....	636
35	Radial Basis Functions for Data Mining	
	<i>Miyoung Shin, Amrit L. Goel</i>	639
35.1	Problem Statement	640
35.2	RBF Model and Parameters	641

35.3	Design Algorithms	642
35.4	Illustrative Example.....	643
35.5	Diabetes Disease Classification	645
35.6	Analysis of Gene Expression Data	647
35.7	Concluding Remarks	648
	References.....	648
36	Data Mining Methods and Applications	
	<i>Kwok-Leung Tsui, Victoria Chen, Wei Jiang, Y. Alp Aslandogan</i>	<i>651</i>
36.1	The KDD Process	653
36.2	Handling Data.....	654
36.3	Data Mining (DM) Models and Algorithms	655
36.4	DM Research and Applications	664
36.5	Concluding Remarks	667
	References.....	667
 Part E Modeling and Simulation Methods		
37	Bootstrap, Markov Chain and Estimating Function	
	<i>Feifang Hu</i>	<i>673</i>
37.1	Overview.....	673
37.2	Classical Bootstrap.....	675
37.3	Bootstrap Based on Estimating Equations.....	678
37.4	Markov Chain Marginal Bootstrap.....	681
37.5	Applications	682
37.6	Discussion	684
	References.....	684
38	Random Effects	
	<i>Yi Li.....</i>	<i>687</i>
38.1	Overview.....	687
38.2	Linear Mixed Models.....	688
38.3	Generalized Linear Mixed Models	690
38.4	Computing MLEs for GLMMs	692
38.5	Special Topics: Testing Random Effects for Clustered Categorical Data	697
38.6	Discussion	701
	References.....	701
39	Cluster Randomized Trials: Design and Analysis	
	<i>Mirjam Moerbeek.....</i>	<i>705</i>
39.1	Cluster Randomized Trials	706
39.2	Multilevel Regression Model and Mixed Effects ANOVA Model	707
39.3	Optimal Allocation of Units	709
39.4	The Effect of Adding Covariates	712
39.5	Robustness Issues.....	713

39.6	Optimal Designs for the Intra-Class Correlation Coefficient	715
39.7	Conclusions and Discussion	717
	References	717
40	A Two-Way Semilinear Model for Normalization and Analysis of Microarray Data	
	<i>Jian Huang, Cun-Hui Zhang</i>	719
40.1	The Two-Way Semilinear Model	720
40.2	Semiparametric M-Estimation in TW-SLM	721
40.3	Extensions of the TW-SLM	724
40.4	Variance Estimation and Inference for β	725
40.5	An Example and Simulation Studies	727
40.6	Theoretical Results	732
40.7	Concluding Remarks	734
	References	734
41	Latent Variable Models for Longitudinal Data with Flexible Measurement Schedule	
	<i>Haiqun Lin</i>	737
41.1	Hierarchical Latent Variable Models for Longitudinal Data	738
41.2	Latent Variable Models for Multidimensional Longitudinal Data	741
41.3	Latent Class Mixed Model for Longitudinal Data	743
41.4	Structural Equation Model with Latent Variables for Longitudinal Data	744
41.5	Concluding Remark: A Unified Multilevel Latent Variable Model	746
	References	747
42	Genetic Algorithms and Their Applications	
	<i>Mitsuo Gen</i>	749
42.1	Foundations of Genetic Algorithms	750
42.2	Combinatorial Optimization Problems	753
42.3	Network Design Problems	757
42.4	Scheduling Problems	761
42.5	Reliability Design Problem	763
42.6	Logistic Network Problems	766
42.7	Location and Allocation Problems	769
	References	772
43	Scan Statistics	
	<i>Joseph Naus</i>	775
43.1	Overview	775
43.2	Temporal Scenarios	776
43.3	Higher Dimensional Scans	784
43.4	Other Scan Statistics	786
	References	788

44 Condition-Based Failure Prediction

<i>Shang-Kuo Yang</i>	791
44.1 Overview	792
44.2 Kalman Filtering	794
44.3 Armature-Controlled DC Motor	796
44.4 Simulation System	797
44.5 Armature-Controlled DC Motor Experiment	801
44.6 Conclusions	804
References	804

45 Statistical Maintenance Modeling for Complex Systems

<i>Wenjian Li, Hoang Pham</i>	807
45.1 General Probabilistic Processes Description	809
45.2 Nonrepairable Degraded Systems Reliability Modeling	810
45.3 Repairable Degraded Systems Modeling	819
45.4 Conclusions and Perspectives	831
45.5 Appendix A	831
45.6 Appendix B	832
References	833

46 Statistical Models on Maintenance

<i>Toshio Nakagawa</i>	835
46.1 Time-Dependent Maintenance	836
46.2 Number-Dependent Maintenance	838
46.3 Amount-Dependent Maintenance	842
46.4 Other Maintenance Models	843
References	847

Part F Applications in Engineering Statistics**47 Risks and Assets Pricing**

<i>Charles S. Tapiero</i>	851
47.1 Risk and Asset Pricing	853
47.2 Rational Expectations, Risk-Neutral Pricing and Asset Pricing	857
47.3 Consumption Capital Asset Price Model and Stochastic Discount Factor	862
47.4 Bonds and Fixed-Income Pricing	865
47.5 Options	872
47.6 Incomplete Markets and Implied Risk-Neutral Distributions	880
References	898

48 Statistical Management and Modeling for Demand of Spare Parts

<i>Emilio Ferrari, Arrigo Pareschi, Alberto Regattieri, Alessandro Persona</i>	905
48.1 The Forecast Problem for Spare Parts	905
48.2 Forecasting Methods	909
48.3 The Applicability of Forecasting Methods to Spare-Parts Demands ...	911

48.4	Prediction of Aircraft Spare Parts: A Case Study	912
48.5	Poisson Models	915
48.6	Models Based on the Binomial Distribution	917
48.7	Extension of the Binomial Model Based on the Total Cost Function ..	920
48.8	Weibull Extension	923
	References	928
49	Arithmetic and Geometric Processes	
	<i>Kit-Nam F. Leung</i>	931
49.1	Two Special Monotone Processes	934
49.2	Testing for Trends	936
49.3	Estimating the Parameters	938
49.4	Distinguishing a Renewal Process from an AP (or a GP)	939
49.5	Estimating the Means and Variances	939
49.6	Comparison of Estimators Using Simulation	945
49.7	Real Data Analysis	946
49.8	Optimal Replacement Policies Determined Using Arithmetico-Geometric Processes	947
49.9	Some Conclusions on the Applicability of an AP and/or a GP	950
49.10	Concluding Remarks	951
49.A	Appendix	953
	References	954
50	Six Sigma	
	<i>Fugee Tsung</i>	957
50.1	The DMAIC Methodology	960
50.2	Design for Six Sigma	965
50.3	Six Sigma Case Study	970
50.4	Conclusion	971
	References	971
51	Multivariate Modeling with Copulas and Engineering Applications	
	<i>Jun Yan</i>	973
51.1	Copulas and Multivariate Distributions	974
51.2	Some Commonly Used Copulas	977
51.3	Statistical Inference	981
51.4	Engineering Applications	982
51.5	Conclusion	987
51.A	Appendix	987
	References	989
52	Queuing Theory Applications to Communication Systems: Control of Traffic Flows and Load Balancing	
	<i>Panlop Zeepongsekul, Anthony Bedford, James Broberg, Peter Dimopoulos, Zahir Tari</i>	991
52.1	Brief Review of Queueing Theory	994
52.2	Multiple-Priority Dual Queue (MPDQ)	1000

52.3	Distributed Systems and Load Balancing.....	1005
52.4	Active Queue Management for TCP Traffic.....	1012
52.5	Conclusion	1020
	References.....	1020
53	Support Vector Machines for Data Modeling with Software Engineering Applications	
	<i>Hojung Lim, Amrit L. Goel</i>	1023
53.1	Overview.....	1023
53.2	Classification and Prediction in Software Engineering	1024
53.3	Support Vector Machines	1025
53.4	Linearly Separable Patterns.....	1026
53.5	Linear Classifier for Nonseparable Classes	1029
53.6	Nonlinear Classifiers	1029
53.7	SVM Nonlinear Regression.....	1032
53.8	SVM Hyperparameters	1033
53.9	SVM Flow Chart.....	1033
53.10	Module Classification.....	1034
53.11	Effort Prediction	1035
53.12	Concluding Remarks	1036
	References.....	1036
54	Optimal System Design	
	<i>Suprasad V. Amari</i>	1039
54.1	Optimal System Design	1039
54.2	Cost-Effective Designs.....	1047
54.3	Optimal Design Algorithms.....	1051
54.4	Hybrid Optimization Algorithms	1055
	References.....	1063
	Acknowledgements	1065
	About the Authors	1067
	Detailed Contents	1085
	Subject Index	1113