

Springer Transactions in Civil and Environmental Engineering

Rajib Maity

Statistical Methods in Hydrology and Hydroclimatology

Second Edition

This second edition focuses on the application of statistical methods in the field of hydrology and hydroclimatology. Among the latest theories being used in these fields, the book introduces the theory of copulas and its applications in this context. The purpose is to develop an understanding and illustrate the usefulness of the statistical techniques with detailed theory and numerous worked out examples. Apart from this, sample scripts based on MATLAB, Python and R for some examples are also provided to assist the readers to handle real life data. Besides serving as a textbook for graduate courses on stochastic modeling in hydrology and related disciplines, the book offers a valuable resource for researchers and professionals involved in the field of hydrology and climatology.

ISSN 2363-7633

ISBN 978-981-16-5516-6



► springer.com



1	Introduction	1
1.1	Definitions and Scope	1
1.2	Role of Statistical Methods	2
1.2.1	Hydrologic and Hydroclimatic Variability	2
1.2.2	Need of Statistical Methods	2
1.3	Organization of the Book	4
2	Basic Concepts of Probability and Statistics	7
2.1	Concepts of Random Experiments and Random Variables	7
2.1.1	Random Experiments, Sample Space and Events	7
2.1.2	Concept of Random Variables and Events	8
2.2	Basic Concepts of Probability	11
2.2.1	The Axioms of Probability	11
2.2.2	Some Elementary Properties on Probability	12
2.3	Conditional Probability Theorem	15
2.4	Total Probability Theorem and Bayes Rule	18
2.5	Univariate and Bivariate Probability Distribution of Random Variables	22
2.5.1	Discrete Random Variable	23
2.5.2	Continuous Random Variable	25
2.6	Marginal and Conditional Probability Distribution	30
2.6.1	Marginal Probability Distribution	30
2.6.2	Conditional Distribution Function	36
2.7	Independence Between Random Variables	41
2.8	Functions of Random Variables	42
2.8.1	Univariate Random Variable	42
2.8.2	Bivariate Random Variables	43
2.9	Coding Examples with MATLAB, PYTHON and R	45
	Exercise	47

3	Basic Statistical Properties of Data	51
3.1	Descriptive Statistics	51
3.1.1	Measures of Central Tendency	52
3.1.2	Measure of Dispersion	55
3.1.3	Measure of Symmetry	56
3.1.4	Measure of Tailedness	57
3.2	Concept of Moments and Expectation	59
3.2.1	Expectation	61
3.3	Moment Generating Functions	67
3.4	Characteristic Functions	69
3.5	Statistical Properties of Jointly Distributed Random Variables	72
3.5.1	Expectation	72
3.5.2	Moment about the Origin	73
3.5.3	Moment about the Mean (Central Moment)	73
3.5.4	Moment Generating Function	73
3.5.5	Covariance	75
3.5.6	Correlation Coefficient	77
3.5.7	Further Properties of Moments	79
3.6	Properties of the Estimator	80
3.6.1	Unbiasedness	80
3.6.2	Consistency	80
3.6.3	Efficiency	80
3.6.4	Sufficiency	81
3.7	Parameter Estimation	82
3.7.1	Method of Moments	82
3.7.2	Maximum Likelihood	82
3.8	Chebyshev Inequality	84
3.9	Law of Large Number	84
3.10	Coding Examples with MATLAB, PYTHON and R	85
	Exercise	86
4	Probability Distributions and Their Applications	91
4.1	Discrete Probability Distributions	91
4.1.1	Binomial Distribution	92
4.1.2	Negative Binomial Distribution	95
4.1.3	Multinomial Distribution	97
4.1.4	Hypergeometric Distribution	99
4.1.5	Geometric Distribution	102
4.1.6	Poisson Distribution	104
4.2	Continuous Probability Distributions	106
4.2.1	Uniform Distribution	106
4.2.2	Exponential Distribution	108
4.2.3	Normal Distribution	111
4.2.4	Lognormal Distribution	115

4.2.5	Gamma Distribution	118
4.2.6	Extreme Value Distribution	121
4.2.7	Beta Distribution	126
4.2.8	Pearson and Log-Pearson Type III Distribution	128
4.3	Mixed Distribution	129
4.4	Some Important Distributions of Sample Statistics	130
4.4.1	Chi-Square Distribution	130
4.4.2	The t-Distribution	131
4.4.3	The F-Distribution	132
4.5	Coding Examples with MATLAB, PYTHON and R	134
Exercise		138
5	Frequency Analysis, Risk and Uncertainty in Hydroclimatic Analysis	141
5.1	Concept of Return Period	141
5.2	Probability Plotting and Plotting Positions Formulae	145
5.3	Probability Paper	148
5.3.1	Mathematical Construction of Probability Paper	148
5.3.2	Graphical Construction of Probability Paper	149
5.4	Frequency Analysis of Hydroclimatic Extremes	153
5.4.1	Normal Distribution	153
5.4.2	Lognormal Distribution	154
5.4.3	Log-Pearson Type III Distribution	155
5.4.4	Extreme Value Type I Distribution	156
5.5	Risk and Reliability in Hydrologic Design	167
5.6	Concept of Uncertainty	168
5.6.1	Analysis of Uncertainty	169
5.6.2	Measures of Uncertainty	170
5.7	Reliability, Resilience and Vulnerability of Hydrologic Time Series	171
5.7.1	Reliability	172
5.7.2	Resilience	172
5.7.3	Vulnerability	173
5.8	Coding Examples with MATLAB, PYTHON and R	175
Exercise		180
6	Hypothesis Testing and Non-parametric Test	185
6.1	Populations and Samples	185
6.2	Random Samples	186
6.3	Sampling Distribution	186
6.3.1	Sampling Distribution of the Mean	187
6.3.2	Sampling Distribution of the Variance	188
6.4	Statistical Inference	188
6.4.1	Point Estimation	189
6.4.2	Interval Estimation	190
6.4.3	Hypothesis Testing	194

6.4.4	Goodness-of-Fit Test	207
6.4.5	Non-parametric Test	210
6.5	Coding Examples with MATLAB, PYTHON and R	217
Exercise		218
7	Regression Analysis and Curve Fitting	221
7.1	Simple Linear Regression	221
7.2	Curvilinear Regression	228
7.2.1	Model Transformable to Linear Regression	228
7.2.2	Model Not Transformable to Linear Regression	232
7.3	Multiple Linear Regression	232
7.4	Evaluation of Regression Model	237
7.5	Correlation and Regression	240
7.6	Correlation and Causality	241
7.7	Confidence Interval	241
7.8	Coding Examples with MATLAB, PYTHON and R	244
Exercise		247
8	Multivariate Analysis	253
8.1	Principal Component Analysis	253
8.1.1	Determination of Principal Components	255
8.2	Supervised Principal Component Analysis	259
8.3	Dimensionality Reduction Using PCA and SPCA	261
8.4	Canonical Correlation Analysis	263
8.5	Empirical Orthogonal Function	266
8.6	Data Generation	268
8.6.1	Univariate Data Generation	268
8.6.2	Multivariate Data Generation	269
8.7	Analysis of Variance in Hydrology and Hydroclimatology	274
8.7.1	One-Way Analysis of Variance	274
8.7.2	Two-Way Analysis of Variance	280
8.7.3	Multiple Comparisons	284
8.8	Coding Examples with MATLAB, PYTHON and R	285
Exercise		290
9	Time Series Analysis	297
9.1	Data Representation in Hydroclimatology	297
9.2	Stationary and Non-stationary Time Series	299
9.3	Ensemble and Realization	299
9.4	Trend Analysis	300
9.4.1	Tests for Randomness and Trend	301
9.4.2	Trend Removal	303
9.5	Analysis of Periodicity	307
9.5.1	Harmonic Analysis	307
9.5.2	Spectral Analysis	308
9.6	Data Transformation	309

	9.6.1	Test for Normal Distribution	311
9.7		Time Series Modeling in Hydroclimatology	313
	9.7.1	Measures of Linear Association in Time Series	313
	9.7.2	Statistical Operators on Time Series	319
	9.7.3	Properties of Time Series Models	321
	9.7.4	Auto-Regressive (AR) Model	323
	9.7.5	Moving Average (MA) Model	329
	9.7.6	Auto-Regressive Moving Average (ARMA) Model	334
	9.7.7	Auto-regressive Integrated Moving Average (ARIMA) Model	337
	9.7.8	Auto-regressive Moving Average Model with Exogenous Inputs (ARMAX)	339
	9.7.9	Forecasting with ARMA/ARMAX	341
	9.7.10	Parsimony of Time Series Models	346
	9.7.11	Diagnostic Check for ARMA Models	348
9.8		Wavelet Analysis	352
	9.8.1	Haar Wavelet	358
	9.8.2	Multi-resolution Analysis	359
9.9		Coding Examples with MATLAB, PYTHON and R	361
	Exercise		368
10		Theory of Copula in Hydrology and Hydroclimatology	373
	10.1	Introduction	373
	10.2	Preliminary Concepts	374
		10.2.1 Definition of Copula	374
		10.2.2 Graphical Representation of Copula	374
	10.3	Sklar's Theorem	374
	10.4	Basic Properties of a Copula Function	376
		10.4.1 Basic Terminologies	376
	10.5	Non-parametric Measures of Association	381
	10.6	Copula and Function of Random Variables	386
	10.7	Survival Copula	386
	10.8	Most Commonly Used Copula Function	387
		10.8.1 Elliptical Copula	387
		10.8.2 Archimedean Copula	389
	10.9	Selection of Best-Fit Copula	396
		10.9.1 Test Using Empirical Copula	396
		10.9.2 Test Using Kendall's Transform	397
		10.9.3 Test Using Rosenblatt Probability Integral Transformation	398
	10.10	Use of Copulas	399
		10.10.1 Data Generation	399
		10.10.2 Probabilistic Prediction Using Copulas	403
	10.11	Coding Examples with MATLAB, PYTHON and R	410

Exercise	413
References	416
Appendix A: Data Set	417
Appendix B: Statistical Tables	423
Index	433