

Jürgen Groß

A Normal Distribution Course



PETER LANG

Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Contents

1	Data Analysis	1
1.1	Describing Data	1
1.1.1	Statistical Software Packages	2
1.1.2	Observations	2
1.1.3	Frequency Distribution	5
1.2	Inference Base	8
1.2.1	Random Variable	8
1.2.2	Distribution of a Random Variable	10
1.2.3	Random Sample	22
1.3	Inference	23
1.3.1	Estimation	24
1.3.2	Sample Mean and Variance	31
1.3.3	Maximum Likelihood Estimation	34
1.3.4	Fitting Probability Density Functions	38
1.4	A Summary Illustration	43
2	The Normal Distribution	45
2.1	Definition of the Normal Distribution	46
2.1.1	Location and Scale Parameter	46
2.1.2	Expectation and Variance	48
2.1.3	Alternative Parametrizations	49
2.2	Standard Normal Distribution	50
2.2.1	The Function $\phi(x)$	50
2.2.2	The Function $\Phi(x)$	51
2.2.3	Related Functions	54

2.3	Moments	56
2.3.1	The 1, 2, 3 standard deviation intervals	56
2.3.2	Higher Moments	57
2.3.3	Skewness and Kurtosis	58
2.3.4	Sample Skewness and Kurtosis	60
2.4	Quantiles	61
2.4.1	Quantiles of the Normal Distribution	62
2.4.2	Sample Quantiles	65
2.5	Parameter Estimation	67
2.5.1	Maximum Likelihood Estimators	67
2.5.2	Properties of Estimators	68
2.6	The Central Limit Theorem	71
2.6.1	Approximate Normality of the Sample Mean	71
2.6.2	Cautionary Notes	75
2.6.3	Approximate Normality of Sample Statistics	76
2.7	Approximate Normality of the Binomial Distribution	79
2.7.1	The Binomial Distribution	79
2.7.2	Approximation by the Normal Distribution	81
2.7.3	The Galton Board	84
2.8	Random Sample Generation	85
3	Checking for Normality	89
3.1	Sample Characteristic Values	89
3.2	Graphics	92
3.2.1	The Histogram	92
3.2.2	The ECDF	95
3.2.3	The Normal Quantile-Quantile Plot	97
3.3	Summary	101
4	Testing for Normality	103
4.1	General Remarks	103
4.1.1	Significance Tests	103
4.1.2	Classification of Tests	106
4.2	The Chi-Square Test	107
4.3	Tests Based on the ECDF	110
4.3.1	The Lilliefors (Kolmogorov-Smirnov) Test	111
4.3.2	The Cramér-von Mises Test	114

4.3.3	The Anderson-Darling Test	115
4.4	Correlation and Regression Tests	116
4.4.1	The Shapiro-Francia Test	117
4.4.2	The Shapiro-Wilk Test	118
4.5	Summary	119
5	Variants of the Normal Distribution	121
5.1	Truncated Normal Distribution	121
5.1.1	Truncation of the Normal Distribution	122
5.1.2	Estimation of Parameters	125
5.1.3	Random Sample Generation	128
5.2	Two-Normals-Mixture Distribution	129
5.2.1	Probability Density Function	130
5.2.2	Estimation of Parameters	132
5.2.3	Random Sample Generation	139
5.3	Skew-Normal Distribution	139
5.3.1	Probability Density Function	140
5.3.2	Estimation of Parameters	143
5.3.3	Random Sample Generation	145
6	Transformations to Normality	147
6.1	Data Transformation	147
6.2	The Johnson System of Distributions	148
6.2.1	Three Types of Distribution	148
6.2.2	Estimation of Parameters	152
6.2.3	Choice of Fit	158
6.3	The Lognormal Distribution	158
6.3.1	Two- and Three-Parameter Lognormal Distribution	160
6.3.2	Estimation of Parameters	161
6.4	Power Transformations	163
6.4.1	Box-Cox Transformation	163
6.4.2	Transformation for Proportions	165
7	Two Normal Variables	169
7.1	The Bivariate Normal Distribution	169
7.1.1	Probability Density Function	170
7.1.2	Some Properties	170

7.1.3	Uncorrelatedness and Independence	173
7.1.4	Estimation	174
7.1.5	Assessing Bivariate Normality	174
7.1.6	Random Sample Generation	176
7.2	Two Normal Variables	177
7.2.1	Are Two Normals Jointly Normal?	177
7.2.2	Are Two Uncorrelated Normals Independent?	180
7.2.3	Does Uncorrelatedness of Two Normals Imply Independence Only If They Are Jointly Normal?	182
7.2.4	Is Any Linear Combination of Two Normals Again Normal?	182
7.3	Vector and Matrix Representation	183
8	Transformations of Normal Variables	185
8.1	Functions of a Random Variable	185
8.2	Functions of a Single Normal Variable	187
8.2.1	Distribution of the Absolute Value	187
8.2.2	Distribution of the Square	189
8.2.3	Distribution of the Reciprocal	191
8.2.4	Distribution of Minimum and Maximum	192
8.3	Functions of Two Independent Normals	193
8.3.1	Distribution of the Sum	194
8.3.2	Distribution of Product and Quotient	194
8.4	Functions Related to a Normal Sample	197
8.4.1	Standard Normal Sample	197
8.4.2	Distribution of Parameter-Sample Statistic Functions	204
8.4.3	Distribution of the Sample z -Score	208
	Bibliography	211
	Index	217