

Contents

Introduction — VII

1	Basic statistical concepts — 1
1.1	Arithmetic mean — 1
1.2	Standard deviation — 1
1.3	Standard error — 2
1.4	Confidence intervals — 3
1.5	Applied example — 4
1.6	P-values — 5
1.7	Testing for normality — 8
1.8	Probability density functions — 10
1.9	Bias and variance — 13
1.10	Normality and confidence intervals revisited — 14
2	Bootstrapping — 17
2.1	Working example — 17
2.2	Requirements — 19
2.3	Diagnostics — 21
2.3.1	Unique values — 22
2.3.2	Distribution of bootstrap-results — 23
2.3.3	Rate of convergence — 25
2.4	The logic of bootstrapping — 26
2.5	Failures of the bootstrap — 29
3	Confidence intervals — 32
3.1	Normal intervals — 32
3.2	Percentile intervals — 33
3.3	Bias corrected intervals (BC) — 34
3.4	Bias corrected and accelerated intervals (BCa) — 37
3.5	Double bootstrap — 39
3.6	Transformations — 42
3.7	Summary and simulation — 44
4	Hypothesis testing — 49
4.1	The logic of hypothesis tests — 49
4.2	Confidence intervals — 51
4.3	Sampling in accordance with the null — 53
4.4	Permutation tests — 54
4.4.1	Paired permutation tests — 59
4.4.2	Boot or permute? — 61

4.5	Power calculations —	63
5	Regression models —	66
5.1	Sampling observations —	66
5.2	Sampling residuals —	70
5.3	Dealing with outliers —	71
5.4	Predicting outcomes and bagging —	75
6	Applied programming examples —	78
6.1	Python —	78
6.1.1	Basics —	78
6.1.2	All confidence intervals —	80
6.1.3	Permutation tests —	85
6.1.4	Multithreading —	89
6.1.5	Importing data —	90
6.1.6	Appendix —	92
6.2	Stata —	94
6.2.1	Descriptive statistics —	95
6.2.2	Estimating models —	97
6.2.3	Writing own programs —	99
6.2.4	Permutation tests —	106
6.2.5	Longitudinal and cluster analyses —	108
6.3	Computational speed —	114
6.4	Further reading —	117
	Bibliography —	119