

Contents

1	Data Collection	1
1.1	Sample Surveys	2
1.2	Fitness Club: Example of a Sample Survey	4
1.3	Experiments	4
1.4	Experiments: An Example	4
1.5	Data Collection	5
1.6	Registers	6
1.7	Questionnaire Surveys	6
1.7.1	Background Questions	7
1.7.2	Study Questions	7
1.8	Sources of Errors in Surveys	9
1.9	Comparing Methods of Data Collection	11
1.10	Example Continued	12
2	Presentation of Data	13
2.1	Bar Charts	13
2.2	Histograms	14
2.3	Pie Charts	17
2.4	Scatter Plots	18
2.5	Line Charts	18
2.6	Bubble Plots	20
2.7	Tables	20
2.7.1	The Ingredients of a Table	21
2.7.2	Percentages	22
3	Description of Data	25
3.1	Systematic and Random Variation	25
3.2	Measures of Location	27
3.2.1	Average	27
3.2.2	Median	28

3.2.3 Mode	29
3.2.4 Choosing a Measure of Location	30
3.3 Measures of Dispersion	32
3.3.1 Range	32
3.3.2 Variance and Standard Deviation	33
3.3.3 Interquartile Range	36
3.3.4 Choosing a Measure of Dispersion	37
3.3.5 Relative Spread (Dispersion)	38
3.4 Example: Statistical Functions in Spreadsheets	39
3.5 Data Type and Descriptive Statistics	40
3.5.1 Data Types	41
3.5.2 Descriptive Statistics and Type of Data	41
4 The Normal Distribution	43
4.1 Characteristics of the Normal Distribution	43
4.2 Density Function and Distribution Function	45
4.3 Fractiles	46
4.4 Calculations in the Normal Distribution	47
4.5 The Normal Distribution and Spreadsheets	48
4.5.1 NORMDIST (X; Mean; Stdev; Cumulative)	49
4.5.2 NORMINV (Probability; Mean; Stdev)	49
4.5.3 Example	49
4.6 Testing for the Normal Distribution	50
4.6.1 Simple Methods	51
4.6.2 Skewness and Kurtosis	52
4.6.3 Normal Plot	55
4.7 Random Numbers	56
4.8 Confidence Intervals	57
4.8.1 Confidence Interval for the Mean	58
4.8.2 Confidence Interval for the Mean in Case of a Small Sample	62
4.8.3 Confidence Interval for the Standard Deviation	66
4.9 More About the Normal Distribution	68
5 Analysis of Qualitative Data	71
5.1 The Binomial Distribution	71
5.1.1 Example	72
5.2 The Binomial Distribution and the Normal Distribution	73
5.3 The Binomial Distribution in Spreadsheets	75
5.3.1 Example	75
5.4 Statistical Uncertainty in Sample Surveys	76
5.4.1 Example	77
5.5 Is the Sample Representative?	80
5.6 Statistical Tests	81

5.6.1 Example	82
5.6.2 Approximation with the Normal Distribution	83
5.6.3 Significance Level	84
5.6.4 Statistical Test or Confidence Interval	85
5.7 Frequency Tables	85
5.7.1 Introduction to Chi-Squared Test	85
5.7.2 Confidence Interval for Difference Between Two Proportions	89
5.7.3 Several Rows and/or Columns	89
5.7.4 Calculations in Spreadsheets	92
5.7.5 Calculations by Calculator	93
6 Error Sources and Planning	95
6.1 Two Kinds of Errors	95
6.2 Random Error and Sample Size	95
6.2.1 A Qualitative Variable	96
6.2.2 A Quantitative Variable	99
6.3 Bias (Systematic Errors)	100
6.3.1 Errors in the Sampling (Sample Selection)	101
6.3.2 Errors in the Definition of the Sample	101
6.3.3 What Is a Representative Sample?	102
6.4 Sampling (Sample Selection)	103
6.4.1 Simple Random Sampling	103
6.4.2 Stratified Sampling	104
6.4.3 Cluster Sampling	105
6.4.4 Systematic Sampling	107
6.4.5 Quota Sampling	107
6.4.6 Purposive Sampling	108
6.4.7 Convenience Sampling	108
7 Assessment of Relationship	111
7.1 Example	112
7.2 Linear Regression with Spreadsheets	115
7.3 Is There a Relationship?	117
7.3.1 Note	119
7.4 Multiple Linear Regression	119
7.5 Final Remarks	120
8 Comparing Two Groups	121
8.1 Matched Pairs: The Paired t -Test	121
8.1.1 Example	121
8.1.2 Description	122
8.1.3 Calculation	123
8.1.4 Spreadsheets	125

8.2 Comparing Two Groups Means	125
8.2.1 Example	125
8.2.2 Description	126
8.2.3 Calculation	127
8.2.4 Spreadsheets	129
8.2.5 Size of an Experiment	129
8.3 Other Statistical Tests for Two Groups	130
8.3.1 Test for the Same Variance in the Two Groups	130
8.3.2 Comparing Two Group Means: Two Samples with Equal Variances	130
8.4 Final Remarks	131
9 Appendices	133
9.1 Probability Theory	133
9.1.1 Sample Space, Events, and Probability	133
9.1.2 Random Variables; the Binomial Distribution	138
9.1.3 Random Variables: Mean and Variance	141
9.1.4 Technical Note: The Binomial Coefficient	142
9.2 Summary of Statistical Methods	143
9.2.1 Quantitative Data	143
9.2.2 Qualitative Data	145
9.3 Statistical Functions in Spreadsheets	146
9.4 Statistical Tables	147
9.4.1 Fractiles in the Normal Distribution	147
9.4.2 Probabilities in the Normal Distribution	148
9.4.3 Table of the <i>t</i> -Distribution	148
9.4.4 Table of the Chi-Squared Distribution	148
9.4.5 Statistical Uncertainty in Sample Surveys	149
9.5 Fitness Club: Data from the Sample Survey	151
9.6 Where to Go from Here	152
9.6.1 Literature	152
9.6.2 Useful Links	153
9.6.3 Overview of Statistical Software	154
9.7 Glossary	155
Index	159