CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

Contents

Preface *page* xv

1 Introduction 1
1.1 Why do earth scientists need to understand experimental design and statistics? 1
1.2 What is this book designed to do? 6

2 "Doing science": hypotheses, experiments and disproof 8
2.1 Introduction 8
2.2 Basic scientific method 8
2.3 Making a decision about a hypothesis 11
2.4 Why can't a hypothesis or theory ever be proven? 11
2.5 "Negative" outcomes 12
2.6 Null and alternate hypotheses 12
2.7 Conclusion 13
2.8 Questions 14

3 Collecting and displaying data 15
3.1 Introduction 15
3.2 Variables, sampling units and types of data 15
3.3 Displaying data 17
3.4 Displaying ordinal or nominal scale data 21
3.5 Bivariate data 21
3.6 Data expressed as proportions of a total 25
3.7 Display of geographic direction or orientation 26
3.8 Multivariate data 26
3.9 Conclusion 27

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

4 Introductory concepts of experimental design 28
4.1 Introduction 28
4.2 Sampling: mensurative experiments 29
4.3 Manipulative experiments 34
4.4 Sometimes you can only do an unreplicated experiment 40
4.5 Realism 41
4.6 A bit of common sense 42
4.7 Designing a "good" experiment 43
4.8 Conclusion 44
4.9 Questions 44

5 Doing science responsibly and ethically 45
5.1 Introduction 45
5.2 Dealing fairly with other people's work 45
5.3 Doing the sampling or the experiment 47
5.4 Evaluating and reporting results 48
5.5 Quality control in science 50
5.6 Questions 50

6 Probability helps you make a decision about your results 51
6.1 Introduction 51
6.2 Statistical tests and significance levels 52
6.3 What has this got to do with making a decision or statistical testing? 57
6.4 Making the wrong decision 57
6.5 Other probability levels 58
6.6 How are probability values reported? 60
6.7 All statistical tests do the same basic thing 60
6.8 A very simple example: the chi-square test for goodness of fit 60
6.9 What if you get a statistic with a probability of exactly 0.05? 64
6.10 Conclusion 65
6.11 Questions 65

7 Working from samples: data, populations and statistics 66
7.1 Using a sample to infer the characteristics of a population 66
7.2 Statistical tests 66

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

7.3 The normal distribution 66
7.4 Samples and populations 71
7.5 Your sample mean may not be an accurate estimate of the population mean 73
7.6 What do you do when you only have data from one sample? 75
7.7 Why are the statistics that describe the normal distribution so important? 78
7.8 Distributions that are not normal 80
7.9 Other distributions 80
7.10 Other statistics that describe a distribution 82
7.11 Conclusion 83
7.12 Questions 84

8 Normal distributions: tests for comparing the means of one and two samples 85
8.1 Introduction 85
8.2 The 95% confidence interval and 95% confidence limits 85
8.3 Using the *Z* statistic to compare a sample mean and population mean when population statistics are known 86
8.4 Comparing a sample mean to an expected value when population statistics are not known 87
8.5 Comparing the means of two related samples 96
8.6 Comparing the means of two independent samples 98
8.7 Are your data appropriate for a *t* test? 100
8.8 Distinguishing between data that should be analyzed by a paired-sample test and a test for two independent samples 102
8.9 Conclusion 103
8.10 Questions 103

9 Type 1 and Type 2 error, power and sample size 105
9.1 Introduction 105
9.2 Type 1 error 105
9.3 Type 2 error 106
9.4 The power of a test 109

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

9.5	What sample size do you need to ensure the risk of Type 2 error is not too high?	111
9.6	Type 1 error, Type 2 error and the concept of risk	113
9.7	Conclusion	113
9.8	Questions	114
10	**Single-factor analysis of variance**	115
10.1	Introduction	115
10.2	Single-factor analysis of variance	116
10.3	An arithmetic/pictorial example	122
10.4	Unequal sample sizes (unbalanced designs)	128
10.5	An ANOVA does not tell you which particular treatments appear to be from different populations	128
10.6	Fixed or random effects	128
10.7	Questions	129
11	**Multiple comparisons after ANOVA**	131
11.1	Introduction	131
11.2	Multiple comparison tests after a Model I ANOVA	131
11.3	An a posteriori Tukey comparison following a significant result for a single-factor Model I ANOVA	134
11.4	Other a posteriori multiple comparison tests	138
11.5	Planned comparisons	138
11.6	Questions	140
12	**Two-factor analysis of variance**	142
12.1	Introduction	142
12.2	What does a two-factor ANOVA do?	145
12.3	How does a two-factor ANOVA analyze these data?	146
12.4	How does a two-factor ANOVA separate out the effects of each factor and interaction?	150
12.5	An example of a two-factor analysis of variance	153
12.6	Some essential cautions and important complications	154
12.7	Unbalanced designs	164
12.8	More complex designs	164
12.9	Questions	165

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

13 Important assumptions of analysis of variance, transformations and a test for equality of variances 166
13.1 Introduction 166
13.2 Homogeneity of variances 166
13.3 Normally distributed data 167
13.4 Independence 171
13.5 Transformations 171
13.6 Are transformations legitimate? 172
13.7 Tests for heteroscedasticity 174
13.8 Questions 176

14 Two-factor analysis of variance without replication, and nested analysis of variance 178
14.1 Introduction 178
14.2 Two-factor ANOVA without replication 178
14.3 A posteriori comparison of means after a two-factor ANOVA without replication 183
14.4 Randomized blocks 184
14.5 Nested ANOVA as a special case of a single-factor ANOVA 185
14.6 A pictorial explanation of a nested ANOVA 187
14.7 A final comment on ANOVA: this book is only an introduction 192
14.8 Questions 192

15 Relationships between variables: linear correlation and linear regression 194
15.1 Introduction 194
15.2 Correlation contrasted with regression 195
15.3 Linear correlation 195
15.4 Calculation of the Pearson *r* statistic 196
15.5 Is the value of *r* statistically significant? 202
15.6 Assumptions of linear correlation 202
15.7 Conclusion 202
15.8 Questions 203

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

16 Linear regression 204
16.1 Introduction 204
16.2 Linear regression 204
16.3 Calculation of the slope of the regression line 205
16.4 Calculation of the intercept with the *Y* axis 208
16.5 Testing the significance of the slope and the intercept of the regression line 211
16.6 An example: school cancellations and snow 217
16.7 Predicting a value of *Y* from a value of *X* 219
16.8 Predicting a value of *X* from a value of *Y* 219
16.9 The danger of extrapolating beyond the range of data available 220
16.10 Assumptions of linear regression analysis 220
16.11 Multiple linear regression 223
16.12 Further topics in regression 224
16.13 Questions 225

17 Non-parametric statistics 227
17.1 Introduction 227
17.2 The danger of assuming normality when a population is grossly non-normal 227
17.3 The value of making a preliminary inspection of the data 229

18 Non-parametric tests for nominal scale data 230
18.1 Introduction 230
18.2 Comparing observed and expected frequencies: the chi-square test for goodness of fit 231
18.3 Comparing proportions among two or more independent samples 234
18.4 Bias when there is one degree of freedom 237
18.5 Three-dimensional contingency tables 242
18.6 Inappropriate use of tests for goodness of fit and heterogeneity 242
18.7 Recommended tests for categorical data 243
18.8 Comparing proportions among two or more related samples of nominal scale data 243
18.9 Questions 245

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

19 Non-parametric tests for ratio, interval or ordinal scale data 247
19.1 Introduction 247
19.2 A non-parametric comparison between one sample and an expected distribution 248
19.3 Non-parametric comparisons between two independent samples 250
19.4 Non-parametric comparisons among more than two independent samples 256
19.5 Non-parametric comparisons of two related samples 259
19.6 Non-parametric comparisons among three or more related samples 262
19.7 Analyzing ratio, interval or ordinal data that show gross differences in variance among treatments and cannot be satisfactorily transformed 264
19.8 Non-parametric correlation analysis 266
19.9 Other non-parametric tests 268
19.10 Questions 268

20 Introductory concepts of multivariate analysis 270
20.1 Introduction 270
20.2 Simplifying and summarizing multivariate data 271
20.3 An *R*-mode analysis: principal components analysis 272
20.4 How does a PCA combine two or more variables into one? 273
20.5 What happens if the variables are not highly correlated? 276
20.6 PCA for more than two variables 277
20.7 The contribution of each variable to the principal components 279
20.8 An example of the practical use of principal components analysis 282
20.9 How many principal components should you plot? 282
20.10 How much variation must a PCA explain before it is useful? 283
20.11 Summary and some cautions and restrictions on use of PCA 283
20.12 *Q*-mode analyses: multidimensional scaling 284
20.13 How is a univariate measure of dissimilarity among sampling units extracted from multivariate data? 285

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

20.14 An example 287
20.15 Stress 289
20.16 Summary and cautions on the use of multidimensional scaling 290
20.17 *Q*-mode analyses: cluster analysis 291
20.18 Which multivariate analysis should you use? 295
20.19 Questions 295

21 Introductory concepts of sequence analysis 297
21.1 Introduction 297
21.2 Sequences of ratio, interval or ordinal scale data 298
21.3 Preliminary inspection by graphing 298
21.4 Detection of within-sequence similarity and dissimilarity 299
21.5 Cross-correlation 307
21.6 Regression analysis 308
21.7 Simple linear regression 309
21.8 More complex regression 311
21.9 Simple autoregression 317
21.10 More complex series with a cyclic component 320
21.11 Statistical packages and time series analysis 322
21.12 Some very important limitations and cautions 322
21.13 Sequences of nominal scale data 323
21.14 Records of the repeated occurrence of an event 327
21.15 Conclusion 331
21.16 Questions 332

22 Introductory concepts of spatial analysis 334
22.1 Introduction 334
22.2 Testing whether a spatial distribution occurs at random 335
22.3 Data for the direction of objects 346
22.4 Prediction and interpolation in two dimensions 352
22.5 Conclusion 362
22.6 Questions 362

23 Choosing a test 364
23.1 Introduction 364

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-76322-6 - Geostatistics Explained: An Introductory Guide for Earth Scientists
Steve McKillup and Melinda Darby Dyar
Table of Contents
More information

Appendices
Appendix A *Critical values of chi-square, t and F* 374
Appendix B *Answers to questions* 380
References 389
Index 391

© in this web service Cambridge University Press www.cambridge.org