

**THE POWER OF CATEGORICAL
GOODNESS-OF-FIT TEST
STATISTICS**

Michael C Steele

**Submitted in fulfilment of the requirements of
the degree of Doctor of Philosophy**

July 2002

Australian School of Environmental Studies

Faculty of Environmental Sciences

Griffith University

ABSTRACT

The relative power of goodness-of-fit test statistics has long been debated in the literature. Chi-Square type test statistics to determine 'fit' for categorical data are still dominant in the goodness-of-fit arena. Empirical Distribution Function type goodness-of-fit test statistics are known to be relatively more powerful than Chi-Square type test statistics for restricted types of null and alternative distributions. In many practical applications researchers who use a standard Chi-Square type goodness-of-fit test statistic ignore the rank of ordinal classes.

This thesis reviews literature in the goodness-of-fit field, with major emphasis on categorical goodness-of-fit tests. The continued use of an asymptotic distribution to approximate the exact distribution of categorical goodness-of-fit test statistics is discouraged. It is unlikely that an asymptotic distribution will produce a more accurate estimation of the exact distribution of a goodness-of-fit test statistic than a Monte Carlo approximation with a large number of simulations.

Due to their relatively higher powers for restricted types of null and alternative distributions, several authors recommend the use of Empirical Distribution Function test statistics over nominal goodness-of-fit test statistics such as Pearson's Chi-Square. In-depth power studies confirm the views of other authors that categorical Empirical Distribution Function type test statistics do not have higher power for some common null and alternative distributions. Because of this, it is not sensible to make a conclusive recommendation to always use an Empirical Distribution Function type test statistic instead of a nominal goodness-of-fit test statistic.

Traditionally the recommendation to determine 'fit' for multivariate categorical data is to treat categories as nominal, an approach which precludes any gain in power which may accrue from a ranking, should one or more variables be ordinal. The presence of multiple criteria through multivariate data may result in partially ordered categories, some of which have equal ranking. This thesis proposes a modification to the currently available Kolmogorov-Smirnov test statistics for ordinal and nominal categorical data to account for situations of partially ordered categories. The new test statistic, called the Combined Kolmogorov-Smirnov, is relatively more powerful than Pearson's Chi-Square and the nominal Kolmogorov-Smirnov test statistic for some null and alternative distributions. A recommendation is made to use the new test statistic with higher power in situations where some benefit can be achieved by incorporating an Empirical Distribution Function approach, but the data lack a complete natural ordering of categories.

The new and established categorical goodness-of-fit test statistics are demonstrated in the analysis of categorical data with brief applications as diverse as familiarity of defence programs, the number of recruits produced by the Merlin bird, a demographic problem, and DNA profiling of genotypes. The results from these applications confirm the recommendations associated with specific goodness-of-fit test statistics throughout this thesis.

ACKNOWLEDGEMENTS

I would like to thank all the staff and students from the Faculty of Environmental Sciences at Griffith University for supporting and funding my work for this PhD. Thanks to Bill Hogarth for the right advice at the right times. To Lynne Bradshaw for help with the presentation of this thesis. To Brendon Holley (The Visual Basic Guru), thank you for your programming help and advice. To Cameron Hurst, Glenn Scarffe and David Appleby, thank you for our regular 'statistics' discussions at the GU pub.

To the School of Mathematical and Physical Sciences at James Cook University thank you for the time to complete this thesis. Thank you to the Forensic Biology Laboratory for the DNA profiling data.

To Mum, Dad and all my family and friends that I am not going to name here, thank you for your support and encouragement when most of you had no idea what I was doing.

To my wife Tracee, thank you for your understanding and support.

Finally to my PhD supervisor Janet Chaseling. Thank you for your time, support, encouragement and enthusiasm for my work, your work, your research students and everything else. It has been a pleasure.

DECLARATION OF ORIGINALITY

This thesis presents the original research of the author. This work has not been previously submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person, except where due reference is made in the thesis itself.

Michael C Steele

TABLE OF CONTENTS

Abstract	i
Acknowledgements	iii
Declaration of Originality	iv
Table of Contents	v
List of Tables	viii
List of Figures	xi
Glossary	xv
Chapter 1: Introduction	1
Chapter 2: A Review of Goodness-of-Fit Test Statistics	9
2.1 Introduction	9
2.2 The Distribution of a Categorical Goodness-of-Fit Test Statistic	11
2.2.1 The Exact Distribution of a Categorical Goodness-of-Fit Test Statistic	11
2.2.2 Monte Carlo Approximation of the Exact Distribution of a Categorical Goodness-of-Fit Test Statistic	12
2.3 Chi-Square Type Test Statistics	14
2.3.1 Pearson's Chi-Square Test Statistic	14
2.3.2 Common Chi-Square Type Test Statistics	15
2.3.3 Conditions where the Chi-Square Distribution is a Valid Approximation of the Exact Distribution of the Chi-Square Test Statistic	16
2.4 Empirical Distribution Function Type Test Statistics	19
2.4.1 Introduction	19
2.4.2 Kolmogorov-Smirnov Test Statistics	20
2.4.3 Cramér-von Mises Test Statistics	23
2.4.4 Riedwyl Test Statistics	25
2.5 Conclusion	25
Chapter 3: Aspects of the Power Studies	27
3.1 Introduction	27
3.2 The GFIT Computer Software	30
3.2.1 Introduction	30
3.2.2 Development of the GFIT Software	31
3.2.3 Use of the GFIT Software	31
3.2.4 Limitations of the GFIT Software	32
3.2.5 Validation of the GFIT Software	32
3.3 Linear Interpolated Power	33
3.4 Consistent Trend Distributions	34
3.5 Step Type Trend Distributions	35
3.6 Triangular Distributions	37
3.7 Platykurtic Distributions	38
3.8 Leptokurtic Distributions	39
3.9 Bimodal Distributions	40
3.10 Conclusion	41

Chapter 4: Power Studies	43
4.1 Introduction	43
4.2 Uniform Null Distributions.....	46
4.2.1 Consistent Trend Alternative Distributions.....	46
4.2.1.1 Alternative Distribution $A_1(0.1)$	47
4.2.1.2 Alternative Distribution $A_1(0.5)$	48
4.2.1.3 Alternative Distribution $A_1(0.8)$	51
4.2.1.4 Alternative Distribution $A_1(0.9)$	53
4.2.1.5 Alternative Distribution $A_1(1.1)$	54
4.2.1.6 Alternative Distribution $A_1(1.2)$	56
4.2.1.7 Alternative Distribution $A_1(1.5)$	57
4.2.1.8 Alternative Distribution $A_1(2)$	59
4.2.1.9 Alternative Distribution $A_1(3)$	61
4.2.2 Step Type Trend Alternative Distributions	62
4.2.2.1 Alternative Distribution $A_2(\delta)$ with $\delta < \frac{1}{2k}$	63
4.2.2.2 Alternative Distribution $A_2(\delta)$ with $\delta \approx \frac{1}{2k}$	65
4.2.2.3 Alternative Distribution $A_2(\delta)$ with $\delta = \frac{1}{k}$	66
4.2.3 Triangular Alternative Distributions	68
4.2.4 Platykurtic Alternative Distributions.....	70
4.2.5 Leptokurtic Alternative Distributions.....	72
4.2.6 Bimodal Alternative Distributions	74
4.2.7 Alternative Distributions with a Small Number of Categories	75
4.2.8 Recommendations for Uniform Null Distributions	80
4.2.9 Application – Uniform Null Distribution.....	81
4.3 Non-Uniform Null Distributions.....	83
4.3.1 Consistent Trend Null Distributions.....	83
4.3.1.1 Null Distribution $A_1(0.5)$	83
4.3.1.2 Null Distribution $A_1(1.5)$	84
4.3.2 Step Type Null Distributions.....	85
4.3.3 Triangular Null Distributions.....	87
4.3.4 Platykurtic Null Distributions	88
4.3.5 Leptokurtic Null Distributions	89
4.3.6 Bimodal Null Distributions	91
4.3.7 Recommendations for Non-Uniform Null Distributions	92
4.3.8 Application - Non-Uniform Null Distribution	93
4.4 Recommendations and Conclusions	94
Chapter 5: A Modified Kolmogorov-Smirnov Test Statistic for Partially Ranked Categorical Data	98
5.1 Introduction	98
5.2 The Effect Category Ordering has on the Power of OrdKS	99
5.2.1 Introduction	99
5.2.2 Results and Discussion.....	100
5.3 A New EDF Type Test Statistic for Partially Ranked Categorical Data	100
5.3.1 Introduction	100
5.3.2 Largest Cumulative Difference for Ordinal Subsets of Categories	102
5.3.3 The Combined Kolmogorov-Smirnov Test Statistic	104
5.4 Power of the Combined Kolmogorov-Smirnov Test Statistic for a Uniform Null Distribution	105
5.4.1 Introduction	105
5.4.2 Consistent Trend Alternative Distributions.....	106
5.4.2.1 Alternative Distribution $A_1(0.1)$	106

5.4.2.2	Alternative Distribution $A_1(0.5)$	106
5.4.2.3	Alternative Distribution $A_1(0.8)$	109
5.4.2.4	Alternative Distribution $A_1(0.9)$	109
5.4.2.5	Alternative Distribution $A_1(1.1)$	112
5.4.2.6	Alternative Distribution $A_1(1.2)$	112
5.4.2.7	Alternative Distribution $A_1(1.5)$	115
5.4.2.8	Alternative Distribution $A_1(2)$	115
5.4.2.9	Alternative Distribution $A_1(3)$	118
5.4.3	Step Type Trend Alternative Distributions	119
5.4.3.1	Alternative Distribution $A_2(\delta)$ with $\delta < \frac{1}{2k}$	119
5.4.3.2	Alternative Distribution $A_2(\delta)$ with $\delta \approx \frac{1}{2k}$	121
5.4.3.3	Alternative Distribution $A_2(\delta)$ with $\delta = \frac{1}{k}$	121
5.4.4	Triangular Alternative Distributions	123
5.4.5	Platykurtic Alternative Distributions	125
5.4.6	Leptokurtic Alternative Distributions	126
5.4.7	Bimodal Alternative Distributions	128
5.4.8	Application – Uniform Null Distribution	128
5.5	Power of the Combined Kolmogorov-Smirnov Test Statistic for Non-Uniform Null Distributions	130
5.5.1	Introduction	130
5.5.2	Consistent Trend Null Distributions	130
5.5.3	Step Type Null Distributions	132
5.5.4	Triangular Null Distributions	133
5.5.5	Platykurtic Null Distributions	133
5.5.6	Leptokurtic Null Distributions	134
5.5.7	Bimodal Null Distributions	134
5.5.8	Application – DNA Profiling of Genotypes	135
5.6	Conclusion	141
	Chapter 6: Conclusion	144
	References	152
	Appendices	156
Appendix 1	Validation of GFIT – Comparing Exact and Simulated p -values	157
Appendix 2	Validation of GFIT – Comparing the Number of Simulations and p -values	162
Appendix 3	Example of Output from GFIT	167
Appendix 4	Graphical Results from Power Studies	168

LIST OF TABLES

Table 2.1	A sample of Chi-Square type test statistics	15
Table 2.2	Conditions where the Chi-Square approximation is valid	19
Table 2.3	Cramér-von Mises test statistics	24
Table 4.1	Description of the range of power values	45
Table 4.2	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.1)$ alternative	47
Table 4.3	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.5)$ alternative	49
Table 4.4	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.8)$ alternative	51
Table 4.5	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.9)$ alternative	53
Table 4.6	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.1)$ alternative	55
Table 4.7	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.2)$ alternative	56
Table 4.8	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.5)$ alternative	58
Table 4.9	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(2)$ alternative	60
Table 4.10	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(3)$ alternative	62
Table 4.11	Powers of OrdKS, NomKS and Chi-Square – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta < \frac{1}{2k}$	64
Table 4.12	Powers of OrdKS, NomKS and Chi-Square – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta \approx \frac{1}{2k}$	65
Table 4.13	Powers of OrdKS, NomKS and Chi-Square – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta = \frac{1}{k}$	67
Table 4.14	Powers of OrdKS, NomKS and Chi-Square – Uniform null against triangular A_3 alternative	68
Table 4.15	Powers of OrdKS, NomKS and Chi-Square – Uniform null against platykurtic A_4 alternative	70
Table 4.16	Powers of OrdKS, NomKS and Chi-Square – Uniform null against leptokurtic A_5 alternative	72
Table 4.17	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(\delta)$ alternative with $0 < \delta < 1$ for a small number of categories	77
Table 4.18	Powers of OrdKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(\delta)$ alternative with $\delta > 1$ for a small number of categories	78
Table 4.19	Powers of OrdKS, NomKS and Chi-Square – Uniform null against step type trend $A_2(\delta)$ alternative for a small number of categories	79
Table 4.20	Powers of OrdKS, NomKS and Chi-Square – Uniform null against triangular A_3 , platykurtic A_4 and leptokurtic A_5 alternative for a small number of categories	80
Table 4.21	Survey data on familiarity of defence programs as reported by Penska and Thai (2000)	81
Table 4.22	Analysis of survey data on familiarity of defence programs	82
Table 4.23	Number of recruits produced by male Merlins for age groups	93

Table 4.24	Analysis of number of recruits produced by male Merlins for age groups..	94
Table 5.1	Rank of the categories for CombKS in the power studies	105
Table 5.2	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.1)$ alternative	107
Table 5.3	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.5)$ alternative	108
Table 5.4	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.8)$ alternative	110
Table 5.5	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.9)$ alternative	111
Table 5.6	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.1)$ alternative.....	113
Table 5.7	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.2)$ alternative	114
Table 5.8	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.5)$ alternative	116
Table 5.9	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(2)$ alternative	117
Table 5.10	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(3)$ alternative	119
Table 5.11	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta < \frac{1}{2k}$	120
Table 5.12	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta \approx \frac{1}{2k}$	122
Table 5.13	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta = \frac{1}{k}$	123
Table 5.14	Powers of CombKS, NomKS and Chi-Square - Uniform null against triangular A_3 alternative	124
Table 5.15	Powers of CombKS, NomKS and Chi-Square - Uniform null against platykurtic A_4 alternative	126
Table 5.16	Powers of CombKS, NomKS and Chi-Square - Uniform null against leptokurtic A_5 alternative	127
Table 5.17	Number of males and females for various age groups	129
Table 5.18	Analysis of number of males and females for various age groups	129
Table 5.19	Allelic probabilities for the HUMTHO1 locus from 509 individuals in the Queensland Forensic Science database in 1996	136
Table 5.20	Genotype frequencies for the HUMTHO1 locus	137
Table 5.21	Genotype frequencies with rank of genotype based on rank of alleles.....	139
Table 5.22	Power of Chi-Square, NomKS and CombKS for various inbreeding coefficients with subsets of genotypes with the same rank	140
Table 6.1	Power recommendations for 6 categorical GOF test statistics – Consistent trend distributions.....	146
Table 6.2	Power recommendations for 6 categorical GOF test statistics – Step type trend distributions.....	147
Table 6.3	Power recommendations for 6 categorical GOF test statistics – Triangular distributions	147
Table 6.4	Power recommendations for 6 categorical GOF test statistics – Platykurtic distributions	148
Table 6.5	Power recommendations for 6 categorical GOF test statistics – Leptokurtic distributions	148
Table 6.6	Power recommendations for 6 categorical GOF test statistics – Bimodal distributions	149
Table A1.1	Exact and simulated p -values for uniform null distributions.....	157

Table A1.2	Exact and simulated p -values for non-uniform null distributions	161
Table A2.1	Number of simulations and p -values of GOF test statistics for 6 categories with uniform expected frequencies of 2 and common OrdKS value.....	162
Table A2.2	Number of simulations and p -values for Kolmogorov-Smirnov test statistics for 7 categories ranked 1223334 with uniform expected frequencies of 2 and common OrdKS value	165

LIST OF FIGURES

Figure 3.1	Consistent trend distribution	29
Figure 3.2	Step type trend distribution	29
Figure 3.3	Triangular distribution	29
Figure 3.4	Platykurtic distribution	30
Figure 3.5	Leptokurtic distribution	30
Figure 3.6	Bimodal distribution.....	30
Figure 3.7	Consistent trend $A_1(0.5)$, $A_1(1.5)$ and uniform distributions for 10 categories	35
Figure 3.8	Step type trend $A_2(0.05)$ and uniform distributions for 10 categories.....	37
Figure 3.9	Triangular A_3 and uniform distributions for 10 categories	38
Figure 3.10	Platykurtic A_4 and uniform distributions for 10 categories	39
Figure 3.11	Leptokurtic A_5 and uniform distributions for 10 categories	40
Figure 3.12	Bimodal A_6 and uniform distributions for 10 categories.....	41
Figure 4.1	Uniform null against consistent trend $A_1(0.1)$ alternative for 10 categories.....	47
Figure 4.2	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(0.1)$ alternative for 10 categories.....	48
Figure 4.3	Uniform null against consistent trend $A_1(0.5)$ alternative for 10 categories.....	48
Figure 4.4	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(0.5)$ alternative for 10 categories	50
Figure 4.5	Uniform null against consistent trend $A_1(0.8)$ alternative for 10 categories.....	51
Figure 4.6	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(0.8)$ alternative for 10 categories	52
Figure 4.7	Uniform null against consistent trend $A_1(0.9)$ alternative for 10 categories.....	53
Figure 4.8	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(0.9)$ alternative for 10 categories	54
Figure 4.9	Uniform null against consistent trend $A_1(1.1)$ alternative for 10 categories.....	54
Figure 4.10	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(1.1)$ alternative for 10 categories	55
Figure 4.11	Uniform null against consistent trend $A_1(1.2)$ alternative for 10 categories.....	56
Figure 4.12	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(1.2)$ alternative for 10 categories.....	57
Figure 4.13	Uniform null against consistent trend $A_1(1.5)$ alternative for 10 categories.....	57
Figure 4.14	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(1.5)$ alternative for 10 categories	59
Figure 4.15	Uniform null against consistent trend $A_1(2)$ alternative for 10 categories.....	59
Figure 4.16	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(2)$ alternative for 10 categories	61
Figure 4.17	Uniform null against consistent trend $A_1(3)$ alternative for 10 categories.....	61

Figure 4.18	Powers of all 6 GOF test statistics – Uniform null against consistent trend $A_1(3)$ alternative for 10 categories	62
Figure 4.19	Uniform null against step type trend $A_2(0.01)$ alternative for 10 categories.....	63
Figure 4.20	Powers of all 6 GOF test statistics – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta < \frac{1}{2k}$ for 10 categories.....	64
Figure 4.21	Uniform null against step type trend $A_2(0.05)$ alternative for 10 categories.....	65
Figure 4.22	Powers of all 6 GOF test statistics – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta \approx \frac{1}{2k}$ for 10 categories.....	66
Figure 4.23	Uniform null against step type trend $A_2(0.1)$ alternative for 10 categories.....	66
Figure 4.24	Powers of all 6 GOF test statistics – Uniform null against step type trend $A_2(\delta)$ alternative with $\delta = \frac{1}{k}$ for 10 categories.....	67
Figure 4.25	Uniform null against triangular A_3 alternative for 10 categories	68
Figure 4.26	Powers of all 6 GOF test statistics – Uniform null against triangular A_3 alternative for 10 categories.....	69
Figure 4.27	Uniform null against platykurtic A_4 alternative for 10 categories	70
Figure 4.28	Powers of all 6 GOF test statistics – Uniform null against platykurtic A_4 alternative for 10 categories.....	71
Figure 4.29	Uniform null against leptokurtic A_5 alternative for 10 categories	72
Figure 4.30	Powers of all 6 GOF test statistics – Uniform null against leptokurtic A_5 alternative for 10 categories.....	73
Figure 4.31	Uniform null against bimodal A_6 alternative for 10 categories.....	74
Figure 4.32	Powers of all 6 GOF test statistics – Uniform null against bimodal A_6 alternative for 10 categories.....	75
Figure 4.33	Consistent trend $A_1(0.5)$ distribution for 3 categories	75
Figure 4.34	Consistent trend $A_1(1.5)$ distribution for 3 categories	75
Figure 4.35	Step type trend $A_2(0.11)$ distribution for 3 categories	76
Figure 4.36	Triangular A_3 distribution for 3 categories	76
Figure 4.37	Platykurtic A_4 distribution for 3 categories	76
Figure 4.38	Leptokurtic A_5 distribution for 3 categories.....	76
Figure 4.39	Powers of all 6 GOF test statistics – Consistent trend $A_1(0.5)$ null against uniform alternative for 10 categories.....	84
Figure 4.40	Powers of all 6 GOF test statistics – Consistent trend $A_1(1.5)$ null against uniform alternative for 10 categories.....	85
Figure 4.41	Powers of all 6 GOF test statistics – Step type $A_2(0.05)$ null against uniform alternative for 10 categories.....	86
Figure 4.42	Powers of all 6 GOF test statistics – Triangular A_3 null against uniform alternative for 10 categories.....	87
Figure 4.43	Powers of all 6 GOF test statistics – Platykurtic A_4 null against uniform alternative for 10 categories.....	89
Figure 4.44	Powers of all 6 GOF test statistics – Leptokurtic A_5 null against uniform alternative for 10 categories.....	90
Figure 4.45	Powers of all 6 GOF test statistics – Bimodal A_6 null against uniform alternative for 10 categories.....	91

Figure 5.1	Power of OrdKS - Uniform null against consistent trend $A_1(0.1)$ alternative with different category orderings	100
Figure 5.2	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.1)$ alternative for 10 categories	107
Figure 5.3	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.5)$ alternative for 10 categories	108
Figure 5.4	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.8)$ alternative for 10 categories	110
Figure 5.5	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(0.9)$ alternative for 10 categories	111
Figure 5.6	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.1)$ alternative for 10 categories.....	113
Figure 5.7	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.2)$ alternative for 10 categories	114
Figure 5.8	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(1.5)$ alternative for 10 categories	116
Figure 5.9	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(2)$ alternative for 10 categories.....	117
Figure 5.10	Powers of CombKS, NomKS and Chi-Square – Uniform null against consistent trend $A_1(3)$ alternative for 10 categories.....	118
Figure 5.11	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta < \frac{1}{2k}$ for 10 categories	120
Figure 5.12	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta \approx \frac{1}{2k}$ for 10 categories	121
Figure 5.13	Powers of CombKS, NomKS and Chi-Square - Uniform null against step type trend $A_2(\delta)$ alternative with $\delta = \frac{1}{k}$ for 10 categories.....	122
Figure 5.14	Powers of CombKS, NomKS and Chi-Square - Uniform null against triangular A_3 alternative for 10 categories	124
Figure 5.15	Powers of CombKS, NomKS and Chi-Square - Uniform null against platykurtic A_4 alternative for 10 categories	125
Figure 5.16	Powers of CombKS, NomKS and Chi-Square - Uniform null against leptokurtic A_5 alternative for 10 categories.....	127
Figure 5.17	Powers of CombKS, NomKS and Chi-Square - Uniform null against bimodal A_6 alternative for 10 categories.....	128
Figure 5.18	Powers of CombKS, NomKS and Chi-Square – Consistent trend $A_1(0.5)$ null against uniform alternative for 10 categories	131
Figure 5.19	Powers of CombKS, NomKS and Chi-Square – Consistent trend $A_1(1.5)$ null against uniform alternative for 10 categories	132
Figure 5.20	Powers of CombKS, NomKS and Chi-Square – Step type $A_2(0.05)$ null against uniform alternative for 10 categories.....	132
Figure 5.21	Powers of CombKS, NomKS and Chi-Square – Triangular A_3 null against uniform alternative for 10 categories.....	133
Figure 5.22	Powers of CombKS, NomKS and Chi-Square – Platykurtic A_4 null against uniform alternative for 10 categories.....	134
Figure 5.23	Powers of CombKS, NomKS and Chi-Square - Leptokurtic A_5 null against uniform alternative for 10 categories.....	135
Figure 5.24	Powers of CombKS, NomKS and Chi-Square – Bimodal A_6 null against uniform alternative for 10 categories.....	135

Figure 5.25	Power of Chi-Square, NomKS and CombKS for various inbreeding coefficients with subsets of genotypes with the same rank	140
Figure 6.1	Steep decreasing trend and uniform distribution	146
Figure 6.2	Steep increasing trend and uniform distribution	146
Figure 6.3	Near uniform trend and uniform distribution	146
Figure 6.4	Step type trend and uniform distribution	147
Figure 6.5	Near uniform step type trend and uniform distribution.....	147
Figure 6.6	Triangular and uniform distribution	147
Figure 6.7	Platykurtic and uniform distribution	148
Figure 6.8	Leptokurtic and uniform distribution	148
Figure 6.9	Bimodal and uniform distribution.....	149
Figure A2.1	Number of simulations and p -values of GOF test statistics for 6 categories with uniform expected frequencies of 2 and $T_{OrdKS} = 4$	163
Figure A2.2	Number of simulations and p -values of GOF test statistics for 6 categories with uniform expected frequencies of 2 and $T_{OrdKS} = 5$	164
Figure A2.3	Number of simulations and p -values of GOF test statistics for 6 categories with uniform expected frequencies of 2 and $T_{OrdKS} = 6$	164
Figure A2.4	Number of simulations and p -values - Kolmogorov-Smirnov test statistics for 7 categories ranked 1223334 with uniform expected frequencies of 2 and $T_{OrdKS} = 4$	165
Figure A2.5	Number of simulations and p -values - Kolmogorov-Smirnov test statistics for 7 categories ranked 1223334 with uniform expected frequencies of 2 and $T_{OrdKS} = 5$	166
Figure A2.6	Number of simulations and p -values - Kolmogorov-Smirnov test statistics for 7 categories ranked 1223334 with uniform expected frequencies of 2 and $T_{OrdKS} = 6$	166
Figure A3.1	Example of output from GFIT.....	167
Figure A4.1	Power - uniform null against $A_1(0.1)$ alternative distribution.....	168
Figure A4.2	Power - uniform null against $A_1(0.5)$ alternative distribution	169
Figure A4.3	Power - uniform null against $A_1(0.8)$ alternative distribution	170
Figure A4.4	Power - uniform null against $A_1(0.9)$ alternative distribution	171
Figure A4.5	Power - uniform null against $A_1(1.1)$ alternative distribution.....	172
Figure A4.6	Power - uniform null against $A_1(1.2)$ alternative distribution.....	173
Figure A4.7	Power - uniform null against $A_1(1.5)$ alternative distribution	174
Figure A4.8	Power - uniform null against $A_1(2)$ alternative distribution	175
Figure A4.9	Power - uniform null against $A_1(3)$ alternative distribution	176
Figure A4.10	Power - uniform null against $A_2(\delta)$ alternative distribution with $\delta < \frac{1}{2k}$	177
Figure A4.11	Power - uniform null against $A_2(\delta)$ alternative distribution with $\delta \approx \frac{1}{2k}$	178
Figure A4.12	Power - uniform null against $A_2(\delta)$ alternative distribution with $\delta = \frac{1}{k}$	179
Figure A4.13	Power - uniform null against A_3 alternative distribution	180
Figure A4.14	Power - uniform null against A_4 alternative distribution	181
Figure A4.15	Power - uniform null against A_5 alternative distribution	182

GLOSSARY

Allele: an alternative form of a genetic locus

Chi-Square type test statistic: a test statistic with an asymptotic chi-square distribution

Combined Kolmogorov-Smirnov test statistic: a modification of the Kolmogorov-Smirnov test statistic for partially ranked categorical data

Cramér-von Mises test statistic: a test statistic based on the square of the cumulative difference between the empirical distribution function and the hypothesised distribution

Empirical distribution function (EDF) type test statistic: a test statistic based on the difference between the empirical distribution function and the hypothesised distribution

Genotype: the particular combination of genes in a specific individual

Goodness-of-fit (GOF) test statistic: a statistical test of how well expected frequencies fit observed frequencies

Hardy-Weinberg Law: a genetic law which proposes that alleles at a single locus are obtained independently from the parents

Heterozygote genotype: a genotype with two non-identical alleles

Homozygote genotype: a genotype with two identical alleles

Inbreeding coefficient: a level of departure from independence of alleles

Kolmogorov-Smirnov test statistic: a test statistic based on the absolute value of the difference between the empirical distribution function and the hypothesised distribution

Locus: the position of a gene on a chromosome

Partial category ranking/ordering: different categories with equivalent rank

Riedwyl test statistic: a goodness-of-fit test statistic based on equally spaced quantiles under the hypothesised distribution