

# ESTIMATING AND CORRECTING THE EFFECTS OF MODEL SELECTION UNCERTAINTY

Dissertation

Presented for the Degree of Doctor of Philosophy at the Faculty of  
Economics and Business Administration of  
the University of Göttingen

by

Georges Lucioni Edison Nguetack Tsague

from Fongo-Tongo, Cameroon.

Göttingen, 2005

# Contents

<b>1</b>	<b>Introduction and Objective</b>	<b>1</b>
1.1	Background and motivation . . . . .	1
1.2	Related work . . . . .	3
1.2.1	Post-model-selection estimators . . . . .	3
1.2.2	Model averaging . . . . .	4
1.3	Specific objectives . . . . .	5
1.4	Outline of the thesis . . . . .	6
<b>2</b>	<b>Model Uncertainty and Model Averaging</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	Model averaging and optimal weights . . . . .	11
2.2.1	Model averaging . . . . .	11
2.2.2	Performance of model averaging estimators . . . . .	12
2.3	Bayesian model averaging . . . . .	14
2.3.1	Description . . . . .	14
2.3.2	Theoretical performance of BMA . . . . .	15
2.3.2.1	Conditioning on the data . . . . .	15
2.3.2.2	Frequentist properties . . . . .	16
2.3.2.3	Predictive performance . . . . .	16
2.3.3	A fully Bayesian model averaging approach . . . . .	17
2.3.3.1	Bayesian decision theory . . . . .	17
2.3.3.2	The Bayesian nature of BMA . . . . .	17
2.3.3.3	Description of a fully BMA . . . . .	18
2.3.3.4	Illustration of a fully BMA method . . . . .	20
2.3.4	Model search strategies . . . . .	23
2.3.4.1	Occam's window method . . . . .	23
2.3.4.2	Markov chain Monte Carlo (MCMC) model de- composition . . . . .	24
2.3.4.3	Other methods . . . . .	24

2.3.4.4	Model space selection uncertainty in BMA . . . .	24
2.4	Frequentist model averaging . . . . .	25
2.4.1	Akaike weights and likelihood weights . . . . .	25
2.4.2	Likelihood of a model and Akaike weights interpretation . . . . .	26
2.4.3	Illustrative example . . . . .	26
<b>3</b>	<b>Model Selection Uncertainty</b>	<b>31</b>
3.1	Introduction . . . . .	31
3.2	Decision theory approach . . . . .	33
3.3	Problem, concepts and definitions . . . . .	34
3.4	Graphical representation and partition . . . . .	37
3.4.1	Direct selection . . . . .	37
3.4.2	Iterative approach to model building . . . . .	37
3.5	Comparing classical and model selection approaches for data analysis . . . . .	39
3.5.1	Frequentist approach to statistical analysis . . . . .	39
3.5.2	Frequentist model selection approach . . . . .	40
3.6	Illustrative examples of PMSEs . . . . .	40
3.6.1	Simple linear regression . . . . .	40
3.6.1.1	Pre-test estimators . . . . .	41
3.6.1.2	Post-information theory approach . . . . .	44
3.6.1.3	Post-Mallows Cp estimators . . . . .	45
3.6.1.4	Optimal selection criteria and optimal significance level $\alpha$ . . . . .	46
3.6.2	Multiple linear regression . . . . .	48
3.6.3	Testing for a given variance . . . . .	51
3.6.4	Testing for a given mean . . . . .	51
3.7	Partition of the sample space . . . . .	53
3.7.1	Model selection probabilities . . . . .	53
3.7.2	Distribution of PMSEs . . . . .	55
3.7.2.1	Distribution of likelihood ratio . . . . .	55
3.7.2.2	The simple linear regression example . . . . .	55
3.7.3	Coverage probabilities . . . . .	56
3.7.4	P-value and goodness of fit tests after model selection . . . . .	57
3.7.5	Consistency and PMSEs . . . . .	58
3.8	Conditional analysis . . . . .	58
3.9	The use of consistent model selection criteria . . . . .	59
3.9.1	Describing consistent criteria . . . . .	59
3.9.2	Asymptotic efficiency . . . . .	60
3.9.3	The Hodges' estimator example . . . . .	61

3.9.4	Linear regression . . . . .	64
3.9.4.1	Model selection probabilities . . . . .	64
3.9.4.2	Moments . . . . .	66
3.9.4.3	Densities . . . . .	66
<b>4</b>	<b>Model Selection and Frequentist Model Averaging</b>	<b>71</b>
4.1	Introduction . . . . .	71
4.2	Similarities and differences . . . . .	72
4.3	Combining model averaging and model selection . . . . .	72
4.3.1	Adjusted Akaike weights . . . . .	73
4.3.2	Adjusted likelihood weights . . . . .	74
4.4	Estimating a multivariate mean . . . . .	74
4.4.1	Variance known . . . . .	74
4.4.1.1	Improvement over James-Stein estimator . . . . .	76
4.4.1.2	Confidence sets for the mean . . . . .	77
4.4.2	Variance unknown . . . . .	78
4.5	Illustrative examples . . . . .	78
4.5.1	A simple linear regression . . . . .	78
4.5.1.1	Likelihood weights . . . . .	80
4.5.1.2	Akaike weights . . . . .	80
4.5.1.3	Adjusted Akaike weights . . . . .	82
4.5.1.4	Adjusted likelihood weights . . . . .	82
4.5.2	Estimation of proportions . . . . .	84
4.5.2.1	A two-model selection problem . . . . .	84
4.5.2.2	Multi-model choice . . . . .	89
<b>5</b>	<b>Bayesian Model Selection and Model Averaging</b>	<b>93</b>
5.1	Introduction . . . . .	93
5.2	Bayesian model selection . . . . .	94
5.2.1	Utility approach and analyst's goal . . . . .	94
5.2.1.1	Identifying the true model . . . . .	94
5.2.1.2	Choosing a model for inference . . . . .	95
5.2.1.3	Other loss functions . . . . .	96
5.2.2	BMA as model selection criterion . . . . .	97
5.2.3	Robustness for prior specification . . . . .	98
5.2.4	M-open framework . . . . .	98
5.3	Applied Bayesian inference and Bayesian model selection inference	99
5.3.1	Bayesian approach to statistical data analysis . . . . .	99
5.3.2	Bayesian model selection approach to statistical data analysis	100

5.4	Model selection uncertainty . . . . .	101
5.4.1	Bayesian post-model-selection estimator . . . . .	101
5.4.2	Long-run performance of Bayes estimators . . . . .	101
5.4.2.1	The frequentist risk . . . . .	102
5.4.2.2	Coverage probabilities . . . . .	102
5.4.2.3	Consistency . . . . .	102
5.4.3	Conditional performance of Bayes estimates . . . . .	103
5.5	Adjusted Bayesian model averaging . . . . .	103
5.5.1	Prior representation of model selection uncertainty . . . . .	103
5.5.2	Posterior representation of model selection uncertainty . . . . .	105
5.5.2.1	Posterior distribution . . . . .	105
5.5.2.2	Posterior mean and variance . . . . .	105
5.5.2.3	A basic property . . . . .	107
5.6	Estimating a multivariate mean . . . . .	107
5.7	Application to one-way ANOVA . . . . .	109
5.8	Estimating a proportion . . . . .	109
5.8.1	Long run evaluation . . . . .	110
5.8.1.1	Two-model choice . . . . .	110
5.8.1.2	Multi-model choice . . . . .	112
5.8.2	Evaluation with integrated risk . . . . .	114
<b>6</b>	<b>Model Selection and Nuisance Parameter Elimination</b>	<b>117</b>
6.1	Introduction . . . . .	117
6.2	Nuisance parameter elimination . . . . .	118
6.2.1	Profile and modified profile likelihood . . . . .	119
6.2.2	Illustrative examples . . . . .	119
6.2.2.1	Univariate normal distribution . . . . .	119
6.2.2.2	Multiple linear regression . . . . .	121
6.2.2.3	Gamma models . . . . .	121
6.3	Deriving basic distributions . . . . .	122
6.3.1	Standard normal distribution and t-distribution . . . . .	123
6.3.2	Equality of variance and $F$ distribution . . . . .	124
6.3.3	Poisson distribution as weighted binomial . . . . .	126
6.3.4	Negative binomial as weighted gamma and Poisson distribution . . . . .	127
6.3.5	Beta-binomial distribution as weighted beta and binomial distributions . . . . .	128

6.3.6	Noncentral chi-squared as weighted central chi-squared and Poisson distributions . . . . .	128
6.4	Consistency of a “within model selection criterion” . . . . .	129
<b>7</b>	<b>Bootstrap after Model Selection</b>	<b>131</b>
7.1	Introduction . . . . .	131
7.2	The complexity of PMSEs . . . . .	131
7.3	Bootstrap model selection . . . . .	132
7.4	Properties of bootstrap model selection . . . . .	134
7.5	Naive bootstrap approximation for model selection . . . . .	135
7.6	Failure of post-model-selection bootstrap . . . . .	135
7.6.1	Bootstrap model selection probability estimates . . . . .	135
7.6.2	One-way ANOVA theoretical example . . . . .	136
7.7	Practical issues . . . . .	138
7.8	An illustrative example . . . . .	139
<b>8</b>	<b>Summary and Conclusion</b>	<b>143</b>
8.1	Summary . . . . .	143
8.2	Suggestions for future research . . . . .	146
	<b>References</b>	<b>147</b>