dandelon.com
© 2008 AGI-Information Management Consultants
May be used for personal purposes only or by
libraries associated to dandelon.com network.

Series/Number 07-152

MODERN METHODS FOR ROBUST REGRESSION

Robert Andersen
University of Toronto

SAGE Publications
Los Angeles • London • New Delhi • Singapore

CONTENTS

List of Figures ix

List of Tables xi

Series Editor's Introduction xiii

Acknowledgments xv

1. Introduction 1

Defining Robustness 2
Defining Robust Regression 3
A Real-World Example: Coital Frequency of Married Couples in the 1970s 4

2. Important Background 6

Bias and Consistency 7
Breakdown Point 7
Influence Function 8
Relative Efficiency 9
Measures of Location 10
The Mean I I
a-Trimmed Mean 12
The Median 13
Measures of Scale 13
Standard Deviation 14
Mean Deviation From the Mean 14
Mean Deviation From the Median 15
Interquartile Range 15
Median A bsolute Deviation 16
M-Estimation 16
M-Estimation of Location 17
Huber Estimates 18
Biweight Estimates 19
^.-*Estimators of Scale* 20

Comparing Various Estimates 23
EXAMPLE 2.1: Simulated Data 23
EXAMPLE 2.2: Public Opinion Toward Pay Inequality in Cross-National Perspective 24'
Notes 28

3. Robustness, Resistance, and Ordinary Least Squares Regression 29

Ordinary Least Squares Regression 29
Implications of Unusual Cases for OLS Estimates and Standard Errors 31
EXAMPLE 3.1: Income Inequality and Public Opinion Toward Pay Equality in 26 New Democracies 35
Detecting Problematic Observations in OLS Regression 37
Detecting Leverage: Hat Values 38
Detecting Regression Outliers: Studentized Residuals and the Bonferroni Adjustment 39
Detecting Influence: DFBETAs, Cook's D, and Partial Regression Plots 41
Some Strategies for Dealing With Influential Cases 46
Notes 47

4. Robust Regression for the Linear Model 47

L-Estimators 47
Least Absolute Values Regression 48
Least Median of Squares Regression 49
Least Trimmed Squares Regression 49
/^-Estimators 50
M-Estimators 51
Iterative!}' Reweighted Least Squares 52
GM-Estimators 53
S-Estimators 55
Generalized ^-Estimators 55
MM-Estimators 56
Comparing the Various Estimators 57
EXAMPLE 4.1: Simulated Data 5 8
EXAMPLE 4.2: Multiple Regression Predicting Public Opinion 59
Diagnostics Revisited: Robust Regression-Related Methods for Detecting Outliers 62
Index Plots of the Weights From the Final IWLS Fit 63

RR-Plots ('•Residual-Residual" Plots) 67
Robust Distances 67
Notes 70

5. Standard Errors for Robust Regression **70**

Asymptotic Standard Errors for Robust Regression Estimators 71
Bootstrapped Standard Errors 71
Random-x Bootstrapping 72
Fixed-x Bootstrapping 72
Constructing Confidence Intervals 73
EXAMPLE 5.1: The Impact of Democracy on the Effects of Income Inequality and Per Capita GDP on Public Opinion 74
Notes 77

6. Influential Cases in Generalized Linear Models **79**

The Generalized Linear Model 79
Detecting Unusual Cases in Generalized Linear Models 82
Residuals From the GLM 82
Hat Values and Leverage 83
Assessing Influence 83
Robust Generalized Linear Models 83
M-Estimation for GLMs 84
EXAMPLE 6.1: Logistic Regression Predicting Vote for the Labour Party in Britain, 2001 85
EXAMPLE 6.2: Robust Poisson Regression Predicting Voluntary Association Membership in Quebec 87
Notes 89

7. Conclusions **90**

Appendix: Software Considerations for Robust Regression **93**

References **95**

Index **101**

About the Author **107**