INFERENCE FOR SOME GLMS AND SURVIVAL REGRESSION MODELS AFTER VARIABLE SELECTION

by

Rasanji Chathumali Rathnayake

M.S., Southern Illinois University Carbondale, 2010

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy Degree

Department of Mathematics
in the Graduate School
Southern Illinois University Carbondale
May, 2019

DISSERTATION APPROVAL

INFERENCE FOR SOME GLMS AND SURVIVAL REGRESSION MODELS AFTER VARIABLE SELECTION

By

Rasanji Chathumali Rathnayake

A Dissertation Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in the field of Mathematics

Approved by:

David Olive, Chair

Jianhong Xu

Dubravka Ban

Michael Sullivan

Jennifer Koran

Graduate School
Southern Illinois University Carbondale
(Date of Approval)

AN ABSTRACT OF THE DISSERTATION OF

Rasanji Chathumali Rathnayake, for the Doctor of Philosophy degree in Mathematics, presented on DATE OF DEFENSE, at Southern Illinois University Carbondale.

TITLE: INFERENCE FOR SOME GLMS AND SURVIVAL REGRESSION MODELS AFTER VARIABLE SELECTION

MAJOR PROFESSOR: Prof. D. J. Olive

Consider a regression model where $Y|\boldsymbol{x} \sim D(h(\boldsymbol{x}), \boldsymbol{\theta})$ for some real valued function such as $h(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta}$ where D is a parametric distribution that depends on $\boldsymbol{x}$ only through $h(\boldsymbol{x})$. Several important generalized linear models, generalized additive models, and survival regression models have this form. To obtain a prediction interval for a future value of the response variable Y_f given a vector of predictors $\boldsymbol{x}_f$, apply the nonparametric shorth prediction interval to $Y_1^*, ..., Y_B^*$ where the Y_i^* are independent and identically distributed from the distribution $D(\hat{h}(\boldsymbol{x}_f), \hat{\boldsymbol{\theta}})$. These prediction intervals can also work after variable or model selection.

This thesis also presents a method for bootstrapping some generalized linear models and survival regression models after variable selection with AIC or BIC. Hypothesis testing is done using three confidence regions.

KEY WORDS: Lasso, Logistic Regression, Model Selection, Poisson Regression, Variable Selection, Weibull Regression, Backward Elimination, Bagging, Bootstrap, Confidence Region, Forward Selection, Prediction Region.

DEDICATION

This thesis work is dedicated to my husband, Dushantha Senadhira, who has been a constant source of support and encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life. This work is also dedicated to my parents, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Dr. David Olive for invaluable support, expert guidance, understanding and encouragements throughout my study and research. Without him this dissertation would have been impossible.

Beside my advisor, I would like to thank the rest of my Ph.D. committee, Dr. Jianhong Xu, Dr. Dubravka Ban, Dr.Michael Sullivan and Dr. Jennifer Koran for their guidance and support throughout my graduate program. Also I am very grateful to all the professors who taught me at SIUC, University of Sri Jayewardenepura, Sri Lanka and special thanks goes to my teachers at Taxila Central College, Sri Lanka for teaching me the basics of Mathematics.

My sincere thanks also goes to Ms. Diane Fritcher, Ms. Linda Gibson and Ms. Naishon Patterson from the Department of Mathematics, SIUC who provided help in numerous ways in order to reach this stage.

Last but not least, I would like to thank my everloving parents, husband, grandmother, sisters and brother-in-law for their unconditional love, support and patience during the last eight years. I would not have been able to complete this dissertation without their continuous love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

In a 1D regression model, the response variable Y is conditionally independent of the $p \times 1$ vector of predictors $\boldsymbol{x}$ given the sufficient predictor $SP = h(\boldsymbol{x})$, written

$$Y \perp\!\!\!\perp \boldsymbol{x} | SP \quad \text{or} \quad \text{Y} \perp\!\!\!\perp \boldsymbol{x} | \text{h}(\boldsymbol{x}), \tag{1}$$

where the real valued function $h : \mathbb{R}^p \to \mathbb{R}$. The estimated sufficient predictor $ESP = \hat{h}(\boldsymbol{x})$. An important special case is a model with a linear predictor $h(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{\beta}$ where $ESP = \boldsymbol{x}^T \hat{\boldsymbol{\beta}}$. This class of models includes the generalized linear model (GLM) and some survival regression models.

Chapter 1 deals with Prediction Intervals for some GLMS, GAMS, and Survival Regression Models. Section 1.5 describes some parametric regression models where the new prediction intervals are useful.

Chapter 2 deals with Bootstrapping some GLMS and Survival Regression Models after Variable Selection. In Section 2.1 we review a variable selection model and some results on bootstrap confidence regions. Section 2.2 gives large sample theory for some variable selection estimators. Section 2.3 considers the parametric bootstrap while Section 2.4 shows how to bootstrap some variable selection estimators for some parametric regression models. We assume the number of predictors, p, is fixed.

Chapter 3 deals with Examples and Simulations for Prediction Intervals for some GLMs, GAMs, and Survival Regression Models.

Chapter 4 deals with Examples and Simulations some GLMs and Survival Regression Models after Variable Selection.

CHAPTER 1

PREDICTION INTERVALS FOR SOME GLMS, GAMS, AND SURVIVAL REGRESSION MODELS

1.1 INTRODUCTION

In a 1D regression model, the response variable Y is conditionally independent of the $p \times 1$ vector of predictors $\boldsymbol{x}$ given the sufficient predictor $SP = h(\boldsymbol{x})$, written

$$Y \perp\!\!\!\perp \boldsymbol{x}|SP \quad \text{or} \quad \text{Y} \perp\!\!\!\perp \boldsymbol{x}|\text{h}(\boldsymbol{x}), \tag{1.1}$$

where the real valued function $h : \mathbb{R}^p \to \mathbb{R}$. The estimated sufficient predictor $ESP = \hat{h}(\boldsymbol{x})$. An important special case is a model with a linear predictor $h(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta}$ where $ESP = \boldsymbol{x}^T\hat{\boldsymbol{\beta}}$. This class of models includes the generalized linear model (GLM). Often $x_1 \equiv 1$ so a constant is in the model. Another important special case is a generalized additive model (GAM), where Y is independent of $\boldsymbol{x} = (1, x_2, ..., x_p)^T$ given the additive predictor AP where, for example, $AP = SP = \alpha + \sum_{j=2}^p S_j(x_j)$ for some (usually unknown) functions S_j. Then the estimated additive predictor $EAP = ESP = \hat{\alpha} + \sum_{j=2}^p \hat{S}_j(x_j)$.

1.2 PREDICTION INTERVAL (PI)

A large sample $100(1-\delta)\%$ prediction interval (PI) for Y_f has the form $[\hat{L}_n, \hat{U}_n]$ where $P(\hat{L}_n \le Y_f \le \hat{U}_n) \to 1-\delta$ as the sample size $n \to \infty$. A PI is asymptotically optimal if its length converges to the shortest population length covering $100(1-\delta)\%$ of the mass where $0 < \delta < 1$.

1.3 SHORTH PI

Consider predicting a future test value Y_f given a $p \times 1$ vector of predictors $\boldsymbol{x}_f$ and training data $(Y_1, \boldsymbol{x}_1), ..., (Y_n, \boldsymbol{x}_n)$. A large sample $100(1-\delta)\%$ *prediction interval* (PI) for Y_f has the form $[\hat{L}_n, \hat{U}_n]$ where $P(\hat{L}_n \le Y_f \le \hat{U}_n) \to 1-\delta$ as the sample size $n \to \infty$.

The shorth(c) estimator is useful for making prediction intervals. Let $Z_{(1)}, ..., Z_{(n)}$ be the order statistics of $Z_1, ..., Z_n$. Then let the shortest closed interval containing at least c of the Z_i be

$$\text{shorth(c)} = [\text{Z}_{(\text{s})}, \text{Z}_{(\text{s+c}-1)}]. \tag{1.2}$$

Let

$$k_n = \lceil n(1-\delta) \rceil \tag{1.3}$$

where $\lceil x \rceil$ is the smallest integer $\geq x$, e.g., $\lceil 7.7 \rceil = 8$. Then Frey (2013) showed that for large $n\delta$ and iid data, the shorth(k_n) PI has maximum undercoverage $\approx 1.12\sqrt{\delta/n}$, and used the shorth($c$) estimator as the large sample $100(1-\delta)\%$ PI where

$$c = \min(n, \lceil n[1 - \delta + 1.12\sqrt{\delta/n}\ ]\ \rceil). \tag{1.4}$$

Example 1.1. Given below were votes for preseason 1A basketball poll from Nov. 22, 2011 WSIL News where the 778 was a typo: the actual value was 78. As shown below, finding shorth(3) from the ordered data is simple. If the outlier was corrected, shorth(3) = [76,78].

```
  111   89   778   78   76
order data: 76 78 89 111 778
                    13 = 89 - 76
                         33 = 111 - 78
                           689 = 778 - 89
shorth(3) = [76,89]
```

The new large sample $100(1-\delta)\%$ PI using $Y_1^*, ..., Y_B^*$ uses the shorth(c) PI with

$$c = \min(B, \lceil B[1 - \delta + 1.12\sqrt{\delta/B}\ ]\ \rceil). \tag{1.5}$$

PI (1.5) is a simple large sample $100(1-\delta)\%$ prediction interval for a future value Y_f given a $p \times 1$ vector of predictors $\boldsymbol{x}_f$ and training data $(\boldsymbol{x}_1, Y_1), ..., (\boldsymbol{x}_n, Y_n)$ for a parametric

1D regression model where $Y|\boldsymbol{x} \sim D(h(\boldsymbol{x}), \boldsymbol{\theta})$, and D is a parametric distribution that depends on $\boldsymbol{x}$ only through $h(\boldsymbol{x})$. Apply the nonparametric shorth prediction interval to $Y_1^*, ..., Y_B^*$ where the Y_i^* are independent and identically distributed (iid) from the distribution $D(\hat{h}(\boldsymbol{x}_f), \hat{\boldsymbol{\theta}})$. If the regression method produces a consistent estimator $(\hat{h}(\boldsymbol{x}), \hat{\boldsymbol{\theta}})$ of $(h(\boldsymbol{x}), \boldsymbol{\theta})$, then this new prediction interval is a large sample $100(1-\delta)\%$ PI that is a consistent estimator of the shortest population interval $[L, U]$ that contains at least $1-\delta$ of the mass as $B, n \to \infty$.

1.4 VARIABLE SELECTION

For models with a linear predictor, we will want prediction intervals after variable selection or model selection. Variable selection is the search for a subset of predictor variables that can be deleted with little loss of information if n/p is large, and so that the model with the remaining predictors is useful for prediction. Following Olive and Hawkins (2005), a *model for variable selection* can be described by

$$\boldsymbol{x}^T\boldsymbol{\beta} = \boldsymbol{x}_S^T\boldsymbol{\beta}_S + \boldsymbol{x}_E^T\boldsymbol{\beta}_E = \boldsymbol{x}_S^T\boldsymbol{\beta}_S \tag{1.6}$$

where $\boldsymbol{x} = (\boldsymbol{x}_S^T, \boldsymbol{x}_E^T)^T$, $\boldsymbol{x}_S$ is an $a_S \times 1$ vector, and $\boldsymbol{x}_E$ is a $(p - a_S) \times 1$ vector. Given that $\boldsymbol{x}_S$ is in the model, $\boldsymbol{\beta}_E = \boldsymbol{0}$ and E denotes the subset of terms that can be eliminated given that the subset S is in the model. Let $\boldsymbol{x}_I$ be the vector of a terms from a candidate subset indexed by I, and let $\boldsymbol{x}_O$ be the vector of the remaining predictors (out of the candidate submodel). Suppose that S is a subset of I and that model (1.6) holds. Then

$$\boldsymbol{x}^T\boldsymbol{\beta} = \boldsymbol{x}_S^T\boldsymbol{\beta}_S = \boldsymbol{x}_S^T\boldsymbol{\beta}_S + \boldsymbol{x}_{I/S}^T\boldsymbol{\beta}_{(I/S)} + \boldsymbol{x}_O^T\boldsymbol{0} = \boldsymbol{x}_I^T\boldsymbol{\beta}_I \tag{1.7}$$

where $\boldsymbol{x}_{I/S}$ denotes the predictors in I that are not in S. Since this is true regardless of the values of the predictors, $\boldsymbol{\beta}_O = \boldsymbol{0}$ if $S \subseteq I$.

Forward selection or backward elimination with the Akaike (1973) AIC criterion are often used for GLM variable selection. The Chen and Chen (2008) EBIC criterion can be useful, especially if n/p is not large. GLM model selection with lasso and the elastic net is also common. See Hastie, Tibshirani, and Wainwright (2015, ch. 3), Tibshirani (1996), Friedman, Hastie, Hoefling, and Tibshirani (2007), and Friedman, Hastie, and Tibshirani (2010). A GLM relaxed lasso applies the GLM to the active predictors with nonzero coefficients selected by lasso. For $n \geq 10p$, Olive and Hawkins (2005) suggested using multiple linear regression variable selection software with C_p to get a subset I, then fit the GLM using Y and $\boldsymbol{x}_I$. For high dimensional GLM variable selection, see Guo, Yang, and Lv (2017).

The modified shorth PI inflates the usual shorth PI to compensate for model selection. Let d be the number of variables $x_1^*, ..., x_d^*$ used by forward selection, lasso, or relaxed lasso. We want n/d large, and the prediction interval length will be increased (penalized) if n/d is not large. Let $q_n = \min(1 - \delta + 0.05, 1 - \delta + d/n)$ for $\delta > 0.1$ and

$$q_n = \min(1 - \delta/2, 1 - \delta + 10\delta d/n), \quad \text{otherwise.}$$

If $1 - \delta < 0.999$ and $q_n < 1 - \delta + 0.001$, set $q_n = 1 - \delta$. Then compute the shorth PI with

$$c_{mod} = \min(B, \lceil B[q_n + 1.12\sqrt{\delta/B}\ ]\ \rceil). \tag{1.8}$$

1.5 PARAMETRIC REGRESSION MODELS

Regression models where $Y|\boldsymbol{x} \sim Y|h(\boldsymbol{x}) \sim Y|SP \sim D(h(\boldsymbol{x}), \boldsymbol{\theta})$ are common. Let $h(\boldsymbol{x}) = SP$. Some important examples follow. See Olive (2010, 2017a: ch. 13, 2019: ch. 4).

i) Nonlinear regression, nonparametric regression, and multiple linear regression are special cases of the *additive error regression* model

$$Y = h(\boldsymbol{x}) + e = SP + e. \tag{1.9}$$

For these models, there is no need to use a parametric distribution. Find the shorth $[\hat{L}_{nr}, \hat{U}_{nr}]$ of the residuals and use PI $[\hat{Y}_f + \hat{L}_{nr}, \hat{Y}_f + \hat{U}_{nr}]$ where $\hat{Y}_f = \hat{h}(\boldsymbol{x}_f)$. See Olive

(2013a) for $n \geq 10d$ where d is the model degrees of freedom. For multiple linear regression see Olive (2007) for $n \geq 10p$ and Pelawa Watagoda and Olive (2019a) for PIs that can work for variable selection and model selection estimators even if n/p is not large. Also see Lei, G'Sell, Rinaldo, Tibshirani, and Wasserman (2017).

ii) The *binary regression model* is $Y \sim$ binomial $\left(1, \rho = \frac{e^{SP}}{1+e^{SP}}\right)$. This model has $E(Y|SP) = \rho = \rho(SP)$ and $V(Y|SP) = \rho(SP)(1-\rho(SP))$. Then $\hat{\rho} = \frac{e^{ESP}}{1+e^{ESP}}$ is the estimated mean function.

iii) The *binomial regression model* is $Y_i \sim$ binomial $\left(m_i, \rho = \frac{e^{SP}}{1+e^{SP}}\right)$. Then $E(Y_i|SP_i)$
$= m_i\rho(SP_i)$ and $V(Y_i|SP_i) = m_i\rho(SP_i)(1-\rho(SP_i))$, and $\hat{E}(Y_i|\boldsymbol{x}_i) = m_i\hat{\rho} = \frac{m_i e^{ESP}}{1+e^{ESP}}$ is the estimated mean function.

iv) The *Poisson regression (PR) model* $Y \sim$ Poisson $\left(\mathrm{e}^{\mathrm{SP}}\right)$ has $E(Y|SP) = V(Y|SP) = \exp(SP)$. The estimated mean and variance functions are $\hat{E}(Y|\boldsymbol{x}) = e^{ESP}$.

v) Suppose Y has a gamma $\mathrm{G}(\nu,\lambda)$ distribution so that $E(Y) = \nu\lambda$ and $V(Y) = \nu\lambda^2$. The *Gamma regression model* $Y \sim G(\nu, \lambda = \mu(SP)/\nu)$ has $E(Y|SP) = \mu(SP)$ and $V(Y|SP) = [\mu(SP)]^2/\nu$. The estimated mean function is $\hat{E}(Y|\boldsymbol{x}) = \mu(ESP)$. The choices $\mu(SP) = SP$, $\mu(SP) = \exp(SP)$ and $\mu(SP) = 1/SP$ are common. Since $\mu(SP) > 0$, Gamma regression models that use the identity or reciprocal link run into problems if $\mu(ESP)$ is negative for some of the cases.

Alternatives to the binomial and Poisson regression models are needed because often the mean function for the model is good, but the variance function is not: there is overdispersion.

A useful alternative to the binomial regression model is a *beta–binomial regression* (BBR) model. Following Simonoff (2003, pp. 93-94) and Agresti (2002, pp. 554-555), let $\delta = \rho/\theta$ and $\nu = (1-\rho)/\theta$, so $\rho = \delta/(\delta+\nu)$ and $\theta = 1/(\delta+\nu)$. Let $B(\delta,\nu) = \frac{\Gamma(\delta)\Gamma(\nu)}{\Gamma(\delta+\nu)}$. If Y has a beta–binomial distribution, $Y \sim \mathrm{BB}(m,\rho,\theta)$, then the probability mass function of Y is $P(Y=y) = \binom{m}{y} \frac{B(\delta+y, \nu+m-y)}{B(\delta,\nu)}$ for $y = 0, 1, 2, ..., m$ where $0 < \rho < 1$ and $\theta > 0$. Hence $\delta > 0$ and $\nu > 0$. Then $E(Y) = m\delta/(\delta+\nu) = m\rho$ and $\mathrm{V}(Y) = m\rho(1-\rho)[1+(m-1)\theta/(1+\theta)]$. If $Y|\pi \sim$ binomial(m, π) and $\pi \sim$ beta(δ, ν), then $Y \sim \mathrm{BB}(m, \rho, \theta)$. As $\theta \to 0$, it can be shown that $V(\pi) \to 0$, and the beta–binomial distribution converges to the binomial distribution.

vi) The BBR model states that $Y_1, ..., Y_n$ are independent random variables where $Y_i|SP_i \sim \mathrm{BB}(m_i, \rho(SP_i), \theta)$. Hence $E(Y_i|SP_i) = m_i\rho(SP_i)$ and

$$V(Y_i|SP_i) = m_i\rho(SP_i)(1-\rho(SP_i))[1+(m_i-1)\theta/(1+\theta)].$$

The BBR model has the same mean function as the binomial regression model, but allows for overdispersion. As $\theta \to 0$, it can be shown that the BBR model converges to the binomial regression model.

A useful alternative to the PR model is a negative binomial regression (NBR) model. If Y has a (generalized) negative binomial distribution, $Y \sim NB(\mu, \kappa)$, then the probability mass function of Y is

$$P(Y=y) = \frac{\Gamma(y+\kappa)}{\Gamma(\kappa)\Gamma(y+1)} \left(\frac{\kappa}{\mu+\kappa}\right)^{\kappa} \left(1 - \frac{\kappa}{\mu+\kappa}\right)^{y}$$

for $y = 0, 1, 2, ...$ where $\mu > 0$ and $\kappa > 0$. Then $E(Y) = \mu$ and $\mathrm{V}(Y) = \mu + \mu^2/\kappa$. (This distribution is a generalization of the negative binomial (κ, ρ) distribution where $\rho = \kappa/(\mu+\kappa)$ and $\kappa > 0$ is an unknown real parameter rather than a known integer.)

vii) The *negative binomial regression (NBR) model* is $Y|SP \sim \mathrm{NB}(\exp(SP), \kappa)$. Thus $E(Y|SP) = \exp(SP)$ and

$$V(Y|SP) = \exp(SP)\left(1 + \frac{\exp(SP)}{\kappa}\right) = \exp(SP) + \tau \exp(2\ SP).$$

The NBR model has the same mean function as the PR model but allows for overdispersion. Following Agresti (2002, p. 560), as $\tau \equiv 1/\kappa \to 0$, it can be shown that the NBR model converges to the PR model.

Olive (2017a: p. 430) describes zero truncated Poisson regression model. Several important survival regression models are parametric 1D regression models with $SP = \boldsymbol{x}^T \boldsymbol{\beta}$. The *accelerated failure time (AFT) model* has $\log(Y) = \alpha + SP_A + \sigma e$ where $SP_A = \boldsymbol{x}^T \boldsymbol{\beta}_A$, $V(e) = 1$, and the e_i are iid from a location scale family.

If the Y_i are Weibull, the e_i are from a smallest extreme value distribution. The Weibull regression model is a proportional hazards model using Y_i and an accelerated failure time model using $\log(Y_i)$ with $\boldsymbol{\beta} = \boldsymbol{\beta}_A/\sigma$.

If the Y_i are lognormal, the e_i are normal.

If the Y_i are loglogistic, the e_i are logistic.

viii) The *Weibull proportional hazards regression (WPH) model* is

$$Y|SP \sim W(\gamma = 1/\sigma, \lambda_0 \exp(SP)),$$

where $\lambda_0 = \exp(-\alpha/\sigma)$, and Y has a Weibull $W(\gamma, \lambda)$ distribution if the pdf of Y is

$$f(y) = \lambda \gamma y^{\gamma-1} \exp[-\lambda y^{\gamma}]$$

for $y > 0$. Note that the PI is for survival times Y, not censored survival times.

1.6 R FUNCTIONS FOR PI COVERAGES AND LENGTHS

The simulations were done in *R*. See R Core Team (2016). We used several *R* functions including lasso with the `cv.glmnet` functions from the Friedman, Hastie, Simon, and Tibshirani (2015) `glmnet` library.

The collection of Olive (2019) *R* functions *slpack*, available from (http://lagrange.math.siu.edu/Olive/slpack.txt), has some useful functions for the inference. Tables 3.1 to 3.6 were made with `prpisim2`.

The Wood (2006) library `mgcv` was used for fitting a GAM, and the Venables and Ripley (2010) library `MASS` is useful for the negative binomial family. The Hastie and Tibshirani (1990) `gam` library is also useful. The Lesnoff and Lancelot (2010) *R* package `aod` has function `betabin` for beta binomial regression and is also useful for fitting negative binomial regression.

CHAPTER 2

BOOTSTRAPPING SOME GLMS AND SURVIVAL REGRESSION MODELS AFTER VARIABLE SELECTION

2.1 THE BOOTSTRAP

Definition 2.1. Suppose that data $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ has been collected and observed. Often the data is a random sample (iid) from a distribution with cdf F. The *empirical distribution* is a discrete distribution where the $\boldsymbol{x}_i$ are the possible values, and each value is equally likely. If $\boldsymbol{w}$ is a random variable having the empirical distribution, then $p_i = P(\boldsymbol{w} = \boldsymbol{x}_i) = 1/n$ for $i = 1, ..., n$. The *cdf of the empirical distribution* is denoted by F_n.

Example 2.1. Let $\boldsymbol{w}$ be a random variable having the empirical distribution given by Definition 2.1. Show that $E(\boldsymbol{w}) = \overline{\boldsymbol{x}} \equiv \overline{\boldsymbol{x}}_n$ and $\text{Cov}(\boldsymbol{w}) = \frac{n-1}{n}\boldsymbol{S} \equiv \frac{n-1}{n}\boldsymbol{S}_n$.

Solution: Recall that for a discrete random vector, the population expected value $E(\boldsymbol{w}) = \sum \boldsymbol{x}_i p_i$ where $\boldsymbol{x}_i$ are the values that $\boldsymbol{w}$ takes with positive probability p_i. Similarly, the population covariance matrix

$$\text{Cov}(\boldsymbol{w}) = E[(\boldsymbol{w} - E(\boldsymbol{w}))(\boldsymbol{w} - E(\boldsymbol{w}))^T] = \sum (\boldsymbol{x}_i - E(\boldsymbol{w}))(\boldsymbol{x}_i - E(\boldsymbol{w}))^T p_i.$$

Hence

$$E(\boldsymbol{w}) = \sum_{i=1}^{n} \boldsymbol{x}_i \frac{1}{n} = \overline{\boldsymbol{x}},$$

and

$$\text{Cov}(\boldsymbol{w}) = \sum_{i=1}^{n} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^T \frac{1}{n} = \frac{n-1}{n}\boldsymbol{S}. \ \square$$

Example 2.2. If $W_1, ..., W_n$ are iid from a distribution with cdf F_W, then the empirical cdf F_n corresponding to F_W is given by

$$F_n(y) = \frac{1}{n} \sum_{i=1}^{n} I(W_i \le y)$$

where the indicator $I(W_i \le y) = 1$ if $W_i \le y$ and $I(W_i \le y) = 0$ if $W_i > y$. Fix n and y. Then $nF_n(y) \sim$ binomial $(n, F_W(y))$. Thus $E[F_n(y)] = F_W(y)$ and $V[F_n(y)] = F_W(y)[1 - F_W(y)]/n$. By the central limit theorem,

$$\sqrt{n}(F_n(y) - F_W(y)) \xrightarrow{D} N(0, F_W(y)[1 - F_W(y)]).$$

Thus $F_n(y) - F_W(y) = O_P(n^{-1/2})$, and F_n is a reasonable estimator of F_W if the sample size n is large.

Suppose there is data $\boldsymbol{w}_1, ..., \boldsymbol{w}_n$ collected into an $n \times p$ matrix $\boldsymbol{W}$. Let the statistic $T_n = t(\boldsymbol{W}) = T(F_n)$ be computed from the data. Suppose the statistic estimates $\boldsymbol{\mu} = T(F)$, and let $t(\boldsymbol{W}^*) = t(F_n^*) = T_n^*$ indicate that t was computed from an iid sample from the empirical distribution F_n: a sample $\boldsymbol{w}_1^*, ..., \boldsymbol{w}_n^*$ of size n was drawn with replacement from the observed sample $\boldsymbol{w}_1, ..., \boldsymbol{w}_n$. This notation is used for von Mises differentiable statistical functions in large sample theory. See Serfling (1980, ch. 6). The empirical distribution is also important for the influence function (widely used in robust statistics). The *empirical bootstrap* or *nonparametric bootstrap* or *naive bootstrap* draws B samples of size n from the rows of $\boldsymbol{W}$, e.g. from the empirical distribution of $\boldsymbol{w}_1, ..., \boldsymbol{w}_n$. Then T_{jn}^* is computed from the jth bootstrap sample for $j = 1, ..., B$.

Example 2.3. Suppose the data is 1, 2, 3, 4, 5, 6, 7. Then $n = 7$ and the sample median T_n is 4. Using R, we drew $B = 2$ bootstrap samples (samples of size n drawn with replacement from the original data) and computed the sample median $T_{1,n}^* = 3$ and $T_{2,n}^* = 4$.

```
b1 <- sample(1:7,replace=T)
b1
[1] 3 2 3 2 5 2 6
median(b1)
[1] 3
b2 <- sample(1:7,replace=T)
b2
[1] 3 5 3 4 3 5 7
median(b2)
[1] 4
```

2.2 VARIABLE SELECTION MODEL AND BOOTSTRAP CONFIDENCE REGIONS

We also want to use bootstrap tests. Consider testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ where $\boldsymbol{\theta}_0$ is a known $g \times 1$ vector. Given training data $\boldsymbol{z}_1, ..., \boldsymbol{z}_n$, a large sample $100(1-\delta)\%$ confidence region for $\boldsymbol{\theta}$ is a set $\mathcal{A}_n$ such that $P(\boldsymbol{\theta} \in \mathcal{A}_n) \to 1 - \delta$ as $n \to \infty$. Then reject H_0 if $\boldsymbol{\theta}_0$ is not in the confidence region $\mathcal{A}_n$.

To bootstrap a confidence region, Mahalanobis distances will be useful. Let the $g \times 1$ column vector T be a multivariate location estimator, and let the $g \times g$ symmetric positive definite matrix $\boldsymbol{C}$ be a dispersion estimator. Then the ith *squared sample Mahalanobis distance* is the scalar

$$D_i^2 = D_i^2(T, \boldsymbol{C}) = D^2_{\boldsymbol{z}_i}(T, \boldsymbol{C}) = (\boldsymbol{z}_i - T)^T \boldsymbol{C}^{-1} (\boldsymbol{z}_i - T) \tag{2.1}$$

for each observation $\boldsymbol{z}_i$. Notice that the Euclidean distance of $\boldsymbol{z}_i$ from the estimate of center T is $D_i(T, \boldsymbol{I}_g)$ where $\boldsymbol{I}_g$ is the $g \times g$ identity matrix. The classical Mahalanobis distance D_i uses $(T, \boldsymbol{C}) = (\overline{\boldsymbol{z}}, \boldsymbol{S})$, the sample mean and sample covariance matrix where

$$\overline{\boldsymbol{z}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{z}_i \quad \text{and} \quad \boldsymbol{S} = \frac{1}{\mathrm{n} - 1} \sum_{\mathrm{i}=1}^{\mathrm{n}} (\boldsymbol{z}_{\mathrm{i}} - \overline{\boldsymbol{z}})(\boldsymbol{z}_{\mathrm{i}} - \overline{\boldsymbol{z}})^{\mathrm{T}}. \tag{2.2}$$

Let $q_n = \min(1 - \delta + 0.05, 1 - \delta + g/n)$ for $\delta > 0.1$ and

$$q_n = \min(1 - \delta/2, 1 - \delta + 10\delta g/n), \quad \text{otherwise.} \tag{2.3}$$

If $1 - \delta < 0.999$ and $q_n < 1 - \delta + 0.001$, set $q_n = 1 - \delta$. Let

$$c = \lceil n q_n \rceil. \tag{2.4}$$

Let $(T, \boldsymbol{C}) = (\overline{\boldsymbol{z}}, \boldsymbol{S})$, and let $D_{(U_n)}$ be the $100 q_n$th sample quantile of the D_i.

Let a statistic T estimate $\boldsymbol{\theta}$. The following confidence region is found by applying a prediction region to the bootstrap sample $T_1^*, ..., T_B^*$. Let $\overline{T}^*$ and $\boldsymbol{S}_T^*$ be the sample mean and sample covariance matrix of the bootstrap sample. The Olive (2017ab, 2018) prediction region method large sample $100(1-\delta)\%$ confidence region for $\boldsymbol{\theta}$ is
$\{\boldsymbol{w} : (\boldsymbol{w} - \overline{T}^*)^T [\boldsymbol{S}_T^*]^{-1} (\boldsymbol{w} - \overline{T}^*) \leq D^2_{(U_B)}\} =$

$$\{\boldsymbol{w} : D^2_{\boldsymbol{w}}(\overline{T}^*, \boldsymbol{S}_T^*) \leq D^2_{(U_B)}\} \tag{2.5}$$

where $D^2_{(U_B)}$ is computed from $D_i^2 = (T_i^* - \overline{T}^*)^T [\boldsymbol{S}_T^*]^{-1} (T_i^* - \overline{T}^*)$ for $i = 1, ..., B$. Note that the corresponding test for $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ rejects H_0 if $(\overline{T}^* - \boldsymbol{\theta}_0)^T [\boldsymbol{S}_T^*]^{-1} (\overline{T}^* - \boldsymbol{\theta}_0) > D^2_{(U_B)}$.

The modified Bickel and Ren (2001) large sample $100(1-\delta)\%$ confidence region is
$\{\boldsymbol{w} : (\boldsymbol{w} - T_n)^T [\boldsymbol{S}_T^*]^{-1} (\boldsymbol{w} - T_n) \leq D^2_{(U_B,T)}\} =$

$$\{\boldsymbol{w} : D^2_{\boldsymbol{w}}(T_n, \boldsymbol{S}_T^*) \leq D^2_{(U_B,T)}\} \tag{2.6}$$

where $D^2_{(U_B,T)}$ is computed from $D_i^2 = (T_i^* - T_n)^T [\boldsymbol{S}_T^*]^{-1} (T_i^* - T_n)$.

Shift region (2.5) to have center T_n, or equivalently, change the cutoff of region (2.6) to $D^2_{(U_B)}$ to get the Pelawa Watagoda and Olive (2018) hybrid large sample $100(1-\delta)\%$ confidence region: $\{\boldsymbol{w} : (\boldsymbol{w} - T_n)^T [\boldsymbol{S}_T^*]^{-1} (\boldsymbol{w} - T_n) \leq D^2_{(U_B)}\} =$

$$\{\boldsymbol{w} : D^2_{\boldsymbol{w}}(T_n, \boldsymbol{S}_T^*) \leq D^2_{(U_B)}\}. \tag{2.7}$$

Under regularity conditions, Olive (2017b, 2018) proved that (2.5) is a large sample confidence region, while Pelawa Watagoda and Olive (2019b) gave simpler proofs, and proved that the shorth(c) interval applied to a bootstrap sample of a random variable gives a large sample confidence interval. They also showed that the three confidence regions were useful for bootstrapping $\hat{\boldsymbol{\beta}}_{I_{min},0}$ after variable selection for multiple linear regression. The shorth confidence interval is a practical implementation of the Hall (1988) shortest bootstrap interval based on all possible bootstrap samples.

The ratio of the volumes of regions (2.5) and (2.6) is

$$\frac{|\boldsymbol{S}_T^*|^{1/2}}{|\boldsymbol{S}_T^*|^{1/2}} \left(\frac{D_{(U_B)}}{D_{(U_B,T)}} \right)^g = \left(\frac{D_{(U_B)}}{D_{(U_B,T)}} \right)^g . \tag{2.8}$$

The volume of confidence region (2.6) tends to be greater than that of (2.5) since the T_i^* are closer to $\overline{T}^*$ than T_n on average. Regions (2.5) and (2.7) have the same volume.

2.3 LARGE SAMPLE THEORY FOR SOME VARIABLE SELECTION ESTIMATORS

Suppose the regression model satisfies $Y \perp\!\!\!\perp \boldsymbol{x} | \boldsymbol{x}^T \boldsymbol{\beta}$, that model (1.6) holds, and that if $S \subseteq I$,then $\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_j} - \boldsymbol{\beta}_{I_j}) \xrightarrow{D} N_{a_j}(\boldsymbol{0}, \boldsymbol{V}_j)$. Also assume that a variable selection criterion, such as AIC or BIC, is used such that $P(S \subseteq I_{min}) \to 1$ as $n \to \infty$. Hence

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_j,0} - \boldsymbol{\beta}) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}_{j,0}) \tag{2.9}$$

where $\boldsymbol{V}_{j,0}$ adds columns and rows of zeroes corresponding to the x_i not in I_j.Hence $\boldsymbol{V}_{j,0}$ is singular unless I_j corresponds to the full model. Then $\hat{\boldsymbol{\beta}}_{I_{min},0}$ is a $\sqrt{n}$ consistent estimator of $\boldsymbol{\beta}$ under model (1.6) if the variable selection criterion is used with forward selection, backward elimination, or all subsets. These assumptions hold for many regression models, including many generalized linear models, some time series models, and some survival regression models. This section will use mixture distributions to find the limiting distribution of $\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_{min},0} - \boldsymbol{\beta})$, generalizing the Pelawa Watagoda and Olive (2019b) theorem for multiple linear regression.

Mixture distributions are useful for variable selection since $\hat{\boldsymbol{\beta}}_{I_{min},0}$ has a mixture distribution of the $\hat{\boldsymbol{\beta}}_{I_j,0}$. A random vector $\boldsymbol{u}$ has a mixture distribution of random vectors $\boldsymbol{u}_j$

with probabilities π_j if $\boldsymbol{u}$ equals random vector $\boldsymbol{u}_j$ with probability π_j for $j = 1, ..., J$. Let $\boldsymbol{u}$ and $\boldsymbol{u}_j$ be $p \times 1$ random vectors. Then the cumulative distribution function (cdf) of $\boldsymbol{u}$ is

$$F_{\boldsymbol{u}}(\boldsymbol{t}) = \sum_{j=1}^{J} \pi_j F_{\boldsymbol{u}_j}(\boldsymbol{t})$$

where the probabilities π_j satisfy $0 \leq \pi_j \leq 1$ and $\sum_{j=1}^{J} \pi_j = 1$, $J \geq 2$, and $F_{\boldsymbol{u}_j}(\boldsymbol{t})$ is the cdf of $\boldsymbol{u}_j$.

Suppose $E(h(\boldsymbol{u}))$ and the $E(h(\boldsymbol{u}_j))$ exist. Then

$$E(h(\boldsymbol{u})) = \sum_{j=1}^{J} \pi_j E[h(\boldsymbol{u}_j)] \text{ and } E(\boldsymbol{u}) = \sum_{j=1}^{J} \pi_j E[\boldsymbol{u}_j].$$

Hence $\text{Cov}(\boldsymbol{u}) = E(\boldsymbol{u}\boldsymbol{u}^T) - E(\boldsymbol{u})E(\boldsymbol{u}^T) = E(\boldsymbol{u}\boldsymbol{u}^T) - E(\boldsymbol{u})[E(\boldsymbol{u})]^T =$
$\sum_{j=1}^{J} \pi_j E[\boldsymbol{u}_j \boldsymbol{u}_j^T] - E(\boldsymbol{u})[E(\boldsymbol{u})]^T =$

$$\sum_{j=1}^{J} \pi_j \text{Cov}(\boldsymbol{u}_j) + \sum_{j=1}^{J} \pi_j E(\boldsymbol{u}_j)[E(\boldsymbol{u}_j)]^T - E(\boldsymbol{u})[E(\boldsymbol{u})]^T.$$

If $E(\boldsymbol{u}_j) = \boldsymbol{\theta}$ for $j = 1, ..., J$, then $E(\boldsymbol{u}) = \boldsymbol{\theta}$ and

$$\text{Cov}(\boldsymbol{u}) = \sum_{j=1}^{J} \pi_j \text{Cov}(\boldsymbol{u}_j).$$

Following Pelawa Watagoda and Olive (2019b), suppose that T_n is equal to the estimator T_{jn} with probability π_{jn} for $j = 1, ..., J$ where $\sum_j \pi_{jn} = 1$, $\pi_{jn} \to \pi_j$ as $n \to \infty$, and $\boldsymbol{u}_{jn} = \sqrt{n}(T_{jn} - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{u}_j$ with $E(\boldsymbol{u}_j) = \boldsymbol{0}$ and $\text{Cov}(\boldsymbol{u}_j) = \boldsymbol{\Sigma}_j$. Then T_n has a mixture distribution of the T_{jn} with probabilities π_{jn}, and the cdf of T_n is $F_{T_n}(\boldsymbol{z}) = \sum_j \pi_{jn} F_{T_{jn}}(\boldsymbol{z})$ where $F_{T_{jn}}(\boldsymbol{z})$ is the cdf of T_{jn}. Hence $\sqrt{n}(T_n - \boldsymbol{\theta})$ has a mixture distribution of the $\sqrt{n}(T_{jn} - \boldsymbol{\theta})$, and

$$\sqrt{n}(T_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{u} \tag{2.10}$$

where the cdf of $\boldsymbol{u}$ is $F_{\boldsymbol{u}}(\boldsymbol{z}) = \sum_j \pi_j F_{\boldsymbol{u}_j}(\boldsymbol{z})$ and $F_{\boldsymbol{u}_j}(\boldsymbol{z})$ is the cdf of $\boldsymbol{u}_j$. Thus $\boldsymbol{u}$ is a mixture distribution of the $\boldsymbol{u}_j$ with probabilities π_j, $E(\boldsymbol{u}) = \boldsymbol{0}$, and $\text{Cov}(\boldsymbol{u}) = \boldsymbol{\Sigma}_{\boldsymbol{u}} = \sum_j \pi_j \boldsymbol{\Sigma}_j$.

Applying the above results makes large sample theory for $\hat{\boldsymbol{\beta}}_{I_{min},0}$ simple. Let $T_n = \hat{\boldsymbol{\beta}}_{I_{min},0}$ and $T_{jn} = \hat{\boldsymbol{\beta}}_{I_j,0}$. Let $T_n = T_{kn} = \hat{\boldsymbol{\beta}}_{I_k,0}$ with probabilities π_{kn} where $\pi_{kn} \to \pi_k$ as

$n \to \infty$. Denote the π_k with $S \subseteq I_k$ by π_j. The other $\pi_k = 0$ since $P(S \subseteq I_{min}) \to 1$ as $n \to \infty$. Then $\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_j} - \boldsymbol{\beta}_{I_j}) \xrightarrow{D} N_{a_j}(\boldsymbol{0}, \boldsymbol{V}_j)$ and $\boldsymbol{u}_{jn} = \sqrt{n}(\hat{\boldsymbol{\beta}}_{I_j,0} - \boldsymbol{\beta}) \xrightarrow{D} \boldsymbol{u}_j \sim N_p(\boldsymbol{0}, \boldsymbol{V}_{j,0})$ by (2.9).

Then Equation (2.10) holds:

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_{min},0} - \boldsymbol{\beta}) \xrightarrow{D} \boldsymbol{u} \tag{2.11}$$

where the cdf of $\boldsymbol{u}$ is $F_{\boldsymbol{u}}(\boldsymbol{z}) = \sum_j \pi_j F_{\boldsymbol{u}_j}(\boldsymbol{z})$. Thus $\boldsymbol{u}$ is a mixture distribution of the $\boldsymbol{u}_j$ with probabilities π_j, $E(\boldsymbol{u}) = \boldsymbol{0}$, and $\text{Cov}(\boldsymbol{u}) = \boldsymbol{\Sigma}_{\boldsymbol{u}} = \sum_j \pi_j \boldsymbol{V}_{j,0}$. The values of π_j depend on the regression variable selection method with AIC or BIC, such as backward elimination, forward selection, and all subsets. Let $\boldsymbol{A}$ be a $g \times p$ full rank matrix with $1 \le g \le p$. Then

$$\sqrt{n}(\boldsymbol{A}\hat{\boldsymbol{\beta}}_{I_{min},0} - \boldsymbol{A}\boldsymbol{\beta}) \xrightarrow{D} \boldsymbol{A}\boldsymbol{u} = \boldsymbol{v} \tag{2.12}$$

where $\boldsymbol{A}\boldsymbol{u}$ has a mixture distribution of the $\boldsymbol{A}\boldsymbol{u}_j \sim N_g(\boldsymbol{0}, \boldsymbol{A}\boldsymbol{V}_{j,0}\boldsymbol{A}^T)$ with probabilities π_j.

Under regularity conditions, $P(S \subseteq I_{min}) \to 1$ as $n \to \infty$ if BIC or AIC is used. $P(S = I_{min}) \to 1$ for BIC while AIC tends to overfit since the BIC penalty is larger than the AIC penalty for adding variables. See Charkhi and Claeskens (2018), Claeskens and Hjort (2008, pp. 70, 101, 102, 114, 232) and Haughton (1988, 1989).

2.4 THE PARAMETRIC BOOTSTRAP

Suppose $Y_i|\boldsymbol{x}_i \sim D(\boldsymbol{x}_i^T\boldsymbol{\beta}, \boldsymbol{\theta})$, $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}(\boldsymbol{\beta}))$, and that $\boldsymbol{V}(\hat{\boldsymbol{\beta}}) \xrightarrow{P} \boldsymbol{V}(\boldsymbol{\beta})$ as $n \to \infty$. These assumptions tend to be mild for a parametric regression model where the maximum likelihood estimator (MLE) $\hat{\boldsymbol{\beta}}$ is used. Then $\boldsymbol{V}(\boldsymbol{\beta}) = \boldsymbol{I}^{-1}(\boldsymbol{\beta})$, the inverse Fisher information matrix. If $\boldsymbol{I}_n(\boldsymbol{\beta})$ is the Fisher information matrix based on a sample of size n, then $\boldsymbol{I}_n(\boldsymbol{\beta})/n \xrightarrow{P} \boldsymbol{I}(\boldsymbol{\beta})$. For GLMs, see, for example, Sen and Singer (1993, p. 309). For the parametric regression model, we regress $\boldsymbol{Y}$ on $\boldsymbol{X}$ to obtain $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}})$ where the $n \times 1$ vector $\boldsymbol{Y} = (Y_i)$ and the ith row of the $n \times p$ design matrix $\boldsymbol{X}$ is $\boldsymbol{x}_i^T$.

The parametric bootstrap uses $\boldsymbol{Y}_j^* = (Y_i^*)$ where $Y_i^*|\boldsymbol{x}_i \sim D(\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}})$ for $i = 1, ..., n$. Regress $\boldsymbol{Y}_j^*$ on $\boldsymbol{X}$ to get $\hat{\boldsymbol{\beta}}_j^*$ for $j = 1, ..., B$. The large sample theory for $\hat{\boldsymbol{\beta}}^*$ is simple.

Note that if $Y_i^*|\boldsymbol{x}_i \sim D(\boldsymbol{x}_i^T\boldsymbol{b}, \hat{\boldsymbol{\theta}})$ where $\boldsymbol{b}$ does not depend on n, then $(\boldsymbol{Y}^*, \boldsymbol{X})$ follows the parametric regression model with parameters $(\boldsymbol{b}, \hat{\boldsymbol{\theta}})$. Hence $\sqrt{n}(\hat{\boldsymbol{\beta}}^* - \boldsymbol{b}) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}(\boldsymbol{b}))$. Now fix large integer n_0, and let $\boldsymbol{b} = \hat{\boldsymbol{\beta}}_{n_o}$. Then $\sqrt{n}(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}}_{n_o}) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}(\hat{\boldsymbol{\beta}}_{n_o}))$. Since $N_p(\boldsymbol{0}, \boldsymbol{V}(\hat{\boldsymbol{\beta}})) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}(\boldsymbol{\beta}))$, we have

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}}) \xrightarrow{D} N_p(\boldsymbol{0}, \boldsymbol{V}(\boldsymbol{\beta})) \tag{2.13}$$

as $n \to \infty$.

Now suppose $S \subseteq I$. Without loss of generality, let $\boldsymbol{\beta} = (\boldsymbol{\beta}_I^T, \boldsymbol{\beta}_O^T)^T$ and $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}(I)^T, \hat{\boldsymbol{\beta}}(O)^T)^T$. Then $(\boldsymbol{Y}, \boldsymbol{X}_I)$ follows the parametric regression model with parameters $(\boldsymbol{\beta}_I, \boldsymbol{\theta})$. Hence $\sqrt{n}(\hat{\boldsymbol{\beta}}_I - \boldsymbol{\beta}_I) \xrightarrow{D} N_a(\boldsymbol{0}, \boldsymbol{V}(\boldsymbol{\beta}_I))$. Now $(\boldsymbol{Y}^*, \boldsymbol{X}_I)$ only follows the parametric regression model asymptotically, since $\hat{\boldsymbol{\beta}}(O) \neq \boldsymbol{0}$. Note that $\sqrt{n}\hat{\boldsymbol{\beta}}(O) = O_P(1)$. However, under regularity conditions, $E(\hat{\boldsymbol{\beta}}_I^*) \approx \hat{\boldsymbol{\beta}}_I$ and $\text{Cov}(\hat{\boldsymbol{\beta}}_I^*) -$ $\text{Cov}(\hat{\boldsymbol{\beta}}_I) \to \boldsymbol{0}$ as $n, B \to \infty$.

To see the above claim for GLMs, consider a GLM with $\eta_i = SP_i = \boldsymbol{x}_i^T\boldsymbol{\beta} = g(\mu_i)$ where $\mu_i = E(Y_i|\boldsymbol{x}_i) = g^{-1}(\eta_i)$. Let $V_i = V(Y_i|\boldsymbol{x}_i)$. Let

$$z_i = g(\mu_i) + g'(\mu_i)(Y_i - \mu_i) = \eta_i + \frac{\partial \eta_i}{\partial \mu_i}(Y_i - \mu_i), \quad \boldsymbol{Z} = (z_i),$$

$$w_i = \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \frac{1}{V_i}, \quad \boldsymbol{W} = diag(w_i), \quad \hat{\boldsymbol{W}} = \boldsymbol{W}|_{\hat{\boldsymbol{\beta}}}, \quad \text{and} \quad \hat{\boldsymbol{Z}} = \boldsymbol{Z}|_{\hat{\boldsymbol{\beta}}}.$$

Then

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\hat{\boldsymbol{W}}\boldsymbol{X})^{-1}\boldsymbol{X}^T\hat{\boldsymbol{W}}\hat{\boldsymbol{Z}} \quad \text{and} \quad \hat{\boldsymbol{\beta}}_{\text{I}} = (\boldsymbol{X}_{\text{I}}^{\text{T}}\hat{\boldsymbol{W}}_{\text{I}}\boldsymbol{X}_{\text{I}})^{-1}\boldsymbol{X}_{\text{I}}^{\text{T}}\hat{\boldsymbol{W}}_{\text{I}}\hat{\boldsymbol{Z}}_{\text{I}}$$

while

$$\hat{\boldsymbol{\beta}}_I^* = (\boldsymbol{X}_I^T\hat{\boldsymbol{W}}_I^*\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\hat{\boldsymbol{W}}_I^*\hat{\boldsymbol{Z}}_I^* \tag{2.14}$$

where $\hat{\boldsymbol{\beta}}_I^*$ is fit as if $(\boldsymbol{Y}^*, \boldsymbol{X}_I)$ follows the GLM with parameters $(\hat{\boldsymbol{\beta}}(I), \hat{\boldsymbol{\theta}})$. Hence $\eta_{iI}^* = \boldsymbol{x}_{iI}^T\hat{\boldsymbol{\beta}}(I) = g(\mu_{iI}^*)$, and $V_{iI}^* = V_M(Y_i^*|\boldsymbol{x}_{iI})$ where V_M is the model variance from the GLM with parameters $(\hat{\boldsymbol{\beta}}(I), \hat{\boldsymbol{\theta}})$. Also, the estimated asymptotic covariance matrices are

$$\widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}^T\hat{\boldsymbol{W}}\boldsymbol{X})^{-1} \quad \text{and} \quad \widehat{Cov}(\hat{\boldsymbol{\beta}}_I) = (\boldsymbol{X}_I^T\hat{\boldsymbol{W}}_I\boldsymbol{X}_I)^{-1}.$$

See, for example, Agresti (2002, pp. 138, 147), Hillis and Davis (1994), and McCullagh and Nelder (1989). From Sen and Singer (1994, p. 307), $n(\boldsymbol{X}_I^T \hat{\boldsymbol{W}}_I \boldsymbol{X}_I)^{-1} \xrightarrow{P} \boldsymbol{I}^{-1}(\boldsymbol{\beta}_I)$ as $n \to \infty$ if $S \subseteq I$.

Let $\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{Z}$. Then $E(\tilde{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ since $E(\boldsymbol{Z}) = \boldsymbol{X}\boldsymbol{\beta}$, and $\text{Cov}(\boldsymbol{Y}) = \text{Cov}(\boldsymbol{Y}|\boldsymbol{X}) = diag(V_i)$. Since

$$\frac{\partial \mu_i}{\partial \eta_i} = \frac{1}{g'(\mu_i)} \text{ and } \frac{\partial \eta_i}{\partial \mu_i} = g'(\mu_i),$$

$\text{Cov}(\boldsymbol{Z}) = \text{Cov}(\boldsymbol{Z}|\boldsymbol{X}) = \boldsymbol{W}^{-1}$. Thus $\text{Cov}(\tilde{\boldsymbol{\beta}}) = (\boldsymbol{X}\boldsymbol{W}\boldsymbol{X})^{-1}$. Although $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = O_P(n^{-1/2})$, we have $n(\boldsymbol{X}^T \hat{\boldsymbol{W}} \boldsymbol{X})^{-1} - n(\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1} \xrightarrow{P} \boldsymbol{I}^{-1}(\boldsymbol{\beta}) - \boldsymbol{I}^{-1}(\boldsymbol{\beta}) = \boldsymbol{0}$ as $n \to \infty$.

Let $\tilde{\boldsymbol{\beta}}_I^* = (\boldsymbol{X}_I^T \boldsymbol{W}_I^* \boldsymbol{X}_I)^{-1} \boldsymbol{X}_I^T \boldsymbol{W}_I^* \boldsymbol{Z}_I^*$ where $\boldsymbol{W}_i^*$ and $\boldsymbol{Z}_I^*$ are evaluated using $\hat{\boldsymbol{\beta}}(I)$. Then $\text{Cov}(\boldsymbol{Y}^*) = diag(V_i^*) \to diag(V_{iI}^*)$. Hence $\text{Cov}(\boldsymbol{Z}_I^*) \to \boldsymbol{W}_I^{*-1}$ and $\text{Cov}(\tilde{\boldsymbol{\beta}}_I^*) \to (\boldsymbol{X}_I^T \boldsymbol{W}_I^* \boldsymbol{X}_I)^{-1}$ as $n, B \to \infty$. Hence $\text{Cov}(\hat{\boldsymbol{\beta}}_I^*) -$ $\text{Cov}(\hat{\boldsymbol{\beta}}_I) \to \boldsymbol{0}$ as $n, B \to \infty$ if $S \subseteq I$.

As an example, consider the Poisson regression model from Section 1.5. Then $\mu_{iI}^* = \exp(\boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}(I)) = \exp(\eta_{iI}^*) = V_{iI}^*$. Hence

$$\frac{\partial \mu_{iI}^*}{\partial \eta_{iI}^*} = \exp(\eta_{iI}^*) = \mu_{iI}^* = V_{iI}^*,$$

$w_{iI}^* = \exp(\boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}(I))$, and $\hat{w}_{iI}^* = \exp(\boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}_I^*)$. Similarly, $\eta_{iI}^* = \log(\mu_{iI}^*)$,

$$z_{iI}^* = \eta_{iI}^* + \frac{\partial \eta_{iI}^*}{\partial \mu_{iI}^*}(Y_i^* - \mu_{iI}^*) = \eta_{iI}^* + \frac{1}{\mu_{iI}^*}(Y_i^* - \mu_{iI}^*), \text{ and}$$

$$\hat{z}_{iI}^* = \boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}_I^* + \frac{1}{\exp(\boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}_I^*)}(Y_i^* - \exp(\boldsymbol{x}_{iI}^T \hat{\boldsymbol{\beta}}_I^*)).$$

Note that for $(\boldsymbol{Y}, \boldsymbol{X}_I)$, the formulas are the same with the asterisks removed and $\mu_{iI} = \exp(\boldsymbol{x}_{iI}^T \boldsymbol{\beta}_I)$.

Next consider the multiple linear regression model $Y_i = \beta_1 + x_{i,2}\beta_2 + \cdots + x_{i,p}\beta_p + e_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + e_i$ for $i = 1, ..., n$ where the random variables e_i are iid with variance $V(e_i) = \sigma^2$. In matrix notation, these n equations become $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$. Let $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$. Assume the maximum leverage $\max_{i=1,...,n} \boldsymbol{x}_{iI}^T (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1} \boldsymbol{x}_{iI} \to 0$ in probability as $n \to \infty$ for each I with $S \subseteq I$. For the (ordinary) least squares (OLS) model with $S \subseteq I$,

$\sqrt{n}(\hat{\boldsymbol{\beta}}_I - \boldsymbol{\beta}_I) \xrightarrow{D} N_{a_I}(\boldsymbol{0}, \boldsymbol{V}_I)$ where $(\boldsymbol{X}_I^T \boldsymbol{X}_I)/(n\sigma^2) \xrightarrow{P} \boldsymbol{V}_I^{-1}$. See, for example, Olive (2017a, p. 39) and Sen and Singer (1993, p. 280).

Consider the parametric bootstrap for the above model with $\boldsymbol{Y}^* \sim N_n(\boldsymbol{X}\hat{\boldsymbol{\beta}}, \hat{\sigma}_n^2 \boldsymbol{I}) \sim N_n(\boldsymbol{HY}, \hat{\sigma}_n^2 \boldsymbol{I})$ where **we are not assuming** that the $e_i \sim N(0, \sigma^2)$, and

$$\hat{\sigma}_n^2 = MSE = \frac{1}{n-p} \sum_{i=1}^n r_i^2$$

where the residuals are from the full OLS model. Then MSE is a $\sqrt{n}$ consistent estimator of σ^2 under mild conditions by Su and Cook (2012). Thus $\hat{\boldsymbol{\beta}}_I^* = (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1} \boldsymbol{X}_I^T \boldsymbol{Y}^* \sim N_{a_I}(\hat{\boldsymbol{\beta}}_I, \hat{\sigma}_n^2 (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1})$ since $E(\hat{\boldsymbol{\beta}}_I^*) = (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1} \boldsymbol{X}_I^T \boldsymbol{HY} = \hat{\boldsymbol{\beta}}_I$ because $\boldsymbol{HX}_I = \boldsymbol{X}_I$, and $\text{Cov}(\hat{\boldsymbol{\beta}}_I^*) = \hat{\sigma}_n^2 (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1}$. Hence

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_I^* - \hat{\boldsymbol{\beta}}_I) \sim N_{a_I}(\boldsymbol{0}, n\hat{\sigma}_n^2 (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1}) \xrightarrow{D} N_{a_I}(\boldsymbol{0}, \boldsymbol{V}_I)$$

as $n, B \to \infty$ if $S \subseteq I$. For the residual bootstrap using residuals from the full OLS model, Pelawa Watagoda and Olive (2019b) showed that $E(\hat{\boldsymbol{\beta}}_I^*) = \hat{\boldsymbol{\beta}}_I$ and $\text{Cov}(\hat{\boldsymbol{\beta}}_I^*) = [(n-p)/n]\hat{\sigma}_n^2 (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1} \xrightarrow{P} \boldsymbol{V}_I$. Note that both the residual bootstrap and above nonparametric bootstrap for OLS are robust to the unknown error distribution of the iid e_i.

As another example, the multivariate linear regression model $\boldsymbol{y}_i = \boldsymbol{B}^T \boldsymbol{x}_i + \boldsymbol{\epsilon}_i$ for $i = 1, ..., n$ has $m \geq 2$ response variables $Y_1, ..., Y_m$ and p predictor variables $x_1, x_2, ..., x_p$. The model is written in matrix form as $\boldsymbol{Z} = \boldsymbol{XB} + \boldsymbol{E}$ where the matrices are defined below. The model has $E(\boldsymbol{\epsilon}_k) = \boldsymbol{0}$ and $\text{Cov}(\boldsymbol{\epsilon}_k) = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = (\sigma_{ij})$ for $k = 1, ..., n$. Then the $p \times m$ coefficient matrix $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \dots & \boldsymbol{\beta}_m \end{bmatrix}$ and the $m \times m$ covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ are to be estimated, and $E(\boldsymbol{Z}) = \boldsymbol{XB}$ while $E(Y_{ij}) = \boldsymbol{x}_i^T \boldsymbol{\beta}_j$. The $\boldsymbol{\epsilon}_i$ are assumed to be iid. Subscripts are needed for the m multiple linear regression models $\boldsymbol{Y}_j = \boldsymbol{X}\boldsymbol{\beta}_j + \boldsymbol{e}_j$ for $j = 1, ..., m$ where $E(\boldsymbol{e}_j) = \boldsymbol{0}$. The $n \times m$ matrices

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 & \dots & \boldsymbol{Y}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}_1^T \\ \vdots \\ \boldsymbol{y}_n^T \end{bmatrix} \text{ and } \boldsymbol{E} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{\epsilon}_1^T \\ \vdots \\ \boldsymbol{\epsilon}_n^T \end{bmatrix}.$$

Considering the ith row of $\boldsymbol{Z}, \boldsymbol{X}$, and $\boldsymbol{E}$ shows that $\boldsymbol{y}_i^T = \boldsymbol{x}_i^T \boldsymbol{B} + \boldsymbol{\epsilon}_i^T$. The least squares estimators

$$\hat{\boldsymbol{B}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Z} = \left[\begin{array}{cccc} \hat{\boldsymbol{\beta}}_1 & \hat{\boldsymbol{\beta}}_2 & \dots & \hat{\boldsymbol{\beta}}_m \end{array}\right] \text{ and } \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \frac{\hat{\boldsymbol{E}}^{\mathrm{T}}\hat{\boldsymbol{E}}}{\mathrm{n-p}} = \frac{1}{\mathrm{n-p}}\sum_{\mathrm{i=1}}^{\mathrm{n}} \hat{\boldsymbol{\epsilon}}_{\mathrm{i}}\hat{\boldsymbol{\epsilon}}_{\mathrm{i}}^{\mathrm{T}}.$$

Assume $n(\boldsymbol{X}^T\boldsymbol{X})^{-1} \xrightarrow{P} \boldsymbol{W}$. By Su and Cook (2012), $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}$ is a $\sqrt{n}$ consistent estimator of $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$, and under mild regularity conditions

$$\sqrt{n}\ \ vec(\hat{\boldsymbol{B}} - \boldsymbol{B}) \xrightarrow{D} N_{pm}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \otimes \boldsymbol{W})$$

where $vec(\boldsymbol{B})$ stacks the columns of a matrix $\boldsymbol{B}$ into a vector and the Kronecker product of an $m \times n$ matrix $\boldsymbol{C}$ and a $p \times q$ matrix $\boldsymbol{D}$ is the $mp \times nq$ matrix $\boldsymbol{C} \otimes \boldsymbol{D}$. Also see Eck (2018), Olive (2017a, ch. 12), and Olive (2017b, ch. 12).

Let $\boldsymbol{A}$ be a full rank $q \times pm$ matrix, $\boldsymbol{\theta} = \boldsymbol{A}vec(\boldsymbol{B})$, and $\hat{\boldsymbol{\theta}} = \boldsymbol{A}vec(\hat{\boldsymbol{B}})$. Consider testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $H_A : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Then

$$T = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T(\boldsymbol{A}[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} \otimes (\boldsymbol{X}^T\boldsymbol{X})^{-1}]\boldsymbol{A}^T)^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} \chi^2_q.$$

The test that rejects H_0 if $T/q > F_{q,n-pm}(1-\delta)$ is a large sample $100(1-\delta)\%$ test since $qF_{q,n-pm} \xrightarrow{D} \chi^2_q$ as $n \to \infty$. We want $n \geq (m+p)^2$, and this test is the Hotelling Lawley test if $rm \times pm$ block diagonal matrix $\boldsymbol{A} = diag(\boldsymbol{L}, ..., \boldsymbol{L})$ is used to test $H_0 : \boldsymbol{LB} = \boldsymbol{0}$ versus $H_A : \boldsymbol{LB} \neq \boldsymbol{0}$ where $\boldsymbol{L}$ is a full rank $r \times p$ matrix. See Olive (2017ab).

Consider the parametric bootstrap for the above model with $\boldsymbol{y}_i^* \sim N_m(\hat{\boldsymbol{B}}^T\boldsymbol{x}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}})$ for $i = 1, ..., n$ where **we are not assuming** that the $\boldsymbol{\epsilon}_i \sim N_m(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})$. Let $\boldsymbol{Z}_j^*$ have ith row $\boldsymbol{y}_i^{*T}$ and regress $\boldsymbol{Z}_j^*$ on $\boldsymbol{X}$ to obtain $\hat{\boldsymbol{B}}_j^*$ for $j = 1, ..., B$. Let $S \subseteq I$, let $\hat{\boldsymbol{B}}_I = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\boldsymbol{Z}^*$, and assume $n(\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1} \xrightarrow{P} \boldsymbol{W}_I$ for any I such that $S \subseteq I$. Then with calculations similar to those for the multiple linear regression model, $E(\hat{\boldsymbol{B}}_I^*) = \hat{\boldsymbol{B}}_I$,

$$\sqrt{n}\ \ vec(\hat{\boldsymbol{B}}_I - \boldsymbol{B}_I) \xrightarrow{D} N_{a_I m}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \otimes \boldsymbol{W}_I),$$

and

$$\sqrt{n}\ \ vec(\hat{\boldsymbol{B}}_I^* - \hat{\boldsymbol{B}}_I) \sim N_{a_I m}(\boldsymbol{0}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} \otimes n(\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}) \xrightarrow{D} N_{a_I m}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \otimes \boldsymbol{W}_I)$$

as $n, B \to \infty$ if $S \subseteq I$. Let $\hat{\boldsymbol{B}}^*_{I,0}$ be formed from $\hat{\boldsymbol{B}}^*_I$ by adding rows of zeroes corresponding to omitted variables.

The nonparametric bootstrap (= empirical bootstrap = naive bootstrap) samples cases $(Y_i, \boldsymbol{x}_i)$ with replacement to form $(\boldsymbol{Y}^*_j, \boldsymbol{X}^*_j)$, and regresses $\boldsymbol{Y}^*_j$ on $\boldsymbol{X}^*_j$ to get $\hat{\boldsymbol{\beta}}^*_j$ for $j = 1, ..., B$. The nonparametric bootstrap can be useful even if heteroscedasticity or overdispersion is present, if the cases are an iid sample from some population, a very strong assumption. See Freedman (1981) and Efron (1982, pp. 18, 19, 35, 36) for the residual bootstrap and nonparametric bootstrap applied to multiple linear regression. See Eck (2018) for using these two methods to bootstrap multivariate linear regression.

For bootstrapping a GLM, see, for example, Chatterjee and Bose (2005), Davison and Hinkley (1997, section 7.2), Friedl (1997), Moulton and Zeger (1991), Shao and Tu (1995), and Simonoff and Tsai (1988). Several of these methods do not make a $\boldsymbol{Y}^*$, and so may not be useful for variable selection.

2.5 BOOTSTRAPPING VARIABLE SELECTION ESTIMATORS

Consider testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ where $\boldsymbol{\theta}$ is $g \times 1$. The Pelawa Watagoda and Olive (2019b) geometric argument is useful. Assume $T_1, ..., T_B$ are iid with nonsingular covariance matrix $\boldsymbol{\Sigma}_{T_n}$. Then the large sample $100(1-\delta)\%$ prediction region $R_p = \{\boldsymbol{w} : D^2_{\boldsymbol{w}}(\overline{T}, \boldsymbol{S}_T) \leq D^2_{(U_B)}\}$ centered at $\overline{T}$ contains a future value of the statistic T_f with probability $1 - \delta_B \to 1 - \delta$ as $B \to \infty$. Assume $\sqrt{n}(T_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{u}$ with $E(\boldsymbol{u}) = \boldsymbol{0}$ and $\text{Cov}(\boldsymbol{u}) = \boldsymbol{\Sigma}_{\boldsymbol{u}}$, and $(n\boldsymbol{S}_T)^{-1}$ to be fairly well behaved (not too ill conditioned) for each $n \geq 20g$, say. Then the region $R_c = \{\boldsymbol{w} : D^2_{\boldsymbol{w}}(T_n, \boldsymbol{S}_T) \leq D^2_{(U_B)}\}$ centered at a randomly selected T_n is a large sample $100(1-\delta)\%$ confidence region for $\boldsymbol{\theta}$ as $n, B \to \infty$. Geometrically, the bootstrap sample data cloud of $T^*_1, ..., T^*_B$ is like the data cloud of iid $T_1, ..., T_B$ shifted to be centered at T_n if $\sqrt{n}(T^*_j - T_n) \xrightarrow{D} \boldsymbol{u}$.

The explanation for why the bootstrap confidence regions (2.5), (2.6), and (2.7) give useful results after variable selection is nearly the same as that given by Pelawa Watagoda and Olive (2019b) for the residual bootstrap for multiple linear regression. Let the variable selection estimator $T_n = \boldsymbol{A}\hat{\boldsymbol{\beta}}_{I_{min},0}$ with $\boldsymbol{\theta} = \boldsymbol{A}\boldsymbol{\beta}$. Then T_n is not smooth since T_n is equal to the estimator T_{jn} with probability π_{jn} for $j = 1, ..., J$. Here $\boldsymbol{A}$ is a known full rank $g \times p$ matrix with $1 \leq g \leq p$. We have $\sqrt{n}(T_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{v}$ by (2.12) where $E(\boldsymbol{v}) = \boldsymbol{0}$, and $\boldsymbol{\Sigma}_{\boldsymbol{v}} = \sum_j \pi_j \boldsymbol{A}\boldsymbol{V}_{j,0}\boldsymbol{A}^T$. Hence the above geometric argument holds: (2.6) and (2.7) are large sample confidence regions for $\boldsymbol{\theta}$. For variable selection, this section will show that the bootstrap sample data cloud $T_1^*, ..., T_B^*$ tends to be slightly more variable than the data cloud of iid $T_1, ..., T_B$ for large n. Hence the coverage of these two confidence regions $\geq 1 - \delta$, asymptotically. Empirically, for a mixture distribution, the bagging estimator $\overline{T}^*$ tends estimate $\boldsymbol{\theta}$ at least as well as T_n. See Breiman (1996) and Yang (2003). Hence the coverage of (2.5) tends to be $\geq$ the coverage of the hybrid region (2.7).

Assume p is fixed and $n \geq 20p$. The response plot has $ESP = \boldsymbol{x}^T\hat{\boldsymbol{\beta}}$ on the horizontal axis and Y on the vertical axis, respectively. Often the model mean function and a scatterplot smoother are added to the plots as visual aids to check the model. See Olive (2013b, 2017a, ch. 13). Olive (2010, ch. 16) gives plots for checking survival regression models.

For the bootstrap, suppose that T_i^* is equal to T_{ij}^* with probability ρ_{jn} for $j = 1, ..., J$ where $\sum_j \rho_{jn} = 1$, and $\rho_{jn} \to \pi_j$ as $n \to \infty$. Let B_{jn} count the number of times $T_i^* = T_{ij}^*$ in the bootstrap sample. Then the bootstrap sample $T_1^*, ..., T_B^*$ can be written as

$$T_{1,1}^*, ..., T_{B_{1n},1}^*, ..., T_{1,J}^*, ..., T_{B_{Jn},J}^*$$

where the B_{jn} follow a multinomial distribution and $B_{jn}/B \xrightarrow{P} \rho_{jn}$ as $B \to \infty$. Denote $T_{1j}^*, ..., T_{B_{jn},j}^*$ as the jth bootstrap component of the bootstrap sample with sample mean $\overline{T}_j^*$ and sample covariance matrix $\boldsymbol{S}_{T,j}^*$. Then

$$\overline{T}^* = \frac{1}{B}\sum_{i=1}^{B} T_i^* = \sum_j \frac{B_{jn}}{B}\frac{1}{B_{jn}}\sum_{i=1}^{B_{jn}} T_{ij}^* = \sum_j \hat{\rho}_{jn}\overline{T}_j^*.$$

Similarly, we can define the jth component of the iid sample $T_1, ..., T_B$ to have sample mean $\overline{T}_j$ and sample covariance matrix $\boldsymbol{S}_{T,j}$.

For the parametric bootstrap, Section 2.4 showed that under regularity conditions, $\text{Cov}(\hat{\boldsymbol{\beta}}^*_I)-$ $\text{Cov}(\hat{\boldsymbol{\beta}}_I) \to \mathbf{0}$ as $n, B \to \infty$ if $S \subseteq I$. Hence $\text{Cov}(T_{jn}) - \text{Cov}(T^*_{jn}) \to \mathbf{0}$ as $n, B \to \infty$ if $S \subseteq I$. Here $T_n = \boldsymbol{A}\hat{\boldsymbol{\beta}}_{I_{min},0}$, $T_{jn} = \boldsymbol{A}\hat{\boldsymbol{\beta}}_{I_j,0}$, $T^*_n = \boldsymbol{A}\hat{\boldsymbol{\beta}}^*_{I_{min},0}$, and $T^*_{jn} = \boldsymbol{A}\hat{\boldsymbol{\beta}}^*_{I_j,0}$. Then $E(T_{jn}) \approx \boldsymbol{A\beta} = \boldsymbol{\theta}$ while the $E(T^*_{jn})$ are more variable than the $E(T_{jn})$ with $E(T^*_{jn}) \approx \boldsymbol{A}\hat{\boldsymbol{\beta}}(I_j, 0)$, roughly, where $\hat{\boldsymbol{\beta}}(I_j, 0)$ is formed from $\hat{\boldsymbol{\beta}}(I_j)$ by adding zeroes corresponding to variables not in I_j.

Hence the jth component of an iid sample $T_1, ..., T_B$ and the jth component of the bootstrap sample $T^*_1, ..., T^*_B$ have the same variability asymptotically. Since $E(T_{jn}) \approx \boldsymbol{\theta}$, each component of the iid sample is approximately centered at $\boldsymbol{\theta}$. The bootstrap components are centered at $E(T_{jn})$. Geometrically, separating the component clouds so that they are no longer centered at one value makes the overall data cloud larger. Thus the variability of T^*_n is larger than that of T_n for a mixture distribution, asymptotically. Hence the prediction region applied to the bootstrap sample is slightly larger than the prediction region applied to the iid sample, asymptotically (we want $n \geq 20p$). Hence cutoff $\hat{D}^2_{1,1-\delta} = D^2_{(U_B)}$ gives coverage close to or higher than the nominal coverage for confidence regions (2.5) and (2.7), using the geometric argument. The deviation $T^*_i - T_n$ tends to be larger in magnitude than the deviation and $T^*_i - \overline{T}^*$. Hence the cutoff $\hat{D}^2_{2,1-\delta} = D^2_{(U_B,T)}$ tends to be larger than $D^2_{(U_B)}$, and region (2.6) tends to have higher coverage than region (2.7) for a mixture distribution.

In simulations for $n \geq 20p$, the coverage tends to get close to $1-\delta$ for $B \geq \max(200, 50p)$ so that $\boldsymbol{S}^*_T$ is a good estimator of $\text{Cov}(T^*)$. In the simulations where S is not the full model, inference with backward elimination with I_{min} using AIC appears to be more precise than inference with the full model if $n \geq 20p$ and $B \geq 50p$. It is possible that $\boldsymbol{S}^*_T$ is singular if a column of the bootstrap sample is equal to $\mathbf{0}$.

Undercoverage can occur if bootstrap sample data cloud is less variable than the iid data cloud, e.g., if $(n-p)/n$ is not close to one. Coverage can be higher than the nominal coverage for two reasons: i) the bootstrap data cloud is more variable than the iid data cloud of $T_1, ..., T_B$, and ii) zero padding.

To see the effect of zero padding, consider $H_0 : \boldsymbol{A\beta} = \boldsymbol{\beta}_O = \boldsymbol{0}$ where $\boldsymbol{\beta}_O = (\beta_{i_1}, ..., \beta_{i_g})^T$ and $O \subseteq E$ in (1.6) so that H_0 is true. Suppose a nominal 95% confidence region is used and $U_{(B)} = 0.96$. Hence the confidence region (2.5) or (2.6) covers at least 96% of the bootstrap sample. If $\hat{\boldsymbol{\beta}}^*_{O,j} = \boldsymbol{0}$ for more than 4% of the $\hat{\boldsymbol{\beta}}^*_{O,1}, ..., \hat{\boldsymbol{\beta}}^*_{O,B}$, then $\boldsymbol{0}$ is in the confidence region and the bootstrap test fails to reject H_0. If this occurs for each run in the simulation, then the observed coverage will be 100%.

Now suppose $\hat{\boldsymbol{\beta}}^*_{O,j} = \boldsymbol{0}$ for $j = 1, ..., B$. Then $\boldsymbol{S}^*_T$ is singular, but the singleton set $\{\boldsymbol{0}\}$ is the large sample $100(1-\delta)\%$ confidence region (2.5), (2.6), or (2.7) for $\boldsymbol{\beta}_O$ and $\delta \in (0, 1)$, and the pvalue for $H_0 : \boldsymbol{\beta}_O = \boldsymbol{0}$ is one. (This result holds since $\{\boldsymbol{0}\}$ contains 100% of the $\hat{\boldsymbol{\beta}}^*_{O,j}$ in the bootstrap sample.) For large sample theory tests, the pvalue estimates the population pvalue. Let I denote the other predictors in the model so $\boldsymbol{\beta} = (\boldsymbol{\beta}_I^T, \boldsymbol{\beta}_O^T)^T$. For the I_{min} model from variable selection, there may be strong evidence that $\boldsymbol{x}_O$ is not needed in the model given $\boldsymbol{x}_I$ is in the model if the "100%" confidence region is $\{\boldsymbol{0}\}$, $n \geq 20p$, and $B \geq 50p$. (Since the pvalue is one, this technique may be useful for data snooping: applying MLE theory to submodel I may have negligible selection bias.)

2.6 R FUNCTIONS FOR THE BACKWARD ELIMINATION

The simulations were done in *R*. See R Core Team (2016). We used several *R* functions including backward elimination computed with the `step` function from the Venables and Ripley (1997) `MASS` library. The collection of Olive (2019) *R* functions *slpack*, available from (http://lagrange.math.siu.edu/Olive/slpack.txt), has some useful functions for the inference. The functions `binregbootsim` and `pregbootsim` are useful for the full binomial regression and full Poisson regression models. The functions `vsbrbootsim` and `vsprbootsim` were used to bootstrap backward elimination for binomial and Poisson regression.

CHAPTER 3

SIMULATIONS FOR PREDICTION INTERVALS

3.1 SIMULATION

Let $\boldsymbol{x} = (1\ \boldsymbol{u}^T)^T$ where $\boldsymbol{u}$ is the $(p-1) \times 1$ vector of nontrivial predictors. In the simulations, for $i = 1, ..., n$, we generated $\boldsymbol{w}_i \sim N_{p-1}(\boldsymbol{0}, \boldsymbol{I})$ where the $m = p-1$ elements of the vector $\boldsymbol{w}_i$ are iid N(0,1). Let the $m \times m$ matrix $\boldsymbol{A} = (a_{ij})$ with $a_{ii} = 1$ and $a_{ij} = \psi$ where $0 \leq \psi < 1$ for $i \neq j$. Then the vector $\boldsymbol{z}_i = \boldsymbol{A}\boldsymbol{w}_i$ so that Cov$(\boldsymbol{z}_i) = \boldsymbol{\Sigma}_{\boldsymbol{z}} = \boldsymbol{A}\boldsymbol{A}^T = (\sigma_{ij})$ where the diagonal entries $\sigma_{ii} = [1+(m-1)\psi^2]$ and the off diagonal entries $\sigma_{ij} = [2\psi+(m-2)\psi^2]$. Hence the correlations are $cor(z_i, z_j) = \rho = (2\psi+(m-2)\psi^2)/(1+(m-1)\psi^2)$ for $i \neq j$. Then $1+\sum_{j=1}^k z_j \sim N(1, k\sigma_{ii}+k(k-1)\sigma_{ij}) = N(1, v^2)$. Let $\boldsymbol{u} = a\boldsymbol{z}/v$. Then $cor(x_i, x_j) = \rho$ for $i \neq j$ where x_i and x_j are nontrivial predictors. If $\psi = 1/\sqrt{cp}$, then $\rho \rightarrow 1/(c+1)$ as $p \rightarrow \infty$ where $c > 0$. As ψ gets close to 1, the predictor vectors $\boldsymbol{u}_i$ cluster about the line in the direction of $(1, ..., 1)^T$. Let $SP = \boldsymbol{x}^T\boldsymbol{\beta} = \beta_1 + 1x_{i,2} + \cdots + 1x_{i,k+1} \sim N(\beta_1, a^2)$ for $i = 1, ..., n$. Hence $\boldsymbol{\beta} = (\beta_1, 1, .., 1, 0, ..., 0)^T$ with β_1, k ones and $p-k-1$ zeros. The default settings for Poisson regression use $\beta_1 = 1 = a$.

The simulation used 5000 runs, so an observed coverage in [0.94, 0.96] gives no reason to doubt that the PI has the nominal coverage of 0.95. The simulation used $B = 1000$; $p = 4, 50, n$, or $2n$; $\psi = 0, 1/\sqrt{p}$, or 0.9; and $k = 1, 19$, or $p-1$. The simulated data sets are rather small since the R estimators are rather slow. For Poisson regression, we only computed the GAM for $p = 4$ with $SP = AP = \alpha + S_2(x_2) + S_2(x_3) + S_4(x_4)$. We only computed the full model GLM if $n \geq 5p$. Lasso and relaxed lasso were computed for all cases.

Tables 3.1 and 3.2 show some simulation results for Poisson regression. Lasso minimized 10-fold CV and relaxed lasso was applied to the selected lasso model. The full model GLM and GAM used PI (1.5) while lasso, relaxed lasso, forward selection using the Olive and Hawkins (2005) method, and backward elimination used PI (1.8). For $n \geq 10p$, coverages tended to be near or higher than the nominal value of 0.95, except for lasso and the Olive and Hawkins (2005) method in Table 3.2. In Table 3.1, coverages were high because the Poisson counts were small and the Poisson distribution is discrete. In Table 3.2, the Poisson counts were not small, so the discreteness of the distribution did not affect the coverage much.

For $n \leq p$, good performance needed stronger regularity conditions. Tables 3.3-3.10 show more results. Relaxed lasso often performed better from lasso if $n \geq 10k$ and $n \leq 5p$.

3.2 CONCLUSION

Since PIs (1.5) and (1.8) are for a parametric regression model, it is crucial to check that the parametric model is appropriate. For example, if a negative binomial regression model is appropriate, but a Poisson regression model is fit, then the PI coverage will likely be poor. The response plot of the ESP on the horizontal axis versus the response on the vertical axis is useful. This plot and a plot for detecting overdispersion are described in Olive (2013b, 2017a: ch. 13). Olive (2019: ch. 4) shows that these plots can be useful for methods such as lasso and elastic net if n/p is not large, although estimation becomes more difficult. Plots for checking survival regression models are given in Olive (2010: ch. 16). In a similar application, Olive (2014, p. 364) used the shorth estimator to estimate Bayesian credible regions. Cai, Tian, Solomon, and Wei (2008) give useful prediction intervals.

Let d be the number of variables $x_1^*, ..., x_d^*$ used by forward selection, lasso, or relaxed lasso. Or let d be the degrees of freedom of the selected model if that model was chosen in advance without model or variable selection. Then we want $n \geq 10d$ in order to use the new PI (1.5). For fixed p, the GLM estimator $\hat{\boldsymbol{\beta}}$ tends to be a $\sqrt{n}$ consistent asymptotically normal estimator of $\boldsymbol{\beta}$. Forward selection with AIC tends to produce a $\sqrt{n}$ consistent estimator $\hat{\boldsymbol{\beta}}_{I_{min},0}$ of $\boldsymbol{\beta}$ if the probability that the subset I_{min} contains S goes to one as $n \rightarrow \infty$ where the $p \times 1$ vector $\hat{\boldsymbol{\beta}}_{I_{min},0}$ has zeroes corresponding to variables that were not selected. For example, if $p = 4$ and $\hat{\boldsymbol{\beta}}_{I_{min}} = (\hat{\beta}_1, \hat{\beta}_3)^T$, then $\hat{\boldsymbol{\beta}}_{I_{min},0} = (\hat{\beta}_1, 0, \hat{\beta}_3, 0)^T$.

Table 3.1. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 4$, $\beta_1 = 1 = a$

n	ψ	k		GLM	GAM	lasso	RL	OHFS	BE
100	0	1	cov	0.9712	0.9714	0.9810	0.9800	0.9792	0.9734
			len	6.6448	6.6118	7.2770	7.2004	7.0680	6.6632
400	0	1	cov	0.9692	0.9694	0.9728	0.9714	0.9722	0.9665
			len	6.6392	6.6474	6.7996	6.7722	6.7588	6.6778
100	0.5	1	cov	0.9642	0.9644	0.9796	0.9786	0.9760	0.9689
			len	6.6922	6.6806	7.3136	7.2824	7.1160	6.7767
400	0.5	1	cov	0.9668	0.9670	0.9722	0.9716	0.9702	0.9754
			len	6.672	6.6896	6.8342	6.8140	6.7992	6.7802
100	0.9	1	cov	0.9672	0.9674	0.9766	0.9768	0.9738	0.9665
			len	6.6038	6.6186	7.1480	7.1214	7.0002	6.5789
400	0.9	1	cov	0.9660	0.9662	0.9734	0.9700	0.9692	0.9798
			len	6.5838	6.5746	6.7526	6.7196	6.7004	6.7443
100	0	3	cov	0.9696	0.9698	0.9848	0.9834	0.9818	0.9654
			len	6.7080	6.7084	7.5632	7.5442	7.5348	6.7408
400	0	3	cov	0.9728	0.9730	0.9750	0.9746	0.9748	0.9657
			len	6.5718	6.5684	6.7690	6.7356	6.7406	6.7063
100	0.5	3	cov	0.9672	0.9674	0.9842	0.9838	0.9736	0.9592
			len	6.6992	6.7044	7.5804	7.5494	7.3810	6.7128
400	0.5	3	cov	0.9682	0.9684	0.9730	0.9722	0.9702	0.9772
			len	6.6794	6.6890	6.8726	6.8520	6.8466	6.7504
100	0.9	3	cov	0.9664	0.9666	0.9804	0.9810	0.9750	0.9678
			len	6.6704	6.6646	7.2880	7.2672	7.0722	6.7635
400	0.9	3	cov	0.9690	0.9692	0.9744	0.9742	0.9736	0.9667
			len	6.7960	6.8092	6.9696	6.9682	6.9120	6.6987

Table 3.2. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 4$, $\beta_1 = 5$, $a = 2$

n	ψ	k		GLM	GAM	lasso	RL	OHFS	BE
100	0	1	cov	0.9500	0.9440	0.7730	0.9664	0.9654	0.9520
			len	77.6072	77.6306	84.1066	81.8374	82.4752	84.1432
400	0	1	cov	0.9580	0.9564	0.7566	0.9622	0.9628	0.9534
			len	82.0126	82.0212	85.5704	83.2692	83.4374	80.9897
100	0.5	1	cov	0.9456	0.9424	0.7646	0.9634	0.9408	0.9512
			len	83.0236	82.9034	90.5822	88.3060	88.6700	79.6887
400	0.5	1	cov	0.9530	0.9500	0.7584	0.9604	0.9566	0.9678
			len	83.8588	83.8292	87.4336	85.1042	85.1434	79.9855
100	0.9	1	cov	0.9492	0.9452	0.7688	0.9646	0.7712	0.9654
			len	78.3554	78.3798	87.0086	84.6072	83.4980	81.5432
400	0.9	1	cov	0.9550	0.9574	0.7606	0.9606	0.7928	0.9513
			len	76.7028	76.7594	80.5070	78.2308	78.2538	80.1298
100	0	3	cov	0.9544	0.9466	0.7798	0.9708	0.9404	0.9487
			len	80.1476	80.1362	92.1372	89.8532	90.3456	79.4565
400	0	3	cov	0.9560	0.9548	0.7514	0.9582	0.9566	0.9567
			len	80.7868	80.8976	85.0642	82.7982	82.7912	79.4522
100	0.5	3	cov	0.9516	0.9478	0.7848	0.9694	0.3324	0.9515
			len	77.1120	77.1130	88.9346	86.4680	85.8634	81.5643
400	0.5	3	cov	0.9568	0.9558	0.7534	0.9636	0.5214	0.9528
			len	80.4226	80.4932	84.7646	82.5590	83.7526	79.9786
100	0.9	3	cov	0.9492	0.9456	0.7882	0.9620	0.7510	0.9554
			len	79.5374	79.6172	91.2052	89.0692	84.5648	81.8544
400	0.9	3	cov	0.9544	0.9546	0.7638	0.9554	0.7384	0.9586
			len	79.7384	79.6906	83.8318	81.6862	81.0882	80.7521

Table 3.3. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 100$, $\beta_1 = 1 = a$

n	ψ	k		lasso	RL
100	0	1	cov	0.9780	0.9646
			len	7.7882	7.5340
100	0.1	1	cov	0.9828	0.9684
			len	7.9222	7.7810
100	0.9	1	cov	0.9820	0.9820
			len	7.5274	7.4780
100	0	19	cov	0.9056	0.8482
			len	7.8496	7.7138
100	0.1	19	cov	0.9736	0.9602
			len	7.7392	7.6224
100	0.9	19	cov	0.9816	0.9792
			len	7.5280	7.4934
100	0	99	cov	0.8356	0.7660
			len	7.9902	7.9322
100	0.1	99	cov	0.9746	0.9630
			len	8.0254	7.9280
100	0.9	99	cov	0.9822	0.9824
			len	7.6210	7.5858

Table 3.4. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 100$, $\beta_1 = 5, a = 2$

n	ψ	k		lasso	RL
100	0	1	cov	0.7552	0.9678
			len	90.8148	88.3212
100	0.1	1	cov	0.7658	0.9700
			len	91.4890	88.8760
100	0.9	1	cov	0.7894	0.9650
			len	93.6124	91.1536
100	0	19	cov	0.2494	0.8134
			len	89.4138	89.2508
100	0.1	19	cov	0.4576	0.7152
			len	92.7120	91.7778
100	0.9	19	cov	0.7774	0.9640
			len	93.2236	90.6284
100	0	99	cov	0.0750	0.1042
			len	100.7164	93.1344
100	0.1	99	cov	0.3584	0.5270
			len	91.6490	89.6920
100	0.9	99	cov	0.7826	0.9624
			len	92.6190	90.0968

Table 3.5. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 200$, $\beta_1 = 1 = a$

n	ψ	k		lasso	RL
100	0	1	cov	0.9816	0.9612
			len	7.8350	7.5660
100	0.07	1	cov	0.9790	0.9696
			len	7.8488	7.6680
100	0.9	1	cov	0.9880	0.9858
			len	7.6380	7.5954
100	0	19	cov	0.8678	0.8038
			len	7.8126	7.5188
100	0.07	19	cov	0.9620	0.9444
			len	7.9010	7.7828
100	0.9	19	cov	0.9838	0.9848
			len	7.6900	7.6454
100	0.07	199	cov	0.9720	0.9548
			len	8.0784	7.9956
100	0.9	199	cov	0.9834	0.9814
			len	7.6728	7.6602

Table 3.6. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 200$, $\beta_1 = 5, a = 2$

n	ψ	k		lasso	RL
100	0	1	cov	0.7620	0.9662
			len	93.7318	91.4898
100	0.07	1	cov	0.7652	0.9706
			len	92.0774	89.7966
100	0.9	1	cov	0.7850	0.9628
			len	98.2158	95.9954
100	0	19	cov	0.1610	0.6754
			len	88.0896	90.6916
100	0.07	19	cov	0.3790	0.5832
			len	92.3918	92.1424
100	0.9	19	cov	0.7884	0.9594
			len	97.744	95.2898
100	0	199	cov	0.0416	0.0588
			len	113.2106	99.9042
100	0.07	199	cov	0.3076	0.4394
			len	90.4494	88.0354
100	0.9	199	cov	0.7888	0.9586
			len	97.0954	94.7604

Table 3.7. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 50$, $\beta_1 = 1 = a$

n	ψ	k		GLM	lasso	RL	OHFS	BE
500	0	1	cov	0.9566	0.9742	0.9688	0.9716	0.9618
			len	6.6532	6.9764	6.8936	7.0210	6.6524
500	0.14	1	cov	0.9600	0.9768	0.9742	0.9758	0.9644
			len	6.7688	7.0956	7.0508	7.1446	6.7742
500	0.9	1	cov	0.9526	0.9742	0.9746	0.9748	0.9608
			len	6.7022	6.9494	6.9244	7.0496	6.7044
500	0	19	cov	0.9602	0.9858	0.9844	0.9858	0.9628
			len	6.5862	7.8326	7.7302	7.6870	6.5986
500	0.14	19	cov	0.9568	0.9854	0.9852	0.9662	0.9590
			len	6.7284	7.9386	7.8970	7.2752	6.7288
500	0.9	19	cov	0.9620	0.9782	0.9772	0.9780	0.9654
			len	6.6730	6.9296	6.9006	7.0076	6.6802
500	0	49	cov	0.9500	0.9796	0.9786	0.9674	0.9506
			len	6.6330	7.7976	7.7888	7.8018	6.6508
500	0.14	49	cov	0.9572	0.9852	0.9834	0.9622	0.9570
			len	6.6826	7.8842	7.8406	7.2322	6.6940
500	0.9	49	cov	0.9572	0.9852	0.9834	0.9622	0.9570
			len	6.6826	7.8842	7.8406	7.2322	6.6940

Table 3.8. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 50$, $\beta_1 = 5, a = 2$

n	ψ	k		GLM	lasso	RL	OHFS	BE
500	0	1	cov	0.9352	0.7564	0.9598	0.9640	0.9476
			len	81.2668	84.3188	81.8934	85.2922	81.1010
500	0.14	1	cov	0.9370	0.7508	0.9580	0.9628	0.9458
			len	81.1820	84.4530	82.1894	85.2304	81.1146
500	0.9	1	cov	0.9368	0.7630	0.9620	0.8994	0.9456
			len	80.4568	86.3506	84.4942	84.1448	80.4202
500	0	19	cov	0.9388	0.7592	0.9756	0.3778	0.9472
			len	81.6922	96.8546	94.6350	99.7436	81.7218
500	0.14	19	cov	0.9368	0.7556	0.9730	0.2770	0.9438
			len	80.0654	95.2964	93.2748	87.3814	80.1276
500	0.9	19	cov	0.935	0.7544	0.9536	0.9480	0.9352
			len	79.7324	86.3448	84.0674	83.2958	79.6172
500	0	49	cov	0.9386	0.7104	0.9666	0.1004	0.9364
			len	81.1422	96.4304	94.8818	108.0518	81.2516
500	0.14	49	cov	0.9396	0.7194	0.9558	0.2858	0.9402
			len	79.7874	94.8908	93.2538	86.4234	79.8692
500	0.9	49	cov	0.9380	0.7640	0.9480	0.9512	0.9430
			len	78.8146	85.5786	83.2812	82.4104	78.8316

Table 3.9. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 50$, $\beta_1 = 1 = a$

n	ψ	k		GLM	lasso	RL	OHFS	BE
1000	0	1	cov	0.9630	0.9734	0.9698	0.9714	0.9666
			len	6.6790	6.8744	6.8138	6.8648	6.6962
1000	0.14	1	cov	0.9680	0.9762	0.9750	0.9752	0.9712
			len	6.7208	6.8990	6.8718	6.9096	6.7354
1000	0.9	1	cov	0.9616	0.9720	0.9708	o.9710	0.9668
			len	6.7140	6.9018	6.8724	6.9064	6.7258
1000	0	19	cov	0.9658	0.9830	0.9808	0.9790	0.9652
			len	6.8096	7.6654	7.6138	7.2908	6.8084
1000	0.14	19	cov	0.9644	0.9806	0.9792	0.9644	0.9654
			len	6.6764	7.3378	7.3192	6.9968	6.6778
1000	0.9	19	cov	0.9668	0.9726	0.9722	0.9708	0.9644
			len	6.7122	6.8712	6.8524	6.8722	6.7080
1000	0	49	cov	0.9630	0.9850	0.9852	0.9836	0.9610
			len	6.7108	7.8856	7.8596	7.8602	6.7092
1000	0.14	49	cov	0.9652	0.9806	0.9810	0.9630	0.9626
			len	6.7066	7.6140	7.5996	7.0104	6.7108
1000	0.9	49	cov	0.9690	0.9784	0.9780	0.9764	0.9714
			len	6.6628	6.8546	6.8288	6.8418	6.6868

Table 3.10. Simulated Large Sample 95% PI Coverages and Lengths for Poisson Regression, $p = 50$, $\beta_1 = 5, a = 2$

n	ψ	k		GLM	lasso	RL	OHFS	BE
1000	0	1	cov	0.9466	0.7582	0.9576	0.9592	0.9488
			len	80.7284	83.3446	81.2210	82.7864	80.7752
1000	0.14	1	cov	0.9426	0.7458	0.9538	0.9576	0.9480
			len	79.3900	82.3612	79.8172	81.3890	79.3798
1000	0.9	1	cov	0.9476	0.7608	0.9576	0.8950	0.9486
			len	80.4882	84.4228	82.3468	82.4528	80.3866
1000	0	19	cov	0.9436	0.7270	0.9612	0.6700	0.9438
			len	80.6532	88.0292	85.7734	88.5528	80.5738
1000	0.14	19	cov	0.9438	0.7452	0.9634	0.2828	0.9460
			len	82.1308	91.1308	89.1170	87.2076	82.2310
1000	0.9	19	cov	0.9432	0.7354	0.9444	0.9466	0.9444
			len	82.2056	86.5944	84.4406	84.1516	82.0824
1000	0	49	cov	0.9488	0.7466	0.9760	0.1532	0.9484
			len	80.6898	96.3376	94.2010	99.9758	80.6404
1000	0.14	49	cov	0.9502	0.7614	0.9752	0.2822	0.9502
			len	79.2666	94.6062	92.5096	85.7218	79.3182
1000	0.9	49	cov	0.9536	0.7380	0.9536	0.9536	0.9542
			len	78.9550	83.5724	81.2776	80.8550	79.0614

Many 1D regression models with $SP = \boldsymbol{x}^T\boldsymbol{\beta}$ do not yet have convenient variable selection programs. For $n \geq 10p$, Olive and Hawkins (2005) suggest using a variable selection program for multiple linear regression with C_p to get a subset I. Then use that subset in the 1D regression software. When n/p is not large, the lasso for multiple linear regression can sometimes give a useful subset I. See Olive (2019, $\oint$ 4.6.2).

GLMs were introduced by Nelder and Wedderburn (1972). Cook and Zhang (2015) show that envelope methods have the potential to significantly improve GLMs. Useful references for generalized additive models include Hastie and Tibshirani (1986, 1990), and Wood (2006). Some large sample theory for the GAM is given by Wang, Liu, Liang, and Carroll (2011). Zhou (2001) is useful for simulating the Weibull regression model. Tables were made with R function prpisim2.

CHAPTER 4

SIMULATIONS FOR BOOTSTRAPPING

4.1 SIMULATION

A simulation was done, using $B = \max(200, n/10, 50p)$ and 5000 runs. The simulation used $p = 4, 6, 7, 8,$ and 10; $n = 25p$, $n = 50p$; $\psi = 0, 1/\sqrt{p}$, and 0.9; and $k = 1$ and $p - 2$ where k and ψ are defined in the following paragraph.

Let $\boldsymbol{x} = (1\ \boldsymbol{u}^T)^T$ where $\boldsymbol{u}$ is the $(p-1) \times 1$ vector of nontrivial predictors. In the simulations, for $i = 1, ..., n$, we generated $\boldsymbol{w}_i \sim N_{p-1}(\boldsymbol{0}, \boldsymbol{I})$ where the $m = p-1$ elements of the vector $\boldsymbol{w}_i$ are iid N(0,1). Let the $m \times m$ matrix $\boldsymbol{A} = (a_{ij})$ with $a_{ii} = 1$ and $a_{ij} = \psi$ where $0 \leq \psi < 1$ for $i \neq j$. Then the vector $\boldsymbol{z}_i = \boldsymbol{A}\boldsymbol{w}_i$ so that $\text{Cov}(\boldsymbol{z}_i) = \boldsymbol{\Sigma}_{\boldsymbol{z}} = \boldsymbol{A}\boldsymbol{A}^T = (\sigma_{ij})$ where the diagonal entries $\sigma_{ii} = [1+(m-1)\psi^2]$ and the off diagonal entries $\sigma_{ij} = [2\psi+(m-2)\psi^2]$. Hence the correlations are $cor(z_i, z_j) = \rho = (2\psi+(m-2)\psi^2)/(1+(m-1)\psi^2)$ for $i \neq j$. Then $\sum_{j=1}^{k} z_j \sim N(0, k\sigma_{ii} + k(k-1)\sigma_{ij}) = N(0, v^2)$. Let $\boldsymbol{u} = a\boldsymbol{z}/v$. Then $cor(x_i, x_j) = \rho$ for $i \neq j$ where x_i and x_j are nontrivial predictors. If $\psi = 1/\sqrt{cp}$, then $\rho \to 1/(c+1)$ as $p \to \infty$ where $c > 0$. As ψ gets close to 1, the predictor vectors $\boldsymbol{u}_i$ cluster about the line in the direction of $(1, ..., 1)^T$. Let $SP = \boldsymbol{x}^T\boldsymbol{\beta} = \beta_1 + 1x_{i,2} + \cdots + 1x_{i,k+1} \sim N(\beta_1, a^2)$ for $i = 1, ..., n$. Hence $\boldsymbol{\beta} = (\beta_1, 1, .., 1, 0, ..., 0)^T$ with β_1, k ones and $p-k-1$ zeros. Binomial regression used $\beta_1 = 0, a = 5/3$, and $m_i = m$ with $m = 1$ or 20. Poisson regression used $\beta_1 = 1 = a$ and $\beta_1 = 5$ with $a = 2$.

The simulation computed the Frey shorth(c) interval for each β_i and used bootstrap confidence regions to test $H_0 : \boldsymbol{\beta}_S = (\beta_1, 1, ..., 1)^T$ where $\beta_2 = \cdots = \beta_{k+1} = 1$, and $H_0 : \boldsymbol{\beta}_E = \boldsymbol{0}$ (whether the last $p-k-1$ $\beta_i = 0$). The nominal coverage was 0.95 with $\delta = 0.05$. Observed coverage between 0.94 and 0.96 would suggest coverage is close to the nominal value. In the tables, see columns pr1, br1 and hyb1 for the 1st test, and columns pr0, br0 and hyb0 for the 2nd test. The headers pr, br, and hyb refer to confidence regions (2.5), (2.6), and (2.7), respectively.

In the tables, there are two rows for each model giving the observed confidence interval coverages and average lengths of the confidence intervals. The term "reg" is for the full model regression, and the term "vs" is for backward elimination. The last six columns give results for the tests. The length and coverage = P(fail to reject H_0) for the interval $[0, D_{(U_B)}]$ or $[0, D_{(U_B,T)}]$ where $D_{(U_B)}$ or $D_{(U_B,T)}$ is the cutoff for the confidence region. The cutoff will often be near $\sqrt{\chi^2_{g,0.95}}$ if the statistic T is asymptotically normal. Note that $\sqrt{\chi^2_{2,0.95}} = 2.448$ is close to 2.45 for the full model regression bootstrap tests. Volumes of the confidence regions can be compared using (2.8). The zeroes in $\hat{\boldsymbol{\beta}}_E$ result in higher than nominal coverage for the variable selection estimator, but can greatly decrease the volume of the confidence region compared to that of the full model.

Variable selection coverage tended to be near 0.95 unless the $\hat{\beta}_i$ could equal 0. Full model coverage tended to be near 0.95. An exception was binary logistic regression with $m = 1$ in Table 4.1 where variable selection and the full model had higher coverage than the nominal 0.95 for the hypothesis tests. With $\boldsymbol{x}^T\boldsymbol{\beta} \sim N(0, (5/3)^2))$, nearly all of the $\boldsymbol{x}_i^T\boldsymbol{\beta}$ were between -5 and 5. Then some probabilities $\hat{\rho}(\boldsymbol{x}_i)$ were computed to be exactly 0 or 1, which causes problems for the GLM MLE. When there are problems with the MLE, the bootstrap confidence regions using smaller a and larger n resulted in coverages closer to 0.95 for the full model.

The confidence intervals for variable selection tended to have shorter average length than those for the full model. The confidence region volumes for variable selection and the full model can not be compared by looking at the cutoffs, since the covariance matrices for the variable selection model and the full model differ.

Table 4.1. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9516	0.9528	0.9524	0.9504	0.9724	0.9872	0.9920	0.9802	0.9838	0.9888
len	1.1605	1.0953	0.7171	0.7151	2.5225	2.5225	2.5476	2.5173	2.5173	2.6893
vs,0	0.9564	0.9322	0.9976	0.9976	0.9960	0.9964	0.9988	0.9774	0.9794	0.9948
len	1.1483	1.0798	0.6143	0.6204	2,7329	2.7329	3.0386	2.5160	2.5160	2.6899
reg,0.5	0.9538	0.9428	0.9440	0.9544	0.9680	0.9854	0.9896	0.9724	0.9828	0.9858
len	1.1622	1.6737	1.4547	1.4588	2.5221	2.5221	2.5475	2.5165	2.5165	2.6037
vs,0.5	0.9528	0.9662	0.9978	0.9982	0.9948	0.9918	0.9978	0.9760	0.9756	0.9872
len	1.1462	1.6714	1.2879	1.2883	2.7230	2.7230	3.0170	2.5379	2.5379	2.6860
reg,0.9	0.9662	0.9578	0.9520	0.9500	0.9690	0.9846	0.9884	0.9724	0.9848	0.9876
len	1.1606	9.4523	9.4241	9.4379	2.5220	2.5220	2.5454	2.5142	2.5142	2.5389
vs,0.9	0.9566	0.9522	0.9960	0.9974	0.9958	0.9972	0.9982	0.9866	0.9932	0.9956
len	1.1502	8.4654	8.4806	8.4951	2.7700	2.7700	3.0182	2.6176	2.6176	2.7644

Table 4.2. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9504	0.9530	0.9506	0.9572	0.9474	0.9490	0.9500	0.9440	0.9426	0.9444
len	0.2284	0.1978	0.1384	0.1384	2.4627	2.4627	2.4693	2.4623	2.4623	2.4768
vs,0	0.9556	0.9560	0.9988	0.9994	0.9944	0.9860	0.9946	0.9538	0.9550	0.9568
len	0.2278	0.1978	0.1169	0.1169	2.7166	2.7166	3.0218	2.4621	2.4621	2.4809
reg,0.5	0.9506	0.9534	0.9528	0.9554	0.9490	0.9510	0.9506	0.9494	0.9494	0.9526
len	0.2284	0.3138	0.2814	0.2806	2.4640	2.4640	2.4702	2.4613	2.4613	2.4713
vs,0.5	0.9536	0.9650	0.9990	0.9980	0.9934	0.9856	0.9944	0.9594	0.9576	0.9634
len	0.2283	0.3132	0.2487	0.2422	2.6951	2.6951	2.9860	2.4658	2.4658	2.5719
reg,0.9	0.9584	0.9600	0.9498	0.9596	0.9498	0.9530	0.9526	0.9548	0.9564	0.9578
len	0.2283	1.8362	1.8282	2.4647	2.4647	2.4708	2.4634	2.4634	2.4634	2.4696
vs,0.9	0.9578	0.9832	0.9990	0.9990	0.9934	0.9420	0.9968	0.9570	0.9322	0.9592
len	0.2285	1.7158	1.6026	1.6149	2.6072	2.6072	2.8477	2.4505	2.4505	2.5816

Table 4.3. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9524	0.9568	0.9554	0.9524	0.9420	0.9452	0.9456	0.9440	0.9448	0.9484
len	0.2815	0.2167	0.2071	0.2064	2.4618	2.4618	2.4687	2.4594	2.4594	2.4778
vs,0	0.9548	0.9546	0.9988	0.9988	0.9838	0.9824	0.9944	0.9498	0.9522	0.9544
len	0.2814	0.2159	0.1755	0.1749	2.7104	2.7104	3.0156	2.4600	2.4600	2.4959
reg,0.5	0.9564	0.9550	0.9568	0.9578	0.9512	0.9514	0.9526	0.9528	0.9530	0.9558
len	0.2813	0.4254	0.4209	0.4201	2.4633	2.4633	2.4699	2.4602	2.4602	2.4782
vs,0.5	0.9560	0.9742	0.9986	0.9988	0.9950	0.9884	0.9958	0.9656	0.9652	0.9716
len	0.2814	0.4212	0.3571	0.3587	2.6979	2.6979	2.9879	2.4548	2.4748	2.6163
reg,0.9	0.9542	0.9544	0.9516	0.9562	0.9446	0.9440	0.9458	0.9460	0.9476	0.9502
len	0.2819	2.7275	2.7352	2.7271	2.4615	2.4615	2.4672	2.4613	2.4613	2.4774
vs,0.9	0.9458	0.9790	0.9996	0.9972	0.9936	0.9794	0.9948	0.9588	0.9552	0.9656
len	0.2812	2.4059	2.3828	2.3887	2.6377	2.6377	2.8682	2.4687	2.4687	2.5839

Table 4.4. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9494	0.9526	0.9576	0.9536	0.9488	0.9498	0.9506	0.9454	0.9466	0.9472
len	0.0342	0.0102	0.0082	0.0082	2.4593	2.4593	2.4656	2.4580	2.4580	2.4647
vs,0	0.9518	0.9582	0.9996	0.9992	0.9960	0.9856	0.9956	0.9536	0.9536	0.9572
len	0.0341	0.0101	0.0069	0.0069	2.7025	2.7025	3.0065	2.4649	2.4649	2.5232
reg,0.5	0.9502	0.9502	0.9490	0.9546	0.9412	0.9406	0.9416	0.9406	0.9432	0.9426
len	0.0343	0.0179	0.0167	0.0168	2.4608	2.4608	2.4681	2.4588	2.4588	2.4655
vs,0.5	0.9566	0.9704	0.9974	0.9998	0.9916	0.9812	0.9930	0.9594	0.9618	0.9664
len	0.0343	0.0176	0.0143	0.0142	2,6932	2.6932	2.9780	2.4826	2.4826	2.6318
reg,0.9	0.9546	0.9500	0.9504	0.9544	0.9446	0.9460	0.9464	0.9488	0.9492	0.9500
len	0.0343	0.1091	0.1091	0.1098	2.4584	2.4584	2.4639	2.4607	2.4607	2.4680
vs,0.9	0.9584	0.9888	0.9990	0.9988	0.9942	0.9844	0.9958	0.9684	0.9684	0.9724
len	0.0341	0.1058	0.0937	0.0932	2.6970	2.6970	2.9805	2.4838	2.4838	2.6364

Table 4.5. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 2$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9528	0.9462	0.9456	0.9550	0.9574	0.9734	0.9766	0.9824	0.9846	0.9890
len	0.0801	1.4052	1.4091	1.1832	1.9722	1.9722	1.9858	2.9317	2.9317	3.0545
vs,0	0.9596	0.9454	0.9482	0.9972	0.9952	0.9836	0.9958	0.9868	0.9876	0.9914
len	1.0703	1.4224	1.4229	1.0196	2.1393	2.1393	2.3842	2.9600	2.9600	3.0654
reg,0.5	0.9546	0.9526	0.9498	0.9520	0.9550	0.9722	0.9748	0.9798	0.9886	0.9904
len	1.0801	3.3424	3.3411	3.2463	1.9729	1.9729	1.9866	2.9184	2.9184	2.9830
vs,0.5	0.9550	0.9366	0.9502	0.9970	0.9960	0.9462	0.9968	0.9904	0.9874	0.9964
len	1.0691	3.0725	3.0700	2.9615	2.1001	2.1001	2.3408	2.9456	2.9456	3.1867
reg,0.9	0.9624	0.9520	0.9504	0.9500	0.9506	0.9668	0.9708	0.9756	0.9908	0.9932
len	1.0779	22.0623	22.0812	22.0166	1.9714	1.9714	1.9845	2.9152	2.9152	2.9426
vs,0.9	0.9512	0.9602	0.9722	0.9966	0.9952	0.9778	0.9972	0.9942	0.9972	0.9982
len	1.0725	19.7097	19.6565	19.6814	2.1096	2.1096	2.3224	3.1001	3.1001	3.3159

Table 4.6. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 2$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9520	0.9548	0.9518	0.9520	0.9460	0.9484	0.9496	0.9484	0.9492	0.9518
len	0.2279	0.2418	0.2407	0.1949	1.9585	1.9585	1.9641	2.8293	2.8293	2.8450
vs,0	0.9598	0.9532	0.9572	0.9988	0.9972	0.9398	0.9972	0.9530	0.9532	0.9546
len	0.2278	0.2414	0.2413	0.1654	2.1475	2.1475	2.4037	2.8298	2.8298	2.8481
reg,0.5	0.9528	0.9570	0.9546	0.9610	0.9528	0.9554	0.9560	0.9476	0.9502	0.9522
len	0.2277	0.5557	0.5552	0.5359	1.9587	1.9587	1.9637	2.8287	2.8287	2.8393
vs,0.5	0.9526	0.9598	0.9628	0.9990	0.9982	0.9456	0.9982	0.9624	0.9646	0.9714
len	0.2280	0.5530	0.5525	0.4552	2.1516	2.1516	2.4052	2.8538	2.8538	2.9513
reg,0.9	0.9510	0.9540	0.9454	0.9496	0.9416	0.9432	0.9436	0.9420	0.9442	0.9468
len	0.2282	3.6478	3.6575	3.6530	1.9566	1.9566	1.9622	2.8283	2.8283	2.8358
vs,0.9	0.9574	0.9696	0.9642	0.9992	0.9986	0.9730	0.9994	0.9844	0.9766	0.9878
len	0.2284	3.1482	3.1488	3.1718	2.0276	2.0276	2.3289	2.8857	2.8857	3.1278

Table 4.7. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 2$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9576	0.9536	0.9546	0.9548	0.9498	0.9494	0.9502	0.9532	0.9526	0.9554
len	0.2815	0.2999	0.2993	0.2923	1.9584	1.9584	1.9645	2.8287	2.8287	2.8491
vs,0	0.9562	0.9522	0.9542	0.9986	0.9964	0.9438	0.9964	0.9460	0.9488	0.9516
len	0.2813	0.2990	0.2979	0.2489	2.1662	2.1662	2.4127	2.8298	2.8298	2.8585
reg,0.5	0.9532	0.9498	0.9532	0.9494	0.9432	0.9448	0.9448	0.9456	0.9454	0.9466
len	0.2820	0.8109	0.8085	0.8078	1.9582	1.9582	1.9636	2.8309	2.8309	2.8485
vs,0.5	0.9552	0.9578	0.9606	0.9994	0.9978	0.9454	0.9978	0.9780	0.9778	0.9834
len	0.2810	0.8100	0 0.8102	0.6865	2.1503	2.1503	7 2.4021	2.9142	2.9142	3.0555
reg,0.9	0.9586	0.9568	0.9526	0.9546	0.9484	0.9492	0.9496	0.9508	0.9526	0.9546
len	0.2813	5.4518	5.4773	5.4653	1.9594	1.9594	1.9635	2.8288	2.8288	2.8459
vs,0.9	0.9528	0.9528	0.9626	0.9982	0.9970	0.9490	0.9984	0.9862	0.9838	0.9900
len	0.2808	4.6804	4.6778	4.6992	2.0050	2.0050	2.1636	2.9844	2.9844	3.2045

Table 4.8. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 100$, $p = 4$, $k = 2$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9532	0.9544	0.9542	0.9520	0.9444	0.9466	0.9462	0.9434	0.9434	0.9442
len	0.0342	0.0130	0.0131	0.0116	1.9563	1.9563	1.9625	2.8219	2.8219	2.8287
vs,0	0.9600	0.9580	0.9532	0.9992	0.9982	0.9422	0.9982	0.9506	0.9520	0.9574
len	0.0342	0.0131	0.0131	0.0098	2.1577	2.1577	2.4107	2.8275	2.8275	2.8647
reg,0.5	0.9568	0.9538	0.9544	0.9528	0.9466	0.9468	0.9464	0.9430	0.9418	0.9440
len	0.0343	0.0327	0.0325	0.0322	1.9589	1.9589	1.9633	2.8260	2.8260	2.8329
vs,0.5	0.9516	0.9632	0.9614	0.9988	0.9978	0.9360	0.9978	0.9692	0.9676	0.9736
len	0.0342	0.0323	0.0324	0.0270	2.1439	2.1439	2.3982	2.9083	2.9083	3.0542
reg,0.9	0.9502	0.9598	0.9570	0.9586	0.9478	0.9478	0.9484	0.9532	0.9550	0.9556
len	0.0343	0.2182	0.2178	0.2179	1.9586	1.9586	1.9639	2.8251	2.8251	2.8326
vs,0.9	0.9554	0.9622	0.9628	0.9990	0.9978	0.9404	0.9978	0.9712	0.9708	0.9750
len	0.0342	0.2170	0.2173	0.1840	2.1391	2.1391	2.3936	2.9281	2.9281	3.0819

Table 4.9. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9526	0.9454	0.9508	0.9498	0.9532	0.9634	0.9642	0.9604	0.9662	0.9706
len	0.7110	0.7217	0.5414	0.5417	2.4857	2.4857	2.4954	2.4780	2.4780	2.5480
vs,0	0.9556	0.9504	0.9978	0.9978	0.9938	0.9874	0.9946	0.9592	0.9654	0.9712
len	0.7058	0.7147	0.4588	0.4626	2.7269	2.7269	3.0337	2.4795	2.4795	2.5532
reg,0.5	0.9480	0.9510	0.9510	0.9532	0.9534	0.9622	0.9650	0.9512	0.9592	0.9626
len	0.7098	1.1945	1.0987	1.1000	2.4859	2.4859	2.4948	2.4799	2.4799	2.5114
vs,0.5	0.9550	0.9582	0.9980	0.9978	0.9930	0.9898	0.9938	0.9688	0.9776	0.9814
len	0.7058	1.2134	0.9581	0.9573	2.7131	2.7131	3.0034	2.5085	2.5085	2.6233
reg,0.9	0.9462	0.9536	0.9522	0.9496	0.9548	0.9642	0.9658	0.9496	0.9610	0.9626
len	0.7546	6.0844	6.0691	6.0800	2.4888	2.4888	2.4990	2.4860	2.4860	2.4967
vs,0.9	0.9540	0.9544	0.9968	0.9958	0.9918	0.9906	0.9956	0.9784	0.9792	0.9838
len	0.7087	6.2974	6.3685	6.3456	2.7779	2.7779	3.0209	2.5907	2.5907	2.7273

Table 4.10. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9526	0.9480	0.9522	0.9578	0.9466	0.9460	0.9484	0.9402	0.9410	0.9418
len	0.1594	0.1381	0.0960	0.0960	2.4603	2.4603	2.4677	2.4591	2.4591	2.4685
vs,0	0.9520	0.9522	0.9980	0.9982	0.9934	0.9816	0.9936	0.9478	0.9494	0.9514
len	0.1595	0.1382	0.0813	0.0813	2.7241	2.7241	3.0251	2.4629	2.4629	2.4762
reg,0.5	0.9550	0.9526	0.9518	0.9526	0.9464	0.9494	0.9508	0.9506	0.9514	0.9524
len	0.1595	0.2189	0.1955	0.1953	2.4617	2.4617	2.4680	2.4616	2.4616	2.4698
vs,0.5	0.9514	0.9728	0.9988	0.9986	0.9958	0.9882	0.9966	0.9586	0.9594	0.9654
len	0.1594	0.2185	0.1666	0.1668	2.6973	2.6973	2.9879	2.4663	2.4663	2.5651
reg,0.9	0.9522	0.9530	0.9574	0.9450	0.9416	0.9418	0.9434	0.9410	0.9426	0.9442
len	0.1594	1.2759	1.2715	1.2707	2.4611	2.4611	2.4682	2.4602	2.4602	2.4664
vs,0.9	0.9550	0.9786	0.9994	0.9984	0.9938	0.9732	0.9942	0.9644	0.9518	0.9662
len	0.1596	1.2756	1.1239	1.1244	2.7313	2.7313	3.0506	2.5057	2.5057	2.6903

Table 4.11. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9544	0.9542	0.9542	0.9548	0.9480	0.9488	0.9498	0.9446	0.9470	0.9490
len	0.1966	0.1458	0.1418	0.1412	2.4614	2.4614	2.4669	2.4626	2.4626	2.4747
vs,0	0.9518	0.9514	0.9992	0.9992	0.9958	0.9852	0.9962	0.9464	0.9498	0.9510
len	0.1962	0.1452	0.1198	0.1201	2.7231	2.7231	3.0289	2.4600	2.4600	2.4801
reg,0.5	0.9496	0.9518	0.9530	0.9564	0.9516	0.9510	0.9514	0.9462	0.9482	0.9480
len	0.1968	0.2895	0.2877	0.2870	2.4594	2.4594	2.4655	2.4606	2.4606	2.4719
vs,0.5	0.9568	0.9736	0.9970	0.9992	0.9918	0.9868	0.9938	0.9654	0.9646	0.9712
len	0.1961	0.2883	0.2451	0.2461	2.6893	2.6893	2.9774	2.4722	2.4722	2.6031
reg,0.9	0.9514	0.9538	0.9514	0.9578	0.9454	0.9452	0.9472	0.9464	0.9474	0.9484
len	0.1965	1.8700	1.8667	1.8722	2.4610	2.4610	2.4669	2.4650	2.4650	2.4751
vs,0.9	0.9518	0.9866	0.9988	0.9990	0.9924	0.9488	0.9968	0.9542	0.9332	0.9584
len	0.1964	1.7425	1.6321	1.6288	2.5882	2.5882	2.8297	2.4433	2.4433	2.5754

Table 4.12. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9572	0.9546	0.9584	0.9542	0.9466	0.9480	0.9486	0.9452	0.9478	0.9496
len	0.0230	0.0063	0.0052	0.0052	2.4599	2.4599	2.4673	2.4602	2.4602	2.4666
vs,0	0.9524	0.9606	0.9982	0.9988	0.9928	0.9818	0.9936	0.9518	0.9510	0.9534
len	0.0230	0.0063	0.0044	0.0044	2.7161	2.7161	3.0174	2.4621	2.4621	2.5043
reg,0.5	0.9542	0.9532	0.9550	0.9560	0.9472	0.9468	0.9490	0.9470	0.9482	0.9496
len	0.0230	0.0112	0.0107	0.0107	2.4594	2.4594	2.4675	2.4614	2.4614	2.4681
vs,0.5	0.9532	0.9724	0.9986	0.9994	0.9942	0.9848	0.9954	0.9670	0.9656	0.9720
len	0.0229	0.0111	0.0091	0.0091	2.6967	2.6967	2.9861	2.4802	2.4802	2.6212
reg,0.9	0.9576	0.9518	0.9576	0.9550	0.9516	0.9510	0.9518	0.9478	0.9484	0.9488
len	0.0230	0.0700	0.0696	0.0701	2.4580	2.4580	2.4641	2.4597	2.4597	2.4656
vs,0.9	0.9490	0.9870	0.9992	0.9992	0.9924	0.9856	0.9944	0.9612	0.9616	0.9694
len	0.0229	0.0676	0.0599	0.0599	2.6908	2.6908	2.9768	2.4790	2.4790	2.6338

Table 4.13. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 2$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9550	0.9502	0.9510	0.9568	0.9534	0.9592	0.9612	0.9676	0.9698	0.9756
len	0.7533	0.8175	0.8187	0.6491	1.9651	1.9651	1.9723	2.8795	2.8795	2.9637
vs,0	0.9582	0.9544	0.9480	0.9984	0.9958	0.9652	0.9962	0.9688	0.9720	0.9788
len	0.7514	0.8157	0.8165	0.5546	2.1450	2.1450	2.3906	2.8799	2.8799	2.96651
reg,0.5	0.9484	0.9496	0.9526	0.9520	0.9508	0.9594	0.9612	0.9628	0.9732	0.9760
len	0.7546	1.8565	1.8557	1.7860	1.9671	1.9671	1.9746	2.8724	2.8724	2.9137
vs,0.5	0.9488	0.9584	0.9574	0.9966	0.9944	0.9434	0.9946	0.9570	0.9324	0.9628
len	0.7541	1.8730	1.8720	1.6319	2.0983	2.0983	2.3343	2.8828	2.8828	3.0225
reg,0.9	0.9590v	0.9578	0.9520	0.9546	0.9524	0.9612	0.9618	0.9658	0.9740	0.9770
len	0.7550	12.1336	12.1544	12.1385	1.9661	1.9661	1.9741	2.8688	2.8688	2.8833
vs,0.9	0.9518	0.9550	0.9646	0.9984	0.9964	0.9698	0.9972	0.9896	0.9922	0.9950
len	0.7513	10.6729	10.6901	10.6865	2.1444	2.1444	2.3519	3.1163	3.1163	3.3271

Table 4.14. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 2$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9530	0.9516	0.9516	0.9536	0.9452	0.9474	0.9480	0.9478	0.9464	0.9476
len	0.1596	0.1687	0.1684	0.1359	1.9571	1.9571	1.9626	2.8255	2.8255	2.8363
vs,0	0.9588	0.9526	0.9474	0.9988	0.9976	0.9448	0.9976	0.9440	0.9438	0.9460
len	0.1596	0.1685	0.1685	0.1154	2.1599	2.1599	2.4080	2.8261	2.8261	2.8394
reg,0.5	0.9568	0.9554	0.9494	0.9572	0.9522	0.9534	0.9530	0.9478	0.9490	0.9502
len	0.1596	0.3880	0.3867	0.3747	1.9602	1.9602	1.9647	2.8265	2.8265	2.8355
vs,0.5	0.9504	0.9526	0.9580	0.9992	0.9988	0.9436	0.9988	0.9598	0.9604	0.9652
len	0.1593	0.3851	0.3844	0.3155	2.1499	2.1499	2.4042	2.8502	2.8502	2.9448
reg,0.9	0.9554	0.9494	0.9456	0.9518	0.9466	0.9478	0.9474	0.9424	0.9418	0.9424
len	0.1595	2.5521	2.5498	2.5461	1.9602	1.9602	1.9652	2.8263	2.8263	2.8334
vs,0.9	0.9576	0.9596	0.9642	0.9978	0.9960	0.9412	0.9962	0.9732	0.9656	0.9808
len	0.1597	2.3406	2.3407	2.3246	2.1457	2.1457	2.3709	2.9504	2.9504	3.2276

Table 4.15. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 2$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9592	0.9552	0.9552	0.9484	0.9420	0.9418	0.9424	0.9546	0.9554	0.9560
len	0.1966	0.2032	0.2033	0.2001	1.9570	1.9570	1.9619	2.8262	2.8262	2.8392
vs,0	0.9606	0.9524	0.9550	0.9980	0.9970	0.9452	0.9974	0.9484	0.9486	0.9514
len	0.1966	0.2030	0.2030	0.1691	2.1485	2.1485	2.3977	2.8268	2.8268	2.8445
reg,0.5	0.9478	0.9522	0.9544	0.9530	0.9432	0.9444	0.9450	0.9476	0.9474	0.9482
len	0.1963	0.5510	0.5529	0.5494	1.9569	1.9569	1.9625	2.8269	2.8269	2.8386
vs,0.5	0.9556	0.9616	0.9648	0.9988	0.9974	0.9466	0.9974	0.9734	0.9742	0.9794
len	0.1961	0.5474	0.5476	0.4646	2.1442	2.1442	2.3935	2.8957	2.8957	3.0335
reg,0.9	0.9524	0.9550	0.9566	0.9548	0.9462	0.9448	0.9470	0.9508	0.9500	0.9506
len	0.1964	3.7523	3.7367	3.7368	1.9571	1.9571	1.9614	2.8245	2.8245	2.8360
vs,0.9	0.9552	0.9564	0.9526	0.9976	0.9962	0.9748	0.9978	0.9824	0.9724	0.9846
len	0.1962	3.1885	3.2044	3.2188	2.0110	2.0110	2.3122	2.8862	2.8862	3.1324

Table 4.16. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 200$, $n = 200$, $p = 4$, $k = 2$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9544	0.9590	0.9524	0.9558	0.9508	0.9516	0.9516	0.948	0.9486	0.9508
len	0.0229	0.0082	0.0082	0.0074	1.9611	1.9611	1.9646	2.8226	2.8226	2.8297
vs,0	0.9576	0.9568	0.9524	0.9992	0.9978	0.9358	0.9976	0.9488	0.9508	0.9542
len	0.0230	0.0082	0.0082	0.0063	2.1530	2.1530	2.4066	2.8250	2.8250	2.8531
reg,0.5	0.9528	0.9538	0.9514	0.9542	0.9452	0.9472	0.9478	0.9424	0.9432	0.9442
len	0.0229	0.0207	0.0208	0.02047	1.9572	1.9572	1.9630	2.8229	2.8229	2.8311
vs,0.5	0.9548	0.9602	0.9608	0.9984	0.9962	0.9394	0.9962	0.9700	0.9694	0.9778
len	0.0229	0.0206	0.02071	0.0173	2.1425	2.1425	2.3992	2.9092	2.9092	3.0562
reg,0.9	0.9546	0.9532	0.9536	0.9556	0.9484	0.9468	0.9484	0.9406	0.9408	0.9430
len	0.0229	0.1390	0.1394	0.1390	1.9575	1.9575	1.9627	2.8246	2.8246	2.8327
vs,0.9	0.9530	0.9600	0.9654	0.9976	0.9962	0.9420	0.9960	0.9676	0.9678	0.9742
len	0.0229	0.1376	0.1382	0.1179	2.1511	2.1511	2.4015	2.9289	2.9289	3.0843

Table 4.17. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9540	0.9264	0.9482	0.9506	0.9776	0.9904	0.9924	0.9658	0.9764	0.9854
len	0.9264	0.8574	0.5649	0.5656	3.2072	3.2072	3.2326	2.4899	2.4899	2.6620
vs,0	0.9580	0.9326	0.9986	0.9974	0.9980	0.9988	0.9998	0.9704	0.9784	0.9856
len	0.9126	0.8471	0.4767	0.4766	3.4687	3.4687	3.8313	2.4888	2.4888	2.6689
reg,0.41	0.9506	0.9404	0.9560	0.9562	0.9826	0.9938	0.9956	0.9634	0.9792	0.9846
len	0.9262	1.2895	1.1160	1.1175	3.2072	3.2072	3.2316	2.4923	2.4923	2.5785
vs,0.41	0.9560	0.9572	0.9982	0.9980	0.9974	0.9966	0.9992	0.9686	0.9754	0.9836
len	0.8600	1.4696	1.1488	1.1476	3.4445	3.4445	3.7908	2.5125	2.5125	2.6286
reg,0.9	0.9554	0.9522	0.9554	0.9518	0.9830	0.9932	0.9952	0.9598	0.9770	0.9790
len	0.9265	10.4449	10.4067	10.4078	3.2095	3.2095	3.2332	2.4960	2.4960	2.5118
vs,0.9	0.9488	0.9676	0.9978	0.9980	0.9966	0.9980	0.9992	0.9740	0.9830	0.9860
len	0.8568	10.9342	10.8993	10.9151	3.4163	3.4163	3.7521	2.5673	2.5673	2.7227

Table 4.18. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9504	0.9510	0.9518	0.9548	0.9494	0.9500	0.9504	0.9462	0.9474	0.9486
len	0.1862	0.1608	0.1123	0.1124	3.1118	3.1118	3.1165	2.4587	2.4587	2.4706
vs,0	0.9494	0.9558	0.9992	0.9986	0.9946	0.9938	0.9962	0.9424	0.9452	0.9468
len	0.1864	0.1609	0.0952	0.0955	3.4309	3.4309	3.7933	2.4609	2.4609	2.4817
reg,0.41	0.9592	0.9568	0.9578	0.9556	0.9484	0.9516	0.9514	0.9534	0.9538	0.9556
len	0.1858	0.2494	0.2219	0.2216	3.1144	3.1144	3.1200	2.4580	2.4580	2.4656
vs,0.41	0.9546	0.9630	0.9984	0.9988	0.9944	0.9940	0.9976	0.9550	0.9554	0.9620
len	0.1864	0.2512	0.1917	0.1920	3.3907	3.3907	3.7320	2.4606	2.4606	2.5381
reg,0.9	0.9560	0.9556	0.9548	0.9560	0.9518	0.9536	0.9540	0.9468	0.9474	0.9478
len	0.1857	2.0687	2.0704	2.0652	3.1136	3.1136	3.1180	2.4569	2.4569	2.4613
vs,0.9	0.9576	0.9820	0.9992	0.9984	0.9952	0.9904	0.9968	0.9564	0.9380	0.9636
len	0.1867	1.9647	1.8311	1.8272	3.3219	3.3219	3.6954	2.4403	2.4403	2.5604

Table 4.19. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9512	0.9558	0.9498	0.9610	0.9466	0.9506	0.9516	0.9474	0.9500	0.9530
len	0.2298	0.1751	0.1679	0.1681	3.1242	3.1242	3.1333	2.4615	2.4615	2.4824
vs,0	0.9530	0.9568	0.9992	0.9992	0.9956	0.9930	0.9972	0.9526	0.9544	0.9592
len	0.2288	0.1743	0.1399	0.1394	3.4077	3.4077	3.7621	2.4569	2.4569	2.4962
reg,0.41	0.9520	0.9532	0.9588	0.9554	0.9464	0.9476	0.9488	0.9486	0.9502	0.9534
len	0.2303	0.3346	0.3312	0.3308	3.1283	3.1283	3.1364	2.4605	2.4605	2.4784
vs,0.41	0.9588	0.9662	0.9990	0.9986	0.9952	0.9930	0.9972	0.9612	0.9644	0.9684
len	0.2288	0.3343	0.2816	0.2816	3.3623	3.3623	3.7051	2.4544	2.4544	2.5632
reg,0.9	0.9492	0.9566	0.9586	0.9550	0.9452	0.9474	0.9476	0.9494	0.9520	0.9536
len	0.2292	3.0850	3.0850	3.0860	3.1127	3.1127	3.1189	2.4554	2.4554	2.4708
vs,0.9	0.9486	0.9654	0.9986	0.9992	0.9954	0.9926	0.9970	0.9632	0.9608	0.9724
len	0.2290	2.7236	2.6749	2.6904	3.3877	3.3877	3.7341	2.4771	2.4771	2.6213

Table 4.20. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9530	0.9536	0.9526	0.9546	0.9438	0.9482	0.9440	0.9470	0.9494	0.9492
len	0.0276	0.0081	0.0066	0.0066	3.1103	3.1103	3.1158	2.4572	2.4572	2.4616
vs,0	0.9618	0.9620	0.9988	0.9992	0.9954	0.9938	0.9982	0.9604	0.9652	0.9696
len	0.0276	0.0081	0.0055	0.0055	3.38526	3.3852	3.7365	2.4605	2.4605	2.5303
reg,0.41	0.9568	0.9524	0.9562	0.9556	0.9442	0.9450	0.9472	0.9424	0.9446	0.9446
len	0.0277	0.0138	0.0130	0.0130	3.1072	3.1072	3.1123	2.4559	2.4559	2.4596
vs,0.41	0.9594	0.9742	0.9986	0.9986	0.9930	0.9924	0.9962	0.9660	0.9638	0.9708
len	0.0275	0.0138	0.0111	0.0111	3.3640	3.3640	3.6992	2.4572	2.4572	2.5797
reg,0.9	0.9562	0.9540	0.9504	0.9556	0.9452	0.9462	0.9462	0.9468	0.9474	0.9472
len	0.0277	0.1215	0.1218	0.1217	3.1077	3.1077	3.1137	2.4550	2.4550	2.4600
vs,0.9	0.9542	0.9880	0.9988	0.9984	0.9932	0.9932	0.9964	0.9592	0.9584	0.9682
len	0.0276	0.1211	0.1050	0.1046	3.3525	3.3525	3.6891	2.4541	2.4541	2.5818

Table 4.21. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 4$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9470	0.9484	0.9488	0.9516	0.9510	0.9648	0.9676	0.9814	0.9882	0.9914
len	0.8644	1.4555	1.4556	1.3230	1.9700	1.9700	1.9800	3.4976	3.4976	3.6072
vs,0	0.9532	0.9338	0.9378	0.9986	0.9972	0.9826	0.9976	0.9890	0.9858	0.9914
len	0.9212	1.3157	1.3146	0.9615	2.1301	2.1301	2.3782	3.5566	3.5566	3.6724
reg,0.41	0.9512	0.9564	0.9504	0.9542	0.9538	0.9666	0.9698	0.9816	0.9902	0.9910
len	0.8686	4.8695	4.8565	4.8204	1.9710	1.9710	1.9821	3.5154	3.5154	3.5794
vs,0.41	0.9542	0.9296	0.9272	0.9976	0.9964	0.9558	0.9974	0.9962	0.9952	0.9990
len	0.9148	3.7078	3.7142	3.6054	2.0933	2.0933	2.3156	3.5527	3.5527	3.8690
reg,0.9	0.9524	0.9518	0.9528	0.9562	0.9548	0.9674	0.9712	0.9840	0.9938	0.9950
len	0.8697	48.9265	48.9738	48.9582	1.9694	1.9694	1.9807	3.5011	3.5011	3.5312
vs,0.9	0.9574	0.9817	0.9825	0.9970	0.9950	0.9772	0.9968	0.9974	0.9990	0.9998
len	0.9112	36.7825	36.7189	36.8016	2.0896	2.0896	2.2948	3.6755	3.6755	3.9792

Table 4.22. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 4$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9512	0.9492	0.9514	0.9532	0.9494	0.9502	0.9498	0.9444	0.9456	0.9476
len	0.1869	0.2540	0.2532	0.2258	1.9575	1.9575	1.9626	3.3939	3.3939	3.4100
vs,0	0.9530	0.9544	0.9548	0.9990	0.9970	0.9448	0.9970	0.9526	0.9540	0.9562
len	0.1856	0.2524	0.2521	0.1877	2.1410	2.1410	2.3938	3.3738	3.3738	3.3891
reg,0.41	0.9568	0.9518	0.9536	0.9530	0.9482	0.9498	0.9506	0.9494	0.9542	0.9540
len	0.1866	0.8259	0.8266	0.8182	1.9591	1.9591	1.9638	3.3948	3.3948	3.4067
vs,0.41	0.9530	0.9562	0.9562	0.9990	0.9978	0.9422	0.9978	0.9750	0.9772	0.9812
len	0.1858	0.8304	0.8270	0.6854	2.1371	2.1371	2.3895	3.4177	3.4177	3.4890
reg,0.9	0.9528	0.9506	0.9542	0.9570	0.9536	0.9546	0.9556	0.9488	0.9510	0.9506
len	0.1862	8.2910	8.3040	8.3091	1.9589	1.9589	1.9645	3.3924	3.3924	3.4012
vs,0.9	0.9538	0.9860	0.9856	0.9990	0.9982	0.9296	0.9988	0.9944	0.9930	0.9972
len	0.1858	7.1675	7.1389	7.1568	2.0720	2.0720	2.2870	3.5155	3.5155	3.8719

Table 4.23. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 4$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9540	0.9514	0.9544	0.9584	0.9520	0.9538	0.9532	0.9530	0.9528	0.9556
len	0.2288	0.3385	0.3381	0.3346	1.9578	1.9578	1.9609	3.3738	3.3738	3.3935
vs,0	0.9592	0.9518	0.9562	0.9992	0.9984	0.9412	0.9984	0.9490	0.9480	0.9534
len	0.2291	0.3370	0.3383	0.2812	2.1509	2.1509	2.4048	3.3747	3.3747	3.4016
reg,0.41	0.9516	0.9568	0.9530	0.9510	0.9458	0.9454	0.9470	0.9528	0.9520	0.9520
len	0.2290	1.2119	1.2168	1.2106	1.9580	1.9580	1.9613	3.3748	3.3748	3.3909
vs,0.41	0.9520	0.9474	0.9460	0.9990	0.9974	0.9454	0.9976	0.9614	0.9278	0.9526
len	0.2292	1.2661	1.2687	1.0484	2.1332	2.1332	2.3881	3.4838	3.4838	3.6125
reg,0.9	0.9496	0.9586	0.9566	0.9572	0.9516	0.9518	0.9524	0.9518	0.9518	0.9536
len	0.2290	12.3085	12.3474	12.3360	1.9595	1.9595	1.9630	3.3743	3.3743	3.3901
vs,0.9	0.9532	0.9836	0.9838	0.9964	0.9956	0.9668	0.9958	0.9918	0.9916	0.9960
len	0.2289	10.6485	10.6339	10.6435	2.1090	2.1090	2.3021	3.6363	3.6363	3.9557

Table 4.24. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 150$, $p = 6$, $k = 4$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9552	0.9542	0.9570	0.9606	0.9528	0.9526	0.9536	0.9474	0.9460	0.9466
len	0.0276	0.0140	0.0140	0.0132	1.9599	1.9599	1.9628	3.3669	3.3669	3.3725
vs,0	0.9476	0.9534	0.9562	0.9984	0.9978	0.9430	0.9978	0.9520	0.9522	0.9558
len	0.0277	0.0141	0.0140	0.0110	2.1500	2.1500	2.4111	3.3704	3.3704	3.4005
reg,0.41	0.9566	0.9576	0.9506	0.9576	0.9530	0.9510	0.9520	0.9500	0.9502	0.9508
len	0.0276	0.0480	0.0482	0.0481	1.9580	1.9580	1.9619	3.3664	3.3664	3.3721
vs,0.41	0.9542	0.9582	0.9562	0.9968	0.9952	0.9392	0.9952	0.9652	0.9660	0.9716
len	0.0276	0.0480	0.0481	0.0402	2.1441	2.1441	2.3959	3.4265	3.4265	3.5317
reg,0.9	0.9582	0.9536	0.9520	0.9522	0.9448	0.9434	0.9440	0.9484	0.9490	0.9490
len	0.0276	0.4893	0.4882	0.4900	1.9572	1.9572	1.9612	3.3659	3.3659	3.3719
vs,0.9	0.9582	0.9618	0.9550	0.9984	0.9976	0.9426	0.9976	0.9668	0.9692	0.9758
len	0.0276	0.4888	0.4872	0.4063	2.1513	2.1513	2.4090	3.4415	3.4415	3.5579

Table 4.25. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9530	0.9572	0.9512	0.9600	0.9636	0.9718	0.9736	0.9536	0.9634	0.9704
len	0.6100	0.5432	0.3688	0.3688	3.1557	3.1557	3.1655	2.4717	2.4717	2.5572
vs,0	0.9592	0.9486	0.9986	0.9990	0.9978	0.9978	0.9990	0.9582	0.9664	0.9718
len	0.6057	0.5384	0.3089	0.3096	3.4300	3.4300	3.7907	2.4694	2.4694	2.5617
reg,0.41	0.9574	0.9472	0.9530	0.9580	0.9670	0.9780	0.9784	0.9534	0.9592	0.9636
len	0.6100	0.8309	0.7280	0.7293	3.1521	3.1521	3.1624	2.4690	2.4690	2.5103
vs,0.41	0.9460	0.9508	0.9978	0.9992	0.9954	0.9956	0.9976	0.9574	0.9696	0.9764
len	0.6069	1.0282	0.8040	0.8074	3.3929	3.3929	3.7323	2.4779	2.4779	2.5765
reg,0.9	0.9528	0.9528	0.9512	0.9498	0.9620	0.9738	0.9736	0.9522	0.9606	0.9626
len	0.6106	6.8273	6.7941	6.7979	3.1518	3.1518	3.1616	2.4745	2.4745	2.4813
vs,0.9	0.9552	0.9618	0.9990	0.9986	0.9970	0.9976	0.9992	0.9722	0.9776	0.9832
len	0.6049	5.9733	5.9861	5.9837	3.4109	3.4109	3.7347	2.5506	2.5506	2.7025

Table 4.26. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9510	0.9564	0.9576	0.9572	0.9474	0.9502	0.9516	0.9436	0.9446	0.9464
len	0.1300	0.1123	0.0782	0.0781	3.1117	3.1117	3.1167	2.4560	2.4560	2.4646
vs,0	0.9530	0.9530	0.9986	0.9992	0.9962	0.9958	0.9978	0.9460	0.9482	0.9486
len	0.1296	0.1122	0.0653	0.0652	3.4066	3.4066	3.7662	2.4546	2.4546	2.4664
reg,0.41	0.9530	0.9552	0.9578	0.9546	0.9536	0.9558	0.9560	0.9482	0.9492	0.9500
len	0.1298	0.1740	0.1542	0.1542	3.1101	3.1101	3.1164	2.4566	2.4566	2.4630
vs,0.41	0.9584	0.9668	0.9988	0.9984	0.9934	0.9932	0.9958	0.9546	0.9578	0.9636
len	0.1297	0.2151	0.1699	0.1694	3.3614	3.3614	3.7018	2.4517	2.4517	2.5346
reg,0.9	0.9538	0.9554	0.9576	0.9530	0.9470	0.9482	0.9498	0.9458	0.9460	0.9476
len	0.1298	1.4429	1.4412	1.4425	3.1094	3.1094	3.1148	2.4570	2.4570	2.4607
vs,0.9	0.9538	0.9798	0.9986	0.9990	0.9926	0.9818	0.9950	0.9520	0.928	0.9542
len	0.1300	1.4516	1.2593	1.2556	3.3824	3.3824	3.7078	2.4671	2.4671	2.5955

Table 4.27. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9466	0.9578	0.9548	0.9520	0.9448	0.9478	0.9474	0.9478	0.9488	0.9508
len	0.1598	0.1177	0.1147	0.1149	3.1083	3.1083	3.1139	2.4562	2.4562	2.4673
vs,0	0.9548	0.9560	0.9988	0.9986	0.9950	0.9950	0.9976	0.9498	0.9512	0.9554
len	0.1594	0.1182	0.0962	0.0960	3.4071	3.4071	3.7627	2.4546	2.4546	2.4774
reg,0.41	0.9570	0.9490	0.9580	0.9554	0.9470	0.9484	0.9498	0.9450	0.9450	0.9466
len	0.1595	0.2279	0.2264	0.2266	3.1107	3.1107	3.1167	2.4576	2.4576	2.4653
vs,0.41	0.9562	0.9720	0.9984	0.9990	0.9924	0.9914	0.9964	0.9570	0.9596	0.9656
len	0.1599	0.2940	0.2497	0.2492	3.3606	3.3606	3.7009	2.4525	2.4525	2.5532
reg,0.9	0.9538	0.9546	0.9572	0.9550	0.9460	0.9478	0.9488	0.9474	0.9508	0.9510
len	0.1598	2.1137	2.1073	2.1136	3.1105	3.1105	3.1157	2.4564	2.4564	2.4652
vs,0.9	0.9542	0.9834	0.9990	0.9990	0.9952	0.9908	0.9966	0.9536	0.9408	0.9626
len	0.1598	1.9709	1.8476	1.8464	3.3061	3.3061	3.6769	2.4344	2.4344	2.5564

Table 4.28. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9510	0.9562	0.9528	0.9542	0.9558	0.9566	0.9470	0.9472	0.9478	0.9480
len	0.0185	0.0088	0.0083	0.0083	3.1105	3.1105	3.1152	2.4546	2.4546	2.4588
vs,0	0.9602	0.9618	0.9980	0.9986	0.9938	0.9926	0.9972	0.9520	0.9546	0.9588
len	0.0185	0.0050	0.0035	0.0035	3.3940	3.3940	3.7476	2.4585	2.4585	2.5071
reg,0.41	0.9536	0.9548	0.9528	0.9508	0.9478	0.9478	0.9484	0.9428	0.9430	0.9444
len	0.0185	0.0783	0.0786	0.0786	3.1094	3.1094	3.1151	2.4555	2.4555	2.4594
vs,0.41	0.9624	0.9746	0.9992	0.9982	0.9940	0.9918	0.9962	0.9620	0.9640	0.9700
len	0.0185	0.0111	0.0092	0.0092	3.3574	3.3574	3.6960	2.4543	2.4543	2.5707
reg,0.9	0.9590	0.9532	0.9576	0.9570	0.9500	0.9508	0.9506	0.9434	0.9454	0.9462
len	0.0185	4 0.0089	0.0089	0.0085	1.9575	1.9575	1.9601	3.3685	3.3685	3.3735
vs,0.9	0.9536	0.9548	0.9528	0.9508	0.9478	0.9478	0.9484	0.9428	0.9430	0.9444
len	0.0185	0.0783	0.0786	0.0786	3.1094	3.1094	3.1151	2.4555	2.4555	2.4594

Table 4.29. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 4$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9530	0.9546	0.9470	0.9986	0.9978	0.9576	0.9976	0.9754	0.9816	0.9852
len	0.6088	0.8393	0.8406	0.6227	2.1361	2.1361	2.3858	3.4410	3.4410	3.5155
vs,0	0.9542	0.9496	0.9540	0.9986	0.9972	0.9606	0.9974	0.9768	0.9826	0.9870
len	0.6097	0.8388	0.8422	0.6238	2.1337	2.1337	2.3876	3.4410	3.4410	3.5153
reg,0.41	0.9586	0.9646	0.9532	0.9980	0.9962	0.9360	0.9960	0.9904	0.9836	0.9948
len	0.6061	3.2242	3.2302	3.1425	2.0995	2.0995	2.3283	3.4860	3.4860	3.8037
vs,0.41	0.9528	0.9502	0.9510	0.9992	0.9982	0.9448	0.9982	0.9766	0.9578	0.9836
len	0.6070	2.5137	2.5211	2.3552	2.1057	2.1057	2.3324	3.4487	3.4487	3.7257
reg,0.9	0.9580	0.9556	0.9514	0.9534	0.9488	0.9582	0.9588	0.9644	0.9750	0.9768
len	0.6094	27.2174	27.2096	27.1877	1.9646	1.9646	1.9694	3.4206	3.4206	3.4320
vs,0.9	0.9534	0.9812	0.9813	0.9982	0.9974	0.9634	0.9978	0.9960	0.9962	0.9976
len	0.6057	23.9571	24.0236	23.8686	2.1185	2.1185	2.3411	3.6534	3.6534	3.9492

Table 4.30. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 4$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9542	0.9572	0.9574	0.9546	0.9466	0.9488	0.9474	0.9516	0.9512	0.9520
len	0.1298	0.1761	0.1762	0.1561	1.9569	1.9569	1.9600	3.3687	3.3687	3.3779
vs,0	0.9556	0.9496	0.9538	0.9994	0.9984	0.9410	0.9984	0.9486	0.9508	0.9516
len	0.1298	0.1757	0.1763	0.1306	2.1456	2.1456	2.3969	3.3686	3.3686	3.3799
reg,0.41	0.9516	0.9524	0.9576	0.9542	0.9496	0.9490	0.9498	0.9426	0.9450	0.9454
len	0.1299	0.5719	0.5733	0.5677	1.9590	1.9590	1.9624	3.3684	3.3684	3.3762
vs,0.41	0.9584	0.9584	0.9558	0.9976	0.9970	0.9428	0.9970	0.9722	0.9744	0.9788
len	0.1300	0.7633	0.7623	0.6308	2.1384	2.1384	2.3931	3.4127	3.4127	3.4971
reg,0.9	0.9524	0.9572	0.9536	0.9600	0.9528	0.9522	0.9536	0.9524	0.9530	0.9544
len	0.1299	5.7585	5.7640	5.7576	1.9580	1.9580	1.9616	3.3692	3.3692	3.3747
vs,0.9	0.9544	0.9998	0.9534	0.9982	0.9972	0.9590	0.9972	0.9904	0.9858	0.9944
len	0.1298	5.0170	5.0331	5.0374	2.1186	2.1186	2.3609	3.5577	3.55771	3.8661

Table 4.31. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 4$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9578	0.9518	0.9578	0.9536	0.9472	0.9466	0.9474	0.9496	0.9490	0.9514
len	0.1595	0.2317	0.2313	0.2298	1.9582	1.9582	1.9613	3.3700	3.3699	3.3835
vs,0	0.9530	0.9566	0.9522	0.9972	0.9962	0.9350	0.9962	0.9484	0.9500	0.9500
len	0.1598	0.2316	0.2310	0.1918	2.1381	2.1381	2.3876	3.3702	3.3702	3.3864
reg,0.41	0.9572	0.9552	0.9564	0.9570	0.9510	0.9518	0.9520	0.9520	0.9512	0.9530
len	0.1599	0.8320	0.8333	0.8324	1.9588	1.9588	1.9619	3.3700	3.3700	3.3806
vs,0.41	0.9538	0.9478	0.9410	0.9990	0.9982	0.9430	0.9982	0.9676	0.9528	0.9646
len	0.1598	1.1583	1.1560	0.9543	2.1359	2.1359	2.3916	3.5019	3.5019	3.6216
reg,0.9	0.9532	0.9520	0.9544	0.9568	0.9508	0.9516	0.9514	0.9476	0.9492	0.9512
len	0.1598	8.4387	8.4381	8.4502	1.9582	1.9582	1.9611	3.3700	3.3700	3.3805
vs,0.9	0.9494	0.9570	0.9584	0.9988	0.9982	0.9668	0.9982	0.9926	0.9918	0.9976
len	0.1596	7.2945	7.3188	7.3289	2.0632	2.0632	2.2713	3.5251	3.5251	3.8799

Table 4.32. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 300$, $n = 300$, $p = 6$, $k = 4$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9608	0.9864	0.9984	0.9978	0.9930	0.9918	0.9964	0.9654	0.9646	0.9730
len	0.0185	0.0778	0.0670	0.0670	3.3601	3.3601	3.6977	2.4535	2.4535	2.5712
vs,0	0.9554	0.9576	0.9600	0.9986	0.9968	0.9416	0.9966	0.9490	0.9508	0.9522
len	0.0185	0.0089	0.0089	0.0071	2.1476	2.1476	2.3983	3.3674	3.3674	3.3904
reg,0.41	0.9530	0.9578	0.9570	0.9570	0.9506	0.9506	0.9506	0.9434	0.9442	0.9450
len	0.0185	0.0309	0.0310	0.0308	1.9579	1.9579	1.9616	3.3656	3.3656	3.3721
vs,0.41	0.9520	0.9554	0.9596	0.9982	0.9970	0.9340	0.9970	0.9642	0.9646	0.9732
len	0.0185	0.0309	0.0307	0.0257	2.1377	2.1377	2.3925	3.4262	3.4262	3.5327
reg,0.9	0.9524	0.9508	0.9502	0.9552	0.9480	0.9490	0.9494	0.9448	0.9454	0.9470
len	0.0185	0.3133	0.3137	0.3137	1.9575	1.9575	1.9606	3.3667	3.3667	3.3729
vs,0.9	0.9484	0.9580	0.9532	0.9988	0.9986	0.9366	0.9986	0.9644	0.9648	0.9712
len	0.0185	4 0.3120	0.3130	0.2617	2.1419	2.1419	2.3871	3.4405	3.4405	3.5523

Table 4.33. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9558	0.9542	0.9548	0.9502	0.9562	0.9620	0.9636	0.9868	0.9922	0.9946
len	0.8523	1.3001	1.3020	1.1608	1.9677	1.9677	1.9775	3.7474	3.7474	3.8776
vs,0	0.9524	0.9340	0.9980	0.9974	0.9978	0.9986	0.9998	0.9666	0.9758	0.9846
len	0.8402	0.7755	0.4319	0.4358	3.7332	3.7332	4.1184	2.4838	2.4838	2.6691
reg,0.38	0.9520	0.9392	0.9508	0.9494	0.9810	0.9920	0.9942	0.9594	0.9762	0.9806
len	0.8514	1.1670	1.0075	1.0086	3.4667	3.4667	3.4915	2.4842	2.4842	2.57230
vs,0.38	0.9494	0.9444	0.9976	0.9984	0.9988	0.9996	0.9998	0.9692	0.9820	0.9870
len	0.8405	1.1628	0.8656	0.8626	3.7007	3.7007	4.0688	2.4949	2.4949	2.6288
reg,0.9	0.9554	0.9476	0.9524	0.9498	0.9818	0.9936	0.9956	0.9548	0.9720	0.9742
len	0.8516	10.6768	10.6667	10.6663	3.4675	3.4675	3.4939	2.4867	2.4867	2.5009
vs,0.9	0.9512	0.9572	0.9972	0.9984	0.9986	0.9994	0.9996	0.9750	0.9840	0.9880
len	0.8400	9.3383	9.3450	9.3656	3.6867	3.6867	4.0459	2.5656	2.5656	2.7239

Table 4.34. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9540	0.9514	0.9508	0.9560	0.9490	0.9512	0.9514	0.9462	0.9498	0.9514
len	0.1721	0.1488	0.1038	0.1039	3.3650	3.3650	3.3696	2.4549	2.4549	2.4673
vs,0	0.9602	0.9582	0.9982	0.9980	0.9972	0.9968	0.9986	0.9568	0.9576	0.9608
len	0.1719	0.1484	0.0863	0.0866	3.6701	3.6701	4.0525	2.4573	2.4573	2.4770
reg,0.38	0.9510	0.9504	0.9546	0.9548	0.9462	0.9478	0.9494	0.9440	0.9458	0.9472
len	0.1722	0.2277	0.2013	0.2011	3.3641	3.3641	3.3690	2.4554	2.4554	2.4626
vs,0.38	0.9596	0.9600	0.9992	0.9982	0.9946	0.9940	0.9970	0.9570	0.9600	0.9658
len	0.1720	0.2285	0.1721	0.1712	3.6259	3.6259	3.9903	2.4555	2.4555	2.5251
reg,0.9	0.9534	0.9574	0.9614	0.9560	0.9518	0.9540	0.9550	0.9486	0.9504	0.9512
len	0.1723	2.1357	2.1306	2.1294	3.3647	3.3647	3.3690	2.4565	2.4565	2.4602
vs,0.9	0.9598	0.9784	0.9990	0.9994	0.9946	0.9918	0.9964	0.9574	0.9434	0.9646
len	0.1721	1.9938	1.8483	1.8479	3.5733	3.5733	3.9660	2.4362	2.4362	2.5520

Table 4.35. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9528	0.9538	0.9532	0.9550	0.9518	0.9528	0.9532	0.9462	0.9484	0.9496
len	0.2120	0.1609	0.1542	0.1549	3.3637	3.3637	3.3689	2.4570	2.4570	2.4749
vs,0	0.9578	0.9566	0.9982	0.9976	0.9958	0.9948	0.9982	0.9540	0.9552	0.9584
len	0.2116	0.1609	0.1301	0.1289	3.6660	3.6660	4.0453	2.4570	2.4570	2.4966
reg,0.38	0.9508	0.9554	0.9544	0.9548	0.9528	0.9534	0.9542	0.9494	0.9482	0.9506
len	0.2118	0.3031	0.3001	0.3001	3.3639	3.3639	3.3689	2.4552	2.4552	2.47152
vs,0.38	0.9558	0.9674	0.9984	0.9988	0.9942	0.9936	0.9964	0.9602	0.9632	0.9694
len	0.2116	0.3044	0.2558	0.2561	3.6266	3.6266	3.9861	2.4539	2.4539	2.5555
reg,0.9	0.9486	0.9576	0.9560	0.9574	0.9530	0.9546	0.9532	0.9486	0.9496	0.9508
len	0.2122	3.1733	3.1718	3.1640	3.3656	3.3656	3.3706	2.4563	2.4563	2.4717
vs,0.9	0.9584	0.9582	0.9982	0.9986	0.9942	0.9934	0.9976	0.9654	0.9646	0.9764
len	0.2116	2.7881	2.7533	2.7475	3.6399	3.6399	4.0048	2.4763	2.4763	2.6313

Table 4.36. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9600	0.9574	0.9484	0.9586	0.9480	0.9504	0.9512	0.9498	0.9504	0.9504
len	0.0255	0.0075	0.0060	0.0060	3.3606	3.3606	3.3658	2.4558	2.4558	2.4594
vs,0	0.9622	0.9682	0.9980	0.9992	0.9950	0.9948	0.9984	0.9652	0.9656	0.9688
len	0.0254	0.0074	0.0051	0.0051	3.6507	3.6507	4.0198	2.4608	2.4608	2.5314
reg,0.38	0.9526	0.9548	0.9544	0.9536	0.9468	0.9486	0.9500	0.9474	0.9492	0.9488
len	0.0255	0.0126	0.0118	0.0118	3.3601	3.3601	3.3650	2.4543	2.4543	2.4579
vs,0.38	0.9638	0.9676	0.9980	0.9984	0.9950	0.9944	0.9978	0.9644	0.9668	0.9738
len	0.0254	0.0126	0.0100	0.0101	3.6185	3.6185	3.9802	2.4562	2.4562	2.5759
reg,0.9	0.9586	0.9532	0.9616	0.9530	0.9438	0.9448	0.9462	0.9512	0.9522	0.9526
len	0.0255	0.1256	0.1252	0.1252	3.3609	3.3609	3.3658	2.4563	2.4563	2.4594
vs,0.9	0.9534	0.9862	0.9990	0.9982	0.9952	0.9940	0.9970	0.9626	0.9610	0.9708
len	0.0254	0.1249	0.1073	0.1082	3.6121	3.6121	3.9680	2.4533	2.4533	2.5722

Table 4.37. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 5$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9496	0.9566	0.9526	0.9588	0.9662	0.9744	0.9752	0.9486	0.9594	0.9668
len	0.5639	0.5009	0.3401	0.3402	3.4061	3.4061	3.4155	2.4661	2.4661	2.5513
vs,0	0.9552	0.9402	0.9392	0.9984	0.9978	0.9802	0.9980	0.9882	0.9820	0.9896
len	0.8511	1.3243	1.3248	0.9848	2.1131	2.1131	2.3650	3.7926	3.7926	3.9026
reg,0.38	0.9528	0.9528	0.9554	0.9542	0.9504	0.9678	0.9710	0.9868	0.9944	0.9958
len	0.8525	4.6109	4.6148	4.5735	1.9682	1.9682	1.9778	3.7318	3.7318	3.8070
vs,0.38	0.9566	0.9222	0.9220	0.9962	0.9956	0.9598	0.9960	0.9986	0.9972	0.9996
len	0.8413	4.0607	4.0649	3.9863	2.1001	2.1001	2.3152	3.8028	3.8028	4.1446
reg,0.9	0.9494	0.9554	0.9506	0.9552	0.9518	0.9672	0.9704	0.9844	0.9958	0.9968
len	0.8519	53.0498	53.2692	53.2381	1.9687	1.9687	1.9775	3.7214	3.7214	3.7486
vs,0.9	0.9578	0.9780	0.9758	0.9984	0.9970	0.9772	0.9982	0.9992	0.9996	0.9998
len	0.8397	46.6775	46.5947	46.6538	2.0855	2.0855	2.29183	3.9133	3.9133	4.2479

Table 4.38. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 5$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9568	0.9564	0.9548	0.9534	0.9470	0.9472	0.9482	0.9508	0.9500	0.9506
len	0.1722	0.2556	0.2561	0.2324	1.9596	1.9596	1.9628	3.5972	3.5972	3.6088
vs,0	0.9542	0.9566	0.9502	0.9990	0.9978	0.9426	0.9978	0.9476	0.9492	0.9514
len	0.1719	0.2556	0.2552	0.1936	2.1462	2.1462	2.3952	3.5963	3.5963	3.6112
reg,0.38	0.9556	0.9506	0.9542	0.9542	0.9496	0.9496	0.9506	0.9488	0.9514	0.9518
len	0.1720	0.9200	0.9198	0.9122	1.9590	1.9590	1.9621	3.5969	3.5969	3.6052
vs,0.38	0.9556	0.9450	0.9480	0.9984	0.9980	0.9474	0.9980	0.9820	0.9822	0.9832
len	0.1722	0.9353	0.9384	0.7647	2.1393	2.1393	2.3908	3.6790	3.6790	3.7424
reg,0.9	0.9500	0.9534	0.9588	0.9526	0.9480	0.9486	0.9492	0.9504	0.9520	0.9532
len	0.1720	10.6577	10.6325	10.6432	1.9591	1.9591	1.9620	3.5972	3.5972	3.6031
vs,0.9	0.9526	0.9626	0.9660	0.9984	0.9972	0.9408	0.9978	0.9954	0.9944	0.9976
len	0.1718	9.2020	9.1980	9.2021	2.0781	2.0781	2.2943	3.7755	3.7755	4.1574

Table 4.39. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 5$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9564	0.9558	0.9588	0.9600	0.9508	0.9502	0.9500	0.9528	0.9512	0.9550
len	0.2122	0.3489	0.3488	0.3448	1.9586	1.9586	1.9615	3.5985	3.5985	3.6182
vs,0	0.9532	0.9564	0.9544	0.9992	0.9990	0.9398	0.9990	0.9552	0.9534	0.9552
len	0.2120	0.3488	0.3491	0.2880	2.1571	2.1571	2.4064	3.5985	3.5985	3.6249
reg,0.38	0.9466	0.9566	0.9542	0.9554	0.9498	0.9500	0.9504	0.9476	0.9486	0.9508
len	0.2119	1.3585	1.3589	1.3589	1.9598	1.9598	1.9628	3.5979	3.5979	3.6149
vs,0.38	0.9498	0.9524	0.9566	0.9988	0.9982	0.9416	0.9982	0.9466	0.9130	0.9470
len	0.2126	1.4181	1.4101	1.1788	2.1298	2.1298	2.3735	3.6933	3.6933	3.8355
reg,0.9	0.9532	0.9550	0.9524	0.9538	0.9428	0.9452	0.9452	0.9486	0.9478	0.9494
len	0.2119	15.8262	15.8462	15.8152	1.9588	1.9588	1.9614	3.5975	3.5975	3.6127
vs,0.9	0.9536	0.9706	0.9786	0.9984	0.9964	0.9670	0.9970	0.9942	0.9946	0.9976
len	0.2116	13.6373	13.7437	13.7113	2.1272	2.1272	2.3442	3.8664	3.8664	4.2048

Table 4.40. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 175$, $p = 7$, $k = 5$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9608	0.9534	0.9548	0.9530	0.9486	0.9466	0.9464	0.9486	0.9494	0.9502
len	0.0255	0.0142	0.0143	0.0136	1.9587	1.9587	1.9612	3.5917	3.5917	3.5967
vs,0	0.9532	0.9536	0.9572	0.9982	0.9966	0.9354	0.9964	0.9516	0.9520	0.9552
len	0.0255	0.0143	0.0143	0.0113	2.1459	2.1459	2.3963	3.5950	3.5950	3.6232
reg,0.38	0.9570	0.9542	0.9576	0.9550	0.9478	0.9490	0.9494	0.9512	0.9516	0.9520
len	0.0255	0.0537	0.0538	0.0537	1.9592	1.9592	1.9616	3.5901	3.5901	3.5954
vs,0.38	0.9576	0.9536	0.9616	0.9986	0.9980	0.9416	0.9980	0.9678	0.9702	0.9742
len	0.0255	0.0541	0.0537	0.0447	2.1435	2.1435	2.3962	3.6403	3.6403	3.7364
reg,0.9	0.9566	0.9530	0.9540	0.9586	0.9506	0.9514	0.9522	0.9504	0.9498	0.9504
len	0.0255	0.6261	0.6256	0.62524	1.9571	1.9571	1.9605	3.5901	3.5901	3.5952
vs,0.9	0.9578	0.9582	0.9540	0.9984	0.9980	0.9374	0.9980	0.9668	0.9688	0.9746
len	0.0255	0.6237	0.6251	0.5232	2.1485	2.1485	2.3981	3.6618	3.6618	3.7634

Table 4.41. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9550	0.9548	0.9526	0.9502	0.9614	0.9712	0.9720	0.9460	0.9608	0.9680
len	0.5635	0.5020	0.3412	0.3411	3.4058	3.4058	3.4157	2.4654	2.4654	2.5532
vs,0	0.9540	0.9442	0.9984	0.9982	0.9964	0.9972	0.9992	0.9574	0.9602	0.9688
len	0.5580	0.4960	0.2851	0.2835	3.7010	3.7010	4.0794	2.4664	2.4664	2.5606
reg,0.38	0.9546	0.9436	0.9548	0.9516	0.9644	0.9756	0.9770	0.9506	0.9598	0.9654
len	0.5623	0.7555	0.6615	0.6605	3.4071	3.4071	3.4176	2.4669	2.4669	2.5080
vs,0.38	0.9572	0.9578	0.9988	0.9996	0.9968	0.9974	0.9988	0.9576	0.9670	0.9724
len	0.5600	0.7538	0.5625	0.5612	3.6524	3.6524	4.0211	2.4656	2.4656	2.5579
reg,0.9	0.9518	0.9550	0.9554	0.9530	0.9664	0.9756	0.9778	0.9534	0.9638	0.9652
len	0.5628	6.9968	6.9772	6.9815	3.4096	3.4096	3.4194	2.4689	2.4689	2.4750
vs,0.9	0.9524	0.9774	0.9992	0.9980	0.9968	0.9976	0.9986	0.9744	0.9724	0.9786
len	0.5596	6.1213	6.0762	6.1192	3.6604	3.6604	4.0204	2.5474	2.5474	2.7030

Table 4.42. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9506	0.9550	0.9520	0.9544	0.9430	0.9444	0.9460	0.9486	0.9472	0.9476
len	0.1202	0.1041	0.0722	0.0723	3.3630	3.3630	3.3681	2.4555	2.4555	2.4623
vs,0	0.9498	0.9590	0.9988	0.9982	0.9960	0.9942	0.9974	0.9538	0.9554	0.9566
len	0.1202	0.1038	0.0599	0.0600	3.6733	3.6733	4.0545	2.4549	2.4549	2.4663
reg,0.38	0.9556	0.9484	0.9528	0.9554	0.9490	0.9500	0.9496	0.9428	0.9448	0.9456
len	0.1202	0.1588	0.1405	0.1407	3.3602	3.3602	3.3650	2.4548	2.4548	2.4600
vs,0.38	0.9614	0.9660	0.9988	0.9992	0.9962	0.9958	0.9976	0.9548	0.9586	0.9634
len	0.1201	0.1599	0.1192	0.1189	3.6240	3.6240	3.9909	2.4532	2.4532	2.5160
reg,0.9	0.9548	0.9566	0.9584	0.9546	0.9564	0.9566	0.9566	0.9542	0.9552	0.9564
len	0.1203	1.4851	1.4849	1.4834	3.3638	3.3638	3.3686	2.4546	2.4546	2.4585
vs,0.9	0.9566	0.9780	0.9984	0.9992	0.9936	0.9894	0.9970	0.9496	0.9180	0.9520
len	0.1204	1.4902	1.2735	1.2796	3.6246	3.6246	3.9830	2.4581	2.4581	2.5706

Table 4.43. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9470	0.9522	0.9532	0.9544	0.9552	0.9548	0.9584	0.9490	0.9500	0.9514
len	0.1476	0.1089	0.1061	0.1060	3.3624	3.3624	3.3673	2.4561	2.4561	2.4668
vs,0	0.9516	0.9550	0.9988	0.9972	0.9950	0.9942	0.9982	0.9500	0.9524	0.9558
len	0.1479	0.1088	0.0884	0.0886	3.6687	3.6687	4.0474	2.4557	2.4557	2.4788
reg,0.38	0.9498	0.9514	0.9516	0.9602	0.9510	0.9524	0.9526	0.9452	0.9464	0.9482
len	0.1476	0.2071	0.2062	0.2058	3.3628	3.3628	3.3669	2.4557	2.4557	2.4649
vs,0.38	0.9550	0.9662	0.9990	0.9986	0.9948	0.9932	0.9968	0.9586	0.9560	0.9634
len	0.1476	0.2079	0.1753	0.1753	3.6217	3.6217	3.9872	2.4541	2.4541	2.5405
reg,0.9	0.9552	0.9512	0.9518	0.9576	0.9484	0.9484	0.9502	0.9498	0.9508	0.9518
len	0.1477	2.1795	2.1756	2.1781	3.3613	3.3613	3.3669	2.4534	2.4534	2.4627
vs,0.9	0.9566	0.9800	0.9992	0.9992	0.9960	0.9932	0.9982	0.9540	0.9444	0.9654
len	0.1475	2.0229	1.8888	1.8908	3.5749	3.5749	3.9622	2.4376	2.4376	2.5589

Table 4.44. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9582	0.9552	0.9562	0.9616	0.9488	0.9484	0.9482	0.9494	0.9498	0.9510
len	0.0171	0.0046	0.0039	0.0039	3.3613	3.3613	3.3654	2.4546	2.4546	2.4585
vs,0	0.9588	0.9646	0.9988	0.9988	0.9952	0.9946	0.9974	0.9586	0.9582	0.9636
len	0.0171	0.0046	0.0032	0.0032	3.6537	3.6537	4.0327	2.4567	2.4567	2.5133
reg,0.38	0.9590	0.9506	0.9580	0.9530	0.9486	0.9498	0.9504	0.9468	0.9496	0.9504
len	0.0171	0.0080	0.0076	0.0076	3.3617	3.3617	3.3665	2.4543	2.4543	2.4579
vs,0.38	0.9528	0.9652	0.9980	0.9978	0.9938	0.9934	0.9978	0.9564	0.9574	0.9656
len	0.0171	0.0080	0.0064	0.0064	3.6182	3.6182	3.9760	2.4547	2.4547	2.5595
reg,0.9	0.9506	0.9602	0.9512	0.9562	0.9488	0.9488	0.9482	0.9484	0.9490	0.9484
len	0.0171	0.0805	0.0804	0.0806	3.3615	3.3615	3.3655	2.4559	2.4559	2.4590
vs,0.9	0.9642	0.9876	0.9982	0.9990	0.9944	0.9940	0.9966	0.9664	0.9676	0.9734
len	0.0171	0.0807	0.0690	0.0691	3.6166	3.6166	3.9739	2.4503	2.4503	2.5625

Table 4.45. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 5$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9542	0.9592	0.9488	0.9548	0.9458	0.9548	0.9558	0.9644	0.9716	0.9756
len	0.5628	0.8469	0.8455	0.7606	1.9632	1.9632	1.9671	3.6615	3.6615	3.7317
vs,0	0.9606	0.9474	0.9522	0.9986	0.9982	0.9550	0.9982	0.9790	0.9854	0.9888
len	0.5623	0.8468	0.8451	0.6413	2.1397	2.1397	2.3827	3.6707	3.6707	3.7414
reg,0.38	0.9524	0.9536	0.9516	0.9510	0.9486	0.9566	0.9578	0.9660	0.9756	0.9802
len	0.5633	3.0237	3.0190	3.0010	1.9652	1.9652	1.9696	3.6527	3.6527	3.6911
vs,0.38	0.9588	0.9534	0.9392	0.9982	0.9972	0.9500	0.9974	0.9818	0.9614	0.9872
len	0.5618	2.7497	2.7592	2.6200	2.1073	2.1073	2.3412	3.6881	3.6881	4.0066
reg,0.9	0.9522	0.9562	0.9502	0.9558	0.9506	0.9584	0.9590	0.9684	0.9786	0.9796
len	0.5638	34.9598	34.9442	34.9259	1.9634	1.9634	1.9676	3.6468	3.6468	3.6584
vs,0.9	0.9598	0.9779	0.9876	0.9982	0.9970	0.9634	0.9980	0.9968	0.9976	0.9988
len	0.55931	30.4471	30.5897	30.4169	2.1061	2.1061	2.3306	3.8777	3.8777	4.2092

Table 4.46. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 5$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9526	0.9564	0.9570	0.9578	0.9504	0.9516	0.9522	0.9500	0.9512	0.9530
len	0.1201	0.1778	0.1778	0.1620	1.9598	1.9598	1.9624	3.5955	3.5955	3.6037
vs,0	0.9566	0.9514	0.9538	0.9986	0.9982	0.9438	0.9982	0.9510	0.9524	0.9542
len	0.1203	0.1783	0.1780	0.1349	2.1480	2.1480	2.3950	3.5961	3.5961	3.6064
reg,0.38	0.9588	0.9526	0.9566	0.9498	0.9442	0.9452	0.9454	0.9522	0.9528	0.9528
len	0.1203	0.6415	0.6413	0.6366	1.9581	1.9581	1.9610	3.5932	3.5932	3.5991
vs,0.38	0.9554	0.9530	0.9588	0.9994	0.9982	0.9436	0.9982	0.9628	0.9654	0.9708
len	0.1201	0.6406	0.6409	0.5309	2.1558	2.1558	2.4088	3.6098	3.6098	3.6675
reg,0.9	0.9536	0.9580	0.9528	0.9570	0.9524	0.9534	0.9536	0.9520	0.9526	0.9530
len	0.1203	7.4088	7.4057	7.4145	1.9576	1.9576	1.9607	3.5924	3.5924	3.5983
vs,0.9	0.9570	0.9865	0.9866	0.9976	0.9964	0.9428	0.9970	0.9932	0.9928	0.9968
len	0.1202	6.4273	6.4154	6.4385	2.1186	2.1186	2.3511	3.7945	3.7945	4.1326

Table 4.47. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 5$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9550	0.9582	0.9562	0.9606	0.9564	0.9560	0.9564	0.9464	0.9450	0.9468
len	0.1476	0.2382	0.2388	0.2374	1.9587	1.9587	1.9618	3.5952	3.5952	3.6074
vs,0	0.9510	0.9546	0.9546	0.9984	0.9972	0.9432	0.9972	0.9544	0.9556	0.9588
len	0.2117	0.3496	0.3486	0.2883	2.1517	2.1517	2.4015	3.5984	3.5984	3.6248
reg,0.38	0.9546	0.9592	0.9566	0.9590	0.9500	0.9516	0.9522	0.9476	0.9500	0.9508
len	0.1476	0.9354	0.9354	0.9335	1.9582	1.9582	1.9617	3.5946	3.5946	3.6052
vs,0.38	0.9520	0.9576	0.9514	0.9984	0.9976	0.9438	0.9976	0.9472	0.9098	0.9446
len	0.2128	1.4107	1.4130	1.1697	2.1318	2.1318	2.3846	3.6951	3.6951	3.8382
reg,0.9	0.9570	0.9562	0.9564	0.9556	0.9474	0.9468	0.9492	0.9462	0.9472	0.9496
len	0.1476	10.8607	10.8714	10.9100	1.9580	1.9580	1.9611	3.5931	3.5931	3.6032
vs,0.9	0.9536	0.9810	0.9808	0.9984	0.9964	0.9670	0.9970	0.9942	0.9946	0.9976
len	0.2116	13.6373	13.7437	13.7113	2.1272	2.1272	2.3442	3.8664	3.8664	4.2048

Table 4.48. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 350$, $n = 350$, $p = 7$, $k = 5$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9564	0.9522	0.9494	0.9542	0.9466	0.9476	0.9474	0.9484	0.9480	0.9484
len	0.0171	0.0091	0.0091	0.0087	1.9580	1.9580	1.9615	3.5923	3.5923	3.5985
vs,0	0.9560	0.9574	0.9568	0.9986	0.9972	0.9382	0.9970	0.9448	0.9476	0.9502
len	0.0171	0.0090	0.0091	0.0072	2.1414	2.1414	2.3927	3.5923	3.5923	3.6126
reg,0.38	0.9526	0.9516	0.9534	0.9602	0.9554	0.9548	0.9550	0.9496	0.9496	0.9508
len	0.0171	0.0346	0.0345	0.0344	1.9579	1.9579	1.9607	3.5914	3.5914	3.5965
vs,0.38	0.9508	0.9516	0.9542	0.9988	0.9980	0.9434	0.9980	0.9596	0.9610	0.9668
len	0.0171	0.0345	0.0345	0.0286	2.1467	2.1467	2.4006	3.6427	3.6427	3.7386
reg,0.9	0.9544	0.9546	0.9524	0.9526	0.9442	0.9450	0.9452	0.9472	0.9492	0.9494
len	0.0171	0.4027	0.4025	0.4031	1.9563	1.9563	1.9595	3.5904	3.5904	3.5959
vs,0.9	0.9552	0.9546	0.9524	0.9978	0.9970	0.9446	0.9970	0.9660	0.9660	0.9730
len	0.0171	0.4030	0.4022	0.3355	2.1516	2.1516	2.4069	3.6542	3.6542	3.7574

Table 4.49. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9508	0.9168	0.9514	0.9546	0.9822	0.9936	0.9956	0.9550	0.9756	0.9858
len	0.7932	0.7311	0.4839	0.4834	3.6974	3.6974	3.7218	2.4782	2.4782	2.6560
vs,0	0.9518	0.9268	0.9980	0.9978	0.9980	0.9994	1.0000	0.9574	0.9758	0.9860
len	0.7819	0.7181	0.4055	0.4052	3.9782	3.9782	4.3825	2.4762	2.4762	2.6670
reg,0.35	0.9530	0.9364	0.9488	0.9538	0.9842	0.9956	0.9970	0.9556	0.9722	0.9804
len	0.7953	1.0620	0.9110	0.9101	3.6973	3.6973	3.7220	2.4792	2.4792	2.5711
vs,0.35	0.9536	0.9406	0.9980	0.9984	0.9980	0.9986	0.9994	0.9672	0.9812	0.9878
len	0.7828	1.0528	0.7731	0.7765	3.9390	3.9390	4.3251	2.4835	2.4835	2.6135
reg,0.9	0.9510	0.9500	0.9552	0.9550	0.9838	0.9932	0.9948	0.9570	0.9716	0.9762
len	0.7937	10.8476	10.8171	10.8288	3.6974	3.6974	3.7220	2.4832	2.48324	2.4963
vs,0.9	0.9578	0.9842	0.9980	0.9982	0.9994	0.9994	1.0000	0.9796	0.9854	0.9882
len	0.7810	9.4113	9.3673	9.3350	3.9205	3.9205	4.3066	2.5580	2.5580	2.7128

Table 4.50. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9542	0.9518	0.9582	0.9524	0.9494	0.9516	0.9518	0.9486	0.9516	0.9534
len	0.1610	0.1387	0.0972	0.0970	3.5892	3.5892	3.5938	2.4539	2.4539	2.4657
vs,0	0.9574	0.9574	0.9984	0.9992	0.9954	0.9956	0.9980	0.9532	0.9568	0.9580
len	0.1607	0.1390	0.0809	0.0804	3.9051	3.9051	4.3051	2.4550	2.4550	2.4744
reg,0.35	0.9564	0.9530	0.9596	0.9596	0.9494	0.9512	0.9514	0.9516	0.9534	0.9542
len	0.1611	0.2080	0.1829	0.1829	3.5897	3.5897	3.5948	2.4557	2.4557	2.4625
vs,0.35	0.9524	0.9600	0.9992	0.9994	0.9972	0.9970	0.9996	0.9506	0.9524	0.9572
len	0.1607	0.2086	0.1547	0.1552	3.8608	3.8608	4.2399	2.4541	2.4541	2.5132
reg,0.9	0.9562	0.9606	0.9556	0.9558	0.9574	0.9590	0.9584	0.9514	0.9530	0.9536
len	0.1610	2.1784	2.1768	2.1761	3.5887	3.5887	3.5932	2.4543	2.4543	2.4584
vs,0.9	0.9514	0.9762	0.9982	0.9990	0.9956	0.9952	0.9986	0.9534	0.9448	0.9654
len	0.1609	2.0247	1.8785	1.8739	3.8183	3.8183	4.2251	2.4415	2.4415	2.5609

Table 4.51. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9488	0.9578	0.9546	0.9528	0.9460	0.9464	0.9470	0.9438	0.9480	0.9500
len	0.1984	0.1505	0.1444	0.1442	3.5871	3.5871	3.5923	2.4566	2.4566	2.4751
vs,0	0.9594	0.9568	0.9982	0.9986	0.9974	0.9968	0.9990	0.9530	0.9542	0.9584
len	0.1979	0.1500	0.1205	0.1198	3.8939	3.8939	4.2948	2.4555	2.4555	2.4987
reg,0.35	0.9480	0.9520	0.9556	0.9548	0.9464	0.9470	0.9474	0.9482	0.9498	0.9526
len	0.1982	0.2755	0.2717	0.2710	3.5887	3.5887	3.5937	2.4552	2.4552	2.4717
vs,0.35	0.9552	0.9700	0.9986	0.9986	0.9956	0.9956	0.9982	0.9596	0.9618	0.9676
len	0.1980	0.2753	0.2297	0.2303	3.8553	3.8553	4.2428	2.4517	2.4517	2.5466
reg,0.9	0.9486	0.9562	0.9530	0.9550	0.9480	0.9490	0.9494	0.9450	0.9482	0.9500
len	0.1981	3.2285	3.2435	3.2292	3.5877	3.5877	3.5923	2.4532	2.4532	2.4685
vs,0.9	0.9580	0.9542	0.9984	0.9980	0.9976	0.9970	0.9992	0.9690	0.9686	0.9764
len	0.1979	2.8449	2.7874	2.7973	3.8607	3.8607	4.2448	2.4771	2.4771	2.6314

Table 4.52. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9512	0.9566	0.9566	0.9520	0.9502	0.9504	0.9518	0.9510	0.9498	0.9506
len	0.0238	0.0070	0.0056	0.0056	3.5879	3.5879	3.5920	2.4535	2.4535	2.4565
vs,0	0.9610	0.9600	0.9980	0.9980	0.9950	0.9942	0.9980	0.9578	0.9616	0.9664
len	0.0237	0.0069	0.0047	0.0047	3.8769	3.8769	4.2736	2.4567	2.4567	2.5306
reg,0.35	0.9596	0.9584	0.9588	0.9568	0.9478	0.9486	0.9502	0.9526	0.9530	0.9538
len	0.0238	0.0114	0.0107	0.0107	3.5850	3.5850	3.5896	2.4528	2.4528	2.4563
vs,0.35	0.9606	0.9666	0.9990	0.9988	0.9960	0.9956	0.9984	0.9616	0.9644	0.9698
len	0.0237	0.0091	0.0091	0.0092	3.8492	3.8492	4.2240	2.4531	2.4531	2.5666
reg,0.9	0.9558	0.9588	0.9558	0.9540	0.9524	0.9524	0.9538	0.9516	0.9516	0.9522
len	0.0238	0.1268	0.1275	0.1271	3.5847	3.5847	3.5889	2.4551	2.4551	2.4589
vs,0.9	0.9564	0.9798	0.9994	0.9984	0.9946	0.9934	0.9970	0.9562	0.9580	0.9630
len	0.0238	0.1272	0.1090	0.1090	3.8381	3.8381	4.2156	2.4530	2.4530	2.5639

Table 4.53. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 6$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9498	0.9470	0.9422	0.9466	0.9418	0.9598	0.9626	0.9880	0.9924	0.9948
len	0.7943	1.3010	1.3062	1.1823	1.9687	1.9687	1.9768	3.9552	3.9552	4.0832
vs,0	0.9490	0.9376	0.9454	0.9984	0.9966	0.9770	0.9970	0.9868	0.9776	0.9872
len	0.7935	1.3261	1.3267	1.0012	2.1223	2.1223	2.3729	3.9998	3.9998	4.1094
reg,0.35	0.9552	0.9536	0.9534	0.9582	0.9572	0.9690	0.9724	0.9852	0.9928	0.9944
len	0.7946	4.9004	4.9052	4.8687	1.9683	1.9683	1.9763	3.9368	3.9368	4.0132
vs,0.35	0.9528	0.9208	0.9130	0.9974	0.9964	0.9554	0.9966	0.9982	0.9970	0.9990
len	0.7848	4.2863	4.2866	4.2164	2.0951	2.0951	2.3211	4.0249	4.0249	4.3910
reg,0.9	0.9528	0.9530	0.9510	0.9584	0.9528	0.9664	0.9684	0.9902	0.9964	0.9974
len	0.7932	64.8408	64.8927	64.7783	1.9674	1.9674	1.9762	3.9253	3.9253	3.9523
vs,0.9	0.9560	0.9754	0.9754	0.9982	0.9974	0.9746	0.9984	1.0000	1.0000	1.0000
len	0.7811	56.1974	56.0379	56.2916	2.0887	2.0887	2.2982	4.1323	4.1323	4.4965

Table 4.54. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 6$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9558	0.9546	0.9544	0.9534	0.9470	0.9494	0.9504	0.9476	0.9512	0.9534
len	0.1611	0.2576	0.2581	0.2373	1.9593	1.9593	1.9612	3.8037	3.8037	3.8150
vs,0	0.9586	0.9542	0.9534	0.9988	0.9976	0.9448	0.9976	0.9512	0.9530	0.9546
len	0.1611	0.2574	0.2578	0.1979	2.1546	2.1546	2.3965	3.8012	3.8012	3.8148
reg,0.35	0.9536	0.9528	0.9574	0.9552	0.9466	0.9476	0.9476	0.9472	0.9508	0.9518
len	0.1609	0.9835	0.9844	0.9786	1.9580	1.9580	1.9613	3.8028	3.8028	3.8103
vs,0.35	0.9548	0.9444	0.9416	0.9976	0.9960	0.9466	0.9960	0.9718	0.9684	0.9744
len	0.1612	1.0093	1.0080	0.8210	2.1391	2.1391	2.3903	3.8981	3.8981	3.9587
reg,0.9	0.9594	0.9506	0.9582	0.9572	0.9534	0.9532	0.9534	0.9506	0.9518	0.9520
len	0.1610	13.0567	13.0281	13.0712	1.9600	1.9600	1.9627	3.8031	3.8031	3.8081
vs,0.9	0.9558	0.9802	0.9806	0.9976	0.9964	0.9462	0.9972	0.9944	0.9944	0.9978
len	0.1609	11.2959	11.2181	11.2937	2.0903	2.0903	2.2896	4.0045	4.0045	4.3990

Table 4.55. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 6$, and $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9510	0.9580	0.9570	0.9566	0.9506	0.9512	0.9512	0.9442	0.9460	0.9496
len	0.1979	0.3563	0.3568	0.3541	1.9585	1.9585	1.9611	3.8040	3.8040	3.8235
vs,0	0.9516	0.9580	0.9604	0.9982	0.9976	0.9462	0.9976	0.9564	0.9554	0.9588
len	0.1981	0.3552	0.3557	0.2929	2.1462	2.1462	2.3946	3.8025	3.8025	3.8285
reg,0.35	0.9500	0.9554	0.9572	0.9564	0.9482	0.9484	0.9502	0.9514	0.9528	0.9544
len	0.1982	1.4570	1.4550	1.4539	1.9593	1.9593	1.9626	3.8014	3.8014	3.8181
vs,0.35	0.9514	0.9530	0.9580	0.9982	0.9972	0.9418	0.9972	0.9412	0.9034	0.9396
len	0.1991	1.5021	1.5029	1.2529	2.1200	2.1200	2.3692	3.8809	3.8809	4.0458
reg,0.9	0.9538	0.9534	0.9516	0.9504	0.9438	0.9426	0.9422	0.9486	0.9502	0.9526
len	0.1982	19.3762	19.4191	19.4233	1.9595	1.9595	1.9623	3.8034	3.8034	3.8188
vs,0.9	0.9548	0.9785	0.9774	0.9972	0.9966	0.9632	0.9972	0.9962	0.9962	0.9988
len	0.1982	16.6798	16.7676	16.7003	2.1136	2.1136	2.3236	4.0659	4.0659	4.4287

Table 4.56. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 200$, $p = 8$, $k = 6$, and $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9564	0.9522	0.9578	0.9528	0.9462	0.9476	0.9478	0.9464	0.9460	0.9484
len	0.0238	0.0145	0.0145	0.0139	1.9572	1.9572	1.9602	3.7972	3.7972	3.8026
vs,0	0.9548	0.9574	0.9568	0.9982	0.9968	0.9416	0.9968	0.9564	0.9574	0.9598
len	0.0238	0.0145	0.0145	0.0115	2.1424	2.1424	2.3992	3.7998	3.7998	3.8252
reg,0.35	0.9608	0.9552	0.9548	0.9522	0.9454	0.9470	0.9452	0.9484	0.9502	0.9500
len	0.0238	0.0575	0.0576	0.0574	1.9586	1.9586	1.9614	3.7953	3.7953	3.8002
vs,0.35	0.9582	0.9564	0.9542	0.9994	0.9988	0.9416	0.9988	0.9626	0.9630	0.9688
len	0.0238	0.0573	0.0575	0.0476	2.1519	2.1519	2.3949	3.8421	3.8421	3.9267
reg,0.9	0.9556	0.9574	0.9470	0.9562	0.9482	0.9496	0.9500	0.9474	0.9480	0.9488
len	0.0238	0.7648	0.7657	0.7644	1.9582	1.9582	1.9608	3.7976	3.7976	3.8023
vs,0.9	0.9614	0.9536	0.9500	0.9980	0.9966	0.9396	0.9968	0.9742	0.9750	0.9788
len	0.0237	0.7696	0.7701	0.6367	2.1429	2.1429	2.3936	3.8832	3.8832	3.9765

Table 4.57. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9512	0.9298	0.9522	0.9548	0.9646	0.9760	0.9774	0.9444	0.9604	0.9700
len	0.5266	0.4686	0.3182	0.3182	3.6350	3.6350	3.6440	2.4640	2.4640	2.5525
vs,0	0.9556	0.9316	0.9978	0.9984	0.9982	0.9984	0.9990	0.9496	0.9626	0.9710
len	0.5223	0.4651	0.2645	0.2640	3.9352	3.9352	4.3372	2.4633	2.4633	2.5565
reg,0.35	0.9526	0.9426	0.9498	0.9510	0.9610	0.9752	0.9768	0.9448	0.9594	0.9650
len	0.5270	0.6910	0.5984	0.5979	3.6347	3.6347	3.6442	2.4645	2.4645	2.5083
vs,0.35	0.9552	0.9532	0.9980	0.9974	0.9970	0.9978	0.9988	0.9522	0.9622	0.9676
len	0.5220	0.6874	0.5081	0.5082	3.8875	3.8875	4.2762	2.4632	2.4632	2.5512
reg,0.9	0.9508	0.9550	0.9554	0.9568	0.9686	0.9784	0.9798	0.9496	0.9598	0.9602
len	0.5260	7.1304	7.1168	7.1214	3.6357	3.6357	3.6456	2.4683	2.4683	2.4736
vs,0.9	0.9516	0.9878	0.9986	0.9976	0.9974	0.9982	0.9994	0.9724	0.9724	0.9800
len	0.5212	6.1651	6.1712	6.1778	3.8942	3.8942	4.2805	2.5431	2.5431	2.7049

Table 4.58. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9554	0.9504	0.9516	0.9516	0.9514	0.9518	0.9520	0.9482	0.9484	0.9502
len	0.1124	0.0971	0.0676	0.0676	3.5854	3.5854	3.5910	2.4537	2.4537	2.4604
vs,0	0.9562	0.9584	0.9988	0.9992	0.9976	0.9976	0.9996	0.9512	0.9544	0.9552
len	0.1123	0.0971	0.0561	0.0563	3.9049	3.9049	4.3026	2.4530	2.4530	2.4638
reg,0.35	0.9552	0.9554	0.9504	0.9568	0.9540	0.9556	0.9564	0.9492	0.9486	0.9488
len	0.1124	0.1453	0.1274	0.1274	3.5870	3.5870	3.5920	2.4548	2.4548	2.4598
vs,0.35	0.9572	0.9598	0.9976	0.9990	0.9962	0.9962	0.9988	0.9526	0.9544	0.9592
len	0.1123	0.1457	0.1077	0.1079	3.8533	3.8533	4.2418	2.4533	2.4533	2.5075
reg,0.9	0.9542	0.9562	0.9572	0.9502	0.9520	0.9532	0.9540	0.9480	0.9502	0.9490
len	0.1125	1.5187	1.5146	1.5168	3.5867	3.5867	3.5920	2.4547	2.4547	2.4575
vs,0.9	0.9556	0.9800	0.9982	0.9982	0.9954	0.9908	0.9974	0.9468	0.9100	0.9480
len	0.1124	1.5190	1.3035	1.3058	3.8547	3.8547	4.2353	2.4592	2.4592	2.5555

Table 4.59. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9550	0.9528	0.9528	0.9578	0.9456	0.9460	0.9470	0.9488	0.9512	0.9524
len	0.1381	0.1017	0.0991	0.0992	3.5878	3.5878	3.5927	2.4548	2.4548	2.4664
vs,0	0.9596	0.9560	0.9992	0.9986	0.9980	0.9978	0.9990	0.9558	0.9562	0.9596
len	0.1380	0.1014	0.0822	0.0823	3.8982	3.8982	4.2998	2.4546	2.4546	2.4788
reg,0.35	0.9558	0.9538	0.9452	0.9542	0.9422	0.9418	0.9440	0.9478	0.9494	0.9492
len	0.1382	0.1878	0.1868	0.1866	3.5876	3.5876	3.5929	2.4536	2.4536	2.4627
vs,0.35	0.9552	0.9630	0.9988	0.9984	0.9972	0.9964	0.9994	0.9576	0.9558	0.9646
len	0.1379	0.1886	0.1592	0.1578	3.8566	3.8566	4.2402	2.4516	2.4516	2.5336
reg,0.9	0.9550	0.9542	0.9510	0.9532	0.9496	0.9480	0.9484	0.9480	0.9490	0.9498
len	0.1381	2.2249	2.2251	2.2233	3.5870	3.5870	3.5917	2.4563	2.4563	2.4642
vs,0.9	0.9562	0.9732	0.9994	0.9990	0.9974	0.9952	0.9990	0.9522	0.9440	0.9612
len	0.1380	2.0502	1.9148	1.9224	3.8191	3.8191	4.2211	2.4366	2.4366	2.5612

Table 4.60. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9560	0.9544	0.9540	0.9570	0.9496	0.9508	0.9512	0.9468	0.9476	0.9480
len	0.01602	0.0043	0.0036	0.0036	3.5841	3.5841	3.5889	2.4540	2.4540	2.4569
vs,0	0.9526	0.9594	0.9988	0.9988	0.9966	0.9958	0.9982	0.9512	0.9538	0.9586
len	0.0159	0.0043	0.0030	0.0030	3.8825	3.8825	4.2780	2.4544	2.4544	2.5123
reg,0.35	0.9586	0.9508	0.9556	0.9594	0.9486	0.9492	0.9498	0.9486	0.9468	0.9484
len	0.0160	0.0072	0.0068	0.0069	3.5848	3.5848	3.5891	2.4528	2.4528	2.4553
vs,0.35	0.9524	0.9676	0.9990	0.9992	0.9958	0.9946	0.9980	0.9614	0.9636	0.9702
len	0.0159	0.0072	0.0058	0.0058	3.8498	3.8498	4.2328	2.4515	2.4515	2.5558
reg,0.9	0.9470	0.9556	0.9550	0.9568	0.9492	0.9502	0.9506	0.9462	0.9458	0.9470
len	0.0160	0.0822	0.0823	0.0821	3.5870	3.5870	3.5910	2.4531	2.4531	2.4558
vs,0.9	0.9528	0.9784	0.9990	0.9982	0.9958	0.9952	0.9982	0.9536	0.9570	0.9656
len	0.0160	0.0820	0.0699	0.0700	3.8421	3.8421	4.2213	2.4520	2.4520	2.5530

Table 4.61. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 6$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9476	0.9492	0.9500	0.9560	0.9502	0.9582	0.9598	0.9670	0.9726	0.9770
len	0.5266	0.8521	0.8515	0.7788	1.9629	1.9629	1.9669	3.8693	3.8693	3.9380
vs,0	0.9528	0.9504	0.9542	0.9982	0.9968	0.9516	0.9968	0.9806	0.9894	0.9914
len	0.5250	0.8516	0.8504	0.6521	2.1327	2.1327	2.3762	3.8815	3.8815	3.9493
reg,0.35	0.9572	0.9500	0.9540	0.9592	0.9550	0.9614	0.9628	0.9704	0.9780	0.9790
len	0.5257	3.2295	3.2212	3.2080	1.9639	1.9639	1.9674	3.8600	3.8600	3.8969
vs,0.35	0.9526	0.9296	0.9224	0.9990	0.9974	0.9476	0.9976	0.9850	0.9638	0.9898
len	0.5245	2.8961	2.9056	2.7669	2.1086	2.1086	2.3379	3.9054	3.9054	4.2495
reg,0.9	0.9588	0.9580	0.9570	0.9562	0.9488	0.9568	0.9574	0.9694	0.9776	0.9786
len	0.5265	42.7113	42.6696	42.7334	1.9647	1.9647	1.9681	3.8518	3.8518	3.8625
vs,0.9	0.9530	0.9756	0.9763	0.9994	0.9980	0.9648	0.9990	0.9980	0.9982	0.9996
len	0.5215	36.8778	37.0413	37.0125	2.1090	2.1090	2.3304	4.0898	4.0898	4.4559

Table 4.62. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 6$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9566	0.9538	0.9584	0.9524	0.9494	0.9488	0.9488	0.9514	0.9530	0.9544
len	0.1124	0.1799	0.1796	0.1655	1.9583	1.9583	1.9600	3.7992	3.7992	3.8071
vs,0	0.9508	0.9548	0.9544	0.9984	0.9970	0.9442	0.9970	0.9508	0.9518	0.9530
len	0.1125	0.1800	0.1798	0.1370	2.1454	2.1454	2.3996	3.7991	3.7991	3.8082
reg,0.35	0.9518	0.9550	0.9582	0.9538	0.9490	0.9484	0.9486	0.9456	0.9476	0.9482
len	0.1124	0.6861	0.6867	0.6832	1.9590	1.9590	1.9616	3.7991	3.7991	3.8051
vs,0.35	0.9564	0.9536	0.9540	0.9984	0.9982	0.9408	0.9982	0.9634	0.9650	0.9702
len	0.1124	0.6849	0.6855	0.5675	2.1510	2.1510	2.4014	3.8173	3.8173	3.8676
reg,0.9	0.9572	0.9546	0.9496	0.9540	0.9482	0.9484	0.9480	0.9488	0.9524	0.9532
len	0.1124	9.0909	9.0883	9.0733	1.9591	1.9591	1.9615	3.7976	3.7976	3.8026
vs,0.9	0.9570	0.9844	0.9841	0.9982	0.9972	0.9408	0.9978	0.9952	0.9940	0.9976
len	0.1123	7.8091	7.8175	7.8497	2.1183	2.1183	2.3597	4.0096	4.0096	4.3813

Table 4.63. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 6$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9574	0.9550	0.9560	0.9614	0.9544	0.9552	0.9556	0.9488	0.9492	0.9516
len	0.1382	0.2443	0.2442	0.2423	1.9574	1.9574	1.9600	3.8000	3.8000	3.8123
vs,0	0.9516	0.9520	0.9534	0.9992	0.9984	0.9428	0.9984	0.9480	0.9462	0.9478
len	0.1381	0.2438	0.2438	0.2014	2.1501	2.1501	2.3960	3.7982	3.7982	3.8136
reg,0.35	0.9516	0.9584	0.9612	0.9508	0.9440	0.9440	0.9452	0.9490	0.9486	0.9504
len	0.1381	1.0028	1.0002	1.0012	1.9599	1.9599	1.9620	3.7978	3.7978	3.8084
vs,0.35	0.9546	0.9396	0.9430	0.9992	0.9978	0.9420	0.9978	0.9710	0.9648	0.9712
len	0.1384	1.0303	1.0288	0.8395	2.1310	2.1310	2.3828	3.9205	3.9205	4.0085
reg,0.9	0.9520	0.9550	0.9540	0.9506	0.9436	0.9434	0.9438	0.9448	0.9446	0.9464
len	0.1380	13.3179	13.3432	13.3094	1.9588	1.9588	1.9608	3.8004	3.8004	3.8096
vs,0.9	0.9538	0.9802	0.9801	0.9986	0.9972	0.9536	0.9976	0.9950	0.9940	0.9972
len	0.1380	11.4526	11.4796	11.5146	2.0913	2.0913	2.2913	4.0117	4.0117	4.4041

Table 4.64. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 400$, $n = 400$, $p = 8$, $k = 6$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9574	0.9572	0.9564	0.9560	0.9488	0.9492	0.9498	0.9468	0.9502	0.9502
len	0.0160	0.0092	0.0092	0.0089	1.9598	1.9598	1.9616	3.7959	3.7959	3.8008
vs,0	0.9546	0.9612	0.9508	0.9984	0.9976	0.9372	0.9976	0.9492	0.9472	0.9508
len	0.0159	0.0093	0.0092	0.0074	2.1511	2.1511	2.4017	3.7978	3.7978	3.8171
reg,0.35	0.9532	0.9570	0.9566	0.9540	0.9456	0.9460	0.9460	0.9448	0.9464	0.9474
len	0.0160	0.0372	0.0370	0.0370	1.9590	1.9590	1.9613	3.7959	3.7959	3.8011
vs,0.35	0.9544	0.9498	0.9600	0.9992	0.9984	0.9464	0.9984	0.9618	0.9620	0.9692
len	0.0159	0.0370	0.0371	0.0306	2.1450	2.1450	2.3986	3.8409	3.8409	3.9262
reg,0.9	0.9546	0.9558	0.9590	0.9600	0.9536	0.9528	0.9544	0.9498	0.9506	0.9516
len	0.0160	0.4925	0.4939	0.4917	1.9591	1.9591	1.9614	3.7963	3.7963	3.8018
vs,0.9	0.9542	0.9552	0.9544	0.9992	0.9986	0.9652	0.9984	0.9642	0.9632	0.9704
len	0.0159	0.4926	0.4921	0.4075	2.1511	2.1511	2.4070	3.8523	3.8523	3.9469

Table 4.65. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9510	0.9080	0.9506	0.9484	0.9850	0.9968	0.9980	0.9476	0.9732	0.9860
len	0.7002	0.6420	0.4239	0.4243	4.0994	4.0994	4.1236	2.4716	2.4716	2.6580
vs,0	0.9534	0.9248	0.9970	0.9976	0.9992	0.9996	1.0000	0.9534	0.9732	0.9858
len	0.6870	0.6281	0.3489	0.3468	4.3871	4.3871	4.8337	2.4710	2.4710	2.6694
reg,0.32	0.9524	0.9288	0.9552	0.9494	0.9866	0.9960	0.9974	0.9500	0.9714	0.9784
len	0.6990	0.9291	0.7951	0.7953	4.0995	4.0995	4.1232	2.4711	2.4711	2.5680
vs,0.32	0.9512	0.9370	0.9980	0.9986	0.9994	1.0000	1.0000	0.9556	0.9758	0.9834
len	0.6897	0.9185	0.6683	0.6698	4.3516	4.3516	4.7751	2.4715	2.4715	2.6032
reg,0.9	0.9504	0.9516	0.9488	0.9526	0.9838	0.9938	0.9956	0.9554	0.9698	0.9702
len	0.6990	10.9565	10.9223	10.9347	4.0990	4.0990	4.1223	2.4764	2.4764	2.4863
vs,0.9	0.9538	0.9819	0.9978	0.9974	0.9998	1.0000	1.0000	0.9762	0.9834	0.9864
len	0.6878	9.3204	9.3276	9.2807	4.3420	4.3420	4.7711	2.5566	2.5566	2.7227

4.2 CONCLUSION

Another way to look at the bootstrap confidence region for variable selection estimators is to consider the estimator $T_{2,n}$ that chooses I_j with probability equal to the observed bootstrap proportion $\hat{\rho}_{jn}$. The bootstrap sample $T_1^*, ..., T_B^*$ tends to be slightly more variable than an iid sample $T_{2,1}, ..., T_{2,B}$, and the geometric argument suggests that the large sample coverage of the nominal $100(1-\delta)\%$ confidence region will be at least as large as the nominal coverage $100(1-\delta)\%$.

The bootstrap confidence regions may work well when the probability that the selected model underfits goes to zero rapidly. Hence the confidence regions can give good results for the relaxed lasso and relaxed elastic net estimators, which fit the MLE to the predictors with nonzero lasso or elastic net coefficients. See Efron et al. (2004), Friedman et al. (2007), Friedman, Hastie, and Tibshirani (2010), Hastie, Tibshirani, and Wainwright (2015, ch. 5), Meinshausen (2007), Tibshirani (1996), and Zou and Hastie (2005).

Table 4.66. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9576	0.9502	0.9520	0.9548	0.9500	0.9528	0.9530	0.9480	0.9496	0.9502
len	0.1428	0.1232	0.0860	0.0860	3.9837	3.9837	3.9876	2.4538	2.4538	2.4653
vs,0	0.9510	0.9510	0.9992	0.9978	0.9980	0.9982	0.9998	0.9412	0.9458	0.9478
len	0.1424	0.1229	0.0706	0.0707	4.3081	4.3081	4.7454	2.4531	2.4531	2.4747
reg,0.32	0.9536	0.9534	0.9514	0.9548	0.9496	0.9524	0.9530	0.9474	0.9490	0.9506
len	0.1426	0.1833	0.1609	0.1610	3.9840	3.9840	3.9884	2.4528	2.4528	2.4589
vs,0.32	0.9534	0.9620	0.9966	0.9976	0.9968	0.9976	0.9988	0.9534	0.9544	0.9582
len	0.1424	0.1837	0.1347	0.1352	4.2607	4.2607	4.6891	2.4527	2.4527	2.5042
reg,0.9	0.9514	0.9432	0.9552	0.9498	0.9434	0.9448	0.9446	0.9430	0.9440	0.9450
len	0.1427	2.2178	2.2170	2.2175	3.9846	3.9846	3.9887	2.4530	2.4530	2.4553
vs,0.9	0.9590	0.9656	0.9982	0.9986	0.9982	0.9978	0.9996	0.9532	0.9478	0.9654
len	0.1425	2.0342	1.8778	1.8862	4.2368	4.2368	4.6742	2.4449	2.4449	2.5661

Table 4.67. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9480	0.9526	0.9526	0.9520	0.9502	0.9512	0.9524	0.9432	0.9454	0.9472
len	0.1752	0.1325	0.1275	0.1276	3.9859	3.9859	3.9901	2.4528	2.4528	2.4740
vs,0	0.9552	0.9574	0.9982	0.9982	0.9984	0.9982	0.9998	0.9524	0.9574	0.9628
len	0.1752	0.1323	0.1051	0.1047	4.3004	4.3004	4.7408	2.4543	2.4543	2.5009
reg,0.32	0.9552	0.9518	0.9520	0.9536	0.9538	0.9536	0.9538	0.9510	0.9532	0.9552
len	0.1752	0.2419	0.2390	0.2386	3.9852	3.9852	3.9894	2.4518	2.4518	2.4689
vs,0.32	0.9562	0.9632	0.9986	0.9992	0.9980	0.9982	0.9992	0.9630	0.9644	0.9712
len	0.1750	0.2419	0.2005	0.2004	4.2618	4.2618	4.6811	2.4520	2.4520	2.5384
reg,0.9	0.9478	0.9530	0.9570	0.9554	0.9458	0.9478	0.9484	0.9448	0.9448	0.9476
len	0.1754	3.2873	3.2859	3.2912	3.9831	3.9831	3.9872	2.4536	2.4536	2.4691
vs,0.9	0.9500	0.9574	0.9984	0.9994	0.9970	0.9966	0.9984	0.9638	0.9626	0.9742
len	0.1752	2.8710	2.7922	2.7879	4.2597	4.2597	4.6886	2.4809	2.4809	2.6402

Table 4.68. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9562	0.9596	0.9510	0.9506	0.9478	0.9504	0.9508	0.9520	0.9540	0.9546
len	0.0210	0.0061	0.0050	0.0050	3.9800	3.9800	3.9838	2.4526	2.4526	2.4549
vs,0	0.9532	0.9622	0.9992	0.9986	0.9984	0.9986	0.9992	0.9560	0.9586	0.9656
len	0.0210	0.0061	0.0041	0.0041	4.2732	4.2732	4.7026	2.4536	2.4536	2.5328
reg,0.32	0.9484	0.9482	0.9580	0.9522	0.9504	0.9508	0.9508	0.9460	0.9472	0.9472
len	0.0210	0.0100	0.0094	0.0094	3.9815	3.9815	3.9859	2.4524	2.4524	2.4553
vs,0.32	0.9584	0.9584	0.9982	0.9990	0.9952	0.9962	0.9982	0.9588	0.9592	0.9668
len	0.0210	0.0100	0.0078	0.0079	4.2455	4.2455	4.6600	2.4508	2.4508	2.5546
reg,0.9	0.9514	0.9506	0.9516	0.9548	0.9520	0.9526	0.9530	0.9460	0.9472	0.9478
len	0.0210	0.1295	0.1293	0.1292	3.9814	3.9814	3.9857	2.4532	2.4532	2.4563
vs,0.9	0.9578	0.9754	0.9990	0.9992	0.9974	0.9966	0.9984	0.9604	0.9642	0.9686
len	0.0210	0.1291	0.1090	0.1095	4.2384	4.2384	4.6540	2.4506	2.4506	2.5539

Table 4.69. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 8$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9484	0.9374	0.9426	0.9422	0.9404	0.9590	0.9604	0.9864	0.9926	0.9956
len	0.7003	1.2962	1.2953	1.2021	1.9670	1.9670	1.9745	4.3220	4.3220	4.4470
vs,0	0.9562	0.9342	0.9342	0.9968	0.9956	0.9740	0.9960	0.9864	0.9776	0.9852
len	0.6981	1.3138	1.3129	1.0009	2.1170	2.1170	2.3677	4.3665	4.3665	4.4702
reg,0.32	0.9534	0.9518	0.9474	0.9536	0.9502	0.9652	0.9670	0.9876	0.9944	0.9960
len	0.7000	5.6304	5.6340	5.6167	1.9666	1.9666	1.9742	4.3068	4.3068	4.3806
vs,0.32	0.9528	0.9888	0.9898	0.9982	0.9972	0.9550	0.9970	0.9992	0.9992	1.0000
len	0.6895	4.8348	4.8084	4.7459	2.0989	2.0989	2.3186	4.4252	4.4252	4.8336
reg,0.9	0.9478	0.9582	0.9522	0.9464	0.9458	0.9604	0.9626	0.9870	0.9972	0.9986
len	0.6995	87.5408	87.5621	87.5218	1.9670	1.9670	1.9733	4.2979	4.2979	4.3239
vs,0.9	0.9504	0.9695	0.9708	0.9970	0.9954	0.9770	0.9974	0.9994	0.9998	1.0000
len	0.6875	74.0361	73.8999	74.5606	2.0912	2.0912	2.3087	4.5206	4.5206	4.9329

Table 4.70. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 8$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9542	0.9518	0.9532	0.9536	0.9496	0.9486	0.9492	0.9542	0.9548	0.9566
len	0.1426	0.2589	0.2588	0.2432	1.9595	1.9595	1.9615	4.1693	4.1693	4.1797
vs,0	0.9562	0.9538	0.9478	0.9970	0.9964	0.9378	0.9962	0.9510	0.9534	0.9548
len	0.1427	0.2584	0.2586	0.1995	2.1463	2.1463	2.3925	4.1699	4.1699	4.1822
reg,0.32	0.9498	0.9562	0.9526	0.9538	0.9516	0.9526	0.9526	0.9446	0.9476	0.9488
len	0.1426	1.1402	1.14009	1.1359	1.9593	1.9593	1.9614	4.1688	4.1688	4.1763
vs,0.32	0.9488	0.9368	0.9388	0.9984	0.9972	0.9438	0.9972	0.9590	0.9360	0.9546
len	0.1430	1.1785	1.1807	0.9528	2.1496	2.1496	2.3985	4.2727	4.2727	4.3389
reg,0.9	0.9486	0.9522	0.9512	0.9516	0.9494	0.9504	0.9492	0.9488	0.9508	.9520
len	0.1428	17.7430	17.7223	17.7145	1.9583	1.9583	1.9599	4.1684	4.1684	4.1727
vs,0.9	0.9584	0.9756	0.9757	0.9988	0.9982	0.9646	0.9986	0.9968	0.9972	0.9994
len	0.1423	15.0633	15.0496	15.0457	2.1015	2.1015	2.3095	4.4005	4.4005	4.8309

Table 4.71. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 8$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9452	0.9556	0.9536	0.9522	0.9464	0.9468	0.9476	0.9508	0.9510	0.9532
len	0.1754	0.3630	0.3628	0.3608	1.9582	1.9582	1.9602	4.1697	4.1697	4.1898
vs,0	0.9482	0.9520	0.9472	0.9984	0.9978	0.9456	0.9976	0.9478	0.9466	0.9494
len	0.1754	0.3628	0.3623	0.2964	2.1506	2.1506	2.4003	4.1707	4.1707	4.1958
reg,0.32	0.9492	0.9550	0.9534	0.9506	0.9476	0.9472	0.9482	0.9526	0.9546	0.9558
len	0.1754	1.6920	1.6895	1.6903	1.9593	1.9593	1.9611	4.1690	4.1690	4.1851
vs,0.32	0.9500	0.9552	0.9566	0.9974	0.9970	0.9422	0.9968	0.9306	0.9456	0.9156
len	0.1763	1.6871	1.6804	1.4432	2.1270	2.1270	2.3678	4.2223	4.2223	4.4581
reg,0.9	0.9498	0.9558	0.9522	0.9494	0.9444	0.9450	0.9444	0.9482	0.9482	0.9502
len	0.1752	26.2946	26.3124	26.2841	1.9588	1.9588	1.9613	4.16890	4.1689	4.1838
vs,0.9	0.9508	0.9756	0.9490	0.9990	0.9986	0.9592	0.9988	0.9988	0.9984	0.9998
len	0.1751	22.4658	22.4060	22.4655	2.1305	2.1305	2.3508	4.4349	4.4349	4.8415

Table 4.72. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 250$, $p = 10$, $k = 8$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9522	0.9468	0.9540	0.9518	0.9496	0.9492	0.9488	0.9474	0.9464	0.9478
len	0.0210	0.0146	0.0146	0.0142	1.9593	1.9593	1.9609	4.1633	4.1633	4.1675
vs,0	0.9544	0.9546	0.9518	0.9980	0.9966	0.9374	0.9966	0.9534	0.9524	0.9552
len	0.0210	0.0146	0.0146	0.0117	2.1470	2.1470	2.3955	4.1655	4.1655	4.1880
reg,0.32	0.9522	0.9510	0.9486	0.9540	0.9494	0.9504	0.9516	0.9460	0.9468	0.9472
len	0.0210	0.0664	0.0664	0.0663	1.9595	1.9595	1.9614	4.1636	4.1636	4.1684
vs,0.32	0.9508	0.9596	0.9496	0.9992	0.9986	0.9434	0.9986	0.9634	0.9646	0.9696
len	0.0210	0.0663	0.0662	0.0541	2.1434	2.1434	2.3960	4.1970	4.1970	4.2703
reg,0.9	0.9536	0.9580	0.9550	0.9584	0.9538	0.9538	0.9548	0.9496	0.9512	0.9524
len	0.0210	1.0357	1.0361	1.0336	1.9585	1.9585	1.9605	4.1603	4.1603	4.1643
vs,0.9	0.9486	0.9484	0.9492	0.9988	0.9982	0.9492	0.9982	0.9688	0.9546	0.9676
len	0.0212	1.0742	1.0745	0.8793	2.1387	2.1387	2.3860	4.2883	4.2883	4.3818

Table 4.73. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 1$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9538	0.9272	0.9588	0.9522	0.9672	0.9780	0.9796	0.9450	0.9616	0.9712
len	0.4651	0.4130	0.2806	0.2807	4.0318	4.0318	4.0415	2.4595	2.4595	2.5544
vs,0	0.9508	0.9386	0.9978	0.9976	0.9986	0.9986	0.9996	0.9532	0.9624	0.9720
len	0.4608	0.4087	0.2293	0.2305	4.3398	4.3398	4.7841	2.4600	2.4600	2.5599
reg,0.32	0.9528	0.9402	0.9500	0.9468	0.9646	0.9748	0.9774	0.9454	0.9606	0.9658
len	0.4645	0.6073	0.5260	0.5254	4.0340	4.0340	4.0432	2.4624	2.4624	2.5071
vs,0.32	0.9532	0.9498	0.9978	0.9994	0.9982	0.9984	0.9996	0.9534	0.9620	0.9690
len	0.4614	0.6034	0.4417	0.4392	4.3011	4.3011	4.7208	2.4608	2.4608	2.5448
reg,0.9	0.9498	0.9502	0.9546	0.9530	0.9726	0.9802	0.9812	0.9508	0.9600	0.9604
len	0.4642	7.2387	7.2328	7.2325	4.0329	4.0329	4.0420	2.4643	2.4643	2.4684
vs,0.9	0.9548	0.9994	0.9968	0.9994	0.9992	0.9988	0.9998	0.9754	0.9780	0.9836
len	0.4611	6.1604	6.1283	6.1385	4.2960	4.2960	4.7299	2.5421	2.5421	2.7100

Table 4.74. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 1$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9526	0.9528	0.9554	0.9530	0.9468	0.9472	0.9474	0.9500	0.9524	0.9532
len	0.0995	0.0861	0.0599	0.0598	3.9831	3.9831	3.9872	2.4522	2.4522	2.4595
vs,0	0.9544	0.9462	0.9984	0.9982	0.9984	0.9982	0.9994	0.9442	0.9462	0.9476
len	0.0995	0.0859	0.0490	0.0491	4.3095	4.3095	4.7484	2.4528	2.4528	2.4634
reg,0.32	0.9516	0.9536	0.9514	0.9562	0.9476	0.9490	0.9504	0.9464	0.9482	0.9498
len	0.0996	0.1281	0.1122	0.1122	3.9837	3.9837	3.9874	2.4531	2.4531	2.4574
vs,0.32	0.9564	0.9562	0.9972	0.9986	0.9984	0.9984	0.9996	0.9534	0.9550	0.9580
len	0.0995	0.1285	0.0938	0.0941	4.2557	4.2557	4.6831	2.4523	2.4523	2.4991
reg,0.9	0.9506	0.9554	0.9540	0.9566	0.9502	0.9530	0.9534	0.9488	0.9488	0.9490
len	0.0997	1.5449	1.5448	1.5417	3.9831	3.9831	3.9870	2.4526	2.4526	2.4556
vs,0.9	0.9550	0.9710	0.9984	0.9986	0.9948	0.9952	0.9980	0.9414	0.9072	0.9414
len	0.0995	1.5370	1.3017	1.3078	4.2523	4.2523	4.6765	2.4551	2.4551	2.5324

Table 4.75. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 1$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9530	0.9506	0.9524	0.9526	0.9440	0.9446	0.9450	0.9474	0.9472	0.9486
len	0.1223	0.09004	0.0877	0.0875	3.9831	3.9831	3.9871	2.4530	2.4530	2.4650
vs,0	0.9544	0.9584	0.9986	0.9986	0.9980	0.9976	0.9988	0.9532	0.9552	0.9572
len	0.1221	0.0897	0.07219	0.0720	4.3045	4.3045	4.74174	2.4526	2.4526	2.4795
reg,0.32	0.9556	0.9496	0.9540	0.9566	0.9500	0.9508	0.9516	0.9528	0.9536	0.9542
len	0.1222	0.1654	0.1644	0.1644	3.9812	3.9812	3.9854	2.4534	2.4534	2.4628
vs,0.32	0.9552	0.9558	0.9992	0.9986	0.9974	0.9978	0.9992	0.9558	0.9572	0.9624
len	0.1221	0.1659	0.1374	0.1374	4.2615	4.2615	4.6853	2.4516	2.4516	2.5248
reg,0.9	0.9486	0.9552	0.9556	0.9490	0.9516	0.9520	0.9516	0.9468	0.9496	0.9508
len	0.1221	2.2607	2.2669	2.2586	3.9804	3.9804	3.9848	2.4509	2.4509	2.4595
vs,0.9	0.9554	0.9988	0.9982	0.9982	0.9976	0.9974	0.9998	0.9486	0.9478	0.9626
len	0.1222	2.0697	1.9198	1.9238	4.2353	4.2353	4.6730	2.4416	2.4416	2.5675

Table 4.76. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 1$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9510	0.9538	0.9532	0.9518	0.9492	0.9496	0.9508	0.9502	0.9484	0.9482
len	0.0141	0.0038	0.0032	0.0032	3.9798	3.9798	3.9838	2.4524	2.4524	2.4552
vs,0	0.9562	0.9626	0.9982	0.9984	0.9964	0.9960	0.9984	0.9604	0.9602	0.9676
len	0.0141	0.0038	0.0026	0.0026	4.2852	4.2852	4.7196	2.4547	2.4547	2.5141
reg,0.32	0.9578	0.9596	0.9488	0.9524	0.9478	0.9472	0.9478	0.9564	0.9558	0.9562
len	0.0141	0.0064	0.0060	0.0060	3.9796	3.9796	3.9839	2.4526	2.4526	2.4552
vs,0.32	0.9524	0.9630	0.9994	0.9988	0.9980	0.9976	0.9992	0.9562	0.9568	0.9658
len	0.01413	0.0064	0.0050	0.0051	4.2504	4.2504	4.6675	2.4518	2.4518	2.5450
reg,0.9	0.9542	0.9558	0.9584	0.9582	0.9492	0.9496	0.9508	0.9508	0.9506	0.9496
len	0.0141	0.0835	0.0831	0.0832	3.9813	3.9813	3.9856	2.4527	2.4527	2.4551
vs,0.9	0.9562	0.9758	0.9990	0.9992	0.9964	0.9964	0.9992	0.9590	0.9594	0.9654
len	0.0141	0.0835	0.0700	0.0707	4.2465	4.2465	4.6636	2.4508	2.4508	2.5425

Table 4.77. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 8$, and $m = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9538	0.9440	0.9460	0.9472	0.9460	0.9540	0.9540	0.9720	0.9788	0.9832
len	0.4645	0.8490	0.8490	0.7938	1.9627	1.9627	1.9661	4.2335	4.2335	4.2981
vs,0	0.9544	0.9492	0.9518	0.9972	0.9962	0.9542	0.9962	0.9828	0.9900	0.9924
len	0.4642	0.8493	0.8510	0.6577	2.1467	2.1467	2.3943	4.2601	4.2601	4.3244
reg,0.32	0.9484	0.9530	0.9460	0.9526	0.9488	0.9554	0.9556	0.9646	0.9732	0.9762
len	0.4647	3.7136	3.7238	3.7066	1.9629	1.9629	1.9662	4.2270	4.2270	4.2633
vs,0.32	0.9482	0.9200	0.9256	0.9980	0.9970	0.9436	0.9970	0.9908	0.9788	0.9952
len	0.4625	3.2758	3.2683	3.1466	2.1063	2.1063	2.3407	4.3015	4.3015	4.6969
reg,0.9	0.9546	0.9516	0.9508	0.9562	0.9524	0.9594	0.9596	0.9698	0.9800	0.9808
len	0.4648	57.8391	57.7643	57.8039	1.9613	1.9613	1.9639	4.2203	4.2203	4.2299
vs,0.9	0.9492	0.9756	0.9704	0.9988	0.9984	0.9610	0.9986	0.9986	0.9988	0.9998
len	0.4611	49.1314	49.0763	49.2604	2.1079	2.1079	2.3291	4.4702	4.4702	4.8812

Table 4.78. Bootstrapping Binomial Logistic Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 8$, and $m = 20$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9534	0.9536	0.9576	0.9578	0.9506	0.9516	0.9518	0.9498	0.9502	0.9510
len	0.0995	0.1802	0.1802	0.1691	1.9589	1.9589	1.9605	4.1660	4.1660	4.1731
vs,0	0.9510	0.9510	0.9458	0.9986	0.9970	0.9416	0.9970	0.9518	0.9544	0.9552
len	0.0995	0.1804	0.1803	0.1390	2.1515	2.1515	2.3993	4.1646	4.1646	4.1726
reg,0.32	0.9570	0.9518	0.9558	0.9530	0.9480	0.9482	0.9492	0.9548	0.9570	0.9566
len	0.0994	0.7929	0.7938	0.7907	1.9592	1.9592	1.9616	4.1644	4.1644	4.1703
vs,0.32	0.9528	0.9506	0.9472	0.9982	0.9974	0.9414	0.9974	0.9774	0.9802	0.9822
len	0.0997	0.7995	0.7991	0.6510	2.1384	2.1384	2.3935	4.2160	4.2160	4.2603
reg,0.9	0.9544	0.9558	0.9538	0.9512	0.9438	0.9446	0.9456	0.9530	0.9544	0.9554
len	0.0995	12.3362	12.3417	12.3316	1.9584	1.9584	1.9607	4.1637	4.1637	4.1685
vs,0.9	0.9526	0.9794	0.9795	0.9988	0.9980	0.9442	0.9986	0.9984	0.9982	0.9996
len	0.0995	10.5014	10.5028	10.4758	2.1139	2.1139	2.3534	4.4007	4.4007	4.8202

Table 4.79. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 8$, $a = 1$, $\beta_1 = 1$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9536	0.9532	0.9558	0.9498	0.9478	0.9478	0.9474	0.9524	0.9544	0.9560
len	0.1222	0.2487	0.2485	0.2478	1.9580	1.9580	1.9602	4.1648	4.1648	4.1771
vs,0	0.9532	0.9576	0.9576	0.9982	0.9976	0.9424	0.9976	0.9472	0.9486	0.9510
len	0.1221	0.2484	0.2484	0.2030	2.1365	2.1365	2.3779	4.1656	4.1656	4.1808
reg,0.32	0.9492	0.9528	0.9548	0.9490	0.9422	0.9426	0.9432	0.9522	0.9524	0.9548
len	0.1223	1.1606	1.15916	1.1605	1.9598	1.9598	1.9613	4.1676	4.1676	4.1774
vs,0.32	0.9488	0.9388	0.9360	0.9986	0.9982	0.9490	0.9982	0.9552	0.9262	0.9470
len	0.1223	1.2037	1.1997	0.9697	2.1407	2.1407	2.3882	4.2821	4.2821	4.3774
reg,0.9	0.9512	0.9568	0.9526	0.9526	0.9508	0.9488	0.9496	0.9468	0.9466	0.9494
len	0.1223	18.0816	18.1049	18.0943	1.9581	1.9581	1.9604	4.1644	4.1644	4.1740
vs,0.9	0.9558	0.9764	0.9769	0.9984	0.9976	0.9658	0.9978	0.9982	0.9980	0.9994
len	0.1221	15.3179	15.4078	15.3875	2.1044	2.1044	2.3140	4.4053	4.4053	4.8336

Table 4.80. Bootstrapping Poisson Regression, Backward Elimination with AIC $B = 500$, $n = 500$, $p = 10$, $k = 8$, $a = 2$, $\beta_1 = 5$

ψ	β_1	β_2	β_{p-1}	β_p	pr0	hyb0	br0	pr1	hyb1	br1
reg,0	0.9464	0.9478	0.9508	0.9544	0.9480	0.9490	0.9488	0.9508	0.9512	0.9538
len	0.0141	0.0093	0.0093	0.0091	1.9599	1.9599	1.9616	4.1627	4.1627	4.1666
vs,0	0.9478	0.9552	0.9552	0.9978	0.9966	0.9494	0.9966	0.9522	0.9534	0.9540
len	0.0141	0.0093	0.0093	0.0074	2.1477	2.1477	2.4100	4.1634	4.1634	4.1816
reg,0.32	0.9498	0.9490	0.9518	0.9510	0.9436	0.9448	0.9442	0.9458	0.9452	0.9464
len	0.0141	0.0429	0.0427	0.0427	1.9596	1.9596	1.9617	4.1607	4.1607	4.1656
vs,0.32	0.9510	0.9536	0.9546	0.9994	0.9992	0.9420	0.9992	0.9610	0.9624	0.9670
len	0.0141	0.0427	0.0427	0.0351	2.1533	2.1533	2.3986	4.1987	4.1987	4.2701
reg,0.9	0.9556	0.9508	0.9544	0.9566	0.9510	0.9534	0.9530	0.9484	0.9478	0.9486
len	0.0141	0.6657	0.6670	0.6673	1.9589	1.9589	1.9610	4.1621	4.1621	4.1670
vs,0.9	0.9538	0.9534	0.9576	0.9992	0.9990	0.9408	0.9988	0.9716	0.9722	0.9780
len	0.0141	0.6686	0.6681	0.5469	2.1483	2.1483	2.3925	4.2204	4.2204	4.2986

For multiple linear regression and multivariate linear regression, the residual bootstrap likely outperforms the parametric bootstrap since the residual bootstrap produces $\boldsymbol{Y}^*$ that is closer to $\boldsymbol{Y}$ unless the error distribution is Gaussian. To bootstrap the multivariate linear regression model with $\boldsymbol{\theta} = \boldsymbol{A}vec(\boldsymbol{B})$, make the confidence region using $\hat{\boldsymbol{\theta}}^*_j = \boldsymbol{A}vec(\hat{\boldsymbol{B}}^*_{I_{min},0,j})$.

Following Claeskens and Hjort (2008, pp. 27-28), if $S \subseteq I_j$, the GLM estimator $\hat{\boldsymbol{\beta}}_{I_j}$ satisfies $\sqrt{n}(\hat{\boldsymbol{\beta}}_{I_j} - \boldsymbol{\beta}_{I_j}) \xrightarrow{D} N_{a_j}(\boldsymbol{0}, \boldsymbol{V}_j)$ where $\boldsymbol{V}_j = \boldsymbol{I}^{-1}(\boldsymbol{\beta}_{I_j})$ if the GLM is valid, and $\boldsymbol{V}_j = \boldsymbol{J}_j^{-1}\boldsymbol{I}^{-1}(\boldsymbol{\beta}_{I_j})\boldsymbol{J}_j^{-1}$ under regularity conditions if the GLM mean function $E(Y|\boldsymbol{x}_{I_j}) = E(Y|\boldsymbol{x})$ is correct but there is overdispersion, e.g., the Poisson regression model is fit but the data follow an NBR model. Then $E(Y|\boldsymbol{x}_{I_j}) = \exp(\boldsymbol{x}_{I_j}^T\boldsymbol{\beta}_{I_j}) = \exp(\boldsymbol{x}^T\boldsymbol{\beta}) = E(Y|\boldsymbol{x})$.

Some bootstrap methods, such as the nonparametric bootstrap, may still give good results when there is overdispersion. The parametric bootstrap is likely affected by overdispersion. Hence check the regression model with plots described in Olive (2013b). Compare the results for the binomial regression model with the results from the BBR model, and the results for the Poisson regression model with the results from the NBR model.

Note that there are several important variable selection models, including the model given by Equation (1.6). Another model is $\boldsymbol{x}^T\boldsymbol{\beta} = \boldsymbol{x}_{S_i}^T\boldsymbol{\beta}_{S_i}$ for $i = 1, ..., K$. Then there are $K \geq 2$ competing "true" nonnested submodels where $\boldsymbol{\beta}_{S_i}$ is $a_{S_i} \times 1$. For example, suppose the $K = 2$ models have predictors x_1, x_2, x_3 for S_1 and x_1, x_2, x_4 for S_2. Then x_3 and x_4 are likely to be selected and omitted often by forward selection for the B bootstrap samples. Hence omitting all predictors x_i that have a $\beta_{ij}^* = 0$ for at least one of the bootstrap samples $j = 1, ..., B$ could result in underfitting, e.g. using just x_1 and x_2 in the above $K = 2$ example. If n and B are large enough, the singleton set $\{\boldsymbol{0}\}$ could still be the "100%" confidence region for a vector $\boldsymbol{\beta}_O$.

For some data sets, $\boldsymbol{S}_T^*$ may be singular due to one or more columns of zeroes in the bootstrap sample for $\beta_1, ..., \beta_p$. The variables corresponding to these columns are likely not needed in the model given that the other predictors are in the model. Confidence intervals can be computed without $\boldsymbol{S}_T^*$ for (2.5), (2.6), and (2.7). Also see Efron (2014).

There is a massive literature on variable selection and a fairly large literature for inference after variable selection. See, for example, Leeb and Pötscher (2006, 2008), Leeb, Pötscher, and Ewald (2015), and Tibshirani, Rinaldo, Tibshirani, and Wasserman (2018).

REFERENCES

[1] Agresti, A. (2002), *Categorical Data Analysis*, Wiley, Hoboken, NJ.

[2] Akaike, H. (1973), "Information Theory as an Extension of the Maximum Likelihood Principle," *Proceedings, 2nd International Symposium on Information Theory*, eds. Petrov, B.N., and Csakim, F., Akademiai Kiado, Budapest, 267-281.

[3] Bickel, P.J., and Ren, J.–J. (2001), "The Bootstrap in Hypothesis Testing," *State of the Art in Probability and Statistics: Festschrift for William R. van Zwet*, eds. de Gunst, M., Klaassen, C., and van der Vaart, A., The Institute of Mathematical Statistics, Hayward, CA, 91-112.

[4] Breiman, L. (1996), "Bagging Predictors," *Machine Learning*, 24, 123-140.

[5] Cai, T., Tian, L., Solomon, S.D., and Wei, L.J. (2008), "Predicting Future Responses Based on Possibly Misspecified Working Models," *Biometrika*, 95, 75-92.

[6] Charkhi, A., and Claeskens, G. (2018), "Asymptotic Post-Selection Inference for the Akaike Information Criterion," *Biometrika*, 105, 645-664..

[7] Chatterjee., S., and Bose, A. (2005), "Generalized Bootstrap for Estimating Equations," *The Annals of Statistics*, 33, 414-436.

[8] Chen, J., and Chen, Z. (2008), "Extended Bayesian Information Criterion for Model Selection With Large Model Spaces," *Biometrika*, 95, 75-92.

[9] Claeskens, G., and Hjort, N.L. (2008), *Model Selection and Model Averaging*, Cambridge University Press, New York, NY.

[10] Cook, R.D., and Zhang, X. (2015), "Foundations of Envelope Models and Methods," *Journal of the American Statistical Association*, 110, 599-611.

[11] Davison, A.C., and Hinkley, D.V. (1997), *Bootstrap Methods and Their Application*, Cambridge University Press, Cambridge, UK.

[12] Eck, D.J. (2018), "Bootstrapping for Multivariate Linear Regression Models," *Statistics & Probability Letters*, 134, 141-149.

[13] Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans,* SIAM, Philadelphia, PA.

[14] Efron, B. (2014), "Estimation and Accuracy After Model Selection," (with discussion), *Journal of the American Statistical Association,* 109, 991-1007.

[15] Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004), "Least Angle Regression," (with discussion), *The Annals of Statistics,* 32, 407-451.

[16] Freedman, D.A. (1981), "Bootstrapping Regression Models," *The Annals of Statistics*, 9, 1218-1228.

[17] Frey, J. (2013), "Data-Driven Nonparametric Prediction Intervals," *Journal of Statistical Planning and Inference*, 143, 1039-1048.

[18] Friedl, H. (1997), "Resampling Methods in Generalized Linear Models Useful in Environmetrics," *Econometrika*, 8, 444-457.

[19] Friedman, J., Hastie, T., Hoefling, H., and Tibshirani, R. (2007), "Pathwise Coordinate Optimization," *Annals of Applied Statistics*, 1, 302332.

[20] Friedman, J., Hastie, T., Simon, N., and Tibshirani, R. (2015), *glmnet: Lasso and Elastic-net Regularized Generalized Linear Models*, *R* Package version 2.0, (http://cran.r-project.org/package=glmnet).

[21] Friedman, J., Hastie, T., and Tibshirani, R. (2010), "Regularization Paths for Generalized Linear Models Via Coordinate Descent," *Journal of Statistical Software*, 33, 1-22.

[22] Guo, C., Yang, H., and Lv, J. (2017), "Robust Variable Selection for Generalized Linear Models With a Diverging Number of Parameters," *Communications in Statistics: Theory and Methods*, 46, 2967-2981.

[23] Hall, P. (1988), "Theoretical Comparisons of Bootstrap Confidence Intervals," (with discussion), *The Annals of Statistics*, 16, 927-985.

[24] Hastie, T.J., and Tibshirani, R.J. (1986), "Generalized Additive Models" (with discussion)," *Statistical Science*, 1, 297-318.

[25] Hastie, T.J., and Tibshirani, R.J. (1990), *Generalized Additive Models*, Chapman & Hall, London, UK.

[26] Hastie, T., Tibshirani, R., and Wainwright, M. (2015), *Statistical Learning with Sparsity: the Lasso and Generalizations*, CRC Press Taylor & Francis, Boca Raton, FL.

[27] Haughton, D.M.A. (1988), "On the Choice of a Model to Fit Data From an Exponential Family," *The Annals of Statistics*, 16, 342-355.

[28] Haughton, D. (1989), "Size of the Error in the Choice of a Model to Fit Data From an Exponential Family," *Sankhyā, A,* 51, 45-58.

[29] Hillis, S.L., and Davis, C.S. (1994), "A Simple Justification of the Iterative Fitting Procedure for Generalized Linear Models," *The American Statistician*, 48, 288-289.

[30] Leeb, H., and Pötscher, B.M. (2006), "Can One Estimate the Conditional Distribution of Post-Model-Selection Estimators?" *The Annals of Statistics*, 34, 2554-2591.

[31] Leeb, H. and Pötscher, B.M. (2008), "Can One Estimate the Unconditional Distribution of Post-Model-Selection Estimators?" *Econometric Theory*, 24, 338-376.

[32] Leeb, H., Pötscher, B.M., and Ewald, K. (2015), "On Various Confidence Intervals Post-Model-Selection," *Statistical Science*, 30, 216-227.

[33] Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J., and Wasserman, L. (2017), "Distribution-Free Predictive Inference for Regression," *Journal of the American Statistical Association*, to appear.

[34] Lesnoff, M., and Lancelot, R. (2010), "aod: Analysis of Overdispersed Data," R package version 1.2, (http://cran.r-project.org/package=aod).

[35] McCullagh, P., and Nelder, J.A. (1989), *Generalized Linear Models*, 2nd ed., Chapman & Hall, London, UK.

[36] Meinshausen, N. (2007), "Relaxed Lasso," *Computational Statistics & Data Analysis*, 52, 374-393.

[37] Moulton, L.H., and Zeger, S.L. (1991), "Bootstrapping Generalized Linear Models," *Computational Statistics & Data Analysis,* 11, 53-63.

[38] Nelder, J.A., and Wedderburn, R.W.M. (1972), "Generalized Linear Models," *Journal of the Royal Statistical Society, A*, 135, 370-380.

[39] Olive, D.J., and Hawkins, D.M. (2005), "Variable Selection for 1D Regression Models," *Technometrics*, 47, 43-50.

[40] Olive, D.J. (2007), "Prediction Intervals for Regression Models," *Computational Statistics & Data Analysis*, 3115-3122.

[41] Olive, D.J. (2010), *Multiple Linear and 1D Regression Models*, online course notes, see (http://lagrange.math.siu.edu/Olive/regbk.htm).

[42] Olive, D.J. (2013a), "Asymptotically Optimal Regression Prediction Intervals and Prediction Regions for Multivariate Data," *International Journal of Statistics and Probability*, 2, 90-100.

[43] Olive, D.J. (2013b), "Plots for Generalized Additive Models," *Communications in Statistics: Theory and Methods*, 42, 2610-2628.

[44] Olive, D.J. (2014), *Statistical Theory and Inference*, Springer, New York, NY.

[45] Olive, D.J. (2017a), *Linear Regression*, Springer, New York, NY.

[46] Olive, D.J. (2017b), *Robust Multivariate Analysis*, Springer, New York, NY.

[47] Olive, D.J. (2018), "Applications of Hyperellipsoidal Prediction Regions," *Statistical Papers*, 59, 913-931.

[48] Olive, D.J. (2019), *Prediction and Statistical Learning*, online course notes, see (http://lagrange.math.siu.edu/Olive/slearnbk.htm).

[49] Olive, D.J., and Hawkins, D.M. (2005), "Variable Selection for 1D Regression Models," *Technometrics*, 47, 43-50.

[50] Pelawa Watagoda, L.C.R., and Olive, D. J. (2019a), "Comparing Shrinkage Estimators With Asymptotically Optimal Prediction Intervals." Unpublished manuscript at (http://lagrange.math.siu.edu/Olive/pppicomp.pdf).

[51] Pelawa Watagoda, L.C.R., and Olive, D.J. (2019b), "Bootstrapping Multiple Linear Regression After Variable Selection." Unpublished manuscript at

(http://lagrange.math.siu.edu/Olive/ppboottest.pdf).

[52] R Core Team (2016), *R: a Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, (www.R-project.org).

[53] Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics,* 6, 461-464.

[54] Sen, P.K., and Singer, J.M. (1993), *Large Sample Methods in Statistics: an Introduction with Applications,* Chapman & Hall, New York, NY.

[55] Serfling, R.J. (1980), Approximation Theorems of Mathematical Statistics, Wiley, New York, NY.

[56] Shao, J., and Tu, D.S. (1995), *The Jackknife and the Bootstrap*, Springer, New York, NY.

[57] Simonoff, J.S. (2003), *Analyzing Categorical Data*, Springer, New York, NY.

[58] Simonoff, J.S., and Tsai, C.L. (1988), "Jacknifing and Bootstrapping Quasi-Likelihood Estimators," *Journal of Statistical Computation and Simulation,* 30, 213-232.

[59] Su, Z., and Cook, R.D. (2012), "Inner Envelopes: Efficient Estimation in Multivariate Linear Regression," *Biometrika*, 99, 687-702.

[60] Tibshirani, R. (1996), "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society, B*, 58, 267-288.

[61] Tibshirani, R.J., Rinaldo, A., Tibshirani, R., and Wasserman, L. (2018), "Uniform Asymptotic Inference and the Bootstrap After Model Selection," *The Annals of Statistics*, 46, 1255-1287.

[62] Venables, W.N., and Ripley, B.D. (1997), *Modern Applied Statistics with S*, 2nd ed. Springer, New York, NY.

[63] Venables, W.N., and Ripley, B.D. (2010), *Modern Applied Statistics with S*, Springer, New York, NY.

[64] Wang, L., Liu, X., Liang, H., and Carroll, R.J. (2011), "Estimation and Variable Selection for Generalized Additive Partial Linear Models," *The Annals of Statistics*,

39, 1827-1851.

[65] Wood, S.N. (2006), *Generalized Additive Models: an Introduction with R*, Chapman & Hall/CRC, Boca Rotan, FL.

[66] Yang, Y. (2003), "Regression with Multiple Candidate Models: Selecting or Mixing?" *Statistica Sinica*, 13, 783-809.

[67] Zou, H., and Hastie, T. (2005), "Regularization and Variable Selection Via the Elastic Net," *Journal of the Royal Statistical Society Series, B,* 67, 301-320.

[68] Zhou, M. (2001), "Understanding the Cox Regression Models with Time–Change Covariates," *The American Statistician*, 55, 153-155.

VITA

Graduate School
Southern Illinois University

Rasanji Chathumali Rathnayake

rasanjic@gmail.com

Department of Mathematics, 1245 Lincoln Drive, MC 4408
Southern Illinois University
Carbondale, IL 62901, USA.

Southern Illinois University
Master of Science, Mathematics, December 2010

University of Sri Jayewardenapura, Nugegoda
Bachelor of Science, Mathematics (Special), May 2005

Dissertation Title:
Inference For Some GLMs and Survival Regression Models after Variable Selection

Major Professor: Dr. D. J. Olive

Publications:

1. Bootstrapping Some GLMs, and Survival Regression Models After variable Selection, Work in Progress, with David J. Olive.

2. Prediction Intervals for some GLMs, GAMs, and Survival Regression Models, Work in Progress, with David J. Olive, Mulubrhan Haile.