DATA SPLITTING INFERENCE

by

Lingling Zhang

B.A., Southern Illinois University Carbondale, 2017

M.S., Southern Illinois University Carbondale, 2019

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

School of Mathematical and Statistical Sciences
in the Graduate School
Southern Illinois University Carbondale
August, 2022

DISSERTATION APPROVAL

DATA SPLITTING INFERENCE

by

Lingling Zhang

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

Bhaskar Bhattacharya

Dashun Xu

Jianhong Xu

Lingguo Bu

Graduate School
Southern Illinois University Carbondale
June 30, 2022

AN ABSTRACT OF THE DISSERTATION OF

Lingling Zhang, for the Doctor of Philosophy degree in Statistics, presented on June,30,2022, at Southern Illinois University Carbondale.

TITLE: DATA SPLITTING INFERENCE

MAJOR PROFESSOR: Dr.Olive David

Data splitting divides the training data set into two sets H and the validation set V. Data splitting can give valid inference for complicated data sets. We suggest a data splitting prediction region and a sequential data splitting method for lasso variable selection that can be used for multiple linear regression, Poisson regression, Binomial regression, and the Cox proportional hazards regression model with right censored survival times.

ACKNOWLEDGMENTS

I would like to thank Dr.Olive for his invaluable assistance and insights leading to the writing of this paper. My sincere thanks also goes to the five members of my graduate committee for their patience and understanding during the three years of effort that went into the production of this paper.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This section reviews data splitting, regression methods, and variable selection. Data splitting divides the training data set of n cases into two sets H and the validation set V where H has n_H of the cases and V has the remaining $n_V = n - n_H$ cases $i_1, ..., i_{n_V}$. A common method of data splitting randomly divides the training data into the two sets H and V. Often $n_H \approx \lceil n/2 \rceil$ where $\lceil x \rceil$ is the ceiling function = least integer function, e.g $\lceil 7.7 \rceil = 8$.

An application of data splitting is to use a variable selection method, such as forward selection or lasso, on H to get the a predictors. On the validation set V, use the standard regression inference from regressing the response on the a predictors found from H. The standard regression methods include least squares (OLS) for multiple linear regression, a generalized linear model (GLM) such as binomial logistic regression or Poisson regression, or the Cox (1972) proportional hazards regression model.

The above regression models have a response variable Y that is independent of the $p \times 1$ vector of predictors $\boldsymbol{x}$ given $\boldsymbol{x}^T\boldsymbol{\beta}$, written $Y \perp\!\!\!\perp \boldsymbol{x}|\boldsymbol{x}^T\boldsymbol{\beta}$. Then there are n cases $(Y_i, \boldsymbol{x}_i)$. Two important quantities for a regression model are the sufficient predictor $SP = \boldsymbol{x}^T\boldsymbol{\beta}$, and the estimated sufficient predictor $ESP = \boldsymbol{x}^T\hat{\boldsymbol{\beta}}$. For the regression models, the conditioning and subscripts, such as i, will often be suppressed. The multiple linear regression model is $Y|\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{\beta} + \boldsymbol{e}$ or $Y_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + e_i$ for $i = 1, ..., n$. Consider a parametric regression model $Y|\boldsymbol{x} \sim D(\boldsymbol{x}^T\boldsymbol{\beta}, \boldsymbol{\gamma})$ where D is a parametric distribution that depends on the $p \times 1$ vector of predictors $\boldsymbol{x}$ only through $\boldsymbol{x}^T\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ is a $q \times 1$ vector of parameters. Three examples used in the simulations follow. The *binomial logistic regression model* is $Y_i \sim \text{binomial}\left(\text{m}_\text{i}, \rho(\text{SP}) = \dfrac{\text{e}^{\text{SP}}}{1 + \text{e}^{\text{SP}}}\right)$. The binary logistic regression model has $m_i \equiv 1$ for $i = 1, ..., n$. A useful *Poisson regression model* is $Y \sim \text{Poisson}\left(\text{e}^{\text{SP}}\right)$. The *Weibull proportional hazards regression model* is

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

where Y has a Weibull $W(\gamma, \lambda)$ distribution if the probability density function of Y is

$$f(y) = \lambda \gamma y^{\gamma - 1} \exp[-\lambda y^{\gamma}] \text{ for } y > 0.$$

Forward selection or backward elimination with the Akaike (1973) AIC criterion or Schwarz (1978) BIC criterion are often used for variable selection when $n \geq 10p$. The Mallows (1973) C_p criterion can be used for multiple linear regression. When n/p is not large, the Chan and Chen (2008) and Luo and Chen (2013) EBIC criterion with forward selection can be useful.

Some shrinkage methods do variable selection: the regression method, such as a GLM, uses the predictors that had nonzero shrinkage estimator coefficients. These methods include least angle regression, lasso, relaxed lasso, and elastic net. Least angle regression variable selection is the LARS-OLS hybrid estimator of Efron, Hastie, Johnstone, and Tibshirani (2004, p. 421). Lasso variable selection is called relaxed lasso by Hastie, Tibshirani, and Wainwright (2015, p. 12), and the relaxed lasso estimator with $\phi = 0$ by Meinshausen (2007, p. 376). Also see Friedman et al. (2007), Friedman, Hastie, and Tibshirani (2010), Simon, Friedman, Hastie, and Tibshirani (2011), Tibshirani (1996), and Zou and Hastie (2005). The Meinshausen (2007) relaxed lasso estimator fits lasso with penalty λ_n to get a subset of variables with nonzero coefficients, and then fits lasso with a smaller penalty ϕ_n to this subset of variables where n is the sample size. These methods can be useful even when n/p is not large, if the model is sparse: the number of predictors with nonzero coefficients is small.

Following Olive and Hawkins (2005), a *model for variable selection* can be described by

$$\boldsymbol{x}^T \boldsymbol{\beta} = \boldsymbol{x}_S^T \boldsymbol{\beta}_S + \boldsymbol{x}_E^T \boldsymbol{\beta}_E = \boldsymbol{x}_S^T \boldsymbol{\beta}_S \tag{1.1}$$

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 (1.1) holds. Then

$$\boldsymbol{x}^T\boldsymbol{\beta} = \boldsymbol{x}_S^T\boldsymbol{\beta}_S = \boldsymbol{x}_I^T\boldsymbol{\beta}_I + \boldsymbol{x}_O^T\boldsymbol{0} = \boldsymbol{x}_I^T\boldsymbol{\beta}_I.$$

Thus $\boldsymbol{\beta}_O = \boldsymbol{0}$ if $S \subseteq I$. The model using $\boldsymbol{x}^T\boldsymbol{\beta}$ is the full model.

To clarify notation, suppose $p = 4$, a constant $x_1 = 1$ corresponding to β_1 is always in the model, and $\boldsymbol{\beta} = (\beta_1, \beta_2, 0, 0)^T$. Then the $J = 2^{p-1} = 8$ possible subsets of $\{1, 2, ..., p\}$ that always contain 1 are $I_1 = \{1\}$, $S = I_2 = \{1, 2\}$, $I_3 = \{1, 3\}$, $I_4 = \{1, 4\}$, $I_5 = \{1, 2, 3\}$, $I_6 = \{1, 2, 4\}$, $I_7 = \{1, 3, 4\}$, and $I_8 = \{1, 2, 3, 4\}$. There are $2^{p-a_S} = 4$ subsets I_2, I_5, I_6, and I_8 such that $S \subseteq I_j$. Let $\hat{\boldsymbol{\beta}}_{I_7} = (\hat{\beta}_1, \hat{\beta}_3, \hat{\beta}_4)^T$ and $\boldsymbol{x}_{I_7} = (x_1, x_3, x_4)^T$.

Let I_{min} correspond to the set of predictors selected by a variable selection method such as forward selection or lasso variable selection. If $\hat{\boldsymbol{\beta}}_I$ is $a \times 1$, use zero padding to form the $p \times 1$ vector $\hat{\boldsymbol{\beta}}_{I,0}$ from $\hat{\boldsymbol{\beta}}_I$ by adding 0s corresponding to the omitted variables. For example, if $p = 4$ and $\hat{\boldsymbol{\beta}}_{I_{min}} = (\hat{\beta}_1, \hat{\beta}_3)^T$, then the observed variable selection estimator $\hat{\boldsymbol{\beta}}_{VS} = \hat{\boldsymbol{\beta}}_{I_{min},0} = (\hat{\beta}_1, 0, \hat{\beta}_3, 0)^T$. As a statistic, $\hat{\boldsymbol{\beta}}_{VS} = \hat{\boldsymbol{\beta}}_{I_k,0}$ with probabilities $\pi_{kn} = P(I_{min} = I_k)$ for $k = 1, ..., J$ where there are J subsets, e.g. $J = 2^p - 1$.

Theory for the variable selection estimator $\hat{\boldsymbol{\beta}}_{VS}$ is complicated. See Pelawa Watagoda and Olive (2021a, 2021b) for multiple linear regression, and Rathnayake and Olive (2021) for models such as GLMs and Cox proportional hazards regression. For fixed p, these three papers show that $\hat{\boldsymbol{\beta}}_{VS}$ is $\sqrt{n}$ consistent with a complicated nonnormal limiting distribution, and suggest a method for bootstrapping $\hat{\boldsymbol{\beta}}_{VS}$.

Data splitting is useful for inference after variable selection provided $S \subseteq I_{min}$. Use H to select a model I_{min}, then fit the selected model to the cases in the validation set V using standard inference. See, for example, Cox (1975), Faraway (1995), Moran (1973), Picard and Berk (1990), and Rinaldo et al. (2019). Hurvich and Tsai (1990) note that for multiple linear regression, n_H can be much smaller that $n/2$, and then the remaining n_V cases in V can be used for inference.

Section 2 gives a data splitting prediction region for multivariate data. Section 3 gives a sequential method of data splitting for regression models that uses lasso.

CHAPTER 2

A DATA SPLITTING PREDICTION REGION

Consider predicting a future test value $\boldsymbol{x}_f$, given past training data $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ where $\boldsymbol{x}_1, ..., \boldsymbol{x}_n, \boldsymbol{x}_f$ are iid. A large sample $100(1-\delta)\%$ prediction region is a set $\mathcal{A}_n$ such that $P(\boldsymbol{x}_f \in \mathcal{A}_n)$ is eventually bounded below by $1 - \delta$ as $n \to \infty$. A prediction region is asymptotically optimal if its volume converges in probability to the volume of the minimum volume covering region or the highest density region of the distribution of $\boldsymbol{x}_f$.

Prediction regions often use an estimator of multivariate location and dispersion $(T, \boldsymbol{C})$. For example, let a) $(T, \boldsymbol{C}) = (\overline{\boldsymbol{x}}, \boldsymbol{S})$ be the sample mean and sample covariance matrix, b) $(T, \boldsymbol{C}) = (\text{MED}(\boldsymbol{W}), \boldsymbol{I}_p)$ where MED($\boldsymbol{W}$) is the coordinatewise median of the $\boldsymbol{x}_i$ and $\boldsymbol{I}_p$ is the $p \times p$ identity matrix, or c) let $(T, \boldsymbol{C})$ be a robust estimator such as the RMVN estimator $(T_{RMVN}, \boldsymbol{C}_{RMVN})$ given by Olive (2017), Olive and Hawkins (2010), and Zhang, Olive, and Ye (2012). Olive (2013a) developed a large sample prediction region using $(\overline{\boldsymbol{x}}, \boldsymbol{S})$ if the data distribution has a nonsingular population covariance matrix, and a large sample prediction region using $(T_{RMVN}, \boldsymbol{C}_{RMVN})$ if the data distribution comes from a large class of elliptically contoured distributions. Also see Olive (2018). If $n \geq 20p$, using $(T, \boldsymbol{C}) = (T_{RMVN}, \boldsymbol{C}_{RMVN})$ might result in a prediction region with smaller volume then using $(T, \boldsymbol{C}) = (\overline{\boldsymbol{x}}, \boldsymbol{S})$ since the robust estimator attempts to estimate a small volume hyperellipsoid. The smaller volume can also occur if outliers are present or if the data distribution is highly skewed, and the new data splitting prediction region given in this section does not need the elliptically contoured distribution assumption.

Data splitting divides the training data $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ into two sets H and V where H has n_H of the cases and V has the remaining $n_V = n - n_H$ cases $i_1, ..., i_{n_V}$. Often $n_h \approx n/2$. The estimator $(T_H, \boldsymbol{C}_H)$ is computed using the data set H. Then the squared validation distances $D_j^2 = D^2_{\boldsymbol{x}_{i_j}}(T_H, \boldsymbol{C}_H) = (\boldsymbol{x}_{i_j} - T_H)^T \boldsymbol{C}_H^{-1} (\boldsymbol{x}_{i_j} - T_H)$ are computed for the $j = 1, ..., n_V$ cases in the validation set V. Let $D^2_{(U_V)}$

be the U_Vth order statistic of the D_j^2 where

$$U_V = \min(n_V, \lceil (n_V + 1)(1 - \delta) \rceil). \tag{2.1}$$

The large sample $100(1-\delta)\%$ data splitting prediction region for $\boldsymbol{x}_f$ is

$$\{\boldsymbol{z} : D^2_{\boldsymbol{z}}(T_H, \boldsymbol{C}_H) \le D^2_{(U_V)}\}. \tag{2.2}$$

To show that (2.2) is a prediction region, suppose the $\boldsymbol{x}_i$ are iid for $i = 1, ..., n, n+1$ where $\boldsymbol{x}_f = \boldsymbol{x}_{n+1}$. Compute $(T_H, \boldsymbol{C}_H)$ from the cases in H. Consider the squared validation distances D_k^2 for $k = 1, ..., n_V$ and the squared validation distance $D^2_{n_V+1}$ for case $\boldsymbol{x}_f$. Since these $n_V + 1$ cases are iid, the probability that D_t^2 has rank j for $j = 1, ..., n_V + 1$ is $1/(n_V + 1)$ for each t, i.e., the ranks follow the discrete uniform distribution. Let $t = n_V + 1$ and let the $D^2_{(j)}$ be the ordered squared validation distances using $j = 1, ..., n_V$. That is, get the order statistics without using the unknown squared validation distance $D^2_{n_V+1}$. Then $D^2_{(i)}$ has rank i if $D^2_{(i)} < D^2_{n_V+1}$ but rank $i+1$ if $D^2_{(i)} > D^2_{n_V+1}$. Thus $D^2_{(U_V)}$ has rank $U_V + 1$ if $D^2_{\boldsymbol{x}_f} < D^2_{(U_V)}$ and

$$P(\boldsymbol{x}_f \in \{\boldsymbol{z} : D^2_{\boldsymbol{z}}(T_H, \boldsymbol{C}_H) \le D^2_{(U_V)}\}) = P(D^2_{\boldsymbol{x}_f} \le D^2_{(U_V)}) \ge U_V/(1 + n_V) \to$$

$1 - \delta$ as $n_V \to \infty$. If there are no tied ranks, then

$$P(D^2_{\boldsymbol{x}_f} \le D^2_{(U_V)}) = P(D^2_{\boldsymbol{x}_f} < D^2_{(U_V)}) = P(\text{rank of } D^2_{\mathbf{x}_f} \le \mathrm{U_V}) = \mathrm{U_V}/(1 + \mathrm{n_V}).$$

Note that we can get the actual coverage $U_V/(1 + n_V)$ close to $1 - \delta$ for $n_V \ge 20$ for $\delta = 0.05$ even if $(T_H, \boldsymbol{C}_H)$ is a bad estimator. The volume of the prediction region tends to be much larger than that of the highest density region, even if $\boldsymbol{C}_H$ is well conditioned. We likely need $U_V \ge 50$ for $D^2_{(U_V)}$ to approximate the population percentile of $D_j^2 = (\boldsymbol{x}_{i_j} - T_H)^T \boldsymbol{C}_H^{-1} (\boldsymbol{x}_{i_j} - T_H)$.

As an example, consider using $(T, \boldsymbol{C}) = (\text{MED}(\boldsymbol{W}), \boldsymbol{I}_p)$. Then the prediction region is a hypersphere centered at the coordinatewise median. The prediction region is good if the iid $\boldsymbol{x}_i \sim N_p(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I}_p)$, but if the $\boldsymbol{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ such that highest density region is a hyperellipsoid tightly clustered about a vector in the direction of $\mathbf{1}$, then the prediction region (3) has huge volume compared to the highest density region.

If $p > n$, prediction region (2.2) can be used as long as $\boldsymbol{C}$ is nonsingular. Then $\boldsymbol{C} = \boldsymbol{I}_p$, $\boldsymbol{C} = diag(S_1^2, ..., S_p^2)$, or

$$\boldsymbol{C} = diag([MAD(x_{11}, ..., x_{n1})]^2, ..., [MAD(x_{1p}, ..., x_{np})]^2)$$

could be used where MAD is the median absolute deviation. Regularized covariance matrices or precision matrices could also be used.

CHAPTER 3

SEQUENTIAL DATA SPLITTING

The sequential data splitting algorithm is simple. Let $\lfloor x \rfloor$ be the integer part of x, e.g., $\lfloor 7.7 \rfloor = 7$. Denote the ceiling function by $\lceil x \rceil$, e.g. $\lceil 7.7 \rceil = 8$. Initially, randomly divide the data set into two sets H_1 with $n_1 \leq n/2$ cases and V_1 with $n - n_1$ cases. Apply lasso on H_1 to get a set of a_1 predictors, including a constant if a constant is in the model. If $n_1 \geq 10a_1$, set $H = H_1$ and $V = V_1$. Otherwise, randomly select n_1 cases from V_1 to add to H_1 to form H_2. Let V_2 have the remaining cases from V_1. Apply lasso on H_2 to get a set of a_2 predictors. If $n_2 \geq 10a_2$, set $H = H_1$ and $V = V_1$. Continue in this manner, forming sets $(H_1, V_1), (H_2, V_2), ..., (H_d, V_d)$ where H_i has $n_i = in_1$. Stop when $n_d \geq 10a_d$ or $n_{d+1} > n/2$. For the second case, use $n_d = \lfloor (n - J)/2 \rfloor$ where J is described below. Then $H = H_d$ and $V = V_d$. Use the model I_d with a_d predictors as the full model for inference with the data in $V = V_d$.

Lasso uses up to $n_d + 1$ active predictors, including a constant. If J is an integer between 0 and 5, set $n_1 = \max(1, \lfloor (n - J)/2 \rfloor)$ if $n < 40$. Otherwise, we often used $n_1 = 30$, but changed n_1 to $\lfloor n/2000 \rfloor$ if initially $\lfloor n/(2n1) \rfloor > 1000$. If $n >> p$, let $n_1 = Kp$ with K a positive integer, such as $K = 10$ or $K = 20$, or use $n_1 \approx Kp \approx n/(2M)$ with $M = \lceil n/(2Kp) \rceil$. If n/p is not large, options include $M = 10$ or $n_1 = Ka_0$ where a_0 is, for example, a guess of a lower bound for the number of active predictors.

This procedure can also be used with elastic net variable selection. If nonsequential data spitting is used where $n_1 = n_d$, then forward selection with the Chen and Chen (2008) EBIC criterion, elastic net, and lasso are useful for finding a reasonable fitted model. BIC (and the Hurvich and Tsai (1989) AIC_C criterion for multiple linear regression) can be useful if $n \geq \max(2p, 10a_d)$. For example, if $n = 500000$ and $p = 90$, using $n_1 = 900$ would result in a much smaller loss of efficiency than $n_1 = 250000$.

Data splitting theory for regression is simple in principle: if the model selected using subset H is good, fit the model with set V and do inference. In the literature, the model is good if a) there is no underfitting: no important predictors are left out, b) $n_v \geq M a_d$ where $M \geq 10$ and a_d is the number of predictors selected, and c) the usual checks on the model are good. For many regression models, two useful checks for model fit are a residual plot and a response plot of $\boldsymbol{x}_I^T \hat{\boldsymbol{\beta}}_I$ on the horizontal axis with the response Y on the horizontal axis where $\boldsymbol{x}_I$ denotes the predictors in the selected model, possibly including a constant. The model mean function and a scatterplot smoother can be added to the plot. See, for example, Olive (2013b).

In the simulations, there was often underfitting, in that a predictor that generated the model was not selected, but the prediction intervals still had good coverage with short length. Hence the selected model was still good for prediction. This result can happen because often there are many linear combinations of the predictors that give a useful model. See the following theorem for multiple linear regression and binary regression. We will write the linear combination as $a + \boldsymbol{x}^T \boldsymbol{\eta}$, for example. That is, the nontrivial predictors are in $\boldsymbol{x}$, but the constant is not.

Theorem 1: Suppose the cases $(Y_i, \boldsymbol{x}_i^T)^T$ are iid from some distribution.
a) If the joint distribution of $(Y, \boldsymbol{x}^T)^T$ is multivariate normal,

$$\begin{pmatrix} Y \\ \boldsymbol{x} \end{pmatrix} \sim N_{p+1}\left(\begin{pmatrix} \mu_Y \\ \boldsymbol{\mu}_{\boldsymbol{x}} \end{pmatrix}, \begin{pmatrix} \Sigma_Y & \boldsymbol{\Sigma}_{Y\boldsymbol{x}} \\ \boldsymbol{\Sigma}_{\boldsymbol{x}Y} & \boldsymbol{\Sigma}_{\boldsymbol{x}} \end{pmatrix} \right),$$

then $Y|\boldsymbol{x} \sim Y|(\alpha_{OLS} + \boldsymbol{\beta}_{OLS}^T \boldsymbol{x}) \sim N(\alpha_{OLS} + \boldsymbol{\beta}_{OLS}^T \boldsymbol{x}, \sigma^2)$ follows a multiple linear regression model, but so does $Y|\boldsymbol{\eta}^T \boldsymbol{x} \sim N(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x}, \sigma_O^2)$ where $\alpha_O = \mu_Y - \boldsymbol{\beta}_O^T \boldsymbol{\mu}_{\boldsymbol{x}}$, $\boldsymbol{\beta}_O = \lambda \boldsymbol{\eta}$, $\sigma_O^2 = \Sigma_Y - \boldsymbol{\beta}_O^T \boldsymbol{\Sigma}_{\boldsymbol{x}Y}$, and

$$\lambda = \frac{\boldsymbol{\Sigma}_{\boldsymbol{x}Y}^T \boldsymbol{\eta}}{\boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}} \boldsymbol{\eta}}.$$

b) If the response Y is binary, then $Y|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x}) \sim \text{binomial}(n = 1, \rho(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x}))$ where $E[Y|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})] = \rho(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x}) = P[Y = 1|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})]$. Hence every linear combination of the predictors satisfies a binary regression model.

Proof. a)

$$\begin{pmatrix} 1 & \boldsymbol{0}^T \\ 0 & \boldsymbol{\eta}^T \end{pmatrix} \begin{pmatrix} Y \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} Y \\ \boldsymbol{\eta}^T \boldsymbol{x} \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_Y \\ \boldsymbol{\eta}^T \boldsymbol{\mu}_{\boldsymbol{x}} \end{pmatrix}, \begin{pmatrix} \Sigma_Y & \boldsymbol{\Sigma}_{\boldsymbol{x}Y}^T \boldsymbol{\eta} \\ \boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}Y} & \boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}} \boldsymbol{\eta} \end{pmatrix} \right).$$

Hence $W = Y|\boldsymbol{\eta}^T \boldsymbol{x} \sim N(\mu_W, \sigma_W^2)$ where

$$\mu_W = \mu_Y + \frac{\boldsymbol{\Sigma}_{\boldsymbol{x}Y}^T \boldsymbol{\eta}}{\boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}} \boldsymbol{\eta}} (\boldsymbol{\eta}^T \boldsymbol{x} - \boldsymbol{\eta}^T \boldsymbol{\mu}_{\boldsymbol{x}}) = \mu_Y - \lambda \boldsymbol{\eta}^T \mu_{\boldsymbol{x}} + \lambda \boldsymbol{\eta}^T \boldsymbol{x},$$

and

$$\sigma_W^2 = \sigma_O^2 = \sigma_Y^2 - \frac{\boldsymbol{\Sigma}_{\boldsymbol{x}Y}^T \boldsymbol{\eta} \boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}Y}}{\boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}} \boldsymbol{\eta}} = \sigma_Y^2 - \lambda \boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}Y}.$$

b) $E[Y|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})] = 0P[Y = 0|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})] + 1P[Y = 1|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})] =$
$P[Y = 1|(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})] = \rho(\alpha_O + \boldsymbol{\beta}_O^T \boldsymbol{x})$. □

Note that $\sigma_O^2 < \sigma_Y^2$ unless $\boldsymbol{\eta}^T \boldsymbol{\Sigma}_{\boldsymbol{x}Y} = 0$. If the parameters are easy to estimate, then the above binary regression model can be visualized with a response plot of $\hat{\alpha}_O + \hat{\boldsymbol{\beta}}_O^T \boldsymbol{x}$ versus Y with a scatterplot smoother added to the plot to estimate ρ. Of course some linear combinations are poor, and if the full model follows a logistic regression model, the linear combination may result in a binary regression model that is not a binary logistic regression model.

CHAPTER 4

SIMULATION

The collection of Olive (2022) *R* functions *slpack*, available from (http://parker.ad.siu.edu/Olive/slpack.txt), has some useful functions for the inference.

4.1 PREDICTION REGIONS

The theory for the new prediction regions is simple, so Tables 1-3 are more of a check that the programs work than that the theory works. The output gives cov = observed coverage, up ≈ actual coverage, and mnhsq = mean cutoff $D^2_{(U_V)}$. With 5000 runs, expect observed coverage $\in [0.94, 0.96]$ if the actual coverage is close to 0.95. The random vector $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w}$ where $\boldsymbol{x} = \boldsymbol{w} \sim N_p(\boldsymbol{0}, \boldsymbol{I}_p)$ for xtype = 3, and $\boldsymbol{x} \sim N_p(\boldsymbol{0}, diag(1, ..., p))$ for xtype = 1. For xtype = 2, $\boldsymbol{w}$ has the w_i iid lognormal(0,1) with $\boldsymbol{A} = diag(1, \sqrt{2}, ..., \sqrt{p})$. The dispersion matrix types are dtype = 1 if $(T, \boldsymbol{C}) = (\overline{\boldsymbol{x}}, \boldsymbol{I}_p)$ and dtype = 2 if $(T, \boldsymbol{C}) = (\text{MED}(\boldsymbol{W}), \boldsymbol{I}_p)$.

When xtype=3 and dtype=1, $(T, \boldsymbol{C}) = (\overline{\boldsymbol{x}}, \boldsymbol{I}_p)$ where $\boldsymbol{x}_i \sim N_p(\boldsymbol{0}, \boldsymbol{I}_p)$. Then $D^2_{(U_V)}$ should estimate the population percentile $\chi^2_{p,0.95}$ if $n \geq \max(20p, 200)$ and $n_V = 100$. This result did occur in the simulations.

4.2 REGRESSION

The programs gives the mean n_d: the number of cases used in H, the mean a_d where a_d is the number of nonzero lasso coefficients including the constant for lasso applied to the n_d cases in H_d, and $k = a_d - 1$. The program also computed large sample Olive, Rathnayake, and Haile (2021) 95% prediction intervals (PIs) for lasso applied to all n cases (lsapi), lasso variable selection applied to all n cases (LVSpi), lasso applied to V_d (lsplitpi), and the model selected using H applied to V_d (splitpi). The coverage and average length of the prediction intervals was given. We also computed the number of times lasso and lasso variable selection did not underfit. The simulations used 5000 runs, and $n_1 = 30$ was used unless stated otherwise.

Table 4.1. Data Splitting Nominal 95% Prediction region

n	p	nv	xtype	dtype	cov
50	10	20	1	1	0.9538
50	10	20	2	1	0.9550
50	10	20	3	1	0.9538
50	10	20	1	2	0.9492
50	10	20	2	2	0.9578
50	10	20	3	2	0.9554
50	50	20	1	1	0.9490
50	50	20	2	1	0.9584
50	50	20	3	1	0.9538
50	50	20	1	2	0.9512
50	50	20	2	2	0.9532
50	50	20	3	2	0.9532
50	100	20	1	1	0.9560
50	100	20	2	1	0.9466
50	100	20	3	1	0.9504
50	100	20	1	2	0.9558
50	100	20	2	2	0.9508
50	100	20	3	2	0.9522
100	10	50	1	1	0.9572
100	10	50	2	1	0.9582
100	10	50	3	1	0.9656
100	10	50	1	2	0.9664
100	10	50	2	2	0.9620
100	10	50	3	2	0.9584
100	10	25	1	1	0.9620
100	10	25	2	1	0.9628
100	10	25	3	1	0.9564
100	10	25	1	2	0.9608
100	10	25	2	2	0.9620
100	10	25	3	2	0.9624
100	50	50	1	1	0.9638
100	50	50	2	1	0.9602
100	50	50	3	1	0.9614
100	50	50	1	2	0.9620
100	50	50	2	2	0.9634
100	50	50	3	2	0.9590
100	50	25	1	1	0.9620
100	50	25	2	1	0.9646
100	50	25	3	1	0.9654
100	50	25	1	2	0.9646
100	50	25	2	2	0.9620
100	50	25	3	2	0.9580

Table 4.2. Data Splitting Nominal 95% Prediction region

n	p	nv	xtype	dtype	cov
100	100	50	1	1	0.9620
100	100	50	2	1	0.9622
100	100	50	3	1	0.9596
100	100	50	1	2	0.9638
100	100	50	2	2	0.9578
100	100	50	3	2	0.9638
100	100	25	1	1	0.9588
100	100	25	2	1	0.9658
100	100	25	3	1	0.9568
100	100	25	1	2	0.9622
100	100	25	2	2	0.9672
100	100	25	3	2	0.9662
200	10	100	1	1	0.9498
200	10	100	2	1	0.9470
200	10	100	3	1	0.9476
200	10	100	1	2	0.9544
200	10	100	2	2	0.9494
200	10	100	3	2	0.9504
200	10	50	1	1	0.9606
200	10	50	2	1	0.9592
200	10	50	3	1	0.9606
200	10	50	1	2	0.9632
200	10	50	2	2	0.9602
200	10	50	3	2	0.9610
200	50	100	1	1	0.9494
200	50	100	2	1	0.9552
200	50	100	3	1	0.9502
200	50	100	1	2	0.9472
200	50	100	2	2	0.9544
200	50	100	3	2	0.9550
200	50	50	1	1	0.9564
200	50	50	2	1	0.9656
200	50	50	3	1	0.9646
200	50	50	1	2	0.9624
200	50	50	2	2	0.9574
200	50	50	3	2	0.9646

Table 4.3. Data Splitting Nominal 95% Prediction region

n	p	nv	xtype	dtype	cov
200	100	100	1	1	0.9516
200	100	100	2	1	0.9488
200	100	100	3	1	0.9518
200	100	100	1	2	0.9540
200	100	100	2	2	0.9538
200	100	100	3	2	0.9492
200	100	50	1	1	0.9596
200	100	50	2	1	0.9602
200	100	50	3	1	0.9600
200	100	50	1	2	0.9532
200	100	50	2	2	0.9568
200	100	50	3	2	0.9584

The prediction intervals were computed roughly as follows. If $Y \sim D(\boldsymbol{x}^T\boldsymbol{\beta}, \boldsymbol{\theta})$, then apply a prediction interval to a bootstrap sample of size B $Y_1^*, ..., Y_B^*$ where the Y_i are iid $D(\boldsymbol{x}^T\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}})$. For multiple linear regression, obtain the n_c residuals r_j and apply a prediction interval to $\hat{Y}_f + r_1, ..., \hat{Y}_f + r_{n_c}$ where $n_c = n$ or $n = n_v$ depending on whether all n cases or data splitting was used for the prediction interval.

The full model was simulated as in Pelawa Watagoda and Olive (2021b) and Olive, Rathnayake, and Haile (2021). For the simulations, generating $\boldsymbol{x}^T\boldsymbol{\beta}$ is important for regression models other than multiple linear regression. For example, for binomial logistic regression, typically $-5 \leq \boldsymbol{x}^T\boldsymbol{\beta} \leq 5$ or there can be problems with the MLE. 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 $\sum_{j=1}^k z_j \sim N(0, k\sigma_{ii} + k(k-1)\sigma_{ij}) = N(0, v^2)$. For multiple linear regression, let $\boldsymbol{u} = \boldsymbol{z}$. For the other regression models, 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. The default settings for Poisson regression use $\beta_1 = 1 = a$. The default settings for binomial regression with $m = 4$ trials use $\beta_1 = 0$ and $a = 5/3$. In the Table 4 caption, these values correspond to int=1, $a = 4/3$, and $m = 4$. The bootstrap sample for the prediction intervals had size $B = 1000$.

For the Weibull regression model, there is no constant since the constant appears in the corresponding accelerated failure time model. The data was generated as for the Poisson and Binomial regression, but replace $\boldsymbol{u}$ by $\boldsymbol{x}$ and $p-1$ by p. Let $SP = \boldsymbol{x}_i^T\boldsymbol{\beta} = 1x_{i,1} + \cdots + 1x_{i,k} \sim N(0, a^2)$ for $i = 1, ..., n$. The simulations use $a = 1$ where $\boldsymbol{\beta} = (1, ..., 1, 0, ..., 0)^T$ with k ones and $p-k$ zeros.

The right censored Weibull regression data was generated in a manner similar to Zhou (2001) with $\gamma = 4$. The caption in Table 4.6 gives $a = 1$. The values gam and clam in the caption control the Weibull distribution and the amount of right censoring.

Data splitting is useful for hypothesis testing and confidence intervals. Two lines per run are shown in each table. The first line gives the average coverage of the prediction intervals while the second line gives the average length. The prediction intervals were used as a check for whether lasso was finding a useful model for prediction (coverage near 0.95) even if undercoverage was present. This result could occur for at least two reasons. First, as ψ increases to 1, the predictor variables are roughly $x_i = x_j + e_{ij}$ where the error magnitude rapidly gets close to 0 as $\psi \to 1$. Hence omitting some good predictors may not be a problem for prediction. Second, for some regression models, there are many linear combinations that give a good fit. For example, for binary regression where Y_i is 1 or 0, $Y_i|\boldsymbol{\eta}^T\boldsymbol{x} \sim \text{binary}(n = 1, \rho(\boldsymbol{\eta}^T\boldsymbol{x}_i))$ where $E(Y_i) = \rho(\boldsymbol{\eta}^T\boldsymbol{x}_i) = P(Y_i = 1)$ if the cases are iid. Hence every linear combination results in a binary regression. A similar result holds for multiple linear regression if $(Y, x_2, ..., x_p)^T$ has a multivariate normal distribution or if Y is linearly related with each x_j selected by lasso. See Theorem 1.

For multiple linear regression, the zero mean errors e_i were iid from five distributions: i) N(0,1), ii) t_3, iii) EXP(1) - 1, iv) uniform$(-1, 1)$, and v) 0.9 N(0,1) + 0.1 N(0,100). Only distribution iii) is not symmetric. The lengths of the asymptotically optimal 95% PIs are i) 3.92 = 2(1.96), ii) 6.365, iii) 2.996, iv) 1.90 = 2(0.95), and v) 13.490.

For Poisson regression, first consider $k = 1$. Then there was little underfit for $\psi = 0$. The amount of underfitting tended to increase with ψ, and to be worse with larger p. With $n = 1000$, not much more than 10% of the cases were used for H. For larger values of k, lasso often underfit, especially if $k = p - 1$ and $n/k < 10$. See Table 4.4.

For Binomial regression, first consider $k = 1$. Then there was little underfit for $\psi = 0$. The amount of underfitting tended to increase with ψ, and to be worse with larger p. With $n = 1000$, not much more than 10% of the cases were used for H. For larger values of k, lasso often underfit, especially if $k = p - 1$ and $n/k < 10$. See Table 4.5.

Table 4.4. prsplit

n	p/k	psi	mnnd/mnad	lsapi	LVSpi	lsplitpi	splitpi	noundfit
100	4	0.0000	34.5757	0.9872	0.9924	0.9955	0.9864	5000
	1		2.9438	7.5355	7.7445	7.1744	7.8245	
100	4	0.8000	32.8250	0.9887	0.9837	0.9967	0.9857	4578
	1		2.4602	7.8907	7.1752	8.3510	8.2722	
100	20	0.0000	40.5895	0.9934	0.9797	0.9819	0.9800	5000
	1		3.7327	8.2845	7.8980	9.0938	7.4559	
100	20	0.6000	38.6768	0.9986	0.9909	0.9898	0.9783	3834
	1		3.3482	8.1733	8.2308	8.7990	8.2294	
100	100	0.0000	43.7107	0.9946	0.9788	0.9943	0.9695	5000
	1		5.3520	8.4802	7.7537	7.6238	7.8035	
100	100	0.3000	44.5573	0.9799	0.9819	0.9870	0.9580	4165
	1		7.1632	8.0935	7.1896	7.8179	7.6183	
100	10	0.0000	45.6543	0.9912	0.9902	0.9800	0.9745	4896
	9		8.8543	8.7440	8.0106	8.1908	7.5920	
100	20	0.0000	43.6966	0.9841	0.9733	0.9685	0.9574	2363
	19		10.6884	8.1987	7.7602	8.7657	7.9307	
1000	4	0.0000	36.8298	0.9823	0.9808	0.9883	0.9840	4993
	1		2.9703	7.5124	7.6328	8.5947	8.0170	
1000	4	0.8000	34.1201	0.9911	0.9794	0.9855	0.9792	3718
	1		2.5867	7.5125	6.9803	7.9171	6.9004	
1000	20	0.0000	56.6908	0.9869	0.9777	0.9908	0.9842	4978
	1		3.3662	7.0329	7.7289	7.9915	8.1688	
1000	20	0.5000	56.4071	0.9834	0.9915	0.9849	1.0000	3418
	1		3.9243	7.9746	8.1575	7.1918	7.2256	
1000	1000	0.0000	95.2273	0.9856	0.9908	0.9924	0.9744	4789
	1		5.2895	8.0867	8.7953	7.8825	7.8767	
1000	1000	0.1000	110.0411	0.9880	0.9884	0.9902	0.9881	3240
	1		8.8354	7.3550	8.5021	8.2190	8.1900	
1000	10	0.3160	74.3965	0.9920	0.9920	0.9899	0.9836	24
	9		7.0638	7.5158	7.9454	8.1729	8.1484	

Table 4.5. brsplit, int=1, a=4/3, m=4, B=1000

n	p/k	psi	mnnd/mnad	lsapi	LVSpi	lsplitpi	splitpi	noundfit
100	4	0.0000	33.9066	0.9914	0.9896	0.9910	0.9872	4996
	1		2.6168	2.6678	2.6014	2.6760	2.5832	
100	4	0.8000	32.6724	0.9894	0.9888	0.9908	0.9884	3685
	1		2.5686	2.6048	2.5800	2.6032	2.5630	
100	20	0.0000	40.9616	0.9932	0.9850	0.9904	0.9752	4991
	1		4.6866	2.8414	2.6524	2.8588	2.5972	
100	20	0.5000	41.4784	0.9934	0.9904	0.9908	0.9820	3163
	1		4.5670	2.7706	2.7130	2.7524	2.6444	
100	100	0.0000	43.4521	0.9854	0.9844	0.9945	0.9843	4987
	1		5.6745	2.6474	2.4433	2.9443	2.5556	
100	100	0.2000	45.6434	0.9948	0.9782	0.9886	0.9624	3881
	1		8.7554	2.8426	2.6696	2.8244	2.5558	
1000	4	0.0000	37.1640	0.9876	0.9862	0.9866	0.9858	4993
	1		2.6016	2.4730	2.4594	2.4752	2.4546	
1000	4	0.8000	34.2360	0.9862	0.9860	0.9870	0.9854	3718
	1		2.5336	2.4468	2.4352	2.4442	2.4312	
1000	20	0.0000	56.7660	0.9856	0.9840	0.9858	0.9824	4978
	1		3.7276	2.4988	2.4460	2.4978	2.4400	
1000	20	0.5000	56.0040	0.9874	0.9872	0.9890	0.9878	3418
	1		4.0666	2.4560	2.4374	2.4614	2.4392	
1000	1000	0.0000	95.0734	0.9902	0.9820	0.9898	0.9822	4789
	1		5.3892	2.6302	2.4922	2.6320	2.4834	
1000	10	0.4000	63.4080	0.9870	0.9856	0.9862	0.9858	4
	9		5.2826	2.5050	2.4854	2.5008	2.4890	

For proportional hazards regression, first consider $k = 1$. Then there was little underfit for $\psi = 0$. The amount of underfitting tended to increase with ψ, and to be worse with larger p. With $n = 1000$, not much more than 10% of the cases were used for H. For larger values of k, lasso often underfit, especially if $k = p - 1$ and $n/k < 10$. Now underfitting sometimes caused poor prediction interval coverage. Usually the data splitting PI and PI using all n cases had similar average lengths, but there are data configurations where using all n cases can give a much smaller length. The prediction intervals were made for Weibull regression, which does not have a lasso analog. Hence only two prediction intervals are given. See Table 4.6.

For multiple linear regression, first consider $k = 1$. Then there was little underfit for $\psi = 0$. The amount of underfitting tended to increase with ψ, and to be worse with larger p. With $n = 1000$, not much more than 10% of the cases were used for H. For larger values of k, lasso often underfit, especially if $k = p - 1$ and $n/k < 10$. Usually the data splitting PI and PI using all n cases had similar average lengths, but there are data configurations where using all n cases can give a much smaller length and better coverage. See Table 4.7-4.9.

Table 4.6. PHsplit, n=100, J=5, a=1, gam=1, B=1000, clam=0.1

n	p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
100	4	0.00	31.7646	0.9550	0.9552	4913
	1		1.8314	5.5483	5.5033	
100	4	0.80	30.7004	0.9574	0.9576	156
	1		1.6076	5.5956	5.5384	
100	20	0.00	36.3172	0.9326	0.9328	4747
	1		2.7178	5.9745	25.4093	
100	20	0.60	33.4238	0.9510	0.9512	1289
	1		2.2570	5.7760	9.8008	
100	10	0.00	39.1528	0.9518	0.9520	608
	9		4.2938	6.9368	6.6001	
1000	4	0.6000	31.8420	0.9572	0.9574	511
	3		2.0530	5.3167	5.3241	
1000	4	0.7000	31.4220	0.9592	0.9594	297
	3		1.8144	5.2454	5.2387	
1000	4	0.8000	31.0980	0.9598	0.9600	145
	3		1.5896	5.2561	5.2668	
1000	10	0.4000	42.0120	0.9582	0.9584	1
	9		3.1224	5.4933	5.4875	
1000	10	0.5000	37.7400	0.9618	0.9620	0
	9		2.5946	5.3586	5.3679	

Table 4.7. mlrsplit, J=5, type=1

n	p/k	psi	mnnd/mnad	lsapi	LVSpi	lsplitpi	splitpi	noundfit
100	4	0.0	34.6002	0.9708	0.9696	0.9784	0.9784	4984
	1		2.6340	4.5802	4.5655	5.1234	5.0961	
100	4	0.8	33.5462	0.9728	0.9712	0.9796	0.9776	4984
	1		2.6732	4.6233	4.6217	5.0891	5.0870	
100	20	0.0	40.5264	0.9716	0.9652	0.9800	0.9696	4984
	1		4.5574	4.8229	4.6799	5.4299	5.1550	
100	20	0.6	44.5248	0.9758	0.9720	0.9824	0.9740	4741
	1		5.7334	4.9434	4.9046	5.5457	5.4506	
100	100	0.0	42.6650	0.9758	0.9450	0.9738	0.9126	4885
	1		8.0380	4.9461	4.4606	5.5054	4.5813	
100	100	0.3	46.4458	0.9810	0.9694	0.9830	0.9464	4977
	1		9.1094	5.1002	4.9011	5.7093	5.2444	
100	10	0.0	47.0000	0.9836	0.9840	0.9892	0.9894	5000
	9		10.0000	5.2952	5.2921	6.0284	6.0207	
100	20	0.0	46.9966	0.9846	0.9848	0.9796	0.9792	4977
	19		19.9960	5.7114	5.7068	6.8113	6.7973	
1000	4	0.0	38.7540	0.9482	0.9476	0.9486	0.9466	4982
	1		2.6142	3.9492	3.9477	3.9543	3.9531	
1000	4	0.8	35.8440	0.9514	0.9494	0.9494	0.9492	4282
	1		2.6770	3.9526	3.9524	3.9552	3.9549	
1000	20	0.0	56.0940	0.9528	0.9494	0.9532	0.9508	4959
	1		3.5208	3.9743	3.9633	3.9792	3.9680	
1000	20	0.5	65.4420	0.9484	0.9484	0.9504	0.9472	4984
	1		4.7162	3.9761	3.9752	3.9820	3.9811	
1000	1000	0.0	81.1200	0.9488	0.9340	0.9498	0.9374	4632
	1		4.4482	4.0714	3.9431	4.0864	3.9472	
1000	10	0.3160	120.0000	0.9552	0.9554	0.9576	0.9580	5000
	9		10.0000	4.0526	4.0392	4.0726	4.0601	

Table 4.8. mlrsplit, J=5, type=3

n	p/k	psi	mnnd/mnad	lsapi	LVSpi	lsplitpi	splitpi	noundfit
100	4	0.8000	33.3354	0.9676	0.9672	0.9768	0.9764	4306
	1		2.6874	4.0502	4.0545	4.6570	4.6614	
100	20	0.6000	44.1712	0.9762	0.9730	0.9780	0.9746	4690
	1		5.4618	4.5475	4.5611	5.2808	5.2408	
100	100	0.3000	46.3812	0.9780	0.9660	0.9804	0.9518	4939
	1		8.9376	4.7967	4.6882	5.5010	5.1129	
100	10	0.0000	47.0000	0.9752	0.9752	0.9786	0.9786	5000
	9		10.0000	5.0524	5.0498	5.8915	5.8846	
100	20	0.0000	46.9966	0.9810	0.9804	0.9794	0.9804	4994
	19		19.9950	5.6460	5.6398	6.8107	6.7928	
1000	4	0.0000	38.1120	0.9556	0.9564	0.9570	0.9576	4962
	1		2.5994	3.1418	3.1398	3.1485	3.1471	
1000	4	0.8000	36.0180	0.9558	0.9564	0.9562	0.9562	4321
	1		2.6694	3.1302	3.1306	3.1333	3.1333	
1000	20	0.0000	55.7820	0.9516	0.9474	0.9514	0.9490	4915
	1		3.5346	3.2048	3.2349	3.2139	3.2449	
1000	20	0.5000	64.9500	0.9548	0.9548	0.9536	0.9528	4951
	1		4.7108	3.1909	3.1942	3.2011	3.2054	
1000	10	0.3160	120.0000	0.9574	0.9570	0.9592	0.9580	5000
	9		10.0000	3.3726	3.3244	3.4032	3.3574	
1000	1000	0.0000	82.3860	0.9558	0.9460	0.9572	0.9440	4551
	1		4.5362	3.3778	3.4472	3.4008	3.4644	

Table 4.9. mlrsplit, J=5, type=4

n	p/k	psi	mnnd/mnad	lsapi	LVSpi	lsplitpi	splitpi	noundfit
100	4	0.0000	34.6614	0.9920	0.9922	0.9946	0.9942	5000
	1		2.6322	2.2646	2.2576	2.5170	2.5020	
100	4	0.8000	33.9100	0.9904	0.9906	0.9954	0.9948	4890
	1		2.7850	2.2580	2.2593	2.4874	2.4894	
100	20	0.0000	40.4686	0.9896	0.9800	0.9954	0.9832	5000
	1		4.5912	2.3707	2.3694	2.7225	2.6662	
100	20	0.6000	44.5928	0.9936	0.9918	0.9960	0.9920	4998
	1		5.3614	2.3754	2.3798	2.7588	2.7613	
100	100	0.0000	42.6412	0.9900	0.9546	0.9832	0.9126	5000
	1		8.1888	2.4628	2.3550	2.8180	2.4481	
100	100	0.3000	46.3846	0.9932	0.9870	0.9950	0.9792	5000
	1		7.7210	2.4764	2.4749	2.9251	2.8447	
100	10	0.0000	47.0000	0.9908	0.9914	0.9926	0.9930	5000
	9		10.0000	2.6146	2.6106	3.1510	3.1441	
100	20	0.0000	47.0000	0.9912	0.9918	0.9864	0.9884	5000
	19		20.0000	2.9678	2.9604	3.7576	3.7425	
1000	4	0.0000	38.3160	0.9610	0.9602	0.9584	0.9588	5000
	1		2.5934	1.9231	1.9225	1.9248	1.9242	
1000	4	0.8000	37.0860	0.9594	0.9570	0.9574	0.9588	4906
	1		2.7552	1.9232	1.9228	1.9245	1.9241	
1000	20	0.0000	55.5900	0.9590	0.9560	0.9610	0.9524	5000
	1		3.5238	1.9324	1.9343	1.9350	1.9372	
1000	20	0.5000	66.1500	0.9602	0.9592	0.9618	0.9616	5000
	1		4.7568	1.9312	1.9304	1.9338	1.9331	
1000	10	0.3160	120.0000	0.9646	0.9632	0.9666	0.9652	5000
	9		10.0000	1.9909	1.9541	2.0005	1.9632	
1000	1000	0.0000	85.4520	0.9610	0.9344	0.9592	0.9330	4999
	1		4.7416	1.9628	1.9778	1.9712	1.9869	

CHAPTER 5

CONCLUSIONS

The new prediction regions can be used for distributions that do not have an expected value if appropriate $(T, \boldsymbol{C})$ is used, e.g. $(T, \boldsymbol{C}) = (\text{MED}(\boldsymbol{W}), \boldsymbol{I}_p)$. Prediction regions have some nice applications besides prediction. Applying a prediction region to data generated from a posterior distribution gives an estimated credible region for Bayesian Statistics. See Chen and Shao (1999). Certain prediction regions applied to a bootstrap sample result in a confidence region. See Pelawa Watagoda and Olive (2021a) and Rathnayake and Olive (2021). Mykland (2003) converts prediction regions into investment strategies.

The methods in this paper are simple. For regression, the regression model should be checked on the validation set V using the usual diagnostics for the regression model. Let I be the selected model. For example, I could be the I_{min} model selected from variable selection using the cases in H. We want $S \subseteq I$, and $n_V \geq 10a$ if I uses a predictors. Tay, Narasimhan, and Hastie (2021) describe lasso for several regression models.

Data splitting can be useful for the following dense multiple linear regression model $Y = \beta_1 + \beta_2 x_2 + \cdots + \beta_p x_p + e$ where $p > n$ is possible. For this model, let $\boldsymbol{x} = (1, \boldsymbol{x}_2, ..., \boldsymbol{x}_p)^T$ where $\boldsymbol{x}_1 = 1$. If $(Y, \boldsymbol{x}_2, ..., \boldsymbol{x}_p)^T$ follows a multivariate normal distribution, then $(Y, \boldsymbol{x}_I)$ follows a multiple linear regression model for every I. Hence the full model need not be sparse, although the selected model I may be suboptimal.

Pelawa Watagoda and Olive (2021b) and Lei et al. (2018) use data splitting to obtain prediction intervals for the multiple linear regression model for estimators such as OLS, lasso, and lasso variable selection.

Data splitting can also be used to build a model. Build a model I using the cases in H, then do inference by fitting the model I to the cases in V. The inference is conditional on I being a useful model, for example we want $S \subseteq I$. Hastie, Tibshirani, and Friedman (2009, p. 245) note that one should not build a model then do cross validation if V is the entire data set.

The "Million Song dataset," from the Lichman (2017) UCI Machine Learning Repository, has been analyzed as a multiple linear regression data set with $n = 515345$ and $p = 90$. See Wieczorek and Lei (2021). Using $n_d = 900$ may work for this data set.

SOFTWARE

Simulations were done in *R*. See R Core Team (2018). The collection of Olive (2022) *R* functions *slpack*, available from (http://parker.ad.siu.edu/Olive/slpack.txt), has some useful functions for the inference. The function `predsim2` computes the data splitting prediction region. The functions for regression data splitting are `mlrsplitsim`, `prsplit`, `brsplitsim`, and `PHsplitsim`. These functions used the Friedman et al. (2015) `glmnet` package.

ACKNOWLEDGMENTS

The Big-Dawg High Performance Computing as Southern Illinois University was used for some of the simulations.

CHAPTER 6

MORE SIMULATION TABLES

Table 6.1. prsplit n=100,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.8000	32.8250	0.9887	0.9837	0.9967	0.9857	4578
1		2.4602	7.8907	7.1752	8.3510	8.2722	
10	0.4000	41.0059	0.9950	0.9965	0.9963	0.9813	4998
1		3.8119	8.5145	8.5086	8.5792	8.1492	
10	0.5000	40.1030	0.9934	0.9914	0.9914	0.9879	4977
1		3.2152	7.6849	8.4825	8.6407	8.1411	
10	0.6000	38.3840	0.9959	0.9856	0.9917	0.9977	4777
1		3.6877	7.9721	7.7809	8.5271	7.4736	
10	0.7000	37.0136	0.9896	0.9907	0.9782	0.9846	4045
1		3.5813	7.7982	7.7517	8.7869	7.0561	
10	0.8000	34.9733	0.9901	0.9961	0.9963	0.9819	2724
1		2.8269	7.8342	7.8603	7.7523	8.3285	
20	0.3000	43.2558	0.9977	0.9965	0.9961	0.9868	4997
1		4.9881	8.6226	8.5183	7.2776	8.3187	
20	0.4000	43.4286	0.9964	0.9879	0.9884	0.9782	4968
1		5.9433	7.6704	9.2146	8.2895	8.7570	
20	0.5000	41.8245	1.0027	0.9864	0.9901	0.9809	4721
1		3.6577	8.8031	7.4641	7.8902	7.5508	
20	0.6000	38.6768	0.9986	0.9909	0.9898	0.9783	3834
1		3.3482	8.1733	8.2308	8.7990	8.2294	
20	0.7000	37.0579	0.9987	0.9967	0.9945	0.9911	2564
1		3.4263	8.1288	8.2603	6.7929	7.7137	
20	0.8000	35.7869	0.9904	0.9878	0.9893	0.9925	1313
1		2.6134	8.0683	7.4654	7.2950	7.9985	
50	0.2000	44.7085	1.0000	0.9872	0.9986	0.9692	4998
1		5.9211	7.9219	7.6886	7.7555	7.3748	
50	0.3000	46.3026	0.9914	0.9874	0.9883	0.9691	4914
1		7.0168	7.8420	8.0090	7.9839	8.3992	
50	0.4000	44.0774	0.9878	0.9806	0.9929	0.9700	4254
1		5.0795	7.9325	8.3321	8.1005	7.5445	
50	0.5000	41.4445	0.9927	0.9815	0.9943	0.9799	2925
1		4.8030	7.4656	7.8853	8.3973	7.5788	
50	0.6000	38.0978	0.9936	0.9847	0.9897	0.9806	1683
1		4.0805	7.5632	7.7298	6.7295	8.2817	
50	0.7000	36.9239	0.9909	0.9883	0.9846	0.9829	805
1		3.1184	7.8446	8.1462	7.9220	8.0380	
50	0.8000	33.7526	0.9947	0.9885	0.9897	0.9772	407
1		2.6575	7.6025	8.3846	8.2538	7.8253	
100	0.2000	45.6960	0.9977	0.9754	0.9834	0.9585	4938
1		9.1142	8.7304	7.2884	8.0505	6.7044	
100	0.3000	44.5573	0.9799	0.9819	0.9870	0.9580	4165
1		7.1632	8.0935	7.1896	7.8179	7.6183	
100	0.4000	41.6811	0.9962	0.9904	0.9853	0.9692	2692
1		6.7690	8.3492	8.1753	7.9727	7.2784	

Table 6.2. prsplit n=100,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
100	0.5000	39.9592	0.9909	0.9893	0.9984	0.9789	1375
1		4.4885	8.2436	7.3627	7.1518	7.4774	
100	0.6000	38.5509	0.9884	0.9860	0.9958	0.9825	661
1		3.6756	9.1602	8.0488	7.6237	7.9224	
100	0.7000	36.6515	0.9980	0.9874	0.9910	0.9897	306
1		2.9038	8.9456	6.9888	8.6393	8.1479	
100	0.8000	33.7393	0.9884	0.9930	0.9878	0.9872	173
1		2.2014	6.8185	7.5080	6.9117	7.5139	
4	0.6000	35.4528	0.9877	0.9853	0.9925	0.9889	3018
3		3.3899	8.1190	8.0537	8.6084	7.8463	
4	0.7000	34.2154	0.9933	0.9854	0.9913	0.9946	1693
3		3.0619	7.8423	8.3084	8.0019	8.0447	
4	0.8000	32.1813	0.9894	0.9930	0.9867	0.9851	617
3		1.6894	7.3352	7.9779	8.4213	8.2483	
10	0.4000	43.9819	0.9963	0.9930	0.9938	0.9816	154
9		5.1604	7.8313	7.9573	8.3333	8.4329	
10	0.5000	41.9265	0.9987	0.9894	0.9951	0.9914	82
9		4.5589	7.7785	8.5400	8.2987	7.7270	
10	0.6000	38.7591	0.9877	0.9816	0.9862	0.9908	65
9		3.8388	8.1026	7.9949	8.3727	8.4309	
10	0.7000	36.1000	0.9902	0.9896	0.9924	0.9880	27
9		2.9655	8.3928	7.8038	7.0782	7.5911	
10	0.8000	34.7807	0.9860	0.9882	0.9854	0.9848	10
9		2.6424	9.1016	8.1466	8.0198	7.5636	
20	0.3000	45.2904	0.9914	0.9941	0.9976	0.9834	4
19		7.3920	7.3511	8.7383	8.6840	7.9903	
20	0.4000	44.7693	0.9982	0.9929	0.9902	0.9806	4
19		5.1904	7.4928	7.6672	7.5880	8.0821	
20	0.5000	41.3372	0.9953	0.9870	0.9867	0.9832	4
19		5.6861	7.8967	8.2484	8.7207	8.4316	
20	0.6000	39.4471	0.9886	0.9921	0.9926	0.9900	3
19		4.3201	8.1819	7.5109	7.6698	7.2951	
20	0.7000	37.5752	1.0000	0.9887	0.9981	0.9925	2
19		3.6314	7.7787	7.8987	7.7658	6.5133	
20	0.8000	36.9796	0.9924	0.9932	0.9818	0.9817	0
19		3.4779	7.8358	8.1303	7.6950	8.2539	
50	0.2000	46.7140	0.9870	0.9807	0.9801	0.9752	0
19		9.4167	7.6761	8.0850	8.0581	7.4161	
50	0.3000	45.7587	0.9889	0.9881	0.9881	0.9741	0
19		7.4472	7.9156	6.5912	8.1456	7.8545	
50	0.4000	43.4519	0.9936	0.9803	0.9913	0.9755	0
19		5.9316	7.9649	7.7911	8.3624	6.8956	
50	0.5000	41.5648	0.9869	0.9831	0.9878	0.9787	0
19		5.2432	7.3867	7.7588	8.0417	7.5612	

Table 6.3. prsplit n=100,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
50	0.6000	39.5888	0.9957	0.9907	0.9814	0.9880	0
19		4.1317	8.9453	7.3774	8.0666	7.7877	
50	0.7000	36.7091	0.9907	0.9853	0.9853	0.9788	0
19		2.4395	8.1980	8.3537	7.3609	9.0914	
50	0.8000	35.1108	0.9917	0.9842	0.9901	0.9872	0
19		1.9741	6.9339	7.9872	7.1737	7.2605	
100	0.2000	47.1805	0.9965	0.9778	0.9929	0.9742	0
19		10.0306	8.1947	8.2807	7.7070	7.7125	
100	0.3000	45.6901	0.9880	0.9800	0.9970	0.9643	0
19		8.3786	8.1600	7.9713	8.9907	7.4556	
100	0.4000	43.5724	0.9846	0.9878	0.9919	0.9689	0
19		6.2307	8.6465	8.5297	8.1599	7.9044	
100	0.5000	41.4412	0.9910	0.9842	0.9785	0.9711	0
19		4.7797	8.0899	7.6862	8.2953	7.8994	
100	0.6000	37.9977	0.9883	0.9855	0.9998	0.9877	0
19		4.2600	7.6056	7.7411	8.0589	7.4466	
100	0.7000	36.6965	0.9724	0.9785	0.9851	0.9790	0
19		2.9266	7.4170	8.3187	7.5507	7.7012	
100	0.8000	32.3983	0.9901	0.9957	0.9886	0.9952	0
19		2.1333	7.3119	8.1252	8.0498	8.1259	

Table 6.4. brsplitsim n=200,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.0000	37.5794	0.9734	0.9744	0.9738	0.9726	5000
1		2.8695	7.0422	7.0084	7.0464	7.0006	
10	0.0000	47.5133	0.9750	0.9702	0.9750	0.9714	5000
1		2.9772	7.2174	7.1324	7.2342	7.1478	
20	0.0000	56.5605	0.9762	0.9714	0.9732	0.9670	4999
1		4.4600	7.2996	7.1978	7.3122	7.1986	
100	0.0000	67.4103	0.9804	0.9742	0.9780	0.9634	4999
1		5.7868	7.4938	7.3434	7.5416	7.3412	
200	0.0000	69.7152	0.9814	0.9698	0.9762	0.9604	4999
1		7.3966	7.5618	7.3752	7.6010	7.3388	
4	0.0000	59.0292	0.9646	0.9648	0.9586	0.9574	4999
3		4.4308	9.2504	9.2852	9.2574	9.2740	
10	0.0000	89.9087	0.8563	0.8768	0.8116	0.8382	4979
9		9.3899	13.1236	13.6486	13.7465	14.7528	
20	0.0000	79.5341	0.6886	0.6852	0.6280	0.6270	4619
19		14.5547	10.8091	10.7764	13.9406	10.9702	
100	0.0000	61.2222	0.6672	0.5842	0.6441	0.4863	1325
19		12.3787	10.4382	11.7881	10.3672	12.4674	
200	0.0000	56.4159	0.8645	0.5985	0.6734	0.5047	4673
19		9.9211	8.6789	8.9024	11.4673	12.4532	
4	0.5000	37.0345	0.9722	0.9710	0.9730	0.9720	5000
1		2.7709	10.1303	12.5342	7.3172	7.2994	
10	0.3000	50.1920	0.9750	0.9728	0.9726	0.9720	5000
1		3.5915	8.1170	8.0964	8.0914	8.0962	
20	0.2000	61.0037	0.9726	0.9696	0.9680	0.9642	5000
1		4.4736	8.1050	8.0816	8.1188	8.0860	
100	0.1000	75.2383	0.9726	0.9572	0.9660	0.9462	4999
1		6.4207	8.4984	8.4688	8.4768	8.4742	
200	0.0700	80.1325	0.9706	0.9540	0.9638	0.9368	5000
1		9.0504	8.7142	8.6896	8.6878	8.6800	
4	0.5000	44.5599	0.8630	0.8628	0.8512	0.8536	4042
3		3.3041	17.3176	17.9854	17.3266	18.2918	
10	0.3000	73.3998	0.5814	0.5794	0.5584	0.5584	1204
9		6.3282	8.6773	10.0090	11.2892	10.5051	
20	0.2000	92.6303	0.5066	0.5038	0.4950	0.4938	17
19		10.2943	9.7823	11.2783	12.7832	13.8872	
100	0.1000	96.7874	0.4712	0.4772	0.4744	0.4650	0
19		17.0214	7.7882	9.6724	9.8923	9.6723	
200	0.0700	96.1824	0.4808	0.4668	0.4638	0.4536	0
19		21.1607	9.7892	10.6344	10.8923	11.6723	
4	0.6000	38.5438	0.9730	0.9714	0.9696	0.9710	4995
1		2.2382	7.7986	7.8086	7.8010	7.7928	
4	0.7000	37.1516	0.9732	0.9716	0.9716	0.9696	4932
1		1.9419	8.0598	8.0510	8.0536	8.0550	

Table 6.5. brsplitsim n=200,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.8000	33.5532	0.9732	0.9746	0.9740	0.9730	4588
1		2.7729	8.4000	8.4082	8.4026	8.4190	
10	0.4000	50.5301	0.9670	0.9662	0.9638	0.9584	4997
1		3.8888	8.6886	8.7136	8.6978	8.7244	
10	0.5000	50.3783	0.9638	0.9612	0.9614	0.9566	4981
1		3.2399	9.4086	9.4336	9.4036	9.4726	
10	0.6000	46.3073	0.9562	0.9534	0.9454	0.9436	4824
1		3.7423	11.1132	11.2122	11.1046	11.2198	
10	0.7000	43.5730	0.9410	0.9400	0.9330	0.9326	4204
1		3.6364	12.2780	12.3926	12.2624	12.4344	
10	0.8000	40.1794	0.9336	0.9330	0.9226	0.9216	2889
1		3.1587	13.9934	14.1862	14.0204	14.2082	
20	0.3000	60.7738	0.9718	0.9614	0.9644	0.9554	4997
1		4.4614	9.1176	9.1718	9.1396	9.2074	
20	0.4000	59.8431	0.9528	0.9496	0.9482	0.9320	4987
1		4.6002	10.6930	10.8194	10.6744	10.8188	
20	0.5000	55.3035	0.9260	0.9126	0.9156	0.9046	4806
1		4.3819	13.6884	13.9094	13.6400	14.0446	
20	0.6000	50.0978	0.9078	0.8994	0.8944	0.8786	4156
1		3.3020	17.7558	18.1504	18.0312	18.5792	
20	0.7000	45.1972	0.8826	0.8824	0.8698	0.8626	2776
1		3.5025	23.4866	24.1978	23.7014	24.4882	
20	0.8000	42.9538	0.8644	0.8634	0.8548	0.8506	1416
1		2.5177	32.1352	33.2378	31.9562	33.1088	
100	0.2000	78.1500	0.9308	0.9012	0.9042	0.8690	4994
1		7.2028	13.0926	13.4474	13.0086	13.5444	
100	0.3000	69.5298	0.8636	0.8274	0.8284	0.7948	4823
1		6.5257	23.7482	24.8486	23.5172	25.5010	
100	0.4000	61.0824	0.7992	0.7678	0.7714	0.7366	3679
1		5.6846	13.7120	18.7816	17.5870	14.6734	
100	0.5000	54.2430	0.7340	0.7274	0.7148	0.6908	1911
1		4.1237	10.4783	9.9209	8.6643	9.7522	
100	0.6000	48.9203	0.7168	0.7118	0.7000	0.6940	781
1		3.7355	8.6246	9.0817	11.0919	5.6295	
100	0.7000	43.0662	0.6798	0.6778	0.6776	0.6774	349
1		2.8000	7.7941	8.1851	9.0826	8.4244	
100	0.8000	36.2306	0.6790	0.6738	0.6732	0.6712	213
1		2.3549	8.6340	11.9071	9.0134	10.6586	
200	0.2000	76.8413	0.8646	0.8216	0.8286	0.7872	4955
1		8.8188	9.5910	6.5307	8.1709	9.8968	
200	0.3000	65.4086	0.7646	0.7356	0.7272	0.6914	3866
1		5.6679	9.7525	10.6117	11.2894	8.6125	
200	0.4000	56.8119	0.7016	0.6966	0.6744	0.6736	1773
1		3.5734	11.9004	7.6760	9.7813	8.5894	

Table 6.6. brsplitsim n=200,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
200	0.5000	52.2227	0.6698	0.6684	0.6580	0.6594	612
1		3.9152	9.1973	10.3339	13.6281	8.7551	
200	0.6000	46.9323	0.6392	0.6368	0.6332	0.6314	224
1		2.8909	10.6358	5.2384	10.4889	6.9919	
200	0.7000	40.3114	0.6424	0.6382	0.6388	0.6386	131
1		2.8625	7.1343	11.5065	11.4228	13.2930	
200	0.8000	33.1410	0.6434	0.6360	0.6416	0.6308	102
1		2.0827	8.8213	12.7863	12.8042	12.3179	
4	0.6000	39.7117	0.8454	0.8468	0.8264	0.8290	3165
3		2.3876	9.9339	9.7016	9.2048	9.5264	
4	0.7000	36.5925	0.8210	0.8228	0.8136	0.8166	1848
3		2.5144	11.3043	9.3631	10.9885	9.7091	
4	0.8000	34.0601	0.8032	0.8014	0.7872	0.7948	616
3		2.9383	12.2496	5.0858	10.9369	7.7308	
10	0.4000	62.8596	0.5502	0.5502	0.5372	0.5370	237
9		4.5775	7.2630	11.3914	7.8740	10.9112	
10	0.5000	54.5363	0.5342	0.5320	0.5184	0.5186	33
9		3.7685	11.2996	8.8400	8.6713	6.1338	
10	0.6000	48.0529	0.5304	0.5290	0.5224	0.5234	5
9		3.3606	10.9749	7.4362	11.7939	10.3613	
10	0.7000	42.5625	0.5342	0.5342	0.5318	0.5310	4
9		2.3880	9.1383	9.5994	12.3762	10.7578	
10	0.8000	39.5214	0.5418	0.5414	0.5400	0.5376	0
9		2.9732	7.9284	11.4838	9.3243	10.0797	
20	0.3000	79.1719	0.4932	0.4900	0.4858	0.4846	13
19		6.5858	8.5454	7.6655	9.6434	13.1024	
20	0.4000	66.4252	0.5044	0.5038	0.4948	0.4950	1
19		5.6648	10.7782	9.6259	5.6230	9.9383	
20	0.5000	57.3323	0.4890	0.4876	0.4850	0.4824	0
19		4.0830	9.3385	8.9516	9.8461	10.4454	
20	0.6000	49.5371	0.5100	0.5102	0.5090	0.5074	2
19		3.4735	9.8040	9.2934	13.2754	11.1704	
20	0.7000	44.5033	0.5080	0.5082	0.5060	0.5058	0
19		2.8376	10.6675	12.7630	5.8529	12.5512	
20	0.8000	41.9317	0.5078	0.5086	0.5058	0.5072	0
19		2.6423	11.3558	10.5684	8.1855	9.3587	
100	0.2000	87.6658	0.4884	0.4846	0.4846	0.4760	0
19		9.5089	8.1582	7.8509	7.2973	6.8276	
100	0.3000	71.0336	0.4850	0.4822	0.4828	0.4744	0
19		7.6141	12.7909	10.6123	8.8549	11.5728	
100	0.4000	61.3473	0.4978	0.4936	0.4970	0.4904	0
19		5.4851	8.6004	11.2182	10.5866	14.1495	
100	0.5000	54.7728	0.5024	0.4994	0.4990	0.4952	0
19		3.4102	8.3693	7.1489	11.3929	11.3969	
100	0.6000	48.8326	0.4948	0.4956	0.4956	0.4938	0
19		3.7063	9.8485	10.7878	10.5564	12.0951	
100	0.7000	43.1747	0.4880	0.4876	0.4858	0.4862	0
19		3.7009	8.9335	9.4613	15.2218	8.8131	

Table 6.7. brsplitsim n=200,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
100	0.8000	35.6197	0.4824	0.4808	0.4816	0.4810	0
19		2.5508	10.4912	8.2168	11.7056	13.8423	
200	0.1000	96.4960	0.4636	0.4576	0.4560	0.4482	0
19		18.0810	8.2318	9.7291	13.5376	10.2221	
200	0.2000	82.8187	0.4850	0.4808	0.4812	0.4732	0
19		8.9898	9.3483	8.3389	12.5001	9.2442	
200	0.3000	67.0268	0.4824	0.4804	0.4806	0.4734	0
19		4.7945	5.0994	8.0976	10.1203	10.3304	
200	0.4000	56.8902	0.5018	0.5010	0.5012	0.5012	0
19		4.7784	10.6205	10.4675	8.5820	10.8434	
200	0.5000	51.9658	0.4988	0.5000	0.4984	0.4982	0
19		3.6967	9.8846	9.8775	11.8087	10.1569	
200	0.6000	47.2670	0.5030	0.5018	0.5008	0.5004	0
19		2.3551	11.4561	13.3504	9.8227	12.3481	
200	0.7000	40.7385	0.4938	0.4938	0.4948	0.4944	0
19		3.1094	7.5776	6.5702	5.9613	11.1288	
200	0.8000	33.4330	0.6756	0.5874	0.5872	0.5034	0
19		2.2364	8.9348	9.9660	8.7834	11.2342	

Table 6.8. prsplit n=400 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.6000	37.9468	0.9824	0.9779	0.9862	0.9858	4990
1		3.2733	8.4573	7.0065	7.4956	7.7048	
4	0.7000	36.3269	0.9872	0.9834	0.9963	0.9777	4933
1		2.1981	7.0159	8.6687	8.6838	8.6525	
4	0.8000	34.5463	0.9766	0.9750	0.9749	0.9782	4575
1		2.8491	6.4520	8.9644	7.2697	8.0324	
10	0.4000	51.5798	0.9937	0.9942	0.9869	0.9869	4998
1		3.5563	7.4781	8.3567	7.6277	7.7522	
10	0.5000	50.8833	0.9860	0.9783	0.9788	0.9821	4975
1		3.4488	7.9032	7.2290	7.6665	7.5084	
10	0.6000	47.3636	0.9875	0.9810	0.9901	0.9909	4814
1		3.8933	8.3139	7.5627	7.6215	7.8541	
10	0.7000	42.6868	0.9932	0.9765	0.9844	0.9913	4185
1		2.8677	8.2170	7.7134	6.5297	8.2743	
10	0.8000	39.7681	0.9927	0.9884	0.9892	0.9897	2812
1		2.5715	7.0037	6.6175	7.7574	7.6114	
20	0.3000	62.8304	0.9920	0.9798	0.9851	0.9828	4999
1		4.8678	7.9271	7.4442	8.5847	8.2124	
20	0.4000	60.1814	0.9876	0.9911	0.9867	0.9891	4991
1		4.4172	8.3998	7.4858	7.9137	7.1141	
20	0.5000	56.7613	0.9820	0.9891	0.9839	0.9810	4827
1		4.2682	8.2147	8.5375	8.1715	8.3550	
20	0.6000	48.8833	0.9800	0.9878	0.9821	0.9762	4210
1		2.7453	7.6271	8.0861	6.8833	7.8918	
20	0.7000	44.9911	0.9903	0.9928	0.9792	0.9817	2783
1		3.4033	6.6344	7.6008	7.1149	6.7139	
20	0.8000	42.8328	0.9826	0.9822	0.9973	0.9799	1436
1		2.3799	8.9329	8.0795	6.9363	7.0361	
200	0.2000	88.0159	0.9894	0.9844	0.9934	0.9891	4959
1		7.6854	7.8455	8.4416	8.0482	7.0335	
200	0.3000	69.3914	0.9974	0.9868	0.9884	0.9880	3948
1		5.0484	7.2466	7.2796	7.9984	8.5572	
200	0.4000	59.9825	0.9889	0.9873	0.9916	0.9899	1798
1		2.8116	9.2927	7.8276	8.6308	8.2686	
200	0.5000	53.3131	0.9822	0.9902	0.9948	0.9896	620
1		3.3426	8.2146	7.0470	7.7207	7.4503	
200	0.6000	45.9545	0.9849	0.9943	0.9884	0.9775	258
1		3.0099	8.0271	7.9776	7.3903	7.9218	
200	0.7000	40.1293	0.9913	0.9810	0.9796	0.9808	129
1		2.3180	8.1077	7.1843	8.5536	7.4977	
200	0.8000	33.3077	0.9850	0.9797	0.9946	0.9888	117
1		2.3034	7.7803	7.8859	7.8075	8.1265	
400	0.1000	106.6700	0.9894	0.9858	0.9901	0.9824	4999
1		8.8715	8.8415	8.0022	7.9132	9.2140	
400	0.2000	81.1140	0.9858	0.9795	0.9943	0.9734	4568
1		6.4030	6.9512	8.2972	8.0630	7.6029	

Table 6.9. prsplit n=400 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
400	0.3000	65.4921	0.9903	0.9839	0.9887	0.9932	1972
1		4.1592	7.9125	8.8123	7.7086	7.8397	
400	0.4000	56.6664	0.9898	0.9870	0.9947	0.9820	557
1		3.5899	8.2089	8.0117	7.6086	6.9404	
400	0.5000	50.6826	0.9827	0.9851	0.9893	0.9914	208
1		2.7399	7.8745	7.3750	7.4729	7.6331	
400	0.6000	43.2253	0.9817	0.9936	0.9878	0.9790	91
1		3.4243	7.4999	7.1897	7.7911	8.4646	
400	0.7000	37.3466	0.9919	0.9880	0.9767	0.9928	76
1		2.0224	7.2630	8.6357	8.1497	6.9705	
400	0.8000	32.5081	0.9788	0.9880	0.9926	0.9832	90
1		3.2016	8.0831	8.0888	7.5436	6.8889	
4	0.6000	39.7879	0.9694	0.9691	0.9687	0.9673	69
3		3.6585	7.3599	8.0161	7.8898	7.9341	
4	0.7000	36.9834	0.9755	0.9831	0.9802	0.9785	30
3		3.8682	8.3324	7.9399	7.5077	7.6960	
4	0.8000	33.9412	0.9884	0.9868	0.9903	0.9833	15
3		1.7998	7.2403	7.6879	7.2713	7.4725	
10	0.4000	62.8984	0.9857	0.9833	0.9816	0.9924	205
9		5.6622	8.0381	7.7630	7.6936	7.2342	
10	0.5000	55.7740	0.9866	0.9930	0.9916	0.9813	21
9		5.0597	8.1224	8.1410	8.3404	8.5091	
10	0.6000	46.7899	0.9925	0.9882	0.9967	0.9964	1
9		3.8913	7.1895	6.9732	8.1701	8.1687	
10	0.7000	42.1899	0.9823	0.9883	0.9842	0.9804	0
9		2.3141	7.7910	8.4756	8.0059	8.9443	
10	0.8000	39.6730	0.9851	0.9887	0.9818	0.9778	0
9		3.0139	8.1072	7.9250	8.0682	8.1545	
20	0.3000	85.3290	0.9922	0.9936	0.9923	0.9956	4
19		7.2573	7.7287	7.7028	7.2917	8.1098	
20	0.4000	68.3068	0.9916	0.9855	0.9939	0.9814	0
19		5.0834	7.7765	7.6775	8.4704	7.8269	
20	0.5000	57.4325	0.9864	0.9896	0.9817	0.9901	0
19		4.7374	8.4413	8.3736	8.0697	7.7108	
20	0.6000	49.9498	0.9936	0.9916	0.9882	0.9830	0
19		3.1656	8.1778	8.4791	8.2753	8.3508	
20	0.7000	45.8567	0.9805	0.9832	0.9908	0.9874	0
19		3.0670	7.0056	8.4156	6.7882	7.4806	
20	0.8000	42.6101	0.9859	0.9864	0.9847	0.9852	0
19		2.3570	7.8449	7.9167	7.1197	7.7765	
200	0.2000	97.1515	0.9836	0.9973	0.9813	0.9908	0
19		8.1493	7.1389	7.6818	7.8337	6.5275	

Table 6.10. prsplit n=400 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
200	0.3000	71.3142	0.9944	0.9946	0.9835	0.9854	0
19		4.5819	7.6420	7.7577	8.4117	7.7125	
200	0.4000	59.8689	0.9929	0.9845	0.9902	0.9871	0
19		4.0209	8.2020	8.5841	7.5725	6.9643	
200	0.5000	52.6699	0.9861	0.9897	0.9955	0.9752	0
19		3.2338	6.9595	8.2858	7.7903	6.8320	
200	0.6000	46.5318	0.9815	0.9821	0.9807	0.9782	0
19		2.6996	8.0077	7.8130	7.9841	9.1634	
200	0.7000	38.8062	0.9821	0.9852	0.9804	0.9947	0
19		2.6799	7.6952	7.1542	8.0753	7.3715	
200	0.8000	32.7557	0.9962	0.9811	0.9970	0.9884	0
19		2.2985	8.2636	7.5150	7.7480	7.5459	
400	0.1000	155.6178	0.9896	0.9945	0.9900	0.9877	0
19		15.2212	8.1532	8.9197	8.2326	7.7718	
400	0.2000	83.4805	0.9907	0.9782	0.9914	0.9881	0
19		6.6353	8.0799	7.4621	7.8052	8.1810	
400	0.3000	64.8805	0.9809	0.9927	0.9863	0.9843	0
19		5.0432	7.7671	8.2502	8.4726	8.3370	
400	0.4000	56.1989	0.9938	0.9795	0.9933	0.9821	0
19		3.1990	7.6589	8.2108	9.2946	7.2336	
400	0.5000	49.5555	0.9832	0.9872	0.9886	0.9788	0
19		3.5322	8.3565	8.8724	8.3249	7.2993	
400	0.6000	43.5944	0.9879	0.9920	0.9781	0.9885	0
19		2.5967	7.8192	8.1613	7.0175	7.8588	
400	0.7000	35.4655	0.9981	0.9861	0.9870	0.9873	0
19		2.4800	8.1736	8.1575	8.2275	8.3343	
400	0.8000	32.7132	0.9813	0.9805	0.9815	0.9818	0
19		3.1469	8.0289	7.8563	7.9712	7.8376	

Table 6.11. prsplit n=1000 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.0000	36.8298	0.9823	0.9808	0.9883	0.9840	4993
1		2.9703	7.5124	7.6328	8.5947	8.0170	
10	0.0000	49.0575	0.9909	0.9875	0.9927	0.9895	4994
1		3.3267	7.8335	8.2737	8.2098	7.5758	
20	0.0000	56.6908	0.9869	0.9777	0.9908	0.9842	4978
1		3.3662	7.0329	7.7289	7.9915	8.1688	
500	0.0000	88.8772	0.9864	0.9844	0.9834	0.9886	4863
1		5.7053	8.1009	7.6515	7.9471	7.7390	
1000	0.0000	95.2273	0.9856	0.9908	0.9924	0.9744	4789
1		5.2895	8.0867	8.7953	7.8825	7.8767	
4	0.0000	59.4637	0.9857	0.9821	0.9843	0.9788	4822
3		3.8614	7.2789	7.4257	7.6053	7.7758	
10	0.0000	112.6137	1.0000	0.9929	0.9872	0.9880	4450
9		8.1819	8.1302	7.3750	6.9965	8.3154	
20	0.0000	163.3744	0.9882	0.9853	0.9862	0.9802	3613
19		15.3036	9.3150	7.8899	7.3503	8.1431	
500	0.0000	103.0775	0.9994	0.9883	0.9920	0.9749	614
19		13.1054	8.9484	8.2658	8.3206	8.2635	
1000	0.0000	78.9262	0.9866	0.9728	0.9894	0.9739	346
19		9.0479	8.2306	8.6517	8.4152	7.8264	
4	0.5000	37.1492	0.9929	0.9906	0.9901	0.9849	4959
1		2.7424	7.5358	7.5492	7.3655	8.5092	
10	0.3160	51.0483	0.9877	0.9863	0.9889	0.9901	4973
1		3.6766	7.4373	8.1452	7.5150	8.3925	
20	0.2240	61.5877	0.9742	0.9917	0.9958	0.9871	4986
1		4.4027	7.5176	7.5921	7.0217	7.8748	
500	0.0400	104.9566	0.9913	0.9825	0.9965	0.9862	4901
1		6.7095	8.0162	7.5070	7.9496	7.3994	
1000	0.0300	116.7117	0.9900	0.9865	0.9877	0.9817	4892
1		8.3360	8.1533	7.1798	7.4977	6.8152	
4	0.5000	43.9203	0.9841	0.9850	0.9822	0.9764	1928
3		4.6157	7.6582	7.5049	7.5613	6.9994	
10	0.3160	74.3965	0.9920	0.9920	0.9899	0.9836	24
9		7.0638	7.5158	7.9454	8.1729	8.1484	
20	0.2240	114.3865	0.9917	0.9916	0.9814	0.9785	1
19		9.7871	8.1534	7.4613	7.8464	7.6417	
500	0.0400	426.0544	0.9965	0.9862	0.9926	0.9769	0
19		41.9470	8.3119	8.5249	7.0565	8.2411	
1000	0.0300	464.0275	0.9948	0.9824	0.9898	0.9912	0
19		49.7073	7.9093	8.9000	7.1584	8.3593	
4	0.6000	36.3861	0.9871	0.9878	0.9819	0.9912	4813
1		3.6599	8.6049	7.3648	8.1081	7.7428	
4	0.7000	36.8088	0.9769	0.9901	0.9860	0.9799	4474
1		3.6323	7.9195	8.2349	6.8977	7.3928	

Table 6.12. prsplit n=1000 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.8000	34.1201	0.9911	0.9794	0.9855	0.9792	3718
1		2.5867	7.5125	6.9803	7.9171	6.9004	
10	0.4000	52.4290	0.9764	0.9819	0.9773	0.9792	4902
1		3.8920	7.2416	7.9435	7.8625	8.2432	
10	0.5000	49.6841	0.9881	0.9944	0.9906	0.9878	4532
1		4.2070	7.9839	7.2459	7.4718	7.9675	
10	0.6000	47.2333	0.9891	1.0000	0.9869	0.9785	3660
1		3.5406	7.7978	7.2433	7.5919	7.3823	
10	0.7000	43.5112	0.9866	0.9771	0.9847	0.9769	2601
1		3.7857	8.0985	7.1402	7.1219	8.6710	
10	0.8000	39.9461	0.9812	0.9860	0.9899	0.9790	1728
1		2.5347	8.1781	7.4448	8.2237	7.0545	
20	0.3000	62.2457	0.9876	0.9771	0.9826	0.9889	4910
1		4.3706	7.8638	7.6618	7.2295	7.6853	
20	0.4000	59.9362	0.9811	0.9830	0.9801	0.9773	4494
1		4.6292	8.2569	7.9760	7.1385	8.4864	
20	0.5000	56.4071	0.9834	0.9915	0.9849	1.0000	3418
1		3.9243	7.9746	8.1575	7.1918	7.2256	
20	0.6000	51.1867	0.9721	0.9872	0.9871	0.9860	2271
1		3.3375	7.3855	8.2086	7.9146	7.2235	
20	0.7000	45.2468	0.9919	0.9830	0.9871	0.9889	1284
1		2.7033	7.3022	6.4302	7.0698	7.4170	
20	0.8000	43.1377	0.9883	0.9788	0.9832	0.9876	755
1		1.9365	7.4286	8.2306	8.2887	8.4295	
500	0.1000	111.4411	0.9837	0.9816	0.9899	0.9886	4664
1		7.6625	8.3230	7.8910	7.5375	8.8045	
500	0.2000	80.2712	0.9740	0.9789	0.9926	0.9831	1219
1		5.9207	7.0859	7.9661	7.8510	6.4874	
500	0.3000	62.6021	0.9956	0.9937	0.9880	0.9894	266
1		3.6019	7.4199	6.8162	7.5114	8.1359	
500	0.4000	56.5252	0.9987	0.9954	0.9951	0.9893	116
1		3.4650	7.7560	8.1852	6.7260	7.4746	
500	0.5000	47.9767	0.9865	0.9825	0.9795	0.9911	55
1		3.9646	7.7424	8.6767	7.5678	7.8299	
500	0.6000	42.1734	0.9958	0.9933	0.9748	0.9832	38
1		2.2591	8.4028	7.9374	8.3461	8.1158	
500	0.7000	35.6188	0.9887	0.9812	0.9879	0.9860	19
1		2.6522	6.8601	7.6227	7.7682	8.6952	
500	0.8000	32.7235	0.9834	0.9867	0.9907	0.9795	19
1		2.5539	8.0232	7.5612	8.0432	7.5649	
1000	0.1000	110.0411	0.9880	0.9884	0.9902	0.9881	3240
1		8.8354	7.3550	8.5021	8.2190	8.1900	
1000	0.2000	72.1917	0.9835	0.9966	0.9898	0.9874	339
1		5.4641	7.7972	7.6897	8.1949	8.4211	

Table 6.13. prsplit n=1000 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
1000	0.3000	59.3060	0.9714	0.9816	0.9789	0.9867	82
1		4.0469	6.8533	7.7621	8.2206	8.0131	
1000	0.4000	51.6666	0.9772	0.9832	0.9806	0.9790	30
1		2.7509	7.6252	7.6112	7.3900	7.7805	
1000	0.5000	46.1104	0.9778	0.9828	0.9872	0.9859	13
1		2.8608	8.0646	8.9271	7.4973	7.6695	
1000	0.6000	38.0103	0.9888	0.9866	0.9900	0.9863	15
1		2.8235	8.6182	8.6069	7.2434	8.1339	
1000	0.7000	32.7950	0.9848	0.9759	0.9751	0.9859	13
1		2.3617	7.0251	8.3526	7.9753	7.6571	
1000	0.8000	34.7974	0.9906	0.9836	0.9875	0.9858	17
1		2.6669	7.7272	7.5282	6.8666	8.0647	
4	0.6000	39.2934	0.9851	0.9752	0.9783	0.9822	1075
3		3.8604	8.3560	7.6025	8.4466	6.9445	
4	0.7000	36.6253	0.9840	0.9863	0.9837	0.9825	460
3		2.0863	7.8106	7.9617	7.3583	7.6827	
4	0.8000	33.9027	0.9913	0.9982	0.9859	0.9794	185
3		2.3709	7.6508	6.9104	7.8287	7.7324	
10	0.4000	63.7907	0.9933	0.9865	0.9913	0.9843	4
9		5.9405	8.1578	8.2959	7.5260	8.2921	
10	0.5000	55.4546	0.9936	0.9967	0.9868	0.9851	1
9		4.5311	7.6004	8.7210	7.5195	6.8551	
10	0.6000	48.5367	0.9833	0.9911	0.9864	0.9877	2
9		3.4938	7.0631	8.1991	9.0657	7.9709	
10	0.7000	43.3798	0.9873	0.9926	0.9831	0.9894	0
9		2.2387	7.1967	7.5455	8.1177	8.7701	
10	0.8000	39.5635	0.9763	0.9916	0.9960	0.9868	2
9		2.5949	8.3362	7.7051	8.0004	8.4142	
20	0.3000	85.9431	0.9846	0.9889	0.9927	0.9760	0
19		6.5828	8.4947	8.1404	6.7265	6.7310	
20	0.4000	67.8207	0.9816	0.9876	0.9843	0.9926	0
19		5.4987	8.2090	7.1631	7.6717	8.0408	
20	0.5000	57.2194	0.9886	0.9793	0.9810	0.9883	0
19		4.3419	8.1174	7.8225	7.9158	8.6044	
20	0.6000	50.6866	0.9831	0.9894	0.9881	0.9838	0
19		3.4483	7.4020	7.2876	7.1413	7.0166	
20	0.7000	45.0602	0.9844	0.9836	0.9843	0.9955	0
19		3.1198	8.2488	7.2705	7.5149	8.0042	
20	0.8000	42.3669	0.9808	0.9921	0.9819	0.9885	0
19		3.1652	8.0925	7.5282	8.0179	7.8440	
500	0.1000	152.4692	0.9913	0.9797	0.9876	0.9825	0
19		12.6269	7.5625	7.9409	8.5135	8.6221	
500	0.2000	80.9184	0.9851	0.9841	0.9839	0.9923	0
19		6.3860	9.4077	7.4053	7.7352	7.0157	

Table 6.14. prsplit n=1000 ,type=1,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
500	0.3000	62.8484	0.9885	0.9922	0.9957	0.9884	0
19		4.0928	8.8133	8.1315	7.0348	8.0525	
500	0.4000	56.3032	0.9802	0.9767	0.9786	0.9966	0
19		4.0795	8.0178	7.1714	8.0622	7.6822	
500	0.5000	49.0593	0.9873	0.9921	0.9848	0.9771	0
19		3.8636	6.8921	7.7831	8.6151	7.4424	
500	0.6000	42.1069	0.9914	0.9914	0.9922	0.9866	0
19		3.4903	7.0869	7.3199	7.5057	7.3422	
500	0.7000	35.5471	0.9864	0.9943	0.9976	0.9780	0
19		2.7485	8.1432	7.8422	7.5289	7.4664	
500	0.8000	32.2594	0.9875	0.9871	0.9856	0.9788	0
19		2.5240	8.0372	7.9699	7.1135	7.3324	
1000	0.1000	121.6233	0.9879	0.9826	0.9877	0.9797	0
19		10.0963	7.7023	7.7092	6.7835	8.0915	
1000	0.2000	71.0325	0.9880	0.9825	0.9817	0.9883	0
19		5.6467	8.1866	8.0372	6.4767	7.6533	
1000	0.3000	59.6453	0.9875	0.9752	0.9867	0.9792	0
19		4.5016	7.9923	7.8321	8.3675	7.8521	
1000	0.4000	52.3941	0.9839	0.9854	0.9789	0.9862	0
19		3.8089	7.6693	8.1382	7.3613	8.7149	
1000	0.5000	45.2140	0.9863	0.9815	0.9830	0.9838	0
19		2.4686	6.4362	8.1005	7.3341	6.7855	
1000	0.6000	38.2262	0.9831	0.9895	0.9800	0.9830	0
19		1.9144	7.6717	7.8793	8.4134	7.8265	
1000	0.7000	33.2980	0.9853	0.9904	0.9845	0.9878	0
19		1.8851	7.0836	7.1032	7.5263	8.1112	
1000	0.8000	34.4032	0.9846	0.9804	0.9894	0.9821	0
19		3.0741	7.2761	7.9806	7.4777	7.4166	

Table 6.15. brsplitsim n=100,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.0000	33.9066	0.9914	0.9896	0.9910	0.9872	4996
1		2.6168	2.6678	2.6014	2.6760	2.5832	
10	0.0000	38.9590	0.9944	0.9892	0.9904	0.9834	4988
1		3.6014	2.7686	2.6246	2.7832	2.5904	
20	0.0000	40.9616	0.9932	0.9850	0.9904	0.9752	4991
1		4.6866	2.8414	2.6524	2.8588	2.5972	
50	0.0000	42.4542	0.9940	0.9824	0.9900	0.9642	4964
1		6.5192	2.8936	2.6314	2.9104	2.5616	
100	0.0000	43.4521	0.9854	0.9844	0.9945	0.9843	4987
1		5.6745	2.6474	2.4433	2.9443	2.5556	
4	0.0000	46.4152	0.9922	0.9904	0.9918	0.9898	4823
3		3.9532	2.7036	2.6720	2.7018	2.6532	
10	0.0000	45.4904	0.9922	0.9892	0.9854	0.9804	3147
9		8.7846	2.8044	2.7408	2.8206	2.6786	
20	0.0000	42.9846	0.9830	0.9772	0.9696	0.9478	159
19		10.7912	2.8156	2.6736	2.9150	2.5984	
50	0.0000	43.3422	0.9786	0.9882	0.9767	0.9564	4996
19		8.6763	2.7664	2.7664	2.9002	2.5893	
100	0.0000	43.4565	0.9934	0.9677	0.9988	0.9784	4986
19		5.6723	2.7664	2.6773	2.8774	2.6653	
4	0.5000	33.9644	0.9926	0.9914	0.9916	0.9900	4965
1		2.7514	2.6498	2.6062	2.6572	2.5994	
10	0.3160	40.5978	0.9926	0.9886	0.9900	0.9840	4980
1		3.9866	2.7670	2.6740	2.7736	2.6364	
20	0.2240	42.8962	0.9946	0.9892	0.9932	0.9808	4965
1		5.1880	2.8134	2.6826	2.8246	2.6360	
50	0.1410	43.5655	0.9945	0.9834	0.9834	0.9785	4966
1		6.7882	2.8756	2.8945	2.5784	2.4534	
100	0.1000	45.1606	0.9914	0.9764	0.9880	0.9586	4940
1		8.8390	2.8772	2.6418	2.8900	2.5438	
4	0.5000	37.9696	0.9934	0.9922	0.9916	0.9900	1889
3		3.3084	2.6874	2.6610	2.6910	2.6328	
10	0.3160	45.9664	0.9938	0.9924	0.9928	0.9866	57
9		5.9494	2.8328	2.7464	2.8322	2.7034	
20	0.2240	46.6294	0.9946	0.9910	0.9920	0.9838	2
19		8.0910	2.8424	2.7158	2.8404	2.6428	
50	0.1410	46.8300	0.9912	0.9828	0.9870	0.9634	0
19		11.1226	2.8362	2.6406	2.8228	2.5168	
100	0.1000	46.8674	0.9916	0.9788	0.9874	0.9536	0
19		13.0646	2.8446	2.5978	2.8340	2.4698	
4	0.6000	33.9848	0.9880	0.9874	0.9880	0.9876	4821
1		2.7442	2.6282	2.5900	2.6338	2.5756	
4	0.7000	33.3592	0.9904	0.9884	0.9890	0.9862	4399
1		2.6866	2.6538	2.6128	2.6462	2.6022	

Table 6.16. brsplitsim n=100,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.8000	32.6724	0.9894	0.9888	0.9908	0.9884	3685
1		2.5686	2.6048	2.5800	2.6032	2.5630	
10	0.4000	40.8120	0.9920	0.9884	0.9902	0.9854	4855
1		3.9956	2.7260	2.6472	2.7396	2.6288	
10	0.5000	40.1456	0.9920	0.9908	0.9896	0.9876	4407
1		3.8720	2.7256	2.6610	2.7332	2.6398	
10	0.6000	38.7108	0.9916	0.9900	0.9902	0.9866	3526
1		3.5768	2.7274	2.6690	2.7198	2.6358	
10	0.7000	36.6198	0.9898	0.9888	0.9876	0.9844	2619
1		3.1880	2.7016	2.6624	2.6888	2.6246	
10	0.8000	35.0660	0.9916	0.9890	0.9892	0.9880	1604
1		2.8968	2.6444	2.6094	2.6360	2.5794	
20	0.3000	43.3212	0.9940	0.9904	0.9916	0.9838	4827
1		5.3398	2.7900	2.6670	2.7896	2.6318	
20	0.4000	42.9370	0.9914	0.9872	0.9904	0.9814	4241
1		5.0932	2.7778	2.6930	2.7686	2.6364	
20	0.5000	41.4784	0.9934	0.9904	0.9908	0.9820	3163
1		4.5670	2.7706	2.7130	2.7524	2.6444	
20	0.6000	39.3330	0.9924	0.9896	0.9904	0.9846	2074
1		4.0242	2.7406	2.6868	2.7220	2.6330	
20	0.7000	37.4358	0.9944	0.9904	0.9920	0.9890	1370
1		3.5214	2.6964	2.6436	2.6720	2.5934	
20	0.8000	35.9432	0.9922	0.9898	0.9910	0.9848	833
1		3.2014	2.6674	2.6312	2.6620	2.6018	
50	0.2000	45.0756	0.9930	0.9824	0.9900	0.9666	4800
1		7.5138	2.8362	2.6752	2.8276	2.5828	
50	0.3000	44.7526	0.9914	0.9864	0.9876	0.9710	3683
1		6.9780	2.8090	2.6826	2.8002	2.6034	
50	0.4000	43.2464	0.9924	0.9854	0.9872	0.9712	2305
1		6.0556	2.7856	2.6758	2.7572	2.5914	
50	0.5000	41.1316	0.9912	0.9858	0.9898	0.9772	1372
1		4.9632	2.7546	2.6632	2.7262	2.5950	
50	0.6000	38.7924	0.9914	0.9850	0.9892	0.9788	746
1		4.0706	2.7064	2.6368	2.6796	2.5748	
50	0.7000	37.1842	0.9916	0.9854	0.9886	0.9796	436
1		3.3496	2.7016	2.6362	2.6820	2.5766	
50	0.8000	34.4098	0.9886	0.9844	0.9894	0.9808	263
1		2.6148	2.6170	2.5662	2.6164	2.5376	
100	0.2000	45.6434	0.9948	0.9782	0.9886	0.9624	3881
1		8.7554	2.8426	2.6696	2.8244	2.5558	
100	0.3000	44.4432	0.9904	0.9792	0.9894	0.9700	2103
1		7.4654	2.8000	2.6602	2.7832	2.5890	
100	0.4000	42.3386	0.9942	0.9836	0.9906	0.9746	1061
1		6.0972	2.7810	2.6636	2.7436	2.5796	

Table 6.17. brsplitsim n=100,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
100	0.5000	40.2646	0.9920	0.9842	0.9886	0.9756	534
1		4.9308	2.7350	2.6512	2.7138	2.5782	
100	0.6000	38.4014	0.9924	0.9860	0.9922	0.9808	309
1		4.0060	2.7362	2.6540	2.7192	2.6028	
100	0.7000	36.4668	0.9908	0.9880	0.9916	0.9846	174
1		3.1580	2.6580	2.6174	2.6550	2.5718	
100	0.8000	33.2062	0.9874	0.9870	0.9866	0.9854	103
1		2.4368	2.5940	2.5744	2.5936	2.5456	
4	0.6000	35.4706	0.9924	0.9910	0.9916	0.9898	1004
3		3.0500	2.6666	2.6362	2.6608	2.6050	
4	0.7000	33.4578	0.9910	0.9906	0.9912	0.9890	446
3		2.8116	2.6678	2.6326	2.6588	2.6162	
4	0.8000	32.4310	0.9894	0.9898	0.9890	0.9868	194
3		2.5688	2.6312	2.6002	2.6214	2.5810	
10	0.4000	44.7764	0.9954	0.9920	0.9932	0.9882	32
9		5.2300	2.8228	2.7490	2.8094	2.6894	
10	0.5000	42.0666	0.9932	0.9908	0.9922	0.9890	35
9		4.3690	2.8168	2.7508	2.7876	2.6896	
10	0.6000	39.2650	0.9924	0.9900	0.9898	0.9870	24
9		3.6876	2.7486	2.6838	2.7210	2.6342	
10	0.7000	36.6946	0.9952	0.9928	0.9920	0.9884	27
9		3.2084	2.7252	2.6760	2.6942	2.6324	
10	0.8000	34.8586	0.9892	0.9876	0.9864	0.9840	17
9		2.8622	2.6692	2.6294	2.6612	2.6040	
20	0.3000	46.1398	0.9912	0.9900	0.9922	0.9850	2
19		6.9516	2.8222	2.7280	2.8332	2.6646	
20	0.4000	44.7050	0.9940	0.9916	0.9924	0.9840	0
19		5.7408	2.8218	2.7326	2.8050	2.6758	
20	0.5000	42.1142	0.9920	0.9890	0.9918	0.9824	2
19		4.7598	2.7996	2.7330	2.7712	2.6496	
20	0.6000	39.3806	0.9926	0.9912	0.9908	0.9848	0
19		3.9632	2.7146	2.6534	2.6846	2.5902	
20	0.7000	37.4222	0.9920	0.9900	0.9906	0.9872	1
19		3.4952	2.6820	2.6302	2.6540	2.5784	
20	0.8000	36.0350	0.9906	0.9882	0.9878	0.9828	0
19		3.1672	2.6782	2.6372	2.6670	2.5950	
50	0.2000	46.7450	0.9916	0.9846	0.9880	0.9714	0
19		9.3824	2.8352	2.6806	2.8208	2.5720	
50	0.3000	45.7250	0.9928	0.9876	0.9912	0.9760	0
19		7.5620	2.8154	2.6954	2.7934	2.6082	
50	0.4000	43.4096	0.9952	0.9886	0.9906	0.9782	0
19		5.9874	2.7698	2.6802	2.7578	2.5990	
50	0.5000	41.0126	0.9928	0.9838	0.9906	0.9780	0
19		4.9702	2.7728	2.6898	2.7436	2.6180	

Table 6.18. brsplitsim n=100,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
50	0.6000	39.0168	0.9926	0.9872	0.9906	0.9814	0
19		4.1152	2.7300	2.6610	2.6972	2.5806	
50	0.7000	36.9700	0.9894	0.9834	0.9876	0.9784	0
19		3.3582	2.6798	2.6224	2.6694	2.5726	
50	0.8000	34.4438	0.9936	0.9888	0.9928	0.9850	0
19		2.6514	2.6522	2.5966	2.6426	2.5592	
100	0.2000	46.5444	0.9926	0.9822	0.9890	0.9690	0
19		9.5246	2.8372	2.6852	2.8316	2.5902	
100	0.3000	44.6982	0.9926	0.9812	0.9918	0.9730	0
19		7.4932	2.8150	2.6934	2.7960	2.6012	
100	0.4000	42.6752	0.9928	0.9840	0.9908	0.9724	0
19		6.0868	2.7926	2.6798	2.7466	2.5850	
100	0.5000	40.3530	0.9912	0.9836	0.9902	0.9766	0
19		4.9642	2.7316	2.6346	2.6918	2.5574	
100	0.6000	38.4626	0.9914	0.9856	0.9906	0.9802	0
19		4.0072	2.7106	2.6302	2.6880	2.5880	
100	0.7000	36.5586	0.9896	0.9862	0.9896	0.9838	0
19		3.1766	2.6526	2.6026	2.6476	2.5590	
100	0.8000	33.0872	0.9894	0.9894	0.9896	0.9872	0
19		2.4460	2.6244	2.6044	2.6108	2.5718	

Table 6.19. brsplitsim n=200,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.0000	36.8700	0.9904	0.9874	0.9896	0.9872	4998
1		2.5962	2.5400	2.4752	2.5366	2.4810	
10	0.0000	48.3858	0.9906	0.9860	0.9906	0.9864	4994
1		3.3038	2.6258	2.5206	2.6232	2.5174	
20	0.0000	55.9998	0.9922	0.9864	0.9908	0.9848	4985
1		3.8384	2.6978	2.5660	2.7046	2.5646	
100	0.0000	67.4134	0.9908	0.9814	0.9920	0.9736	4950
1		5.9906	2.8088	2.5924	2.8338	2.5668	
200	0.0000	70.1800	0.9942	0.9816	0.9942	0.9700	4900
1		7.3360	2.8210	2.5764	2.8482	2.5340	
4	0.0000	58.9920	0.9876	0.9878	0.9872	0.9868	4830
3		3.9582	2.5272	2.5098	2.5290	2.5072	
10	0.0000	90.8728	0.9916	0.9928	0.9902	0.9896	4429
9		9.2304	2.7838	2.7628	2.7824	2.7310	
20	0.0000	80.4662	0.9922	0.9906	0.9848	0.9826	2093
19		14.7138	2.8132	2.7570	2.8136	2.6884	
100	0.0000	61.0392	0.9870	0.9564	0.9770	0.9296	12
19		12.2818	2.9892	2.5684	3.0300	2.5164	
200	0.0000	55.9436	0.9876	0.9340	0.9730	0.9102	0
19		9.1532	3.0644	2.4972	3.0956	2.5224	
4	0.5000	37.3140	0.9880	0.9862	0.9872	0.9862	4951
1		2.7316	2.5232	2.4916	2.5220	2.4912	
10	0.3000	51.1004	0.9906	0.9892	0.9896	0.9860	4981
1		3.6750	2.6184	2.5594	2.6166	2.5366	
20	0.2000	60.4538	0.9918	0.9898	0.9904	0.9852	4982
1		4.4298	2.6540	2.5642	2.6690	2.5472	
100	0.1000	74.7792	0.9942	0.9856	0.9918	0.9782	4947
1		7.0926	2.7920	2.6292	2.8060	2.5916	
200	0.0700	79.8188	0.9948	0.9818	0.9934	0.9750	4936
1		9.0404	2.8100	2.6158	2.8148	2.5624	
4	0.5000	44.2080	0.9876	0.9876	0.9868	0.9866	1939
3		3.3098	2.5640	2.5428	2.5632	2.5368	
10	0.3000	73.9934	0.9924	0.9910	0.9908	0.9890	53
9		6.4762	2.7576	2.7136	2.7368	2.6654	
20	0.2000	92.1246	0.9934	0.9922	0.9918	0.9882	0
19		10.2068	2.8174	2.7370	2.8236	2.6916	
100	0.1000	96.4652	0.9954	0.9882	0.9910	0.9750	0
19		17.3780	2.8530	2.6908	2.8338	2.5898	
200	0.0700	96.5938	0.9948	0.9860	0.9930	0.9750	0
19		20.7338	2.8462	2.6636	2.8432	2.5472	
4	0.6000	37.0080	0.9862	0.9848	0.9850	0.9846	4830
1		2.7268	2.5114	2.4880	2.5120	2.4842	

Table 6.20. brsplitsim n=200,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.7000	36.2280	0.9852	0.9838	0.9842	0.9838	4418
1		2.6746	2.5050	2.4822	2.5054	2.4862	
4	0.8000	34.0860	0.9880	0.9884	0.9872	0.9866	3710
1		2.5368	2.5172	2.5044	2.5190	2.5006	
10	0.4000	51.3456	0.9896	0.9880	0.9902	0.9884	4887
1		3.7272	2.5816	2.5362	2.5808	2.5292	
10	0.5000	50.2078	0.9890	0.9868	0.9874	0.9860	4552
1		3.6520	2.6076	2.5708	2.6158	2.5750	
10	0.6000	47.1288	0.9892	0.9882	0.9884	0.9878	3679
1		3.4248	2.5812	2.5544	2.5818	2.5476	
10	0.7000	42.9348	0.9912	0.9896	0.9892	0.9876	2610
1		3.0818	2.5662	2.5444	2.5628	2.5306	
10	0.8000	39.6112	0.9872	0.9858	0.9878	0.9856	1663
1		2.7822	2.5708	2.5430	2.5688	2.5426	
20	0.3000	60.8708	0.9892	0.9864	0.9892	0.9854	4910
1		4.4882	2.6654	2.5942	2.6718	2.5788	
20	0.4000	60.3592	0.9890	0.9860	0.9870	0.9836	4524
1		4.5208	2.6390	2.5884	2.6396	2.5712	
20	0.5000	54.9690	0.9910	0.9896	0.9888	0.9860	3429
1		4.1214	2.6190	2.5726	2.6110	2.5542	
20	0.6000	49.9906	0.9910	0.9902	0.9914	0.9882	2239
1		3.6196	2.6152	2.5878	2.6202	2.5744	
20	0.7000	45.1894	0.9896	0.9886	0.9896	0.9874	1307
1		3.1620	2.5988	2.5654	2.5864	2.5524	
20	0.8000	42.4038	0.9902	0.9884	0.9892	0.9874	834
1		2.8564	2.5594	2.5406	2.5600	2.5342	
100	0.2000	77.6564	0.9918	0.9872	0.9888	0.9810	4643
1		7.4916	2.7430	2.6400	2.7456	2.6026	
100	0.3000	69.4304	0.9934	0.9868	0.9920	0.9830	2814
1		6.2226	2.7116	2.6364	2.7096	2.6052	
100	0.4000	60.9350	0.9918	0.9886	0.9918	0.9860	1286
1		4.9810	2.6966	2.6308	2.6828	2.5922	
100	0.5000	53.6914	0.9896	0.9842	0.9886	0.9838	590
1		3.9510	2.6270	2.5636	2.6222	2.5456	
100	0.6000	48.7500	0.9906	0.9870	0.9902	0.9804	315
1		3.3370	2.6136	2.5656	2.6032	2.5348	
100	0.7000	43.4286	0.9892	0.9864	0.9914	0.9860	162
1		2.8480	2.5720	2.5312	2.5838	2.5342	
100	0.8000	35.6280	0.9876	0.9870	0.9866	0.9870	99
1		2.4142	2.5350	2.5140	2.5294	2.5150	
200	0.2000	77.9492	0.9922	0.9840	0.9890	0.9774	3367
1		8.2394	2.7694	2.6606	2.7612	2.6110	
200	0.3000	65.4682	0.9920	0.9848	0.9906	0.9814	1258
1		5.8834	2.7374	2.6560	2.7152	2.6084	

Table 6.21. brsplitsim n=200,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
200	0.4000	57.2340	0.9876	0.9834	0.9882	0.9804	466
1		4.5450	2.6724	2.6026	2.6532	2.5666	
200	0.5000	52.3108	0.9892	0.9848	0.9894	0.9848	222
1		3.6810	2.6248	2.5734	2.6234	2.5608	
200	0.6000	46.7652	0.9918	0.9892	0.9918	0.9890	127
1		3.0710	2.5604	2.5330	2.5660	2.5318	
200	0.7000	40.1894	0.9844	0.9830	0.9848	0.9822	72
1		2.6160	2.5268	2.5032	2.5242	2.4946	
200	0.8000	33.2640	0.9896	0.9894	0.9888	0.9882	59
1		2.3362	2.5444	2.5374	2.5418	2.5346	
4	0.6000	39.7800	0.9886	0.9882	0.9886	0.9876	1051
3		3.0648	2.5426	2.5322	2.5580	2.5354	
4	0.7000	36.6420	0.9888	0.9874	0.9882	0.9874	504
3		2.8066	2.5414	2.5220	2.5446	2.5208	
4	0.8000	34.1160	0.9878	0.9878	0.9888	0.9874	185
3		2.5690	2.4832	2.4682	2.4856	2.4638	
10	0.4000	63.2606	0.9918	0.9912	0.9912	0.9888	4
9		5.3008	2.7138	2.6714	2.7026	2.6372	
10	0.5000	54.3768	0.9910	0.9904	0.9910	0.9886	4
9		4.3712	2.6692	2.6372	2.6486	2.6084	
10	0.6000	47.9798	0.9896	0.9892	0.9890	0.9882	3
9		3.6514	2.5960	2.5676	2.5890	2.5542	
10	0.7000	42.5918	0.9864	0.9844	0.9860	0.9842	3
9		3.1098	2.5802	2.5534	2.5770	2.5548	
10	0.8000	39.3204	0.9880	0.9866	0.9882	0.9866	2
9		2.7608	2.5460	2.5286	2.5544	2.5300	
20	0.3000	79.9066	0.9932	0.9920	0.9924	0.9874	0
19		7.4126	2.7946	2.7332	2.7684	2.6690	
20	0.4000	66.5602	0.9936	0.9912	0.9924	0.9912	0
19		5.5810	2.7492	2.6910	2.7104	2.6344	
20	0.5000	56.4638	0.9922	0.9904	0.9894	0.9876	0
19		4.4206	2.6796	2.6338	2.6580	2.5992	
20	0.6000	49.8248	0.9912	0.9892	0.9896	0.9894	0
19		3.6670	2.6298	2.5988	2.6242	2.5816	
20	0.7000	45.0218	0.9900	0.9890	0.9894	0.9870	0
19		3.1312	2.6174	2.5864	2.5992	2.5608	
20	0.8000	42.1772	0.9842	0.9850	0.9866	0.9840	0
19		2.8570	2.5398	2.5298	2.5560	2.5264	
100	0.2000	88.2544	0.9940	0.9908	0.9922	0.9828	0
19		10.1834	2.8464	2.7380	2.8254	2.6740	
100	0.3000	71.7182	0.9922	0.9866	0.9906	0.9822	0
19		6.6290	2.7536	2.6750	2.7196	2.6152	
100	0.4000	60.7708	0.9928	0.9908	0.9910	0.9866	0
19		4.9294	2.6860	2.6228	2.6602	2.5870	

Table 6.22. brsplitsim n=200,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
100	0.5000	54.1136	0.9932	0.9876	0.9930	0.9864	0
19		3.9888	2.6274	2.5834	2.6156	2.5470	
100	0.6000	48.7712	0.9880	0.9846	0.9884	0.9816	0
19		3.3442	2.6130	2.5590	2.5944	2.5324	
100	0.7000	42.9126	0.9902	0.9884	0.9914	0.9848	0
19		2.8134	2.5576	2.5250	2.5674	2.5090	
100	0.8000	35.3820	0.9878	0.9862	0.9876	0.9858	0
19		2.4024	2.4956	2.4848	2.5022	2.4860	
200	0.1000	96.1946	0.9924	0.9836	0.9900	0.9744	0
19		16.9204	2.8262	2.6638	2.8234	2.5824	
200	0.2000	82.3692	0.9938	0.9874	0.9918	0.9814	0
19		9.2094	2.8132	2.7152	2.7728	2.6316	
200	0.3000	66.4130	0.9902	0.9850	0.9908	0.9802	0
19		5.9886	2.7142	2.6370	2.6928	2.5812	
200	0.4000	56.9656	0.9868	0.9850	0.9880	0.9804	0
19		4.6174	2.6656	2.6116	2.6464	2.5622	
200	0.5000	52.1430	0.9908	0.9880	0.9904	0.9842	0
19		3.6640	2.6220	2.5824	2.6214	2.5570	
200	0.6000	46.4852	0.9890	0.9886	0.9888	0.9864	0
19		3.0786	2.5574	2.5346	2.5550	2.5208	
200	0.7000	39.9120	0.9886	0.9878	0.9884	0.9874	0
19		2.6062	2.5244	2.5096	2.5240	2.5038	
200	0.8000	33.1440	0.9870	0.9864	0.9866	0.9864	0
19		2.3414	2.5210	2.5030	2.5192	2.5036	

Table 6.23. brsplitsim n=400,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.6000	37.0200	0.9862	0.9850	0.9852	0.9854	4824
1		2.7296	2.4704	2.4534	2.4720	2.4536	
4	0.7000	36.2640	0.9860	0.9854	0.9870	0.9848	4479
1		2.6772	2.4740	2.4588	2.4858	2.4624	
4	0.8000	34.5240	0.9824	0.9824	0.9818	0.9816	3687
1		2.5554	2.4512	2.4336	2.4512	2.4340	
10	0.4000	51.5700	0.9872	0.9866	0.9880	0.9868	4901
1		3.7192	2.4986	2.4626	2.4982	2.4640	
10	0.5000	50.5560	0.9840	0.9844	0.9860	0.9826	4514
1		3.6590	2.4958	2.4808	2.5004	2.4718	
10	0.6000	47.2860	0.9882	0.9878	0.9886	0.9876	3744
1		3.4290	2.5012	2.4798	2.5044	2.4752	
10	0.7000	42.9180	0.9876	0.9866	0.9860	0.9860	2611
1		3.0794	2.4998	2.4818	2.4998	2.4866	
10	0.8000	39.7080	0.9872	0.9850	0.9860	0.9852	1729
1		2.7876	2.4938	2.4834	2.4964	2.4854	
20	0.3000	62.5560	0.9858	0.9828	0.9864	0.9844	4914
1		4.5240	2.5372	2.4874	2.5334	2.4898	
20	0.4000	60.4440	0.9890	0.9890	0.9888	0.9862	4504
1		4.4020	2.5330	2.4986	2.5290	2.4954	
20	0.5000	56.8140	0.9874	0.9862	0.9874	0.9860	3470
1		4.0968	2.5452	2.5214	2.5470	2.5194	
20	0.6000	49.6980	0.9858	0.9858	0.9852	0.9850	2181
1		3.5580	2.5088	2.4912	2.5102	2.4850	
20	0.7000	45.2580	0.9844	0.9834	0.9844	0.9840	1341
1		3.1178	2.5000	2.4832	2.5044	2.4832	
20	0.8000	42.5820	0.9874	0.9862	0.9856	0.9858	779
1		2.7954	2.4862	2.4748	2.4830	2.4748	
200	0.2000	88.4250	0.9908	0.9852	0.9920	0.9850	3542
1		6.7322	2.6310	2.5620	2.6416	2.5528	
200	0.3000	69.6900	0.9890	0.9864	0.9888	0.9846	1221
1		5.1226	2.5976	2.5588	2.6044	2.5438	
200	0.4000	60.2498	0.9862	0.9856	0.9876	0.9854	480
1		4.1062	2.5876	2.5566	2.5922	2.5394	
200	0.5000	53.2500	0.9892	0.9856	0.9882	0.9838	220
1		3.4470	2.5674	2.5342	2.5698	2.5250	
200	0.6000	46.4220	0.9858	0.9846	0.9852	0.9838	112
1		2.9924	2.5086	2.4902	2.5070	2.4912	
200	0.7000	40.1460	0.9886	0.9886	0.9872	0.9892	72
1		2.6352	2.4912	2.4804	2.4906	2.4830	
200	0.8000	33.1980	0.9860	0.9852	0.9858	0.9862	57
1		2.3292	2.4672	2.4620	2.4726	2.4548	
400	0.1000	107.2836	0.9888	0.9830	0.9888	0.9840	4828
1		8.4772	2.6754	2.5750	2.6790	2.5564	

Table 6.24. brsplitsim n=400,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
400	0.2000	82.1174	0.9878	0.9836	0.9878	0.9812	1695
1		6.3594	2.6418	2.5658	2.6308	2.5510	
400	0.3000	65.8894	0.9894	0.9850	0.9888	0.9840	415
1		4.5530	2.5716	2.5214	2.5686	2.5038	
400	0.4000	57.5304	0.9862	0.9874	0.9888	0.9866	152
1		3.7102	2.5262	2.5094	2.5220	2.4974	
400	0.5000	50.2200	0.9868	0.9854	0.9872	0.9872	72
1		3.2156	2.5122	2.4974	2.5050	2.4942	
400	0.6000	43.7700	0.9862	0.9860	0.9862	0.9848	42
1		2.8066	2.4978	2.4872	2.5024	2.4858	
400	0.7000	36.8160	0.9864	0.9856	0.9850	0.9852	34
1		2.4506	2.4984	2.4878	2.4962	2.4880	
400	0.8000	32.5500	0.9870	0.9854	0.9868	0.9858	20
1		2.3234	2.4670	2.4526	2.4678	2.4574	
4	0.6000	40.2000	0.9700	0.9700	0.9700	0.9700	23
3		3.1500	2.3400	2.3800	2.3800	2.3500	
4	0.7000	36.4920	0.9832	0.9816	0.9820	0.9826	465
3		2.8176	2.4880	2.4788	2.4912	2.4818	
4	0.8000	34.2900	0.9880	0.9870	0.9880	0.9888	177
3		2.5610	2.4726	2.4620	2.4790	2.4734	
10	0.4000	63.0780	0.9868	0.9850	0.9864	0.9848	4
9		5.2822	2.5620	2.5472	2.5602	2.5320	
10	0.5000	54.7080	0.9876	0.9876	0.9886	0.9864	3
9		4.3496	2.5562	2.5432	2.5592	2.5340	
10	0.6000	47.4540	0.9896	0.9886	0.9892	0.9890	1
9		3.6238	2.5238	2.5078	2.5170	2.4980	
10	0.7000	42.7860	0.9858	0.9856	0.9868	0.9852	0
9		3.1156	2.5098	2.4994	2.5084	2.4934	
10	0.8000	39.5040	0.9830	0.9836	0.9838	0.9822	0
9		2.7614	2.5058	2.4918	2.5084	2.4978	
20	0.3000	85.6020	0.9892	0.9896	0.9898	0.9890	0
19		7.3414	2.6598	2.6200	2.6456	2.6044	
20	0.4000	67.7160	0.9916	0.9900	0.9902	0.9896	0
19		5.5410	2.6038	2.5786	2.5964	2.5596	
20	0.5000	56.9400	0.9876	0.9866	0.9870	0.9870	0
19		4.3884	2.5742	2.5454	2.5632	2.5386	
20	0.6000	50.1000	0.9906	0.9898	0.9906	0.9892	0
19		3.6124	2.5232	2.5050	2.5240	2.4964	
20	0.7000	45.3300	0.9830	0.9834	0.9834	0.9838	0
19		3.1428	2.5126	2.4968	2.5140	2.4958	
20	0.8000	42.7140	0.9862	0.9850	0.9860	0.9862	0
19		2.8238	2.4858	2.4744	2.4944	2.4764	
200	0.2000	96.4966	0.9934	0.9868	0.9916	0.9882	0
19		8.0320	2.7200	2.6574	2.6950	2.6244	

Table 6.25. brsplitsim n=400,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
200	0.3000	70.6536	0.9916	0.9880	0.9910	0.9882	0
19		5.2720	2.6024	2.5594	2.5934	2.5484	
200	0.4000	59.8112	0.9900	0.9864	0.9892	0.9856	0
19		4.1006	2.5726	2.5284	2.5672	2.5334	
200	0.5000	52.5274	0.9894	0.9868	0.9886	0.9852	0
19		3.4454	2.5484	2.5130	2.5426	2.5016	
200	0.6000	46.8720	0.9848	0.9830	0.9836	0.9830	0
19		3.0178	2.4814	2.4660	2.4810	2.4648	
200	0.7000	39.9240	0.9864	0.9860	0.9852	0.9858	0
19		2.6138	2.5074	2.4874	2.4876	2.4850	
200	0.8000	33.1260	0.9856	0.9858	0.9864	0.9852	0
19		2.3418	2.4752	2.4620	2.4680	2.4582	
400	0.1000	155.5798	0.9938	0.9886	0.9926	0.9850	0
19		15.5978	2.7976	2.6962	2.7760	2.6422	
400	0.2000	83.5198	0.9910	0.9870	0.9896	0.9852	0
19		6.5686	2.6818	2.6242	2.6678	2.6010	
400	0.3000	65.3076	0.9900	0.9872	0.9894	0.9870	0
19		4.5888	2.6010	2.5564	2.5920	2.5414	
400	0.4000	55.8538	0.9856	0.9852	0.9870	0.9848	0
19		3.5890	2.5276	2.5050	2.5246	2.5016	
400	0.5000	49.9260	0.9856	0.9856	0.9864	0.9854	0
19		3.1966	2.5160	2.5022	2.5138	2.4954	
400	0.6000	43.5000	0.9876	0.9880	0.9876	0.9882	0
19		2.7864	2.5146	2.5010	2.5046	2.4994	
400	0.7000	36.6540	0.9870	0.9874	0.9864	0.9858	0
19		2.4236	2.5052	2.4978	2.5116	2.5004	
400	0.8000	32.4780	0.9852	0.9844	0.9860	0.9850	0
19		2.3346	2.4978	2.4808	2.5034	2.4804	

Table 6.26. brsplitsim n=1000,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.0000	37.1640	0.9876	0.9862	0.9866	0.9858	4993
1		2.6016	2.4730	2.4594	2.4752	2.4546	
10	0.0000	48.4800	0.9858	0.9838	0.9846	0.9842	4994
1		3.2824	2.5020	2.4614	2.4962	2.4614	
20	0.0000	56.7660	0.9856	0.9840	0.9858	0.9824	4978
1		3.7276	2.4988	2.4460	2.4978	2.4400	
500	0.0000	88.5240	0.9904	0.9822	0.9900	0.9814	4863
1		5.1026	2.5822	2.4540	2.5910	2.4630	
1000	0.0000	95.0734	0.9902	0.9820	0.9898	0.9822	4789
1		5.3892	2.6302	2.4922	2.6320	2.4834	
4	0.0000	58.9320	0.9848	0.9840	0.9854	0.9834	4822
3		3.9560	2.4456	2.4380	2.4444	2.4382	
10	0.0000	111.3060	0.9886	0.9886	0.9890	0.9880	4450
9		9.2280	2.4986	2.4904	2.4964	2.4884	
20	0.0000	162.8160	0.9866	0.9862	0.9850	0.9852	3613
19		15.0950	2.5364	2.5270	2.5382	2.5222	
500	0.0000	102.6794	0.9952	0.9782	0.9948	0.9768	614
19		12.4270	2.9876	2.6734	3.0108	2.6690	
1000	0.0000	79.4832	0.9964	0.9768	0.9948	0.9726	346
19		8.9226	3.0314	2.6630	3.0402	2.6468	
4	0.5000	37.0800	0.9838	0.9830	0.9848	0.9832	4959
1		2.7322	2.4422	2.4212	2.4426	2.4252	
10	0.3160	51.7620	0.9870	0.9870	0.9880	0.9854	4973
1		3.6982	2.4580	2.4326	2.4608	2.4298	
20	0.2240	61.4400	0.9832	0.9824	0.9838	0.9836	4986
1		4.3448	2.4936	2.4516	2.4906	2.4572	
500	0.0400	106.0860	0.9904	0.9860	0.9900	0.9856	4901
1		7.4554	2.5642	2.4736	2.5668	2.4628	
1000	0.0300	116.9940	0.9896	0.9826	0.9906	0.9842	4892
1		8.3288	2.5916	2.4976	2.5924	2.4930	
4	0.5000	43.9980	0.9854	0.9830	0.9846	0.9836	1928
3		3.3196	2.4596	2.4512	2.4566	2.4476	
10	0.3160	74.6760	0.9884	0.9880	0.9886	0.9872	24
9		6.4568	2.5044	2.4992	2.5100	2.4960	
20	0.2240	114.6180	0.9848	0.9844	0.9852	0.9822	1
19		10.3204	2.5324	2.5182	2.5342	2.5138	
500	0.0400	426.8602	0.9952	0.9926	0.9948	0.9878	0
19		42.4614	2.8484	2.7384	2.8346	2.6870	
1000	0.0300	464.3454	0.9948	0.9904	0.9936	0.9874	0
19		50.1826	2.8528	2.7324	2.8380	2.6650	
4	0.6000	36.9000	0.9868	0.9864	0.9856	0.9852	4813
1		2.7362	2.4408	2.4324	2.4408	2.4208	
4	0.7000	35.8560	0.9858	0.9868	0.9852	0.9858	4474
1		2.6806	2.4438	2.4274	2.4376	2.4274	

Table 6.27. brsplitsim n=1000,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
4	0.8000	34.2360	0.9862	0.9860	0.9870	0.9854	3718
1		2.5336	2.4468	2.4352	2.4442	2.4312	
10	0.4000	51.6900	0.9842	0.9852	0.9844	0.9842	4902
1		3.7164	2.4428	2.4274	2.4420	2.4214	
10	0.5000	50.0580	0.9882	0.9880	0.9878	0.9882	4532
1		3.6680	2.4420	2.4292	2.4500	2.4342	
10	0.6000	46.8780	0.9846	0.9848	0.9860	0.9854	3660
1		3.4494	2.4150	2.4048	2.4160	2.4036	
10	0.7000	43.2000	0.9846	0.9836	0.9852	0.9842	2601
1		3.1016	2.4554	2.4378	2.4506	2.4416	
10	0.8000	39.2220	0.9864	0.9854	0.9858	0.9850	1728
1		2.7662	2.4370	2.4216	2.4322	2.4246	
20	0.3000	62.3820	0.9868	0.9862	0.9870	0.9852	4910
1		4.4828	2.4346	2.4106	2.4380	2.4062	
20	0.4000	60.4380	0.9844	0.9848	0.9840	0.9842	4494
1		4.3770	2.4722	2.4568	2.4780	2.4586	
20	0.5000	56.0040	0.9874	0.9872	0.9890	0.9878	3418
1		4.0666	2.4560	2.4374	2.4614	2.4392	
20	0.6000	50.5980	0.9848	0.9854	0.9850	0.9846	2271
1		3.5888	2.4584	2.4454	2.4616	2.4428	
20	0.7000	45.7260	0.9882	0.9874	0.9870	0.9872	1284
1		3.1282	2.4422	2.4336	2.4390	2.4258	
20	0.8000	42.9480	0.9864	0.9874	0.9866	0.9864	755
1		2.8250	2.4652	2.4612	2.4660	2.4574	
500	0.1000	112.2240	0.9874	0.9836	0.9876	0.9834	4664
1		8.3686	2.5248	2.4790	2.5220	2.4742	
500	0.2000	80.4480	0.9860	0.9820	0.9862	0.9846	1219
1		5.8938	2.4938	2.4648	2.5010	2.4704	
500	0.3000	63.0300	0.9876	0.9844	0.9872	0.9842	266
1		4.2766	2.4938	2.4728	2.4946	2.4670	
500	0.4000	56.3400	0.9882	0.9880	0.9878	0.9886	116
1		3.6444	2.4820	2.4676	2.4788	2.4642	
500	0.5000	48.4440	0.9844	0.9822	0.9840	0.9822	55
1		3.1250	2.4656	2.4498	2.4584	2.4532	
500	0.6000	41.9280	0.9854	0.9850	0.9848	0.9848	38
1		2.6866	2.4464	2.4404	2.4440	2.4408	
500	0.7000	35.4240	0.9874	0.9878	0.9862	0.9880	19
1		2.4150	2.4584	2.4540	2.4602	2.4514	
500	0.8000	32.9400	0.9860	0.9842	0.9866	0.9846	19
1		2.3360	2.4554	2.4454	2.4550	2.4376	
1000	0.1000	109.7700	0.9904	0.9858	0.9896	0.9868	3240
1		8.4434	2.5440	2.4896	2.5444	2.4896	
1000	0.2000	72.2640	0.9884	0.9868	0.9884	0.9872	339
1		5.0394	2.5012	2.4780	2.5118	2.4812	

Table 6.28. brsplitsim n=1000,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
1000	0.3000	59.5920	0.9830	0.9826	0.9828	0.9836	82
1		3.8614	2.4984	2.4794	2.4940	2.4794	
1000	0.4000	51.5820	0.9830	0.9830	0.9842	0.9840	30
1		3.3024	2.4714	2.4520	2.4698	2.4542	
1000	0.5000	45.3300	0.9842	0.9842	0.9848	0.9842	13
1		2.8558	2.4528	2.4430	2.4584	2.4450	
1000	0.6000	38.2020	0.9864	0.9872	0.9874	0.9870	15
1		2.4974	2.4694	2.4568	2.4780	2.4664	
1000	0.7000	33.2700	0.9840	0.9836	0.9840	0.9836	13
1		2.3572	2.4552	2.4398	2.4598	2.4450	
1000	0.8000	34.1340	0.9876	0.9870	0.9872	0.9856	17
1		2.4098	2.4708	2.4584	2.4884	2.4672	
4	0.6000	39.5820	0.9830	0.9826	0.9832	0.9838	1075
3		3.0638	2.4486	2.4294	2.4508	2.4396	
4	0.7000	36.3060	0.9842	0.9844	0.9856	0.9848	460
3		2.8138	2.4500	2.4414	2.4544	2.4408	
4	0.8000	34.0320	0.9852	0.9850	0.9834	0.9846	185
3		2.5648	2.4252	2.4176	2.4200	2.4162	
10	0.4000	63.4080	0.9870	0.9856	0.9862	0.9858	4
9		5.2826	2.5050	2.4854	2.5008	2.4890	
10	0.5000	54.7260	0.9908	0.9896	0.9904	0.9900	1
9		4.3610	2.4824	2.4710	2.4868	2.4698	
10	0.6000	48.3000	0.9884	0.9866	0.9870	0.9872	2
9		3.6546	2.4402	2.4362	2.4430	2.4340	
10	0.7000	43.0320	0.9886	0.9876	0.9878	0.9882	0
9		3.1120	2.4598	2.4506	2.4556	2.4488	
10	0.8000	39.5520	0.9856	0.9868	0.9870	0.9864	2
9		2.7538	2.4446	2.4352	2.4456	2.4390	
20	0.3000	85.4700	0.9854	0.9844	0.9848	0.9846	0
19		7.3694	2.5160	2.4904	2.5166	2.4998	
20	0.4000	67.7880	0.9884	0.9874	0.9894	0.9878	0
19		5.5192	2.5130	2.4928	2.5032	2.4918	
20	0.5000	57.5580	0.9856	0.9842	0.9850	0.9852	0
19		4.4238	2.4980	2.4872	2.4980	2.4820	
20	0.6000	49.9380	0.9858	0.9842	0.9852	0.9850	0
19		3.6314	2.4894	2.4746	2.4822	2.4744	
20	0.7000	45.1620	0.9876	0.9858	0.9872	0.9866	0
19		3.1332	2.4514	2.4388	2.4586	2.4440	
20	0.8000	43.1520	0.9872	0.9872	0.9870	0.9866	0
19		2.8302	2.4368	2.4298	2.4398	2.4262	
500	0.1000	153.6060	0.9922	0.9912	0.9916	0.9892	0
19		13.3678	2.6720	2.6244	2.6494	2.6026	
500	0.2000	80.8920	0.9882	0.9844	0.9876	0.9866	0
19		6.1108	2.5390	2.5074	2.5466	2.5076	

Table 6.29. brsplitsim n=1000,alpha=0.05,int=1,a=4/3,m=4,B=1000,nruns=5000

p/k	psi	mnnd/mnad	laspi	LVSpi	lsplitpi	splitpi	noundfit
500	0.3000	64.0680	0.9882	0.9864	0.9878	0.9870	0
19		4.3336	2.5048	2.5004	2.5066	2.4882	
500	0.4000	55.1940	0.9850	0.9840	0.9856	0.9852	0
19		3.5622	2.4680	2.4608	2.4724	2.4676	
500	0.5000	49.6440	0.9868	0.9866	0.9862	0.9856	0
19		3.1650	2.4890	2.4870	2.4874	2.4824	
500	0.6000	42.0600	0.9872	0.9872	0.9878	0.9878	0
19		2.6948	2.4488	2.4396	2.4574	2.4374	
500	0.7000	35.5740	0.9884	0.9868	0.9886	0.9880	0
19		2.4150	2.4722	2.4620	2.4704	2.4620	
500	0.8000	32.5620	0.9836	0.9828	0.9834	0.9830	0
19		2.3376	2.4384	2.4256	2.4322	2.4260	
1000	0.1000	121.4280	0.9900	0.9876	0.9880	0.9856	0
19		10.1198	2.6348	2.5948	2.6276	2.5802	
1000	0.2000	71.4060	0.9868	0.9862	0.9880	0.9850	0
19		5.0754	2.5164	2.5000	2.5188	2.4932	
1000	0.3000	59.3520	0.9850	0.9846	0.9864	0.9844	0
19		3.8356	2.4834	2.4778	2.4864	2.4732	
1000	0.4000	51.9900	0.9852	0.9852	0.9846	0.9842	0
19		3.2812	2.4624	2.4532	2.4646	2.4506	
1000	0.5000	45.2100	0.9872	0.9878	0.9880	0.9874	0
19		2.8650	2.4442	2.4294	2.4538	2.4372	
1000	0.6000	38.0580	0.9852	0.9846	0.9860	0.9850	0
19		2.4990	2.4758	2.4534	2.4750	2.4552	
1000	0.7000	33.3840	0.9872	0.9870	0.9874	0.9868	0
19		2.3546	2.4414	2.4284	2.4508	2.4386	
1000	0.8000	34.0020	0.9854	0.9848	0.9864	0.9860	0
19		2.3912	2.4838	2.4750	2.5028	2.4928	

Table 6.30. PHsplitsim n=100, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
4	0.0000	31.7646	0.9550	0.9552	4913
1		1.8314	5.5483	5.5033	
10	0.0000	35.0694	0.9512	0.9514	4834
1		2.3408	5.6938	5.7305	
20	0.0000	36.3172	0.9326	0.9328	4747
1		2.7178	5.9745	25.4093	
50	0.0000	37.4732	0.9298	0.9300	4556
1		3.1644	11.4079	96.6708	
100	0.0000	37.9084	0.9056	0.9058	4303
1		3.4020	22.4533	169.9087	
4	0.0000	37.1740	0.9590	0.9592	3668
3		2.9188	6.0138	6.0320	
10	0.0000	39.1970	0.9474	0.9476	622
9		4.2902	6.2835	6.0921	
20	0.0000	36.8816	0.9102	0.9102	0
19		3.1822	6.2632	549.8349	
50	0.0000	35.1850	0.7750	0.7752	0
19		2.1098	346.3883	332.0434	
100	0.0000	34.4812	0.8246	0.8248	0
19		1.6708	242.1134	232.2679	
4	0.5000	31.3634	0.9528	0.9530	4567
1		1.9872	5.6027	5.6367	
10	0.3160	35.1646	0.9514	0.9516	4557
1		2.6112	5.8565	5.7596	
20	0.2240	37.2148	0.9446	0.9448	4485
1		3.1526	6.0353	31.3664	
50	0.1410	39.4486	0.9368	0.9370	4332
1		3.9124	9.9618	93.7964	
100	0.1000	40.0538	0.9254	0.9256	4098
1		4.3076	18.9567	124.8790	
4	0.5000	31.7782	0.9578	0.9580	884
3		2.2608	6.0849	5.9433	
10	0.3160	38.5102	0.9622	0.9624	10
9		3.5834	6.4668	6.3234	
20	0.2240	42.0496	0.9582	0.9584	0
19		4.9144	6.4463	39.0744	
50	0.1410	43.6986	0.9420	0.9422	0
19		6.3944	11.0519	235.7118	
100	0.1000	44.2800	0.9182	0.9184	0
19		7.4530	34.0540	507.4744	
4	0.6000	31.1730	0.9518	0.9520	175
1		1.9138	5.7709	5.7574	
4	0.7000	30.9554	0.9554	0.9556	180
1		1.7904	5.8691	5.8575	

Table 6.31. PHsplitsim n=100, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
4	0.8000	30.7004	0.9574	0.9576	156
1		1.6076	5.5956	5.5384	
10	0.4000	34.6546	0.9512	0.9514	4184
1		2.5930	5.8873	5.8523	
10	0.5000	33.9848	0.9534	0.9536	3547
1		2.4306	5.8484	5.7320	
10	0.6000	32.9648	0.9576	0.9578	2802
1		2.2184	5.9171	5.9371	
10	0.7000	32.2100	0.9550	0.9552	4132
1		1.9484	5.8899	5.8810	
10	0.8000	31.6796	0.9566	0.9568	3358
1		1.6920	5.6975	5.6283	
20	0.3000	37.0992	0.9522	0.9524	2487
1		3.1498	6.0266	23.3893	
20	0.4000	35.9160	0.9526	0.9528	1746
1		2.8924	5.9835	17.1799	
20	0.5000	34.4608	0.9548	0.9548	1309
1		2.5586	5.9973	15.3929	
20	0.6000	32.7846	0.9524	0.9526	3998
1		2.0132	6.0766	11.4839	
20	0.7000	32.0570	0.9574	0.9576	2975
1		1.6680	5.7627	7.8844	
20	0.8000	39.4894	0.9384	0.9386	2016
1		3.9378	9.4670	80.3992	
50	0.2000	37.8336	0.9500	0.9502	950
1		3.4958	8.2650	50.2275	
50	0.3000	33.4068	0.9450	0.9452	639
1		2.1610	9.2957	34.0216	
50	0.4000	31.8428	0.9612	0.9614	3720
1		1.7122	5.4637	24.2605	
50	0.5000	30.3570	0.9616	0.9618	2293
1		1.4274	5.6046	5.5374	
50	0.6000	39.7342	0.9390	0.9392	486
1		4.0962	19.6923	83.6910	
50	0.7000	37.2998	0.9430	0.9432	313
1		3.3406	14.4151	62.7034	
50	0.8000	35.1952	0.9466	0.9468	220
1		2.6844	14.4545	59.0050	
100	0.2000	33.7332	0.9430	0.9432	2427
1		2.3076	27.0150	34.4229	
100	0.3000	32.9478	0.9514	0.9516	1073
1		1.9366	10.5212	8.9234	
100	0.4000	30.9928	0.9546	0.9548	508
1		1.5616	5.6326	5.7635	

Table 6.32. PHsplitsim n=100, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
100	0.5000	30.3502	0.9560	0.9562	318
1		1.4122	5.6396	5.6261	
100	0.6000	31.0574	0.9614	0.9616	179
1		2.0532	5.9299	5.8632	
100	0.7000	30.8738	0.9530	0.9532	130
1		1.8176	5.8545	5.8320	
100	0.8000	30.6256	0.9576	0.9578	110
1		1.6138	5.6779	5.7250	
4	0.6000	36.4328	0.9586	0.9588	530
3		3.0848	6.5062	6.2952	
4	0.7000	34.2228	0.9610	0.9612	291
3		2.6054	6.2433	6.0339	
4	0.8000	32.7506	0.9550	0.9552	186
3		2.2248	6.1179	6.0784	
10	0.4000	32.1522	0.9512	0.9514	9
9		1.9540	5.7559	5.7118	
10	0.5000	31.7238	0.9548	0.9550	6
9		1.6892	5.6690	5.5957	
10	0.6000	39.7546	0.9584	0.9586	2
9		4.0126	6.4520	22.2245	
10	0.7000	36.6946	0.9576	0.9578	1
9		3.1952	6.2901	15.9410	
10	0.8000	34.4914	0.9560	0.9562	3
9		2.6366	6.0942	11.4012	
20	0.3000	33.2436	0.9508	0.9510	0
19		2.2262	5.9552	9.9235	
20	0.4000	32.6622	0.9540	0.9542	1
19		1.9864	5.9395	13.9925	
20	0.5000	32.0128	0.9550	0.9552	0
19		1.6972	5.8343	7.2834	
20	0.6000	42.5324	0.9434	0.9434	0
19		5.3034	9.3749	92.9933	
20	0.7000	38.8672	0.9502	0.9504	0
19		3.8040	8.8210	47.0759	
20	0.8000	35.7664	0.9534	0.9536	0
19		2.9626	8.9539	36.8299	
50	0.2000	34.2806	0.9484	0.9486	0
19		2.4954	7.8986	28.2586	
50	0.3000	33.4272	0.9522	0.9524	0
19		2.1256	8.0160	25.5739	
50	0.4000	31.6864	0.9474	0.9476	0
19		1.7124	6.6366	16.8449	

Table 6.33. PHsplitsim n=100, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
50	0.5000	30.3808	0.9584	0.9586	0
19		1.4200	5.7996	6.5653	
50	0.6000	41.5498	0.9438	0.9438	0
19		4.9046	12.9189	92.7801	
50	0.7000	37.7656	0.9444	0.9446	0
19		3.4848	12.7862	57.5085	
50	0.8000	35.2904	0.9410	0.9412	0
19		2.7618	17.0762	55.5593	
100	0.2000	33.7808	0.9458	0.9460	0
19		2.3670	25.0165	34.8553	
100	0.3000	32.9682	0.9576	0.9578	0
19		1.9548	10.0288	10.7122	
100	0.4000	31.0880	0.9586	0.9588	0
19		1.5510	5.8776	5.8639	
100	0.5000	30.3468	0.9532	0.9534	0
19		1.4292	5.7605	5.7362	
100	0.6000	31.7834	0.9562	0.9672	0
19		1.4673	6.9902	6.8923	
100	0.7000	32.8934	0.9723	0.9782	0
19		2.7823	7.8923	7.2832	
100	0.8000	35.6723	0.9876	0.9673	0
19		1.9774	6.9232	13.8923	

Table 6.34. PHsplitsim n=200, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
4	0.0000	32.8500	0.9570	0.9572	4904
1		1.7968	5.4771	5.4616	
10	0.0000	39.4800	0.9530	0.9532	4831
1		2.2296	5.6325	5.6656	
20	0.0000	43.0700	0.9508	0.9510	4748
1		2.4710	5.2759	5.2382	
100	0.0000	49.1420	0.9310	0.9312	4349
1		2.8422	5.5385	9.8062	
200	0.0000	50.0400	0.9228	0.9230	4087
1		2.9730	5.9103	15.3853	
4	0.0000	42.6780	0.9572	0.9574	3749
3		2.9466	5.5367	5.5326	
10	0.0000	61.4150	0.9668	0.9670	1823
9		5.0460	6.3783	6.3652	
20	0.0000	51.0578	0.9528	0.9530	87
19		4.4876	6.6277	6.4543	
100	0.0000	41.3440	0.8524	0.8526	0
19		1.6920	107.0777	209.8225	
200	0.0000	39.0056	0.8462	0.8464	0
19		1.1932	126.8648	142.1632	
4	0.5000	32.5560	0.9530	0.9532	4599
1		1.9612	5.5771	5.5680	
10	0.3000	39.9268	0.9528	0.9530	4599
1		2.5222	5.7156	5.6879	
20	0.2000	45.0938	0.9514	0.9516	4582
1		2.9330	5.7950	5.7364	
100	0.1000	54.9906	0.9474	0.9476	4239
1		3.9592	5.4562	9.4407	
200	0.0700	58.1400	0.9396	0.9398	4125
1		4.3178	5.8943	18.4859	
4	0.5000	32.8440	0.9622	0.9624	870
3		2.2658	5.6456	5.6661	
10	0.3000	48.2636	0.9614	0.9616	5
9		3.8332	5.9431	5.8418	
20	0.2000	65.7932	0.9602	0.9604	0
19		5.8450	6.4730	6.2230	
100	0.1000	82.0352	0.9474	0.9476	0
19		9.9920	6.1613	24.8866	
200	0.0700	84.7238	0.9412	0.9414	0
19		11.7648	7.9501	68.6064	
4	0.6000	32.2500	0.9528	0.9530	4157
1		1.9064	5.3233	5.3050	
4	0.7000	31.5360	0.9544	0.9546	3597
1		1.7832	5.3806	5.4004	

Table 6.35. PHsplitsim n=200, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
4	0.8000	31.3200	0.9502	0.9504	2786
1		1.6028	5.4862	5.4992	
10	0.4000	38.3654	0.9568	0.9570	4156
1		2.4788	5.6743	5.6309	
10	0.5000	37.0122	0.9528	0.9530	3371
1		2.3774	5.7325	5.7165	
10	0.6000	35.4840	0.9564	0.9566	2557
1		2.1768	5.5901	5.5483	
10	0.7000	34.2360	0.9576	0.9578	1825
1		1.9412	5.6134	5.5763	
10	0.8000	33.1140	0.9554	0.9556	1227
1		1.6646	5.5068	5.5064	
20	0.3000	43.8922	0.9516	0.9518	4082
1		2.9342	5.6378	5.6054	
20	0.4000	41.0720	0.9484	0.9486	3091
1		2.7528	5.6278	5.5947	
20	0.5000	38.1902	0.9506	0.9508	2088
1		2.4434	6.0789	6.0436	
20	0.6000	36.6352	0.9586	0.9588	1391
1		2.1778	5.4250	5.4135	
20	0.7000	35.2344	0.9622	0.9624	904
1		1.9192	5.5508	5.5611	
20	0.8000	33.6060	0.9546	0.9548	652
1		1.5958	5.2658	5.2556	
100	0.2000	53.4370	0.9560	0.9562	2989
1		4.0228	5.8436	7.1314	
100	0.3000	45.4614	0.9526	0.9528	1258
1		3.1924	5.7086	7.3362	
100	0.4000	39.9992	0.9548	0.9550	584
1		2.5540	5.5978	5.5662	
100	0.5000	37.8458	0.9554	0.9556	295
1		2.1882	5.6628	6.9756	
100	0.6000	35.3460	0.9622	0.9624	180
1		1.8576	5.6904	5.7109	
100	0.7000	31.7400	0.9498	0.9500	127
1		1.5278	5.5689	5.5383	
100	0.8000	30.6780	0.9564	0.9566	116
1		1.4214	5.4465	5.4252	
200	0.2000	50.4208	0.9466	0.9468	1477
1		3.7964	6.0233	7.5375	
200	0.3000	42.8938	0.9484	0.9486	464
1		2.8936	9.6364	8.2378	
200	0.4000	38.9224	0.9616	0.9618	226
1		2.3556	7.1008	6.0001	

Table 6.36. PHsplitsim n=200, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
200	0.5000	37.0362	0.9576	0.9578	112
1		1.9824	5.7501	5.7306	
200	0.6000	33.1380	0.9584	0.9586	79
1		1.6684	5.5959	5.5653	
200	0.7000	31.0620	0.9582	0.9584	57
1		1.4774	5.3984	5.3666	
200	0.8000	30.9540	0.9654	0.9656	58
1		1.4582	5.5483	5.5292	
4	0.6000	31.8900	0.9520	0.9522	496
3		2.0360	5.4046	5.3732	
4	0.7000	31.4520	0.9530	0.9532	306
3		1.8206	5.4499	5.4416	
4	0.8000	31.1580	0.9562	0.9564	171
3		1.5978	5.4494	5.4625	
10	0.4000	41.6968	0.9556	0.9558	0
9		3.1070	5.6825	5.5610	
10	0.5000	37.1548	0.9614	0.9616	1
9		2.5776	5.8778	5.8434	
10	0.6000	35.2380	0.9610	0.9612	0
9		2.2146	5.7666	5.7190	
10	0.7000	33.7048	0.9626	0.9628	1
9		1.9250	5.5196	5.4817	
10	0.8000	32.9460	0.9588	0.9590	0
9		1.6672	5.5143	5.4891	
20	0.3000	52.3512	0.9630	0.9632	0
19		4.1916	6.2588	6.1321	
20	0.4000	43.2588	0.9602	0.9604	0
19		3.1460	6.0495	5.9673	
20	0.5000	38.8700	0.9610	0.9612	0
19		2.5996	5.7527	5.6857	
20	0.6000	36.4898	0.9580	0.9582	0
19		2.2096	5.6028	5.5733	
20	0.7000	35.1518	0.9628	0.9630	0
19		1.9232	5.7401	5.7523	
20	0.8000	33.8654	0.9562	0.9564	0
19		1.6044	5.4177	5.3981	
100	0.2000	60.2800	0.9592	0.9594	0
19		5.1268	6.5582	9.5048	
100	0.3000	46.1742	0.9588	0.9590	0
19		3.3670	5.9739	7.0273	
100	0.4000	40.5756	0.9568	0.9570	0
19		2.6216	5.8375	6.0123	
100	0.5000	37.6464	0.9608	0.9610	0
19		2.2004	5.6356	7.4359	

Table 6.37. PHsplitsim n=200, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
100	0.6000	35.0580	0.9618	0.9620	0
19		1.8378	5.5837	5.5017	
100	0.7000	31.7760	0.9588	0.9590	0
19		1.5422	5.4223	5.4214	
100	0.8000	30.5400	0.9600	0.9602	0
19		1.4036	5.4973	5.4670	
200	0.1000	80.9012	0.9482	0.9484	0
19		9.1104	7.2378	23.2208	
200	0.2000	53.8350	0.9566	0.9568	0
19		4.2258	6.2509	6.8314	
200	0.3000	42.9776	0.9546	0.9548	0
19		2.9090	8.5931	6.1495	
200	0.4000	38.9782	0.9604	0.9606	0
19		2.3734	7.0705	5.7963	
200	0.5000	36.7288	0.9620	0.9622	0
19		1.9786	5.5377	5.5133	
200	0.6000	33.5820	0.9574	0.9576	0
19		1.6834	5.5151	5.4575	
200	0.7000	31.0140	0.9606	0.9608	0
19		1.5060	5.4609	5.4237	
200	0.8000	30.8040	0.9584	0.9586	0
19		1.4370	5.6815	5.6726	

Table 6.38. PHsplitsim n=400, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
4	0.6000	35.9583	0.9573	0.9524	4813
1		2.0959	5.4992	5.0720	
4	0.7000	36.5988	0.9530	0.9518	4387
1		3.1883	4.8564	5.6653	
4	0.8000	33.7352	0.9483	0.9533	3689
1		2.9499	5.2137	5.0292	
10	0.4000	51.9534	0.9392	0.9442	4892
1		3.6620	4.7988	5.0550	
10	0.5000	49.4864	0.9500	0.9599	4512
1		4.0322	5.1584	4.7034	
10	0.6000	46.8349	0.9509	0.9493	3792
1		3.5440	5.6184	4.3913	
10	0.7000	42.1900	0.9591	0.9554	2608
1		3.6791	5.1909	4.9938	
10	0.8000	40.0792	0.9574	0.9680	1727
1		2.9821	4.5897	5.1418	
20	0.3000	63.1867	0.9642	0.9685	4915
1		5.4268	5.1416	5.4455	
20	0.4000	59.9666	0.9541	0.9586	4519
1		5.3621	4.4628	5.7715	
20	0.5000	57.1933	0.9530	0.9505	3462
1		3.1756	5.1446	5.1318	
20	0.6000	49.1657	0.9532	0.9477	2024
1		4.4434	4.7011	4.8779	
20	0.7000	45.7937	0.9545	0.9550	1342
1		3.1719	4.5841	5.7318	
20	0.8000	41.8555	0.9544	0.9499	784
1		2.8712	5.2650	5.7691	
200	0.2000	88.3722	0.9543	0.9544	3546
1		6.7037	4.6305	5.2808	
200	0.3000	70.2416	0.9630	0.9606	1223
1		5.3450	4.9522	4.6000	
200	0.4000	60.5977	0.9676	0.9678	485
1		3.7123	4.9635	5.0730	
200	0.5000	53.4756	0.9634	0.9475	225
1		3.1198	5.9475	5.8403	
200	0.6000	46.7443	0.9602	0.9729	115
1		2.7679	6.1076	5.1000	
200	0.7000	39.9315	0.9671	0.9632	75
1		3.0650	5.5272	5.8056	
200	0.8000	32.9286	0.9704	0.9608	67
1		2.2979	5.1620	5.4804	

Table 6.39. PHsplitsim n=400, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
400	0.1000	107.0517	0.9489	0.9651	4824
1		8.3844	5.3952	4.6881	
400	0.2000	81.2099	0.9667	0.9631	1673
1		6.9854	5.6388	5.0837	
400	0.3000	66.6224	0.9650	0.9632	416
1		4.8212	5.9544	4.9340	
400	0.4000	57.6196	0.9584	0.9643	153
1		3.0432	5.3696	5.2909	
400	0.5000	50.3759	0.9595	0.9669	73
1		3.1474	5.8175	6.0231	
400	0.6000	43.8617	0.9661	0.9634	41
1		3.5737	5.2121	4.5457	
400	0.7000	37.4573	0.9547	0.9586	34
1		2.1943	5.2627	5.6291	
400	0.8000	31.5515	0.9628	0.9800	21
1		2.3575	4.0288	5.0604	
4	0.6000	40.3422	0.9615	0.9546	22
3		2.6537	4.9952	4.7753	
4	0.7000	35.4840	0.6455	0.9517	460
3		3.3182	5.8131	4.6898	
4	0.8000	34.6017	0.9423	0.9584	180
3		2.8644	6.3050	4.7707	
10	0.4000	62.9789	0.9613	0.9579	4
9		5.4999	5.7427	4.9679	
10	0.5000	55.4304	0.9572	0.9810	3
9		3.5183	5.4159	5.4712	
10	0.6000	47.9541	0.9542	0.9652	1
9		3.1733	4.8013	5.2207	
10	0.7000	42.9820	0.9702	0.9640	0
9		3.1835	4.2868	4.8375	
10	0.8000	39.1539	0.9566	0.9679	0
9		3.0802	5.8546	5.9552	
20	0.3000	84.5681	0.9750	0.9789	0
19		6.9762	5.8347	5.7993	
20	0.4000	67.5579	0.9773	0.9764	0
19		5.6813	5.7935	6.4888	
20	0.5000	56.8724	0.9777	0.9764	0
19		4.5360	5.9893	6.3930	
20	0.6000	49.0808	0.9668	0.9773	0
19		3.8711	6.0582	5.9623	
20	0.7000	45.2213	0.9846	0.9743	0
19		3.4574	6.5790	5.1213	

Table 6.40. PHsplitsim n=400, J=5,a=1,gam= 1,B=1000,clam=0.1,nruns=5000

p/k	psi	mnnd/mnad	LVSpi	splitpi	noundfit
20	0.8000	42.8074	0.9666	0.9746	0
19		3.8698	5.7532	6.7064	
200	0.2000	97.1514	0.9773	0.9733	0
19		7.1899	6.2578	6.7959	
200	0.3000	71.5239	0.9805	0.9739	0
19		5.3360	5.8611	5.7592	
200	0.4000	59.4892	0.9752	0.9772	0
19		4.9311	5.5091	6.8335	
200	0.5000	51.6466	0.9735	0.9626	0
19		2.7323	6.7896	7.0533	
200	0.6000	46.8124	0.9683	0.9687	0
19		2.4311	6.2746	8.2902	
200	0.7000	39.8618	0.9589	0.9656	0
19		2.8682	6.9956	7.3759	
200	0.8000	34.0654	0.9468	0.9581	0
19		2.0083	7.4511	7.6629	
400	0.1000	157.3363	0.9691	0.9788	0
19		15.6468	6.1858	7.0630	
400	0.2000	82.7962	0.9707	0.9780	0
19		7.6328	5.8508	6.6952	
400	0.3000	65.2919	0.9884	0.9789	0
19		4.2008	5.7190	6.0051	
400	0.4000	55.8235	0.9813	0.9754	0
19		3.4090	5.6063	6.6838	
400	0.5000	50.0755	0.9641	0.9596	0
19		2.7371	7.0574	7.8690	
400	0.6000	43.9504	0.9567	0.9563	0
19		2.5201	7.7426	8.7326	
400	0.7000	36.1858	0.9476	0.9561	0
19		2.0597	8.5167	9.2376	
400	0.8000	33.1759	0.9507	0.9517	0
19		2.1316	7.4790	7.9543	

REFERENCES

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

[2] Chen, J., and Chen, Z. (2008), "Extended Bayesian Information Criterion for Model Selection with Large Model Spaces," *Biometrika*, 95, 759-771.

[3] Chen, M.H., and Shao, Q.M. (1999), "Monte Carlo Estimation of Bayesian Credible and HPD Intervals, *Journal of Computational and Graphical Statistics,* 8, 69-92.

[4] Cox, D.R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, B,* 34, 187-220.

[5] Cox, D.R. (1975), "A Note on Data-Splitting for the Evaluation of Significance Levels," *Biometrika*, 62, 441-444.

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

[7] Faraway, J.J. (1995), "Data Splitting Strategies for Reducing the Effect of Model Selection on Inference," Technical Report # 259, Department of Statistics, University of Michigan.

[8] Friedman, J., Hastie, T., Hoefling, H., and Tibshirani, R. (2007), "Pathwise Coordinate Optimization," *Annals of Applied Statistics*, 1, 302-332.

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

[10] 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).

[11] Hastie, T., Tibshirani, R., and Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York, NY.

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

[13] Hurvich, C., and Tsai, C.L. (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.

[14] Hurvich, C., and Tsai, C.L. (1990), "The Impact of Model Selection on Inference in Linear Regression," *The American Statistician*, 44, 214-217.

[15] Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J., and Wasserman, L. (2018), "Distribution-Free Predictive Inference for Regression," *Journal of the American Statistical Association*, 113, 1094-1111.

[16] Lichman, M. (2017), UCI Machine Learning Repository,

[17] Luo, S., and Chen, Z. (2013), "Extended BIC for Linear Regression Models with Diverging Number of Relevant Features and High or Ultra-High Feature Spaces," *Journal of Statistical Planning and Inference,* 143, 494-504.

[18] Mallows, C. (1973), "Some Comments on C_p," *Technometrics,* 15, 661-676.

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

[20] Moran, P.A.P. (1973), "Dividing a Sample Into Two Parts a Statistical Dilemma," *Sankhyā, the Indian Journal of Statistics, A*, 35, 329-333.

[21] Mykland, P.A. (2003), "Financial Options and Statistical Prediction Intervals," *The Annals of Statistics*, 31, 1413-1438.

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

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

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

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

[26] Olive, D.J. (2022), *Prediction and Statistical Learning*, online course notes, see (http://parker.ad.siu.edu/Olive/slearnbk.htm).

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

[28] Olive, D.J., and Hawkins, D.M. (2010), "Robust Multivariate Location and Dispersion," Preprint, see (http://parker.ad.siu.edu/Olive/pphbmld.pdf).

[29] Olive, D.J., Rathnayake, R.C., and Haile, M.G. (2021), "Prediction Intervals for GLMs, GAMs, and Some Survival Regression Models," *Communications in Statistics: Theory and Methods*, to appear. See (http://parker.ad.siu.edu/Olive/pppigam.pdf).

[30] Pelawa Watagoda, L. C. R., and Olive, D.J. (2021a), "Bootstrapping Multiple Linear Regression after Variable Selection," *Statistical Papers*, 62, 681-700.

[31] Pelawa Watagoda, L.C.R., and Olive, D.J. (2021b), "Comparing Six Shrinkage Estimators with Large Sample Theory and Asymptotically Optimal Prediction Intervals," *Statistical Papers*, 62, 2407-2431.

[32] Picard, R.R., and Berk, K.N. (1990), "Data Splitting," *The American Statistician*, 44, 140-147.

[33] Rathnayake, R.C., and Olive, D.J. (2021), “Bootstrapping Some GLMs and Survival Regression Models after Variable Selection,” *Communications in Statistics: Theory and Methods*, to appear. Preprint at (http://parker.ad.siu.edu/Olive/ ppbootglm.pdf).

[34] R Core Team (2018), “R: a Language and Environment for Statistical Computing,” R Foundation for Statistical Computing, Vienna, Austria, (www.R-project.org).

[35] Rinaldo, A., Wasserman, L., and G’Sell, M. (2019), “Bootstrapping and Sample Splitting for High-Dimensional, Assumption-Lean Inference,” *The Annals of Statistics*, 47, 3438-3469.

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

[37] Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011), “Regularization Paths for Cox’s Proportional Hazards Model Via Coordinate Descent,” *Journal of Statistical Software*, 39, 1-13.

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

[39] Tay, J.K., Narasimhan, B. and Hastie, T. (2021), “Elastic Net Regularization Paths for All Generalized Linear Models,” *Journal of Statistical Software*, to appear.

[40] Wieczorek, J., and Lei, J. (2021), “Model-Selection Properties of Forward Selection and Sequential Cross-Validation for High-Dimensional Regression,” *Canadian Journal of Statistics*, to appear.

[41] Zhang, J., Olive, D.J., and Ye, P. (2012), “Robust Covariance Matrix Estimation with Canonical Correlation Analysis,” *International Journal of Statistics and Probability*, 1, 119-136.

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

[43] 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.

VITA

Graduate School
Southern Illinois University

Lingling Zhang

lingling.zhang@siu.edu

Southern Illinois University at Carbondale
Bachelor of Science, Mathematics, August,2017

Southern Illinois University Carbondale
Master of Science, Mathematics, August, 2019

Dissertation Paper Title:
Data Splitting Inference

Major Professor: Dr. David. Olive