Construction of simultaneous confidence bands using conditional Monte Carlo

S. Zhou[a*], K. Yao[a], W. Liu[b], F. Bretz[c]

[a] Institute of Finance and Economics Research, School of Urban and Regional Science

Shanghai University of Finance and Economics, Shanghai 200433, China

[b] S3RI and School of Mathematics

University of Southampton, SO17 1BJ, UK

[c] Novartis Pharma AG

Basel, 4002, Switzerland

[*] sanyu.zhou@foxmail.com

Abstract

A method based on conditional Monte Carlo is introduced to construct the one-sided and two-sided simultaneous confidence bands of the constant width and hyperbolic shapes.

Keywords: regression; statistical inference; simulation.

1 Introduction

Consider the standard normal-error linear regression model with n observations $(y_i, x_{i1}, \ldots, x_{ik})$, $i = 1, \ldots, n$, which can be written in the following matrix form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{1}$$

where $\mathbf{Y} = (y_1, \ldots, y_n)^T$, $\mathbf{X}$ is the full column rank design matrix whose ith row is given by $(1, x_{i1}, \ldots, x_{ik})$, $i = 1, \ldots, n$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^T$, and $\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$.

The estimators of $\boldsymbol{\beta}$ and σ^2 are given by $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ and $\widehat{\sigma}^2 = \mathbf{Y}^T[\mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T]\mathbf{Y}/\nu$, respectively, with $\nu = n - k - 1$, $\widehat{\boldsymbol{\beta}} \sim N(\mathbf{0}, \sigma^2(\mathbf{X}^T\mathbf{X})^{-1})$, $\widehat{\sigma}^2/\sigma^2 \sim \chi^2_\nu/\nu$, and $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma}^2$ being independent. S-ince the matrix $(\mathbf{X}^T\mathbf{X})^{-1}$ is a positive definite matrix, exists a symmetric matrix $\mathbf{P}$, such that $(\mathbf{X}^T\mathbf{X})^{-1} = \mathbf{P}^2$. Let $\mathbf{N} = [\mathbf{P}^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})]/\sigma$ and $\mathbf{T} = [\mathbf{P}^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})/\sigma]/(\widehat{\sigma}/\sigma) = \mathbf{N}/(\widehat{\sigma}/\sigma)$. Then $\mathbf{N} \sim N(\mathbf{0}, \mathbf{I})$ and $\mathbf{T}$ has a multivariate t-distribution with mean $\mathbf{0}$, covariance $\mathbf{I}$, and ν degrees of freedom. Furthermore, $F = \|\mathbf{T}\|^2/(k+1)$ has an F-distribution with $k+1$ and ν degrees of freedom.

The two-sided confidence band for the linear regression function $\mathbf{x}^T\boldsymbol{\beta}$ has the following form:

$$\mathbf{x}^T\boldsymbol{\beta} \in \mathbf{x}^T\widehat{\boldsymbol{\beta}} \pm c\widehat{\sigma}\Upsilon(\widetilde{\mathbf{x}}), \quad \widetilde{\mathbf{x}} \in \Theta \tag{2}$$

where $\mathbf{x} = (1, x_1, \ldots, x_k)^T$, $\widetilde{\mathbf{x}} = (x_1, \ldots, x_k)^T$, c is a critical constant, $\Theta = \{(x_1, \ldots, x_m) : \ x_1 \in (a_1, b_1), \ldots, \ x_m \in (a_m, b_m)\}$, and $\Upsilon(\widetilde{\mathbf{x}})$ is a function of $\widetilde{\mathbf{x}}$. The confidence band in (2) is a constant width confidence band when $\Upsilon(\widetilde{\mathbf{x}}) = 1$, and a hyperbolic confidence band when $\Upsilon(\widetilde{\mathbf{x}}) = \sqrt{\mathbf{x}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}}$.

The one-sided upper confidence band has the form

$$\mathbf{x}^T\boldsymbol{\beta} < \mathbf{x}^T\widehat{\boldsymbol{\beta}} + c\widehat{\sigma}\Upsilon(\widetilde{\mathbf{x}}), \quad \widetilde{\mathbf{x}} \in \Theta \tag{3}$$

and the lower confidence band has the form

$$\mathbf{x}^T\boldsymbol{\beta} > \mathbf{x}^T\widehat{\boldsymbol{\beta}} - c\widehat{\sigma}\Upsilon(\widetilde{\mathbf{x}}), \quad \widetilde{\mathbf{x}} \in \Theta \tag{4}$$

For a polynomial regression function $\mathbf{x}^T\boldsymbol{\beta}$, $\mathbf{x} = (1, x^1, \ldots, x^k)^T$, $\widetilde{\mathbf{x}} = x$, $\Theta = \{x : \ x \in (a, b)\}$, the ith row of the design matrix $\mathbf{X}$ becomes $(1, x_i^1, \ldots, x_i^k)$, and the confidence bands in (2), (3) and (4) become the corresponding confidence bands for the polynomial regression function $\mathbf{x}^T\boldsymbol{\beta}$ over the covariate interval (a, b).

The main idea of using simulation to construct confidence bands is to express c as the $(1-\alpha)$ quantile of a random variable, which can be approximated by the $(1-\alpha)$ sample quantile of a sample of the random variable. The sample quantile can be a highly accurate approximation to the population quantile c if the sample size is sufficiently large. Generation of the random variable involves different algorithms of maximization for multiple regression and polynomial regression, and so these two cases need to be treated differently.

The purpose of this paper is to propose a new simulation-based method for computing c in (2), (3) and (4) for either multiple or polynomial regressions, and to demonstrate that the new method is more efficient than the currently used simulation-based method (the HM method). The new simulation-based method uses conditional Monte Carlo (referred to as CM-C), which was introduced by Trotter and Tukey (1956) and generalized by Hammersley (1956). In the proposed method, the conditional Monte Carlo is designed for estimating conditional expectations of functions by sampling from unconditional distributions obtained by certain weighting schemes. A suitable weighting scheme will reduce variances of the estimators. A simple proof is given in Section 3.1.

This paper is divided as follows. Section 2 describes some preliminary results. Section 3 computes the confidence levels. Section 4 computes the critical constants. Section 5 offers some summary comments.

2 Preliminary results

The transformation from cartesian to polar coordinates is central to the new simulation method, which is introduced briefly in this section.

Define the polar coordinates of the vector $\mathbf{N} = (N_0, N_1, \ldots, N_k)^T \sim N(\mathbf{0}, \mathbf{I})$ to be $(R, \theta_1 \ldots, \theta_k)^T$ given by $N_j = R\mathrm{sin}\theta_1 \ldots \mathrm{sin}\theta_{j-1}\mathrm{cos}\theta_j$ for $0 \le j \le k-1$ and $N_k = R\mathrm{sin}\theta_1, \ldots, \mathrm{sin}\theta_{k-1}\mathrm{sin}\theta_k$ where $0 \le \theta_j \le \pi$ for $1 \le j \le k-1$, $0 \le \theta_k \le 2\pi$ and $R \ge 0$. The Jacobian of the transformation is $|J| = R^k \sin^{k-1}\theta_1 \sin^{k-2}\theta_2 \ldots \sin\theta_{k-1}$.

It follows directly from the transformation that $\|\mathbf{N}\| = R$ and the joint density function of $R, \theta_1, \ldots, \theta_k$ is given by

$$f(R, \theta_1, \ldots, \theta_k) = (2\pi)^{-(k+1)/2} \mathrm{e}^{-R^2/2} R^k \sin^{k-1}\theta_1 \sin^{k-2}\theta_2 \cdots \sin\theta_{k-1} \quad (5)$$

which implies that R and $\theta_1, \cdots, \theta_k$ are independent. From the joint density function of $R, \theta_1, \ldots, \theta_k$ in (5), it is easy to derive the joint density function of $\theta_1, \ldots, \theta_k$:

$$g(\theta_1, \cdots, \theta_k) = \frac{1}{2}\pi^{-(k+1)/2}\Gamma\left((k+1)/2\right) \sin^{k-1}\theta_1 \sin^{k-2}\theta_2 \cdots \sin\theta_{k-1}$$

where $\Gamma(\cdot)$ denotes the gamma function.

3 Computation of the confidence level

In this section, we consider the computation of the confidence level for a given critical constant c. We provide the details for two-sided hyperbolic and constant width confidence bands, while one-sided hyperbolic and constant width confidence bands can be dealt with in a similar way, and so the details are omitted.

3.1 Two-sided hyperbolic confidence bands

The two-sided hyperbolic confidence band from (2) is given by

$$\mathbf{x}^T\boldsymbol{\beta} \in \mathbf{x}^T\widehat{\boldsymbol{\beta}} \pm c\widehat{\sigma}\sqrt{\mathbf{x}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}}, \quad \widetilde{\mathbf{x}} \in \Theta. \tag{6}$$

Hence its confidence level $1-\alpha$ can be expressed as follows:

$$\begin{aligned}
CL_{\mathrm{H2}}(c) &= \mathrm{Prob}\left\{\sup_{\widetilde{\mathbf{x}}\in\Theta} \frac{|\mathbf{x}^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})|}{\widehat{\sigma}\sqrt{\mathbf{x}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}}} < c\right\} = \mathrm{Prob}\left\{\frac{\|\mathbf{N}\|}{\widehat{\sigma}/\sigma}\sup_{\widetilde{\mathbf{x}}\in\Theta}\frac{\left|(\mathbf{P}\mathbf{x})^T\mathbf{N}\right|}{\|\mathbf{P}\mathbf{x}\|\|\mathbf{N}\|} < c\right\} \\
&= \mathrm{Prob}\left\{\|\mathbf{T}\|Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k) < c\right\} \qquad (7) \\
&= \mathrm{Prob}\left\{F < c^2/[(k+1)Q_{\mathrm{H2}}^2(\theta_1,\ldots,\theta_k)]\right\} \\
&= \int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi} g(\theta_1,\cdots,\theta_k)F_{k+1,\nu}\left(\frac{c^2}{(k+1)Q_{\mathrm{H2}}^2(\theta_1,\ldots,\theta_k)}\right)d\theta_1\ldots d\theta_k \\
&= E_{\theta_1,\ldots,\theta_k}\left\{F_{k+1,\nu}\left(\frac{c^2}{(k+1)Q_{\mathrm{H2}}^2(\theta_1,\ldots,\theta_k)}\right)\right\} \qquad (8)
\end{aligned}$$

where $Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)$ in (7) is defined as

$$Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k) = \sup_{\widetilde{\mathbf{x}}\in\Theta}\frac{\left|(\mathbf{P}\mathbf{x})^T\mathbf{N}\right|}{\|\mathbf{P}\mathbf{x}\|\|\mathbf{N}\|} \tag{9}$$

whose value can be computed quickly by using the method of Liu *et al.* (2005a, b) for multiple linear regression, and the method of Liu *et al.* (2008) for polynomial regression with any $k > 1$; and $F_{k+1,\nu}(\cdot)$ in (8) is the cdf of an F-distribution with $k+1$ and ν degrees of freedom.

Using the expression in (7), the value of $CL_{\mathrm{H2}}(c)$ can be computed directly by the following way.

- Step 1: draw $\mathbf{N} \sim N(\mathbf{0},\mathbf{I})$ and $\widehat{\sigma}/\sigma \sim \sqrt{\chi_\nu^2/\nu}$ independently;

- Step 2: compute $Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)$ using (9), $\mathbf{T} = \mathbf{N}/(\widehat{\sigma}/\sigma)$ and $S = \|\mathbf{T}\|Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)$;

- Step 3: repeat Steps 1 and 2 for a large number v times to get v values of S, and use the proportion $\widehat{P}_{\mathrm{H2}}(c)$ of times that $S < c$ as $CL_{\mathrm{H2}}(c)$.

It is well known that $\widehat{P}_{\mathrm{H2}}(c) \to CL_{\mathrm{H2}}(c)$ a.s. as $v \to \infty$, and the standard error of $\widehat{P}_{\mathrm{H2}}(c)$ is given approximately by $s.e._{\widehat{P}_{\mathrm{H2}}} = \sqrt{\widehat{P}_{\mathrm{H2}}(c)(1 - \widehat{P}_{\mathrm{H2}}(c))/v}$. This method is usually called the hit-and-miss (HM) method (cf. Jones *et al.*, 2009).

Next we propose a new method of computing $CL_{\mathrm{H2}}(c)$ by using the expression (8) in the following way. In expression 8, $CL_{H2}(c)$ is expressed as the expectation of a random variable which is a suitable function of $(\theta_1, \ldots, \theta_k)$; this new method is based on conditional Monte Carlo.

- Step 1: draw $\mathbf{N} \sim N(\mathbf{0}, \mathbf{I})$;
- Step 2: compute $Q_{\mathrm{H2}}(\theta_1, \ldots, \theta_k)$ and $M = F_{k+1,\nu}\left(c^2/[(k+1)Q^2_{\mathrm{H2}}(\theta_1, \ldots, \theta_k)]\right)$;
- Step 3: repeat Steps 1 and 2 for a large number u times to get u values of M, and use the average of the u values of M, $\widehat{M}_{\mathrm{H2}}$, as $CL_{\mathrm{H2}}(c)$.

It is clear from the expression in (8) that $\widehat{M}_{\mathrm{H2}}(c) \to CL_{\mathrm{H2}}(c)$ a.s. as $u \to \infty$. The s.e. of $\widehat{M}_{\mathrm{H2}}$ is given approximately by $s.e._{\widehat{\mathrm{M}}_{\mathrm{H2}}} = \sqrt{\frac{1}{u(u-1)} \sum_{i=1}^{u} (M_i - \widehat{M}_{\mathrm{H2}}(c))^2}$.

In fact, the variance of the proposed estimator is not greater than that produced by the HM method. For convenience, let $\mathrm{Var}(\widehat{CL}_{H2}|\boldsymbol{\theta})$ denote the variance of the proposed estimator and $\mathrm{Var}(\widehat{CL}_{H2})$ the variance of the estimator given by the HM method, with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$. It is easy to prove that $\mathrm{Var}(\widehat{CL}_{H2}|\boldsymbol{\theta}) < \mathrm{Var}(\widehat{CL}_{H2})$. Note that $\mathrm{Var}(\widehat{CL}_{H2}|\boldsymbol{\theta}) = E(\widehat{CL}^2_{H2}|\boldsymbol{\theta}) - \left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]^2$. Taking expectations of both sides of this equation with respect to $\boldsymbol{\theta}$ gives

$$
\begin{aligned}
E\left[\mathrm{Var}(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] &= E\left[E(\widehat{CL}^2_{H2}|\boldsymbol{\theta})\right] - E\left\{\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]^2\right\} \\
&= E(\widehat{CL}^2_{H2}) - E\left\{\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]^2\right\} \qquad (10)
\end{aligned}
$$

Also, because $\mathrm{Var}\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] = E\left\{\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]^2\right\} - \left\{E\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]\right\}^2$ and $E\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] = E(\widehat{CL}_{H2})$, it follows that we have

$$\mathrm{Var}\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] = E\left\{\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]^2\right\} - \left[E(\widehat{CL}_{H2})\right]^2 \tag{11}$$

Upon adding Equations (10) and (11), we obtain the identity $\mathrm{Var}(\widehat{CL}_{H2}) = E\left[\mathrm{Var}(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] + \mathrm{Var}\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right]$. Clearly, we have that $\mathrm{Var}\left[E(\widehat{CL}_{H2}|\boldsymbol{\theta})\right] \leq \mathrm{Var}(\widehat{CL}_{H2})$. It indicates that the proposed method has an advantage of reducing the estimator's variance over the HM method.

3.2 Two-sided constant width confidence bands

The two-sided constant width confidence band from (2) is given by

$$\mathbf{x}^T\boldsymbol{\beta} \in \mathbf{x}^T\widehat{\boldsymbol{\beta}} \pm c\widehat{\sigma}, \quad \widetilde{\mathbf{x}} \in \Theta. \tag{12}$$

Its confidence level can be expressed as

$$\begin{aligned} CL_{\mathrm{C2}} &= \mathrm{Prob}\left\{\sup_{\widetilde{\mathbf{x}}\in\Theta}\left|\mathbf{x}^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})/\widehat{\sigma}\right| < c\right\} \\ &= \mathrm{Prob}\left\{\|\mathbf{T}\|Q_{\mathrm{C2}}(\theta_1,\ldots,\theta_k) < c\right\} \qquad (13) \\ &= E_{\theta_1,\ldots,\theta_k}\left\{F_{k+1,\nu}\left(\frac{c^2}{(k+1)Q^2_{\mathrm{C2}}(\theta_1,\ldots,\theta_k)}\right)\right\} \qquad (14) \end{aligned}$$

where $Q_{\mathrm{C2}}(\theta_1,\ldots,\theta_k)$ in (13) is defined as

$$Q_{\mathrm{C2}}(\theta_1,\ldots,\theta_k) = \sup_{\widetilde{\mathbf{x}}\in\Theta}\frac{\left|(\mathbf{P}\mathbf{x})^T\mathbf{N}\right|}{\|\mathbf{N}\|},$$

the value of $Q_{\mathrm{C2}}(\theta_1,\ldots,\theta_k)$ can be computed efficiently by using the methods in Liu *et al.* (2010) for either multiple or polynomial regressions.

The HM and CMC methods for computing $CL_{\mathrm{C2}}(c)$ can be developed in similar ways as for $CL_{\mathrm{H2}}(c)$ in sub-section 3.1, and so the details are omitted.

Now we use two examples to demonstrate that the proposed method has much greater computational efficiency than the HM method for achieving a similar *s.e.*. All the computation in this paper is done on an ordinary Window's PC (Intel(R) Core(TM) i7-7820HK CPU @ 2.90GHz, RAM 32.0 GB). The first example is about the change of conversion of n-Heptane to acetylene (%) (y) at different reactor temperatures (°C) (x_1) and different ratios of H_2 to n-Heptane (mole ratio) (x_2). The related data were taken from Liu *et al.* (2005a). From the data set, the fitted multiple linear regression model was given by $\widehat{y} = -130.69 + 0.134x_1 + 0.35x_2$ with $R^2 = 0.92$.

Table 1: Computation of the confidence levels for the pre-specified c and $v = u = 100000$.

Regression	SCB†	1 or 2 sided	Pre-specified c	Method	CL	s.e.	Computational time (second)
Multiple	HB◇	2	3.1049	HM	0.9509	5.9338×10^{-3}	4.36
			3.1049	CMC	0.9501	4.1100×10^{-5}	4.39
		1	2.7383	HM	0.9498	7.0943×10^{-3}	2.51
			2.7383	CMC	0.9499	1.4322×10^{-4}	2.50
	CWB♡	2	1.7196	HM	0.9516	6.2922×10^{-3}	1.21
			1.7196	CMC	0.9500	1.3536×10^{-4}	1.28
		1	1.5105	HM	0.9514	6.3686×10^{-3}	1.22
			1.5105	CMC	0.9501	2.0597×10^{-4}	1.43
Polynomial	HB	2	2.9803	HM	0.9501	7.1156×10^{-3}	32.33
			2.9867	CMC	0.9501	1.2066×10^{-4}	32.89
		1	2.6924	HM	0.9499	6.6072×10^{-3}	31.58
			2.6939	CMC	0.9499	1.9764×10^{-4}	32.22
	CWB	2	1.6902	HM	0.9501	6.9029×10^{-3}	24.06
			1.6885	CMC	0.9500	2.7181×10^{-4}	23.88
		1	1.4562	HM	0.9501	6.2021×10^{-3}	23.29
			1.4591	CMC	0.9501	3.3126×10^{-4}	24.15

† SCB denotes the simultaneous confidence band;
◇ HB denotes the hyperbolic confidence band;
♡ CWB denotes the constant width band.

The second example is about the relation between the probabilities of perinatal death (fetal deaths plus deaths within the first month of life) (p) and birth weight (BW). The related data were taken from Liu *et al.* (2008).

Based on the data, we fitted a fourth-order polynomial regression model between $y = log(-log(p))$ and $x = \text{BW}$. The fitted model was given by $\widehat{y} = -2.86 + 4.81x - 2.32x^2 + 0.568x^3 - 0.054x^4$ with $R^2 = 0.99$.

For the given critical constants, the confidence levels were computed by using $u = v = 100000$ simulations. The results were summarized in Table 1. From the last column of Table 1, we can see that the computational time used by the two methods are similar.

However, the s.e.'s of the two methods given in the penultimate column of Table 1 are markedly different. We can see that the s.e.'s of the CMC method are about 1/20, at least, of those of the HM method. Since the s.e. of the HM method depends on the number of replications v through $\frac{1}{\sqrt{v}}$, v should be about $400u$, in order to achieve a similar s.e. of the CMC method with u replications. Hence the HM method takes about 400 times of the CMC computation-time to achieve a similar accuracy of the confidence level.

4 Computation of the critical constant

In this section, we consider the computation of the critical constant c, so that the confidence band has the pre-specified confidence level $1 - \alpha$. Again we focus only on two-sided hyperbolic confidence bands, while one-sided hyperbolic confidence bands, one- and two-sided constant width confidence bands can be treated in a similar way and so the details are omitted.

The currently available method that works for multiple and polynomial regressions for a general $k > 1$ (cf. Liu, 2010) utilizes the expression in (7) to find the critical constant c by solving the equation

$$1 - \alpha = \text{Prob}\left\{\|\mathbf{T}\| Q_{\text{H2}}(\theta_1, \ldots, \theta_k) < c\right\} \tag{15}$$

Note that the critical constant c is the $(1-\alpha)$ quantile of the random variable

$\|\mathbf{T}\| Q_{\rm H2}$, which can be approximated by the $(1-\alpha)$ sample quantile in the following way (see Liu, 2010, pp 65, and the references therein for more details).

- Step 1: draw independent $\mathbf{N} \sim N(\mathbf{0}, \mathbf{I})$ and $\widehat{\sigma}/\sigma \sim \sqrt{\chi^2_\nu/\nu}$;
- Step 2: compute $S = \|\mathbf{T}\| Q_{\rm H2}$ as before;
- Step 3: repeat Steps 1 and 2 for a large number m times to get $S_1, \ldots, S_m$.

Let $\langle (1-\alpha)m \rangle$ denote the integer part of $(1-\alpha)m$. It is well known that $S_{\langle (1-\alpha)m \rangle} \to c$ a.s. as $m \to \infty$ and the s.e. of $S_{\langle (1-\alpha)m \rangle}$ can be computed approximately by using $S_1, \ldots, S_m$ (cf. Liu, 2010, pp 243-244).

For convenience, let $\widehat{c} = S_{\langle (1-\alpha)m \rangle}$. It is known (see e.g., Serfling, 1980) that, under quite weak conditions, $\widehat{c}$ is asymptotically normal with mean c and standard error $s.e. = \sqrt{\alpha(1-\alpha)/[mG^2(c)]}$, where $G(c)$ is the density function of S evaluated at c. Liu (2010) gave a detailed discussion about determining the value of $G(c)$. From these, the $1-\alpha$ "asymptotic" confidence bounds $(\widehat{c} - z_{\alpha/2} \cdot s.e., \widehat{c} + z_{\alpha/2} \cdot s.e.)$ can be derived.

Now we introduce a new method for computing c from (16)

$$1 - \alpha = E_{\theta_1, \ldots, \theta_k} \left\{ F_{k+1,\nu} \left(\frac{c^2}{(k+1) Q^2_{\rm H2}(\theta_1, \ldots, \theta_k)} \right) \right\} \qquad (16)$$

by using the efficient CMC method of computing $CL_{\rm H2}(c)$ given in Section 3. Since this probability is monotonically increasing in c, we use a searching algorithm that solves the equation in (16) to find c.

The bi-section searching algorithm is used in our R code for finding c. Note that (cf. Liu, 2010) the solution c is contained in the interval $(t_\nu^{\alpha/2}, \sqrt{(k+1) f^\alpha_{k+1,\nu}})$ where t^β_ν and $f^\beta_{k+1,\nu}$ denote the β quantiles of the t-distribution with ν degrees of freedom and the F-distribution with $k+1$

and ν degrees of freedom, respectively. Furthermore, the $1-\alpha$ confidence bounds of c can be produced by the bootstrap method.

The subtlety in this method is that, when computing $CL_{\mathrm{H2}}(c)$ using the expression (8) for different values of c, the same u values of $Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)$ are used. Hence, after computing $CL_{\mathrm{H2}}(c)$ for one c, the extra computation for a new c is just to compute u values of the cdf $F_{k+1,\nu}(c^2/[(k+1)Q^2_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)])$ without computing the u values of $Q_{\mathrm{H2}}(\theta_1,\ldots,\theta_k)$.

To assess the s.e. of the critical constant c computed using this new CMC method, we compute w values of c using w different random seeds, and approximate the s.e. by $s.e. = \sqrt{\frac{1}{w-1}\sum_{i=1}^{w}(c_i-\bar{c})}$ where $\bar{c}=\sum_{i=1}^{w}c_i$.

Table 2: Computation of the critical constant for $1-\alpha=0.95$, $m=u=100000$ and $w=100$.

Regression	SCB	1 or 2 sided	Method	c	Confidence bounds	Length of confidence bounds	s.e.	Computational time (second)
Multiple	HB	2	HM	3.0940	(3.0783,3.1097)	0.0314	8.0024×10^{-3}	4.31
			CMC	3.1038	(3.1028,3.1047)	0.0019	4.9017×10^{-4}	7.64
		1	HM	2.7412	(2.7250,2.7573)	0.0323	8.2460×10^{-3}	2.81
			CMC	2.7395	(2.7361,2.7428)	0.0067	1.4870×10^{-3}	6.37
	CWB	2	HM	1.7081	(1.6989,1.7174)	0.0185	4.7158×10^{-3}	1.23
			CMC	1.7200	(1.7180,1.7219)	0.0038	1.0763×10^{-3}	6.34
		1	HM	1.5012	(1.4917,1.5107)	0.0190	4.8445×10^{-3}	1.40
			CMC	1.5098	(1.5070,1.513)	0.0058	1.3197×10^{-3}	6.25
Polynomial	HB	2	HM	2.9795	(2.9679,2.9911)	0.0232	5.9179×10^{-3}	30.87
			CMC	2.9858	(2.9837,2.9879)	0.0043	9.0958×10^{-4}	8.20
		1	HM	2.6932	(2.6807,2.7057)	0.0250	6.3825×10^{-3}	30.97
			CMC	2.6948	(2.6912,2.6984)	0.0072	1.7962×10^{-3}	7.74
	CWB	2	HM	1.6897	(1.6808,1.6986)	0.0178	4.5357×10^{-3}	20.27
			CMC	1.6886	(1.6851,1.6920)	0.0069	1.6662×10^{-3}	7.44
		1	HM	1.4555	(1.4462,1.4647)	0.0186	4.7379×10^{-3}	20.81
			CMC	1.4584	(1.4540,1.4629)	0.0089	2.0044×10^{-3}	7.64

To demonstrate the advantage of the new CMC method over the HM method, we computed the critical constants and their s.e.'s using the two methods, respectively, for the same two examples in Section 3. The computational results were summarized in Table 2.

In search of the critical constant c, we used the resolution 0.001 as a stop-search condition, that is, if the length of an interval was greater than 0.001, the search on the interval would continue, until its length was less than 0.001. This guarantees that the searched critical constant would be accurate, due to searching, to the third decimal place. The entries in the last two columns of Table 2 can be used to compare the computation times of the HM and CMC methods for achieving similar accuracy (i.e., s.e.) of the critical constant c. Note that the s.e. of the HM method depends on m through $1/\sqrt{m}$ (cf. Liu, 2010, pp.243), and the computation time of the CMC method depends on m linearly. Hence from the last two entries in each of the first two rows of Table 2, for example, to achieve a similar s.e. as the CMC method, the HM method needs a value of m approximately equal to $(8.0024 \times 10^{-3}/4.9017 \times 10^{-4})^2 \approx 266$ multiplying the currently used $m - value$ 100,000. This results in an approximate computation time of 266×4.31 seconds, which is $266 \times 4.31/7.64 \approx 150$ times of the computation time 7.64 seconds of CMC computation time. For the eight cases given in Table 4 from top to bottom, the multiples, corresponding to the 150 above, are given respectively by 150, 13, 3, 3, 159, 50 20, 15. Therefore the HM method requires at least three times of the computation time of the IS method to achieve similar accuracy. For the hyperbolic bands, the multiples are 150, 13, 159, 50, which are much more pronounced than for the constant width bands. Moreover, for the confidence level $1 - \alpha = 0.95$, the proposed method gives narrower confidence intervals and their lengths are at most half of those given by the HM method. These demonstrate the superiority of the CMC method over the HM method.

5 Concluding remarks

The proposed method is much more efficient at computing the confidence level and the critical constant than the currently used HM method, and therefore we recommend using it.

The R programs for implementing the computation in this manuscript are available at the website `https://ife.sufe.edu.cn/31/47/c3531a78151/page.htm` for downloading.

References

[1] Hammersley, J., 1956. Conditional Monte Carlo. Journal of the ACM (JACM) 3(2), 73-76.

[2] Jones, O., Maillardet, R., Robinson, A., 2009. Introduction to Scientific Programming and Simulation using r. CRC Press.

[3] Liu, W., 2010. Simultaneous Inference in Regression. CRC Press.

[4] Liu, W., Jamshidian, M., Zhang, Y., Bretz, F., 2005a. Constant width simultaneous confidence bands for multiple linear regression when predictor variables are constrained in intervals. J. Statist. Comput. Simulation 75, 425-436. https://doi.org/10.1080/00949650410001701715

[5] Liu, W., Jamshidian, M., Zhang, Y., Donnelly, J., 2005b. Simulation-based simultaneous confidence bands in multiple linear regression with predictor variables constrained in intervals. J. Computat. Graph. Statist. 14, 459-484. http://www.jstor.org/stable/27594124

[6] Liu, W., Wynn, H.P., Hayter, A.J., 2008. Statistical inferences for linear regression models when the covariates have functional relationships: polynomial regression. J. Stat. Comput. Sim. 78(4), 315-324. http://dx.doi.org/10.1080/00949650601100981

[7] Serfling, R.J., 1980. Approximation theorems of mathematical statistics. Wiley.

[8] Trotter, H., Tukey, J., 1956. Conditional Monte Carlo for normal samples. In H. Meyer (Ed.), Proc. Symp. on Monte Carlo Methods, pp. 64-79. John Wiley and Sons, New York.