Name: Jared Yu   STA243 HW6   Date: Jun 6, 2018

Question 1

(a) The sample variance of LSAT scores in the data is $Var(Y) = 1746.781$, the sample variance of GPA in the data is $Var(Z) = 0.1067267$ and the sample covriance of LSAT and GPA is $Cov(Y, Z) = 7.453905$. Then the estimate of the correlation coefficient is

$$\hat{\rho}_{YZ} = \frac{Cov(Y,Z)}{\sqrt{Var(Y)Var(Z)}} = 0.5459189.$$

(b) Let $\hat{\rho}_{(-i)}$ be the estimate of the correlation coefficient using the data without the i-th observation. Then the jackknife estimate of the standard error of $\hat{\rho}_{YZ}$ is

$$\hat{se}_{jack} = \Big[\frac{n-1}{n}\sum_{i=1}^{15}(\hat{\rho}_{(-i)} - \hat{\rho}_{YZ})^2\Big]^{\frac{1}{2}} = 0.2552259.$$

Let $\hat{\rho}_{(b)}$ be the estimate of the correlation coefficient using the b-th bootstrapped sample. Then the bootstrap estimate of the standard error of $\hat{\rho}_{YZ}$ using $B = 200$ bootstrapped samples is

$$\hat{se}_B = \Big[\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\rho}_b - \text{mean}(\hat{\rho}_b))^2\Big]^{\frac{1}{2}} = 0.2071401.$$

(c) Using the normal theory, we take $B = 2000$ to estimate the standard error of $\hat{\rho}_{YZ}$. And $\hat{se}_B = 0.1971840$, hence the estimated 95% confidence interval for $\hat{\rho}_{YZ}$ is

$$\hat{\rho}_{YZ} \pm 1.96\hat{se}_B = (0.1571196,\ 0.9347182).$$

Using the bootstrap t-intervals method, we take $B = 2000$ to get 2000 estimates of $\hat{\rho}_b$'s. And in each bootstrap sample we bootstrap 200 samples for this bootstap sample to estimate $\hat{se}(b)$ in order to get the estimated "t-table" values $z(b) = (\hat{\rho}_b - \hat{\rho}_{YZ})/\hat{se}(b)$ for $b = 1, \cdots, 2000$. Let $\hat{t}^{(\alpha}$ be the $\alpha 2000$-th largest value of the $z(b)$ values, $\hat{t}^{(1-\alpha)}$ be the $(1-\alpha)2000$-th largest value of the $z(b)$ values and denote the standard error of $\hat{\rho}_b$'s by $\hat{se}$, then the estimated 95% confidence interval is

$$(\hat{\rho}_{YZ} - \hat{t}^{(1-\alpha)}\hat{se},\ \hat{\rho}_{YZ} - \hat{t}^{(\alpha)}\hat{se}) = (-1.280569,\ 0.8557734).$$

We can see that this method gives us a verey small 95% confidence lower bound. This confirms the drawbacks mentioned in class, which is this approach can give erratic results and can be influenced by a few outliers.

Question 2

(a) Let $X_1, \cdots, X_n \sim$ Uniform$(0,3)$ and $\hat{\theta} = X_{max} = \max(X_1, \cdots, X_n)$. The distribution of $\hat{\theta}$ is given below:

$$P(\hat{\theta} \leq x) = P(\max(X_1, \cdots, X_n) \leq x) = P(X_1 \leq x, \cdots, X_n \leq x) = [P(X_1 \leq x)]^n = (\frac{x}{3})^n.$$

Taking the derivative with respect to x, we can obtain the p.d.f of $\hat{\theta}$ as $\frac{nx^{n-1}}{3^n}$.

(b) $E(\hat{\theta}) = \int_0^\theta \frac{nx^n}{\theta^n}\, dx = \frac{n}{n+1}\left[\frac{x^{n+1}}{\theta^n}\right]_0^\theta = \frac{n\theta}{n+1}$
$E(\hat{\theta}^2) = \int_0^\theta \frac{nx^{x+1}}{\theta^n}\, dx = \frac{n\theta^2}{n+2}$.
$Var(\hat{\theta}) = E(\hat{\theta}^2) - E(\hat{\theta})^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$

(c,d) Given $n = 50$ and $\theta = 3$, $Var_{F_\theta}(\hat{\theta}) = 0.00333$. The one approximated by parametric and nonparametric bootstrap are 0.00317 and 0.00115 respectively. Obviously, the one approximated by parametric bootstrap is more accurate.

(e) The required histograms are given in Figure 1.

(f) The true distribution of $\hat{\theta}$ is also given in Figure 1. We can see the parametric bootstrap makes a better approximation for the density function of $\hat{\theta}$ than the nonparametric bootstrap. One of the reasons is because the distribution is very skewed. In order to estimate the shape, nonparametric bootstrap requires more large observations. If the values are not "large" enough, then the density of $\hat{\theta}$ will be readily underestimated. Since we only have 25 data points, this is almost impossible to get a good estimation for the density function using nonparametric bootstrap in this case. A suggestion is to draw much more samples from $U(0,3)$. On the other hand, we can see parametric bootstrap can get a reasonable approximation to the density function given a small number of data points.

Question 3

(a) The regression curve estimates for both test functions using the genetic algorithm are given in Figure 2. In assignment 2, we already fitted the piecewise constant function to test function 1. So the result looks exactly the same. For test function 2, we can see the regression curve estimate is a step function with many jumps. The reason is crystally clear that the test function 2 is not a piecewise constant function. Besides, we can notice that more jumps occurs around the peak.

(b) The 95% point wise confidence bands for both curves estimates using the bootstrap are given in Figure 3, included both "bootstrapping residual" and "bootstrapping pairs" approaches.

According to Figure 3, we can see “bootstrapping residual” is more appropriate for test function 1. It is not surprise because the piecewise constant regression is an appropriate model for the test function. Thus, its residuals should be independent and identically distributed. On the other hand, we know the test function 2 is not piecewise constant. So, the piecewise constant regression function is not an appropriate model. So its residuals are not suitable to be used in this case. Furthermore, the “bootstrapping pairs” seems more suitable to find the CI band for test function 2 in this case.

Comment on the shape of the confidence bands near jump points for the test function from Assignment 2:

We can see that the confidence bands near jump points are sharp and large in the “bootstrapping pairs” approach. This is because the resampling may represent a jump location with other jump points (mismatch).

(c) Now we are trying to address a problem “how to obtain a confidence interval for the location of a jump point using the bootstrap”. Let us first consider a simple model with one jump.

Model:

Assume that a sample of n data pairs $X = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ is observed and generated by the model

$$Y_i = g(X_i) + \epsilon_i, \qquad q \leq i \leq n, \tag{1}$$

where X_i is the order statistics of a random sample a distribution having density f supported on $\mathcal{I} = [0, 1]$. g is continuous on $[0, x_0]$ and on $[x_0, 1]$, where $0 \leq x_0 \leq 1$ and $g(x_0-) \neq g(x_0+)$ and the errors ϵ_i are independent and identically distributed with zero mean and finite variance. Now, let us consider a estimation method for x_0.

Point estimation of x_0:

Suppose we have determined a small interval $[z_0, z_1]$ with $x_0 \in [z_0, z_1]$. If the interval is small enough, we might reasonably consider the following local approximation to (1):

$$Y_i \approx \begin{cases} \theta_1 + \epsilon_i & \text{if} X_i \in [z_0, x_0], \\ \theta_2 + \epsilon_i & \text{if} X_i \in [x_0, z_1], \end{cases} \tag{2}$$

where $\theta_1 \approx g(x_0-)$ and $\theta_2 \approx g(x_0+)$. Then we are approximately in the parametric setting of estimating the change point of a piecewise-constant function. Using (2), we may estimate $i_0 = \max\{i : X_i \leq x_0\}$ by minimizing the sum of squares,

$$S(i, \theta_1, \theta_2) = \sum_{X_j \in [z_0, X_i]} (Y_j - \theta_1)^2 + \sum_{X_j \in [X_{i+1}, z_1]} (Y_j - \theta_2)^2, \tag{3}$$

producing a vector $(\hat{i}_0, \theta_1, \theta_2)$ of parameter estimators. We then estimate x_0 by simply taking the midpoint of the estimated interval $[X_{\hat{i}_0}, X_{\hat{i}_0+1}]$:

$$\hat{x}_0 = \frac{1}{2}(X_{\hat{i}_0} + X_{\hat{i}_0+1}). \tag{4}$$

To determine the interval $[z_0, z_1]$, we first obtain a preliminary estimator $\tilde{x}_0$ of x_0, and then define $[z_0, z_1]$ as the interval concentrated around $\tilde{x}_0$, which is constructed by using a diagnostic D to identify the approximate location of x_0. Specifically, define D by

$$D(x,h) = \frac{\partial}{\partial x}\left(\frac{\sum_{i=1}^{n} K\{(x - X_i)/h\}Y_i}{\sum_{i=1}^{n} K\{(x - X_i)/h\}}\right), \tag{5}$$

where K is a compactly supported, differentiable kernel function and h is a bandwidth; and define $\tilde{x}_0$ to be the value of x that maximizes $|D(x,h)|$ in $(vh, 1 - vh)$, where $[-v, v]$ denotes the support of K. The interval $[z_0, z_1]$ is then defined as $[\tilde{x}_0 - th, \tilde{x}_0 + th]$, with t typically between v and $2v$. In summary, we can put it in a two-step procedure for estimating x_0 as follows:

Step 1: Locate $\tilde{x}_0$, the global maximum of $|D(x,h)|$ on $(vh, 1 - vh)$.

Step 2: Put $[z_0, z_1] = [\tilde{x}_0 - th, \tilde{x}_0 + th]$ and determine the least-squares estimate $\hat{i}_0$ of i_0 by minimizing (3). Then use (4) to determine the final estimator $\hat{x}_0$ of x_0.

Now, let us talk about the interval estimation for x_0 by using bootstrap method.

Interval estimation for x_0:

We can use the bootstrap methods to estimate the distribution of ϵ_i. In this way, an interval estimator of i_0 and, hence, of x_0 could be constructed. The bootstrap algorithm can be stated as three parts:

Part 1. (Estimation of g and computation of residuals.) Let $\hat{x}_0 = \frac{1}{2}(X_{\hat{i}_0} + X_{\hat{i}_0+1})$ denote the estimator introduced above. Using local linear regression construct an estimator, $\hat{g}$, of g on $[0, \hat{x}_0]$, and another on $[\hat{x}_0, 1]$. Define residuals $\tilde{\epsilon}_i = Y_i - \hat{g}(X_i)$ for $1 \leq i \leq n$; calculate their mean, $\tilde{\epsilon}$; and put $\hat{\epsilon}_i = \hat{\epsilon}_i - \tilde{\epsilon}$.

Part 2. (Monte Carlo simulation.) Conditional on X, let $\epsilon_1^*, \ldots, \epsilon_n^*$ denote a resample drawn randomly, with replacement, from the set $\{\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n\}$, and define

$$Y_i^* = \hat{g}(X_i) + \epsilon_i^*, \qquad 1 \leq i \leq n.$$

Then $X^* = \{(X_1, Y_1^*), \ldots, (X_n, Y_n^*)\}$ is the bootstrap version of X.

Part 3. (Constructing an interval estimator of x_0.) Using the abovementioned method, the analogues $\hat{i}_0^*$ and $\hat{x}_0^* = \frac{1}{2}(X_{\hat{i}_0^*} + X_{\hat{i}_0^*+1})$ of $\hat{i}_0$ and $\hat{x}_0$, for the resample X^* rather than the sample X. For $m = 0, -1, 1, -2, 2, -3, 3, \ldots$, determine the bootstrap probabilities

$$p_m = P(\hat{i}_0^* - \hat{i}_0 = m|X).$$

Given $\alpha \in (0,1)$, determine integers m_1^α and m_2^α with $m_1^\alpha < m_2^\alpha$, and with minimal distance $m_2^\alpha - m_1^\alpha$, such that, for some $\beta \leq \alpha$,

$$\sum_{m=m_1^\alpha}^{m_2^\alpha} p_m = P(m_1^\alpha \leq \hat{i}_0^* - \hat{i}_0 \leq m_2^\alpha|X) = 1 - \beta \geq 1 - \alpha.$$

Then $[\hat{i}_0 - m_2^\alpha, \hat{i}_0 - m_1^\alpha]$ is a bootstrap confidence interval for i_0 with nominal level $\beta \leq \alpha$. The corresponding bootstrap confidence interval for x_0 is given by $[X_{\hat{i}_0 - m_2^\alpha}, X_{\hat{i}_0 - m_1^\alpha + 1}]$.

The algorithm is roughly given a procedure to get the confidence interval. More issues still have to be addressed such as the bandwidth choices.

Question 4

(a)

$$\begin{aligned}
f_n(x|\lambda) &= \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) \qquad \text{for} \qquad x_i \geq 0 \\
\log f_n(x|\lambda) &= n \log \lambda - \lambda \sum_{i=1}^n x_i \\
\frac{\partial \log f_n(x|\lambda)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n x_i
\end{aligned}$$

Set it equals to zero and solve for λ, we get $\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$.

(b) By C.L.T., $\sqrt{n}(\bar{x} - \frac{1}{\lambda}) \xrightarrow{D} N(0, \frac{1}{\lambda^2})$. By Delta method with $g(\theta) = \frac{1}{\theta}$, $(g'(\theta))^2 = \frac{1}{\theta^2}$.

$$\begin{aligned}
\sqrt{n}(g(\bar{x}) - g(\frac{1}{\lambda})) &\xrightarrow{D} N(0, \frac{1}{\lambda^2} g'(\frac{1}{\lambda})^2) \\
\sqrt{(n)} \left(\frac{1}{\bar{x}} - \lambda \right) &\xrightarrow{D} N\left(0, \frac{1}{\lambda^2} \left(\frac{1}{(1/\lambda)^2} \right)^2 \right) \\
\sqrt{n}(\hat{\lambda} - \lambda) &\xrightarrow{D} N(0, \lambda^2).
\end{aligned}$$

By delta method with $h(\theta) = \log\theta$, $h'(\theta) = \frac{1}{\theta}$,

$$\begin{aligned}
\sqrt{n}(h(\hat{\lambda}) - h(\lambda)) &\xrightarrow{D} N(0, \lambda^2(h'(\lambda))^2) \\
\sqrt{n}(\log\hat{\lambda} - \log\lambda) &\xrightarrow{D} N(0, 1).
\end{aligned}$$

(c) $100(1-\alpha)\%$ confidence interval for $\log\lambda$ is $\log\hat{\lambda} \pm z_{\alpha/2}\frac{1}{n}$
$100(1-\alpha)\%$ confidence interval for λ is $\exp[\log\hat{\lambda} \pm z_{\alpha/2}\frac{1}{n}] = \hat{\lambda}\exp(\pm z_{\alpha/2}\frac{1}{n})$.

(d)

$$\begin{aligned}
P(\lambda(x_1 + \cdots + x_n) < G^{-1}(\alpha/2)) &= \alpha/2 \\
P(\lambda < \frac{G^{-1}(\alpha/2)}{x_1 + \cdots + x_n}) &= \alpha/2 \\
P(\lambda < \frac{\hat{\lambda}_n G^{-1}(\alpha/2)}{n}) &= \alpha/2.
\end{aligned}$$

$$\begin{aligned}
P(\lambda > \frac{G^{-1}(1-\alpha/2)}{x_1 + \cdots + x_n}) &= 1-\alpha/2 \\
P(\lambda > \frac{\hat{\lambda}_n G^{-1}(1-\alpha/2)}{n}) &= 1-\alpha/2.
\end{aligned}$$

So, $100(1-\alpha)\%$ confidence interval of λ is $\left(\frac{\hat{\lambda}_n G^{-1}(\alpha/2)}{n}, \frac{\hat{\lambda}_n G^{-1}(1-\alpha/2)}{n}\right)$.

(e) A simulation with $n = 50$ and $B = 1000$, $\alpha = 0.05$ is performed. To apply BC_a method, we can start from $\sqrt{n}(\log\hat{\lambda}_n - \log\lambda) \xrightarrow{D} N(0, 1)$, then perform the transformation $\exp(\cdot)$ since BC_a is transformation respecting.

The coverage rate for asymptotic, exact and BC_a confidence interval are 95.1%, 95.3% and 95.4% respectively. The average lengths for the intervals are 0.5615, 0.5534 and 0.5413 respectively. Apparently, the BC_a confidence interval provided a better result.

Question 5

(a) Set $x = G^{-1}(\alpha)$

$$
\begin{aligned}
P\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} + b \le G^{-1}(\alpha)\right) &= \alpha \\
P(g(\hat{\theta}_n) - g(\theta) + b + abg(\theta) \le G^{-1}(\alpha) + aG^{-1}(\alpha)g(\theta)) &= \alpha \\
P(g(\hat{\theta}_n) - (G^{-1}(\alpha) - b) \le g(\theta)[1 - ab + aG^{-1}(\alpha)]) &= \alpha \\
P\left(\frac{g(\hat{\theta}_n) - (G^{-1}(\alpha) - b)}{1 + a[G^{-1}(\alpha) - b]} \ge g(\theta)\right) &= \alpha \\
P\left(g^{-1}\left[\frac{g(\hat{\theta}_n) - (G^{-1}(\alpha) - b)}{1 + a[G^{-1}(\alpha) - b]}\right] \ge \theta\right) &= \alpha
\end{aligned}
$$

Similarly, set $x = G^{-1}(1 - \alpha)$,

$$
\begin{aligned}
P\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} + b \le G^{-1}(1 - \alpha)\right) &= 1 - \alpha \\
P\left(g^{-1}\left[\frac{g(\hat{\theta}_n) - (G^{-1}(1 - \alpha) - b)}{1 + a[G^{-1}(1 - \alpha) - b]}\right] \le \theta\right) &= 1 - \alpha
\end{aligned}
$$

(b)

$$
\begin{aligned}
H(L_n) &= P_{\hat{F}_n}\left(\hat{\theta}_{n*} \le g^{-1}\left(\frac{g(\hat{\theta}_n) - [G^{-1}(1 - \alpha) - b]}{1 + a[G^{-1}(1 - \alpha) - b]}\right)\right) \\
&= P_{\hat{F}_n}\left(g(\hat{\theta}_{n*}) \le \frac{g(\hat{\theta}_n) - [G^{-1}(1 - \alpha) - b]}{1 + a[G^{-1}(1 - \alpha) - b]}\right) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_{n*}) - g(\hat{\theta}_n)}{\sigma_{g(\hat{\theta}_n)}} \le \frac{g(\hat{\theta}_n) - g(\hat{\theta}_n)\{1 + a[G^{-1}(1 - \alpha) - b]\} - [G^{-1}(1 - \alpha) - b]}{\{1 + a[G^{-1}(1 - \alpha) - b]\}[1 + ag(\hat{\theta}_n)]}\right) \\
&\approx P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} \le \frac{-[1 + ag(\hat{\theta}_n)][G^{-1}(1 - \alpha) - b]}{\{1 + a[G^{-1}(1 - \alpha) - b]\}[1 + ag(\hat{\theta}_n)]}\right) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} + b \le b - \frac{G^{-1}(1 - \alpha) - b}{1 + a[G^{-1}(1 - \alpha) - b]}\right) \\
&\approx G\left(b - \frac{G^{-1}(1 - \alpha) - b}{1 + a[G^{-1}(1 - \alpha) - b]}\right).
\end{aligned}
$$

$$
\begin{aligned}
H(U_n) &= P_{\hat{F}_n}\left(\hat{\theta}_{n*} \leq g^{-1}\left(\frac{g(\hat{\theta}_n) - [G^{-1}(\alpha) - b]}{1 + a[G^{-1}(\alpha) - b]}\right)\right) \\
&= P_{\hat{F}_n}\left(g(\hat{\theta}_{n*}) \leq \frac{g(\hat{\theta}_n) - [G^{-1}(\alpha) - b]}{1 + a[G^{-1}(\alpha) - b]}\right) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_{n*}) - g(\hat{\theta}_n)}{\sigma_{g(\hat{\theta}_n)}} \leq \frac{g(\hat{\theta}_n) - g(\hat{\theta}_n)\{1 + a[G^{-1}(\alpha) - b]\} - [G^{-1}(\alpha) - b]}{\{1 + a[G^{-1}(\alpha) - b]\}[1 + ag(\hat{\theta}_n)]}\right) \\
&\approx P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} \leq \frac{-[1 + ag(\hat{\theta}_n)][G^{-1}(\alpha) - b]}{\{1 + a[G^{-1}(\alpha) - b]\}[1 + ag(\hat{\theta}_n)]}\right) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} + b \leq b - \frac{G^{-1}(\alpha) - b}{1 + a[G^{-1}(\alpha) - b]}\right) \\
&\approx G\left(b - \frac{G^{-1}(\alpha) - b}{1 + a[G^{-1}(\alpha) - b]}\right).
\end{aligned}
$$

(c) Given $G(x)$ is the standard normal distribution,

$$
\begin{aligned}
H(L_n) &\approx \Phi\left(b - \frac{\Phi^{-1}(1-\alpha) - b}{1 + a[\Phi^{-1}(1-\alpha) - b]}\right) \\
&= \Phi\left(b + \frac{b - z^{(1-\alpha)}}{1 - a(b - z^{(1-\alpha)})}\right) \\
&= \Phi\left(b + \frac{b + z^{(\alpha)}}{1 - a(b + z^{(\alpha)})}\right).
\end{aligned}
$$

So, $L_n \approx H^{-1}(\alpha_1)$ where $\alpha_1 = \Phi\left(b + \frac{b+z^{(\alpha)}}{1-a(b+z^{(\alpha)})}\right)$

$$
\begin{aligned}
H(U_n) &\approx \Phi\left(b - \frac{\Phi^{-1}(\alpha) - b}{1 + a[\Phi^{-1}(\alpha) - b]}\right) \\
&= \Phi\left(b + \frac{b - z^{(\alpha)}}{1 - a(b - z^{(\alpha)})}\right) \\
&= \Phi\left(b + \frac{b + z^{(1-\alpha)}}{1 - a(b + z^{(1-\alpha)})}\right).
\end{aligned}
$$

So, $U_n \approx H^{-1}(\alpha_2)$ where $\alpha_2 = \Phi\left(b + \frac{b+z^{(1-\alpha)}}{1-a(b+z^{(1-\alpha)})}\right)$.

(d)

$$
\begin{aligned}
H(\hat{\theta}_n) &= P_{\hat{F}_n}(\hat{\theta}_{n*} \le \hat{\theta}_n) \\
&= P_{\hat{F}_n}(g(\hat{\theta}_{n*}) \le g(\hat{\theta}_n)) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_{n*}) - g(\hat{\theta}_n)}{\sigma_{g(\hat{\theta}_n)}} \le 0\right) \\
&\approx P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} \le 0\right) \\
&= P_{\hat{F}_n}\left(\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_{g(\theta)}} + b \le b\right) \\
&\approx G(b).
\end{aligned}
$$

$H(\hat{\theta}_n)$ can be calculated with bootstrap method. So, $b = G^{-1}(H(\hat{\theta}_n))$ where G^{-1} is known and exist, because it is a continuous distribution.

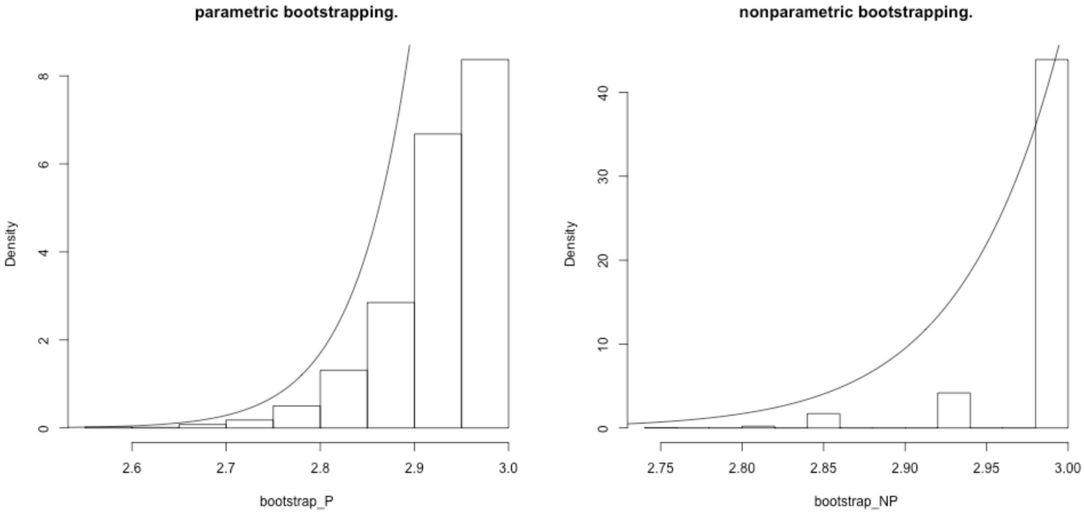

Figure 1: The histograms of $\hat{\theta}^*$ obtained from the parametric (left) and nonparametric (right) bootstraps.

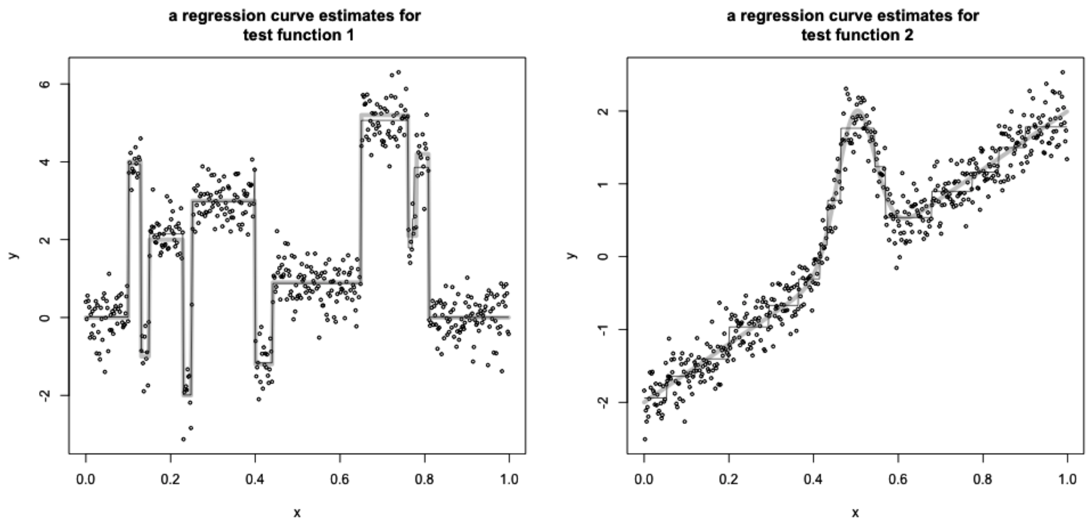

Figure 2: The regression curve estimates for both test functions. The grey curve is the true function. The black curve is the regression curve estimate.

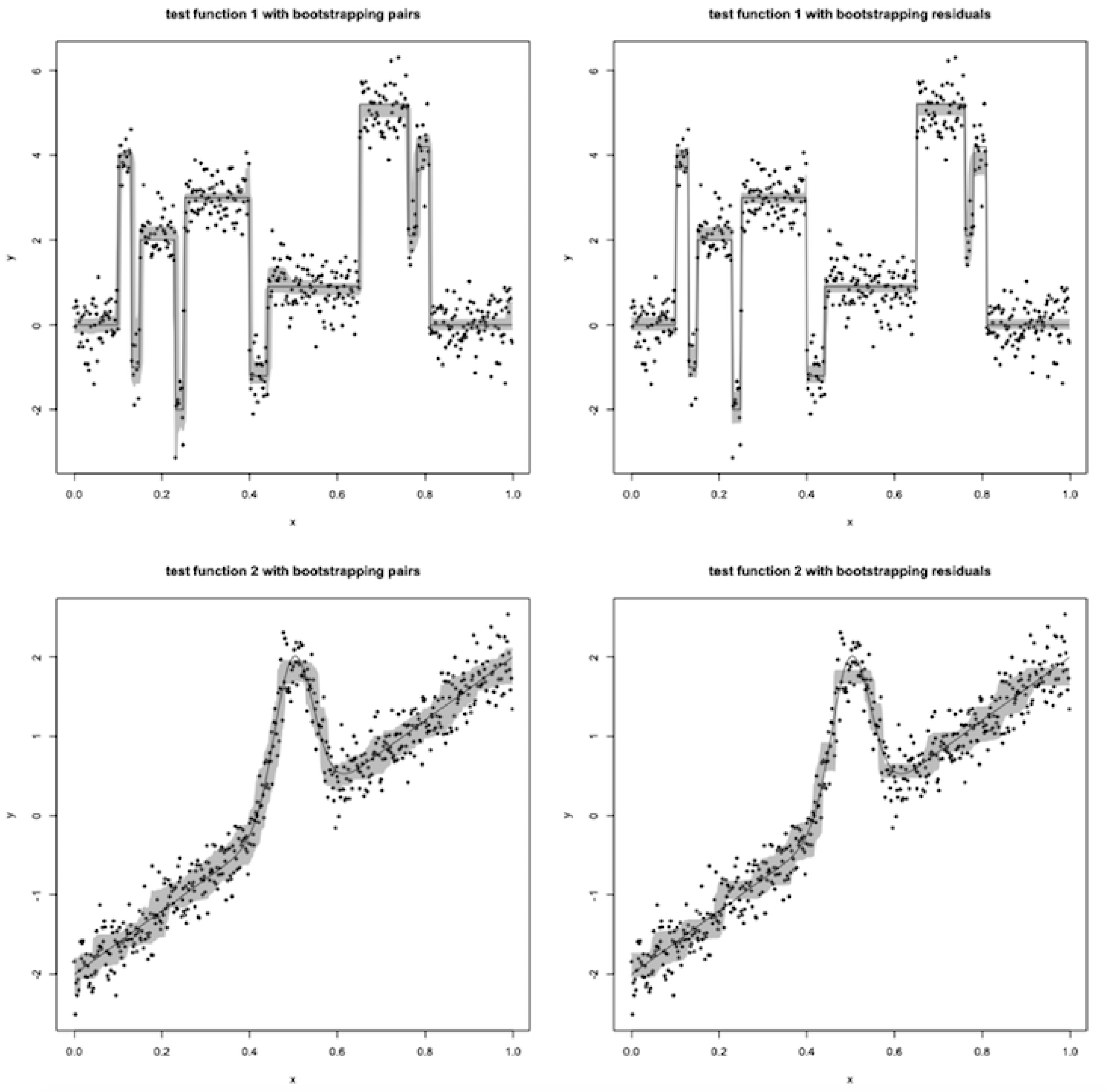

Figure 3: The test functions (black line) with the estimated confidence bands (grey area).