Till now, we studied iid settings.

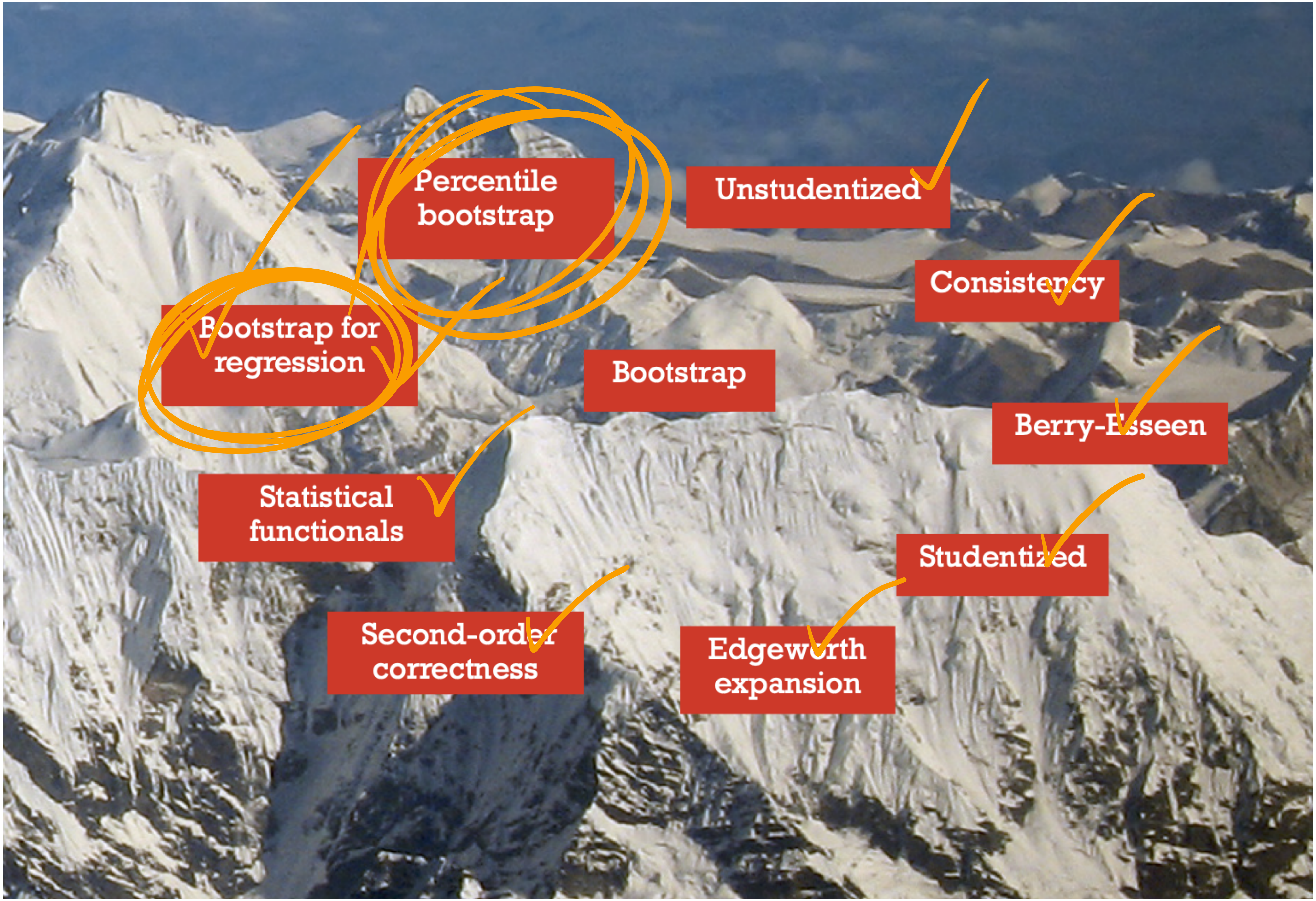
STAT 824 sp 2025 Lec 11 slides

Bootstrap for regression

Karl Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.


Percentile bootstrap
Unstudentized
Consistency
Bootstrap for regression
Bootstrap
Berry-Esseen
Statistical functionals
Studentized
Second-order correctness
Edgeworth expansion

Linear regression model

Let $(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$ be data pairs such that

could be non-normal

$$Y_i = \mathbf{x}_i^T \beta + \varepsilon_i, \quad i = 1, \dots, n,$$

with $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$ $\boxed{\text{deterministic,}}$ $\varepsilon_1, \dots, \varepsilon_n$ iid with $\boxed{\mathbb{E}\varepsilon_1 = 0,}$ $\boxed{\mathbb{E}\varepsilon_1^2 = \sigma^2} < \infty$.

Let $\underset{n\times p}{\mathbf{X}} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T$ and

E.g. $\underset{\sim}{c} = (0 \dots 1\ 0 \cdots 0)^T$, j^{th} position

$$\hat{\beta}_n = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

$$\hat{\sigma}_n^2 = (n-p)^{-1}\|\mathbf{Y} - \mathbf{X}\hat{\beta}_n\|_2^2$$

Then $\underset{\sim}{c}^T \underset{\sim}{\beta} = \beta_j$.

Let $\underset{\sim}{c} \in \mathbb{R}^p$ be a known vector. Estimate $\underset{\sim}{c}^T \underset{\sim}{\beta}$.

Exercise: Consider estimating contrasts $\mathbf{c}^T\beta$ for $\mathbf{c} \in \mathbb{R}^p$ with $\mathbf{c}^T\hat{\beta}_n$.

1. Come up with pivot quantities relevant for making inferences.
2. Give the form of a confidence interval for $\mathbf{c}^T\beta$.

Build C.I. for $\underset{\sim}{c}^T \underset{\sim}{\beta}$? Like $\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}} \pm z_{\alpha/2} \sqrt{\widehat{Var}(\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}})}$

We have

$$\begin{aligned} Var(\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}}) &= Var\left(\underset{\sim}{c}^T (X^T X)^{-1} X^T \underset{\sim}{Y} \right) \\ &= \underset{\sim}{c}^T (X^T X)^{-1} X^T \underbrace{Cov(\underset{\sim}{Y})}_{\sigma^2 I_n} X (X^T X)^{-1} \underset{\sim}{c} \\ &= \sigma^2 \, \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{c} \end{aligned}$$

Consider

$$Z_n = \frac{\underset{\sim}{c}^T \hat{\beta}_n - \underset{\sim}{c}^T \beta}{\sqrt{\sigma^2 \, \underset{\sim}{c}^T (X^T X) \underset{\sim}{c}}} \xrightarrow{D} N(0,1) \text{ as } n \to \infty.$$

(as long as $\max_{1 \le i \le n} h_{ii} \to 0$.)

Corollary to the Lindeberg CLT

For each $n \geq 1$, let $\xi_1, \ldots, \xi_n$ be iid with zero mean and unit variance and let $a_1, \ldots, a_n \in \mathbb{R}$. Then

$$\left(\sum_{i=1}^n a_i^2 \right)^{-1/2} \sum_{i=1}^n a_i \xi_i \xrightarrow{D} \mathcal{N}(0,1)$$

as $n \to \infty$ provided

$$\left(\sum_{j=1}^n a_j^2 \right)^{-1/2} \max_{1 \le i \le n} |a_i| \to 0$$

as $n \to \infty$.

$$\begin{aligned} \hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta} &= (X^T X)^{-1} X \underset{\sim}{Y} - \underset{\sim}{\beta} \\ &= (X^T X)^{-1} X (X \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}) - \underset{\sim}{\beta} \\ &= (X^T X)^{-1} X^T \underset{\sim}{\varepsilon} \end{aligned}$$

$$= \frac{\underset{\sim}{c}^T (X^T X)^{-1} X^T \underset{\sim}{\varepsilon}}{\sqrt{\sigma^2 \, \underset{\sim}{c}^T (X^T X) \underset{\sim}{c}}}$$

$$\underset{n \times p}{X} = \begin{bmatrix} \underset{\sim}{x}_1^T \\ \vdots \\ \underset{\sim}{x}_n^T \end{bmatrix} \qquad \frac{\sum_{i=1}^n a_i \xi_i}{\left(\sum_{j=1}^n a_j^2 \right)^{1/2}}$$

$$\underset{(n \times p)^T}{X^T} = [\underset{\sim}{x}_1 \quad \underset{\sim}{x}_n]$$

$$= \frac{\sum_{i=1}^n \overbrace{\boxed{\underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{x}_i}}^{a_i} (\varepsilon_i / \sigma)}{\sqrt{\underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{c}}}$$

$$\underset{p \times 1}{X^T \underset{\sim}{\varepsilon}} = [\underset{\sim}{x}_1 \cdots \underset{\sim}{x}_n] \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \underbrace{\sum_{i=1}^n \underset{\sim}{x}_i \varepsilon_i}_{p \times 1}.$$

zero mean, unit variance

$$= \frac{\sum_{i=1}^{n} a_i \boxed{(\varepsilon_i/\sigma)}}{\left(\sum_{i=1}^{n} a_i^2\right)^{1/2}},$$

Let $a_i = \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{x}_i$.

Then $\sum_{i=1}^{n} a_i^2 = \sum_{i=1}^{n} \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{x}_i \underset{\sim}{x}_i^T (X^T X)^{-1} \underset{\sim}{c}$

$= \underset{\sim}{c}^T (X^T X)^{-1} X^T X (X^T X)^{-1} \underset{\sim}{c}$

$= \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{c}$

where $a_i = \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{x}_i$.

$\xrightarrow{D} N(0,1)$

$\sum_{i=1}^{n} \underset{\sim}{x}_i \underset{\sim}{x}_i^T = X^T X$

provided

$$\frac{\max_{1 \le i \le n} |a_i|}{\left(\sum_{j=1}^{n} a_j^2\right)^{1/2}} \longrightarrow 0 \quad \text{as} \quad n \to \infty.$$

With

$$\frac{\max_{1 \le i \le n} |a_i|}{\left(\sum_{j=1}^{n} a_j^2\right)^{1/2}} = \frac{\max_{1 \le i \le n} \left| \underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{x}_i \right|}{\sqrt{\underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{c}}}$$

$\left| \underset{\sim}{c}^T (X^T X)^{-1/2} (X^T X)^{-1/2} \underset{\sim}{x}_i \right| = \left| \left[(X^T X)^{-1/2} \underset{\sim}{c}\right]^T \left[(X^T X)^{-1/2} \underset{\sim}{x}_i\right] \right|$

Cauchy-Schwarz

$\le \left\| (X^T X)^{-1/2} \underset{\sim}{c} \right\|_2 \left\| (X^T X)^{-1/2} \underset{\sim}{x}_i \right\|_2$

C-S

$|\underset{\sim}{a}^T \underset{\sim}{b}| \le \|\underset{\sim}{a}\|_2 \|\underset{\sim}{b}\|_2$

$$\le \max_{1 \le i \le n} \left\| (X^T X)^{-1/2} \underset{\sim}{x}_i \right\|_2$$

$$= \max_{1 \le i \le n} \sqrt{\underset{\sim}{x}_i^T (X^T X)^{-1} \underset{\sim}{x}_i} = \max_{1 \le i \le n} \sqrt{h_{ii}}.$$

Try: Let $h_{11}, \dots, h_{nn}$ be the diagonal entries of $X(X^T X)^{-1} X^T$.

Require $\max_{1 \le i \le n} h_{ii} \to 0$ as $n \to \infty$.

Could build $(1-\alpha)100\%$ C.I. for $\underset{\sim}{c}^T \underset{\sim}{\beta}$ as

$$P\left(-z_{\alpha/2} \le \frac{\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}} - \underset{\sim}{c}^T \underset{\sim}{\beta}}{\sqrt{\sigma^2 \underset{\sim}{c}^T (X^T X) \underset{\sim}{c}}} \le z_{\alpha/2} \right) \longrightarrow 1-\alpha$$

as $n \to \infty$.

So

$$\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}} \pm z_{\alpha/2} \, \sigma \sqrt{\underset{\sim}{c}^T (X^T X)^{-1} \underset{\sim}{c}}$$

replace with consistent estimator.

$n^{-1} X^T X = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \to \Sigma$.

Define the quantities

$\mathrm{Var}\, T_n \longrightarrow c^T \Sigma c$.

$$Q_n = \sqrt{n}\cdot \mathbf{c}^T(\hat{\beta}_n - \beta) \quad \text{and} \quad T_n = \sqrt{n}\cdot \mathbf{c}^T(\hat{\beta}_n - \beta)/\hat{\sigma}_n$$

$$\mathrm{Var}\, Q_n = n\,\mathrm{Var}(c^T\hat{\beta}_n) = n\sigma^2 c^T(X^TX)^{-1}c = \sigma^2 c^T(n^{-1}X^TX)^{-1}c \longrightarrow \sigma^2 c^T\Sigma c \text{ as } n\to\infty.$$

Theorem (Asymptotically Normal pivots for least-squares coefficients)

For any $\mathbf{c} \in \mathbb{R}^p$ *we have*

1. $[\mathbf{c}^T(n^{-1}\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}]^{\cancel{1/2}\;-1/2} Q_n \xrightarrow{D} \mathcal{N}(0, \sigma^2)$ *and*
2. $[\mathbf{c}^T(n^{-1}\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}]^{\cancel{1/2}\;-1/2} T_n \xrightarrow{D} \mathcal{N}(0, 1)$

as $n \to \infty$, *provided*

$$\max_{1\leq i\leq n} h_{ii} \to 0 \quad \text{as } n \to \infty,$$

where h_{ii}, $i = 1, \ldots, n$ *is the* i*th diagonal entry of* $\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$.

Exercise: Proof the result using the Corollary to the Lindeberg CLT on next slide.

Corollary to the Lindeberg CLT

For each $n \geq 1$, let $\xi_1, \dots, \xi_n$ be iid with zero mean and unit variance and let $a_1, \dots, a_n \in \mathbb{R}$. Then

$$\Big(\sum_{i=1}^n a_i^2\Big)^{-1/2} \sum_{i=1}^n a_i \xi_i \xrightarrow{\text{D}} \mathcal{N}(0,1)$$

as $n \to \infty$ provided

$$\Big(\sum_{j=1}^n a_j^2\Big)^{-1/2} \max_{1 \leq i \leq n} |a_i| \to 0$$

as $n \to \infty$.

$\underset{\sim}{Y}^* = (Y_1^*, \ldots, Y_n^*)^T$

Residual bootstrap for linear regression

1. Draw $\varepsilon_1^*, \ldots, \varepsilon_n^*$ with repl. from $\hat{\varepsilon}_i = Y_i - \mathbf{x}_i^T \hat{\beta}_n$, $i = 1, \ldots, n$
2. Set $Y_i^* = \mathbf{x}_i^T \hat{\beta}_n + \varepsilon_i^*$ for $i = 1, \ldots, n$.
3. Compute $\hat{\beta}_n^* = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}^*$ and $(\hat{\sigma}_n^*)^2 = (n-p)^{-1}\|\mathbf{Y}^* - \mathbf{X}\hat{\beta}_n^*\|_2^2$.
4. Compute the bootstrap versions of Q_n and T_n given by

$$Q_n^* = \sqrt{n} \cdot \mathbf{c}^T(\hat{\beta}_n^* - \hat{\beta}_n) \quad \text{and} \quad T_n^* = \sqrt{n} \cdot \mathbf{c}^T(\hat{\beta}_n^* - \hat{\beta}_n)/\hat{\sigma}_n^*.$$

If $T_n \sim G_{T_n}$, $P\left(G_{T_n}(\alpha/2) \leq \sqrt{n}\,\dfrac{\underset{\sim}{c}^T(\hat{\beta}_n - \beta)}{\hat{\sigma}_n} \leq G_{T_n}(1-\alpha/2) \right) = 1-\alpha.$

Then

$$\left[\underset{\sim}{c}^T \hat{\beta}_n - G_{T_n}(1-\alpha/2) \frac{\hat{\sigma}_n}{\sqrt{n}} \,,\; \underset{\sim}{c}^T \hat{\beta}_n - G_{T_n}(\alpha/2) \frac{\hat{\sigma}_n}{\sqrt{n}} \right]$$

is a $(1-\alpha)100\%$ C.I. for $\underset{\sim}{c}^T \underset{\sim}{\beta}$

Let $\hat{G}_{T_n}(x) = P_*\left(T_n^* \le x\right)$

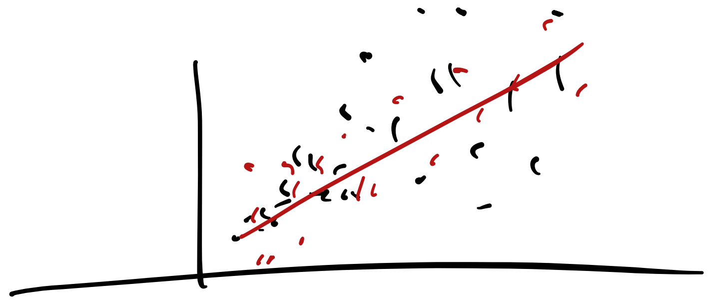
$Q_n = \sqrt{n}\ \underset{\sim}{c}^T(\hat{\beta}_n - \underset{\sim}{\beta})$ $\quad T_n = \sqrt{n}\ \underset{\sim}{c}^T(\hat{\beta}_n - \beta)/\hat{\sigma}_n$

Monte Carlo implementation of residual bootstrap

Given MC realizations $Q^{*(1)} \leq \cdots \leq Q_n^{*(B)}$ of Q_n^* and $T^{*(1)} \leq \cdots \leq T_n^{*(B)}$ of T_n^*, B large, $(1-\alpha)100\%$ bootstrap CIs for $\mathbf{c}^T\beta$ based on Q_n and T_n are

1. $\left[\mathbf{c}^T\hat{\beta}_n - Q_n^{*(\lceil(\alpha/2)B\rceil)}n^{-1/2},\ \mathbf{c}^T\hat{\beta}_n - Q_n^{*(\lceil(1-\alpha/2)B\rceil)}n^{-1/2}\right]$
2. $\left[\mathbf{c}^T\hat{\beta}_n - T_n^{*(\lceil(\alpha/2)B\rceil)}n^{-1/2}\hat{\sigma}_n,\ \mathbf{c}^T\hat{\beta}_n - T_n^{*(\lceil(1-\alpha/2)B\rceil)}n^{-1/2}\hat{\sigma}_n\right]$

Exercise: Simulate performance of residual bootstrap CIs for (i) $\mathbf{c}^T\beta_0$ and (ii) β_{0j}.



Linear regression model with heteroscedasticity

Let $(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$ be data pairs such that

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta}_0 + \varepsilon_i, \quad i = 1, \dots, n,$$

with $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$ deterministic, $\mathbb{E}\varepsilon_i = 0$ and $\mathbb{E}\varepsilon_i^2 = \sigma_i^2 \in (0, \infty)$, $i = 1, \dots, n$.

Exercise: Give an expression for $\mathrm{Var}(\mathbf{c}^T \hat{\boldsymbol{\beta}}_n)$ for $\mathbf{c} \in \mathbb{R}^p$.

$$\mathrm{Var}(\underset{\sim}{c}^T \hat{\underset{\sim}{\beta}}_n) = \mathrm{Var}\left(\underset{\sim}{c}^T (X^T X)^{-1} X^T \underset{\sim}{Y}\right)$$

$$= \underset{\sim}{c}^T (X^T X)^{-1} X^T (\mathrm{Cov}\, \underset{\sim}{Y}) X (X^T X)^{-1} \underset{\sim}{c}$$

$$= \underset{\sim}{c}^T (X^T X)^{-1} X^T \mathrm{diag}(\sigma_1^2, \dots, \sigma_n^2) X (X^T X)^{-1} \underset{\sim}{c}$$

Define $\hat{\sigma}^2_{\mathbf{c},n} = n \cdot \mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \cdot \text{diag}(\hat{\varepsilon}_1^2, \dots, \hat{\varepsilon}_n^2) \cdot \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}$ and let

$$Q_n = \sqrt{n} \cdot \mathbf{c}^T(\hat{\beta}_n - \beta) \quad \text{and} \quad H_n = \sqrt{n} \cdot \mathbf{c}^T(\hat{\beta}_n - \beta_0)/\hat{\sigma}_{c,n}$$

Theorem (Asymptotically Normal pivots for LS coefs under hetsc.)

Let $\sigma^2_{\mathbf{c},n} = n\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\sigma_1^2, \dots, \sigma_n^2)\,\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}$ and assume $\sigma^2_{\mathbf{c},n} \to \sigma^2_{\mathbf{c}}$ as $n \to \infty$ for some $\sigma^2_{\mathbf{c}} \in (0, \infty)$. Then we have

1. $Q_n \xrightarrow{D} \mathcal{N}(0, \sigma^2_{\mathbf{c}})$
2. $H_n \xrightarrow{D} \mathcal{N}(0, 1)$ $\Longrightarrow$ An asymp. $(1-\alpha)100\%$ C.I. for $\underset{\sim}{c}^T\underset{\sim}{\beta}$ is

$$\underset{\sim}{c}^T\hat{\underset{\sim}{\beta}} \pm z_{\alpha/2}\frac{\hat{\sigma}_{c,n}}{\sqrt{n}}$$

as $n \to \infty$ provided

$$\max_{1 \le i \le n} h^{\sigma}_{ii}/\sigma_i^2 \to 0$$

as $n \to \infty$, where h^{σ}_{ii} is diagonal entry i of the matrix $\mathbf{X}(\mathbf{X}^T(\sigma_1^2, \dots, \sigma_n^2)\,\mathbf{X})\mathbf{X}^T$ and under some additional moment conditions (see [5]).

Wild bootstrap for linear regression

1. Generate indep. bootstrap ~~residuals~~ error terms $\varepsilon_1^{*W}, \dots, \varepsilon_n^{*W}$ satisfying $\boxed{\mathbb{E}_*[\varepsilon_i^{*W}] = 0}$, $\boxed{\mathbb{E}_*[(\varepsilon_i^{*W})^2] = \hat{\varepsilon}_i^2}$, and $\mathbb{E}_*[(\varepsilon_i^{*W})^3] = \hat{\varepsilon}_i^3$, where $\hat{\varepsilon}_i = Y_i - \mathbf{x}_i^T\hat{\beta}_n$, $i = 1, \dots, n$.
2. Set $Y_i^{*W} = \mathbf{x}_i^T\hat{\beta}_n + \varepsilon_i^*$, $i = 1, \dots, n$.
3. Compute $\hat{\beta}_n^{*W} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}^{*W}$, $\hat{\varepsilon}_i^{*W} = Y_i^{*W} - \mathbf{x}_i^T\hat{\beta}_n^{*W}$, $i = 1, \dots, n$.
4. Compute wild bootstrap versions of the pivots Q_n and H_n given by

$$Q_n^{*W} = \sqrt{n}\cdot\mathbf{c}^T(\hat{\beta}_n^{*W} - \hat{\beta}_n) \quad \text{and} \quad H_n^{*W} = \sqrt{n}\cdot\mathbf{c}^T(\hat{\beta}_n^{*W} - \hat{\beta}_n)/\hat{\sigma}_{\mathbf{c},n}^{*W},$$

where $(\hat{\sigma}_{\mathbf{c},n}^{*W})^2 = n\cdot\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\cdot\operatorname{diag}((\hat{\varepsilon}_1^{*W})^2, \dots, (\hat{\varepsilon}_n^{*W})^2)\cdot\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}$.

Monte Carlo implementation of wild bootstrap

Given MC realizns $Q^{*W(1)} \leq \cdots \leq Q_n^{*W(B)}$ of Q_n^{*W} and $H^{*W(1)} \leq \cdots \leq H_n^{*W(B)}$ of H_n^{*W}, B large, $(1-\alpha)100\%$ bootstrap CIs for $\mathbf{c}^T\beta$ based on Q_n and H_n are

1. $\left[\mathbf{c}^T\hat{\beta}_n - Q_n^{*W(\lceil(\alpha/2)B\rceil)} n^{-1/2},\ \mathbf{c}^T\hat{\beta}_n - Q_n^{*W(\lceil(1-\alpha/2)B\rceil)} n^{-1/2}\right]$
2. $\left[\mathbf{c}^T\hat{\beta}_n - H_n^{*W(\lceil(\alpha/2)B\rceil)} n^{-1/2}\hat{\sigma}_{\mathbf{c},n},\ \mathbf{c}^T\hat{\beta}_n - H_n^{*W(\lceil(1-\alpha/2)B\rceil)} n^{-1/2}\hat{\sigma}_{\mathbf{c},n}\right]$

Two ways to obtain wild bootstrap ~~residuals~~ error terms [1], [2]

1. Mammen (1993): For $i = 1, \dots, n$, get $V_{i,1}, V_{i,2} \stackrel{\text{ind}}{\sim} \text{Normal}(0, 1)$. Then set
$$U_i = (\delta_1 + V_{i,1}/\sqrt{2})(\delta_2 + V_{i,2}/\sqrt{2}) - \delta_1\delta_2,$$
where $\delta_1 = (3/4 + \sqrt{17}/12)^{1/2}$, $\delta_2 = (3/4 - \sqrt{17}/12)^{1/2}$. Then let
$$\varepsilon_i^{*W} = \hat{\varepsilon}_i \cdot U_i.$$
2. Das et al. (2019): For $i = 1, \dots, n$, generate $U_i \sim \text{Beta}(1/2, 3/2)$. Then set
$$\varepsilon_i^{*W} = \hat{\varepsilon}_i \cdot 4(U_i - 1/4).$$

Exercise: Simulate performance of wild bootstrap CIs for (i) $\mathbf{c}^T\boldsymbol{\beta}_0$ and (ii) β_{0j}.

Discuss: Comparison of methods in simulation.

```
r <- .7
R <- r^abs( outer(1:8,1:8,"-"))
P <- 2*sin( R * pi / 6)
X <- cbind(1,(pnorm( matrix(rnorm(n*8),ncol = 8) %*% chol(P)) - .5) * 5)
beta <- c(-1,c(4:1)*(-1)^(4:1),0,0,0,0)
sigma <- 1/4 + abs(X[,2] + 2.5)^2
error <- rnorm(n,0,sigma)
Y <- as.numeric(X %*% beta) + error
```

depends on x_2

$\sigma_i^2 = \sigma^2 \; \forall i$

Coverage of 95% confidence intervals for β_3 at sample sizes $n = 10, 20, \dots, 100$.

C.I. for β_3

Too high.

method	10	20	30	40	50	60	70	80	90	100
Q_n^*	0.34	0.81	0.88	0.91	0.93	0.95	0.93	0.95	0.95	0.95
T_n^*	0.98	0.99	0.98	0.97	0.97	0.97	0.97	0.98	0.97	0.96
T_n as $N(0,1)$	0.66	0.92	0.94	0.96	0.95	0.97	0.96	0.97	0.96	0.95
Q_n^{*W}	0.31	0.79	0.87	0.90	0.93	0.95	0.93	0.95	0.93	0.93
H_n^{*W}	0.90	0.91	0.92	0.93	0.93	0.95	0.94	0.94	0.94	0.93
H_n as $N(0,1)$	0.30	0.79	0.86	0.90	0.93	0.94	0.93	0.93	0.93	0.93

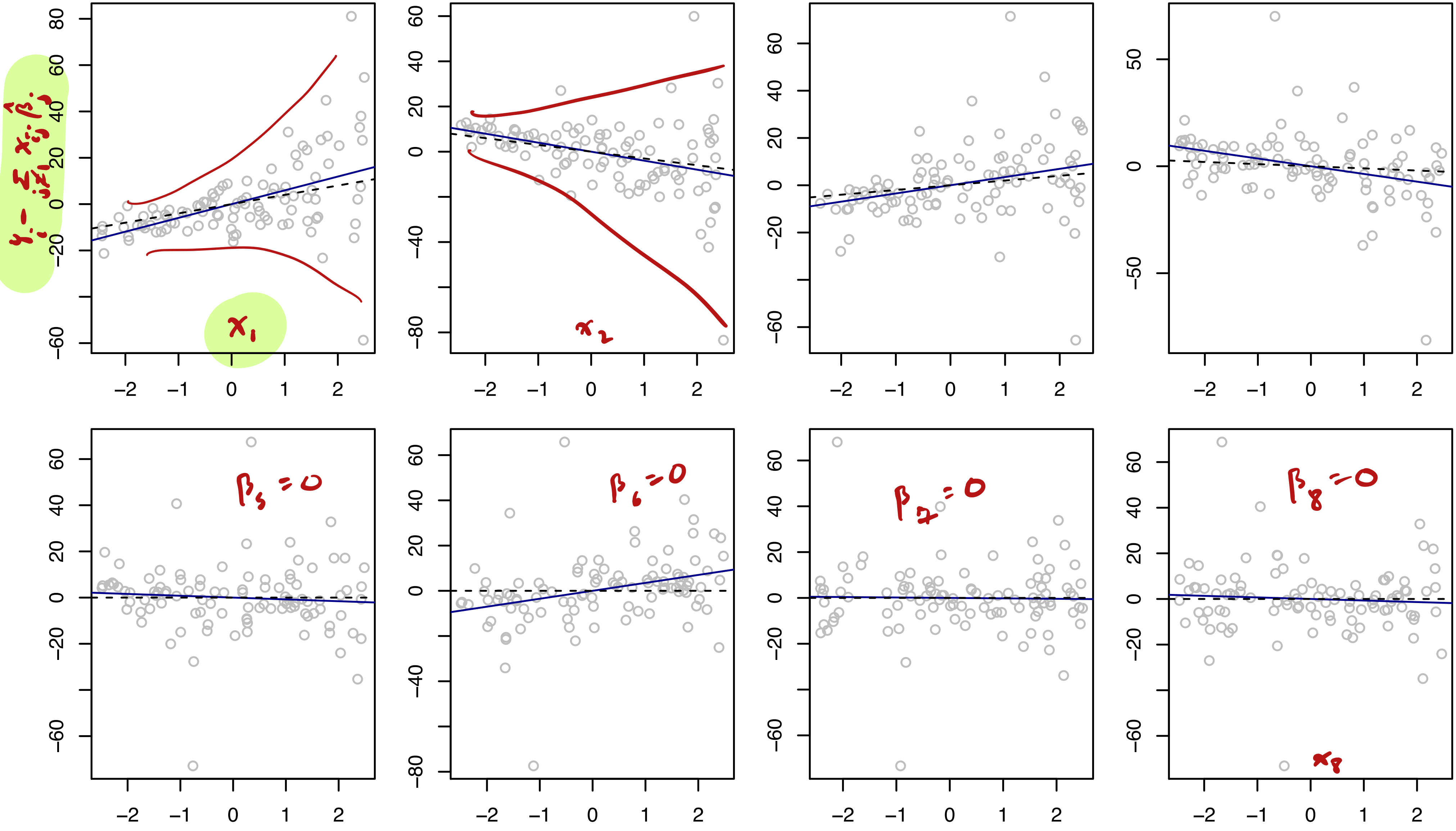
Asymptotic, assumes

Residual boot

Wild

Asymptotic, accommodates $\sigma_i^2, i=1,\dots,n$

$p = 8$


$y_i - \sum_{j \neq 1} x_{ij}\hat{\beta}_j$
x_1
x_2
$\beta_5 = 0$
$\beta_6 = 0$
$\beta_7 = 0$
$\beta_8 = 0$
x_8

Linear regression model with heteroscedasticity and a random design

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be data pairs such that

$$Y_i = X_i^T \beta_0 + \varepsilon_i, \quad i = 1, \dots, n,$$

with $X_1, \dots, X_n \in \mathbb{R}^p$ rvs, $\mathbb{E}[\varepsilon_i | X_i] = 0$ and $\mathbb{E}[\varepsilon_i^2 | X_i] = \sigma_i^2 \in (0, \infty)$, $i = 1, \dots, n$.

A random design is often more realistic (but does it really matter?).

Mammen (1993) showed that the wild bootstrap works in the above setting.

The resampling pairs bootstrap

1. Draw $(X_1^*, Y_1^*), \dots, (X_n^*, Y_n^*)$ with replacement from $(X_1, Y_1), \dots, (X_n, Y_n)$.
2. Then let $\hat{\beta}_n^* = \boxed{(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}\mathbf{Y}^*.}$

Note that we must compute the inverse $(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}$ for every bootstrap resample!

Random X

This is taken from Mammen (1993). "Bootstrap" is the resampling pairs bootstrap.

TABLE 1

Rates of convergence of the bootstrap procedures and the mean zero normal approximation under the assumption $E(\varepsilon_i|X_i) = 0$

Estimation of	$\mathscr{L}(\overbrace{\boxed{\sqrt{n}\,c^T(\hat\beta - \beta)}}^{Q_n})$	$\mathscr{L}(\overbrace{\sqrt{n}\,c^T(\hat\beta - \beta)/\hat\sigma_c}^{H_n})$
Normal approximation $N(0, \hat\sigma_c^2)$	$O_P(n^{-1/2} + pn^{-1})$	
Wild bootstrap	$O_P(n^{-1/2} + pn^{-1})$	$O_P(n^{-1} + pn^{-3/2})$
Bootstrap ← Resample pairs (X_i^*, y_i^*)	$O_P(n^{-1/2} + pn^{-1})$	$O_P(pn^{-1})$

Wild is better

In this paper the affect of the dimension p is tracked along with that of n.

The studentized resampling pairs bootstrap is more adversely affected by high dimension than the wild bootstrap.

Nonparametric regression model

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be data pairs such that

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, \ldots, n,$$

with $X_1, \ldots, X_n \in [0, 1]$ deterministic, $\mathbb{E}\varepsilon_i = 0$, $\mathbb{E}\varepsilon_i^2 = \sigma_i^2 \in (0, \infty)$, $i = 1, \ldots, n$.

Consider *linear estimators*, i.e. estimators of the form

$$\hat{m}_n^{NW}(x) = \frac{\sum_{i=1}^n K\left(\frac{X_i - x}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)} = \sum_{i=1}^n W_{ni}(x) Y_i,$$

$$\hat{m}_n(x) = \sum_{i=1}^n W_{ni}(x) Y_i, \quad \text{for } x \in [0, 1].$$

$$W_{ni}(x) = \frac{K\left(\frac{X_i - x}{h}\right)}{\sum_{j=1}^n K\left(\frac{X_j - x}{h}\right)}$$

Exercise: Discuss estimators of $\mathrm{Var}\, \hat{m}_n(x)$ in the cases

1. $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2 \in (0, \infty)$.
2. $\sigma_1^2, \ldots, \sigma_n^2$ are heteroscedastic.

Build C.I. for $m(x)$ at one point x?

We have

$$\text{Var}\, \hat{m}_n(x) = \text{Var}\left(\sum_{i=1}^n W_{ni}(x) Y_i \right) = \sum_{i=1}^n W_{ni}^2(x)\, \sigma_i^2$$

Consider a pivotal quantity:

$$\frac{\hat{m}_n(x) - m(x)}{SE\{\hat{m}_n(x)\}} = \frac{\hat{m}_n(x) - m(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\, \sigma_i^2}}$$

$$= \frac{\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\, \sigma_i^2}} + \frac{\mathbb{E}\,\hat{m}_n(x) - m(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\, \sigma_i^2}}$$

$$\mathbb{E}\,\hat{m}_n(x) = \mathbb{E} \sum_{i=1}^n W_{ni}(x) Y_i$$

$$= \sum_{i=1}^n W_{ni}(x)\, m(X_i)$$

$$= \underbrace{\frac{\sum_{i=1}^n W_{ni}(x)\, \sigma_i \left(\varepsilon_i / \sigma_i\right)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\, \sigma_i^2}}}_{\xrightarrow{D} N(0,1) \text{ under mild conditions}} + \frac{\text{Bias}}{\sqrt{\text{Var}}}$$

$$\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)$$

$$= \sum W_{ni}(x) Y_i - \sum W_{ni}(x)\, m(X_i)$$

$$= \sum W_{ni}(x)\, \varepsilon_i$$

$\longrightarrow 0$ if we undersmooth.

(Choose bandwidth smaller than MSE-optimal)

Suppose we wish to build a confidence interval for $m(x)$ at some $x \in [0,1]$.

Consider our discussions from Lecture 4: We have

$$\frac{\hat{m}_n(x) - m_n(x)}{\sqrt{\operatorname{Var} \hat{m}_n(x)}} = \underbrace{\frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\sqrt{\operatorname{Var} \hat{m}_n(x)}}}_{\to^D N(0,1)} + \underbrace{\frac{\mathbb{E}\hat{m}_n(x) - m_n(x)}{\sqrt{\operatorname{Var} \hat{m}_n(x)}}}_{\to^p \text{ 0 if } \hat{m}_n \text{ undersmoothed}} .$$

Strategy: Undersmooth and pretend $\mathbb{E}\hat{m}_n(x)$ is equal to $m(x)$.

Studentized pivots under constant and heteroscedastic variances

$$T_{n,x} = \frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\hat{\sigma}_n \sqrt{\sum_{i=1}^n W_{ni}^2(x)}} \to N(0,1) \quad \text{and} \quad T_{n,x}^{\text{het}} = \frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\hat{\varepsilon}_i^2}} \to N(0,1)$$

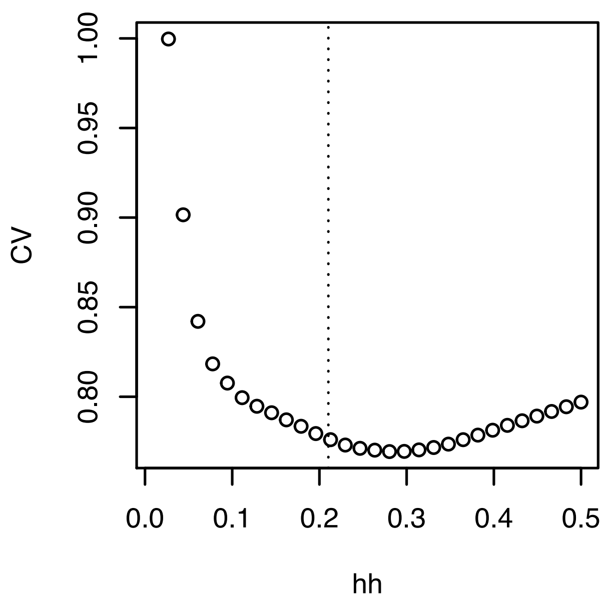
$\sigma_1^2 = \dots = \sigma_n^2 = \sigma^2$

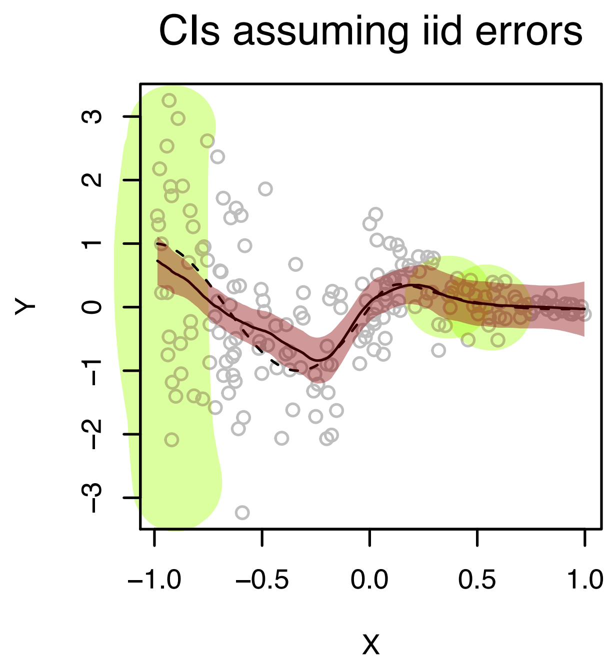
$\sigma_1^2, \dots, \sigma_n^2$ don't need to be same

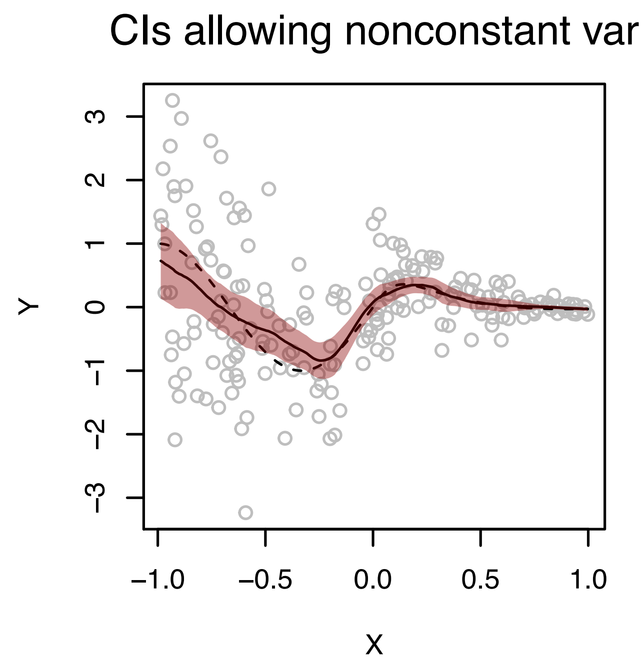
```
m <- function(x){sin(3* pi * x / 2)/(1 + 18 * x^2*(sign(x) + 1))}
n <- 200
X <- runif(n,-1,1)
sigma <- (1.5 - X)^2/4
Y <- m(X) + rnorm(n,0,sigma)
```

The asymptotic Normality of the pivots suggests the pointwise CIs

$$\left[\hat{m}_n(x) \pm z_{\alpha/2}\hat{\sigma}_n\sqrt{\textstyle\sum_{i=1}^n W_{ni}^2(x)}\right] \quad \text{and} \quad \left[\hat{m}_n(x) \pm z_{\alpha/2}\sqrt{\textstyle\sum_{i=1}^n W_{ni}^2(x)\hat{\varepsilon}_i^2}\right].$$







Consider the constant variances case $\boxed{\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2 \in (0, \infty)}$

Assume $\hat{m}_n$ is undersmoothed and pretend $\mathbb{E}\hat{m}_n(x)$ equals $m(x)$ (i.e. ignore bias).

A pivot for building confidence bands

Define

$$T_n = \sup_{x \in [0,1]} \left| \frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\hat{\sigma}_n \sqrt{\sum_{i=1}^n W_{ni}^2(x)}} \right|$$

and denote by G_n the distribution of T_n.

Exercise: Propose a confidence band for $m(x)$, $x \in [0, 1]$ of the form

$$\mathcal{B}_{n,\alpha} = \{(x, y) : l_{n,\alpha}(x) \leq y \leq u_{n,\alpha}(x), x \in [0, 1]\}$$

assuming the distribution G_n is known.

$$P\left(\underset{\substack{\uparrow \\ 1-\alpha \text{ quantile of } T_n}}{T_n \le G_n^{-1}(1-\alpha)} \right) = 1-\alpha$$

$\iff$

$$P\left(\sup_{x\in[0,1]} \left| \frac{\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)}{\sqrt{\hat{\sigma}_n^2 \sum_{i=1}^n W_{ni}^2(x)}} \right| \le G_n^{-1}(1-\alpha) \right) = 1-\alpha$$

$\iff$

$$P\left(\bigcap_{x\in[0,1]} \left\{ \left| \frac{\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)}{\sqrt{\hat{\sigma}_n^2 \sum_{i=1}^n W_{ni}^2(x)}} \right| \le G_n^{-1}(1-\alpha) \right\} \right) = 1-\alpha$$

$\iff$

$$P\left(\bigcap_{x\in[0,1]} \left\{ \left| \hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x) \right| \le \sqrt{\hat{\sigma}^2 \sum_{i=1}^n W_{ni}^2(x)} \; G_n^{-1}(1-\alpha) \right\} \right) = 1-\alpha$$

$\iff$

$$P\left(\bigcap_{x\in[0,1]} \left\{ \mathbb{E}\,\hat{m}_n(x) \in \left[\hat{m}_n(x) \pm \hat{\sigma} \sqrt{\sum_{i=1}^n W_{ni}^2(x)} \; G_n^{-1}(1-\alpha) \right] \right\} \right) = 1-\alpha$$

So an asymptotic $(1-\alpha)$ Conf. Band is

replace with c_α

$$[L(x), U(x)] = \left[\hat{m}_n(x) \pm \hat{\sigma} \sqrt{\sum_{i=1}^n W_{ni}^2(x)} \; G_n^{-1}(1-\alpha) \right], \quad x \in [0,1]$$

Asymptotic result of Sun and Loader (1994) [3], [4].

For $Z_1, \dots, Z_n \overset{\text{ind}}{\sim} \text{Normal}(0,1)$, we have

$$P\left(\sup_{x\in[0,1]}\left|\sum_{i=1}^n M_{ni}(x)Z_i\right| > c\right) \approx 2(1-\Phi(c)) + \frac{\kappa_0}{\pi}e^{-c^2/2}$$

for large c and large enough n, where $\kappa_0 = \int_0^1 \sqrt{\sum_{i=1}^n \left[\frac{\partial}{\partial x}M_{ni}(x)\right]^2}dx$.

The above result is sometimes called a *tube formula*.

Exercise:

1. Find $M_{ni}(x)$, $i = 1, \dots, n$, such that we may write

$$T_n = \sup_{x\in[0,1]}\left|\sum_{i=1}^n M_{ni}(x)(\varepsilon_i/\hat{\sigma}_n)\right|$$

2. Propose a conf. band for $m(x), x \in [0,1]$ based on the above result.

$$T_n = \sup_{x \in [0,1]} \left| \frac{\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)}{\hat{\sigma}_n \sqrt{\sum_{i=1}^n W_{ni}^2(x)}} \right|$$

$$= \sup_{x \in [0,1]} \left| \frac{\sum_{i=1}^n W_{ni}(x)\,(\varepsilon_i/\sigma)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)}} \right| \underbrace{\left(\frac{\sigma}{\hat{\sigma}_n}\right)}_{\xrightarrow{P} 1}$$

$$= \sup_{x \in [0,1]} \left| \sum_{i=1}^n M_{ni}(x)\,(\varepsilon_i/\sigma) \right| \left(\frac{\sigma}{\hat{\sigma}_n}\right),$$

where

$$M_{ni}(x) = \frac{W_{ni}(x)}{\sqrt{\sum_{j=1}^n W_{nj}^2(x)}}$$

$$P\left(T_n > c\right) \approx 2\left(1 - \Phi(c)\right) + \frac{\kappa_0}{\pi} e^{-c^2/2} \overset{\text{set}}{=} \alpha,$$

solve for c.

$$\kappa_0 = \int_0^1 \sqrt{\sum_{i=1}^n \left[\frac{\partial}{\partial x} M_{ni}(x)\right]^2}\, dx$$

Can compute from $W_{ni}(x)$, $i=1,\dots,n$.

$G_n(1-\alpha) \approx c_\alpha$ which solves

Approximation to κ_0

To compute κ_0 in practice is hard. We can get an approximation to it as

$$\kappa_0 \approx \sum_{j=1}^{N-1} \sqrt{\sum_{i=1}^{n} [M_{ni}(x_{j+1}) - M_{ni}(x_j)]^2}$$

given a grid of values $x_1, \dots, x_N \in [0, 1]$, for some large N.

Exercise:

1. Justify the above approximation to κ_0.
2. Construct the asymptotic $(1-\alpha)100\%$ confidence band

$$\left\{(x, y) : y \in \left[\ \hat{m}_n(x) \pm c_\alpha \hat{\sigma}_n \sqrt{\textstyle\sum_{i=1}^{n} W_{ni}^2(x)}\ \right], \quad x \in [0, 1]\right\}$$

on a simulated data set, where c_α is from the Sun and Loader method.

Consider the case in which $\sigma_1^2, \dots, \sigma_n^2 \in (0, \infty)$ are heteroscedastic.

Assume $\hat{m}_n$ is undersmoothed and pretend $\mathbb{E}\hat{m}_n(x)$ equals $m(x)$ (i.e. ignore bias).

A pivot for building confidence bands under heteroscedasticity

Define

$$T_n^{\text{het}} = \sup_{x \in [0,1]} \left| \frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x) \hat{\varepsilon}_i^2}} \right|$$

and denote by G_n^{het} the distribution of T_n^{het}.

Exercise: Propose a confidence band for $m(x)$, $x \in [0,1]$ of the form

$$\mathcal{B}_{n,\alpha}^{\text{het}} = \{(x, y) : l_{n,\alpha}^{\text{het}}(x) \leq y \leq u_{n,\alpha}^{\text{het}}(x), x \in [0,1]\}$$

assuming the distribution G_n^{het} is known.

$$\frac{\hat{m}_n(x) - \mathbb{E}\hat{m}_n(x)}{\sqrt{\operatorname{Var} \hat{m}_n(x)}} = \frac{\sum_{i=1}^{n} W_{ni}(x)\, \sigma_i \,(\varepsilon_i/\sigma_i)}{\sqrt{\sum_{i=1}^{n} W_{ni}^2(x)\, \sigma_i^2}}$$

$$= \sum_{i=1}^{n} \bar{M}_{ni}(x)\, (\varepsilon_i/\sigma_i),$$

$$\bar{M}_{ni}(x) = \frac{W_{ni}(x)\, \sigma_i}{\sqrt{\sum_{i=1}^{n} W_{ni}^2(x)\, \sigma_i^2}}.$$

$\Rightarrow$ Replace $\bar{M}_{ni}(x)$ with $\hat{M}_{ni}(x) = \dfrac{W_{ni}(x)\, |\hat{\varepsilon}_i|}{\sqrt{\sum_{i=1}^{n} W_{ni}^2(x)\, \hat{\varepsilon}_i^2}}$

Proceed as before.

Exercise:

1. To build a conf. band à la Sun and Loader, we need $\bar{M}_{ni}(x)$ such that

$$T_n^{\text{het}} = \sup_{x\in[0,1]} \left| \sum_{i=1}^n \bar{M}_{ni}(x) Z_i \right|$$

for some rvs $Z_1, \dots, Z_n \overset{\text{ind}}{\sim} \text{Normal}(0,1)$. Find $\bar{M}_{ni}(x)$, $i = 1, \dots, n$.

2. Construct the asymptotic $(1-\alpha)100\%$ confidence band

$$\left\{(x,y) : y \in \left[\ \hat{m}_n(x) \pm c_\alpha \sqrt{\sum_{i=1}^n W_{ni}^2(x)\hat{\varepsilon}_i^2}\ \right],\ x \in [0,1]\right\}$$

on a simulated data set, where c_α is from the S&L method with κ_0 based on

$$\hat{M}_{ni}(x) = \frac{W_{ni}(x)|\hat{\varepsilon}_i|}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\hat{\varepsilon}_i^2}}.$$

Consider the constant variances case $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2 \in (0, \infty)$.

Residual bootstrap for nonparametric regression

1. Draw $\varepsilon_1^*, \dots, \varepsilon_n^*$ w/repl from $\hat{\varepsilon}_i = Y_i - \hat{m}_n(X_i)$, $i = 1, \dots, n$.
2. Set $Y_i^* = \boxed{\hat{m}_n(X_i)} + \varepsilon_i^*$, $i = 1, \dots, n$.
3. Compute $\hat{m}_n^*(x) = \sum_{i=1}^n W_{ni}(x) Y_i^*$ and $\hat{\sigma}_n^*$ based on $(X_1, Y_1^*), \dots, (X_n, Y_n^*)$.
4. Compute bootstrap version of $T_{n,x}$ given by

$$T_{n,x} = \frac{\hat{m}_n(x) - \mathbb{E}\,\hat{m}_n(x)}{\hat{\sigma}_n \sqrt{\sum_{i=1}^n W_{ni}^2(x)}}$$

$$T_{n,x}^* = \frac{\hat{m}_n^*(x) - \mathbb{E}_* \hat{m}_n^*(x)}{\hat{\sigma}_n^* \sqrt{\sum_{i=1}^n W_{ni}^2(x)}}.$$

$$\mathbb{E}_* \hat{m}_n^*(x) = \sum_{i=1}^n W_{ni}(x)\hat{m}_n(X_i)$$

$$\hat{m}_n^*(x) = \sum_{i=1}^n W_{ni}(x) Y_i^*$$

Exercise: Show that $\hat{m}_n^*(x) - \mathbb{E}_* \hat{m}_n^*(x) = \sum_{i=1}^n W_{ni}(x)\varepsilon_i^*$, provided $\sum_{i=1}^n \hat{\varepsilon}_i = 0$.

Consider the case in which $\sigma_1^2, \dots, \sigma_n^2 \in (0, \infty)$ are heteroscedastic.

Wild bootstrap for nonparametric regression

1. Generate indep. bootstrap error terms ~~residuals~~ $\varepsilon_1^{*W}, \dots, \varepsilon_n^{*W}$ satisfying $\mathbb{E}_*[\varepsilon_i^{*W}] = 0$, $\mathbb{E}_*[(\varepsilon_i^{*W})^2] = \hat{\varepsilon}_i^2$, and $\mathbb{E}_*[(\varepsilon_i^{*W})^3] = \hat{\varepsilon}_i^3$, where $\hat{\varepsilon}_i = Y_i - \hat{m}_n(X_i)$, $i = 1, \dots, n$.
2. Set $Y_i^{*W} = \hat{m}_n(X_i) + \varepsilon_i^{*W}$, $i = 1, \dots, n$.
3. Compute $\hat{m}_n^{*W}(x) = \sum_{i=1}^n W_{ni}(x) Y_i^{*W}$.
4. Compute bootstrap version of $T_{n,x}^{\text{het}}$ given by

$$T_{n,x}^{\text{het}*} = \frac{\hat{m}_n^{*W}(x) - \mathbb{E}_* \hat{m}_n^{*W}(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)(\hat{\varepsilon}_i^{*W})^2}}$$

$$T_{n,x}^{\text{het}} = \frac{\hat{m}_n(x) - \mathbb{E}\, \hat{m}_n(x)}{\sqrt{\sum_{i=1}^n W_{ni}^2(x)\, \hat{\varepsilon}_i^2}}$$

Consider constructing bootstrap versions of the pivots T_n and T_n^{het}.

Residual/wild bootstrap versions of T_n and T_n^{het}

Given a grid $x_1, \dots, x_N$ of values in $[0,1]$, define

$$T_n^* = \max_{1\leq j\leq N}\left|\frac{\sum_{i=1}^n W_{ni}(x_j)\varepsilon_i^*}{\hat\sigma_n^*\sqrt{\sum_{i=1}^n W_{ni}^2(x_j)}}\right| \quad \text{and} \quad T_n^{\text{het}*} = \max_{1\leq j\leq N}\left|\frac{\sum_{i=1}^n W_{ni}(x_j)\varepsilon_i^{*W}}{\sqrt{\sum_{i=1}^n W_{ni}^2(x_j)(\hat\varepsilon_i^{*W})^2}}\right|$$

and denote by $\hat G_n$ and $\hat G_n^{\text{het}}$ the dists of T_n^* and $T_n^{\text{het}*}$ conditional on the data.

Exercise:

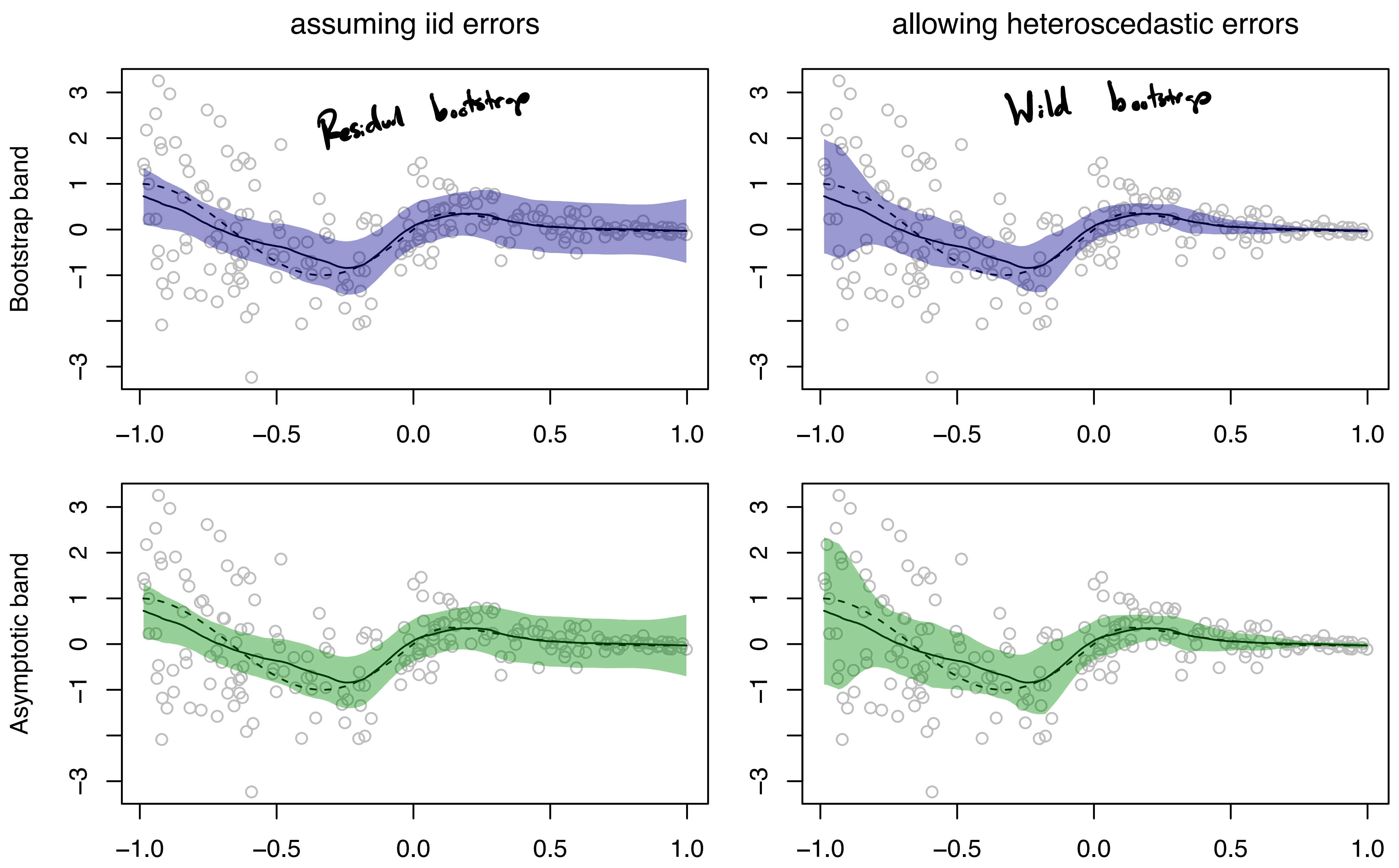
1. Give the form of $(1-\alpha)100\%$ bootstrap confidence bands under constant variances and heteroscedasticity.
2. Demonstrate on simulated data.

Obtain $T_n^{(1)*} < \ldots < T_n^{(B)*}$ (sorted) MC realizations of T_n^*.

replace with $T_n^{(\lceil (1-\alpha) B \rceil)}$

$$\left\{(x, y) : y \in \left[\ \hat{m}_n(x) \pm c_\alpha \hat{\sigma}_n \sqrt{\textstyle\sum_{i=1}^n W_{ni}^2(x)}\ \right],\quad x \in [0,1]\right\}$$

Likewise in non-const. variance case.


assuming iid errors
allowing heteroscedastic errors
Residual bootstrap
Wild bootstrap
Bootstrap band
Asymptotic band
−1.0
−0.5
0.0
0.5
1.0
3
2
1
0
−1
−3

Debraj Das, Karl Gregory, and Soumendra Lahiri.
Perturbation bootstrap in adaptive lasso.
The Annals of Statistics, 47(4):2080–2116, 2019.

Enno Mammen.
Bootstrap and wild bootstrap for high dimensional linear models.
The Annals of Statistics, pages 255–285, 1993.

Jiayang Sun, Clive R Loader, et al.
Simultaneous confidence bands for linear regression and smoothing.
The Annals of Statistics, 22(3):1328–1345, 1994.

Larry Wasserman.
All of nonparametric statistics.
Springer Science & Business Media, 2006.

Halbert White.
A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity.
Econometrica: Journal of the Econometric Society, pages 817–838, 1980.

$\hat{\theta}_n(X_1, \dots, X_n)$ is an estimator of θ.

%-tile bootstrap: Collect $\hat{\theta}_n^{(1)*}, \dots, \hat{\theta}_n^{(B)*}$.

Then make a $(1-\alpha)100\%$ C.I. for θ as

$$\left[\hat{\theta}_n^{*(\lceil(\alpha/2)B\rceil)} , \hat{\theta}_n^{*(\lceil(1-\alpha/2)B\rceil)} \right]$$