PART

Statistical Inference

2

Models, Statistical Inference and Learning

6.1 $\mathbb{P}(X_i = x) = e^{-\lambda}\dfrac{\lambda^x}{x!} \quad \forall x \in \mathbb{Z}_{\geq 0}$.

$\hat\lambda = n^{-1}\sum_{i=1}^n X_i$.

$\mathbb{E}[\hat\lambda] = \lambda \quad (\because \mathbb{E}[X_i] = \lambda)$

$\text{bias}\,\hat\lambda = \mathbb{E}[\hat\lambda] - \lambda = 0$

$\mathbb{V}(\hat\lambda) = \mathbb{E}[\hat\lambda^2] - \mathbb{E}[\hat\lambda]^2$

$\mathbb{E}[\hat\lambda^2] = \dfrac{1}{n^2}\sum \mathbb{E}[X_iX_j]$

$= \dfrac{1}{n^2}\Big(\sum_{i\neq j}\mathbb{E}[X_i]\mathbb{E}[X_j] + \sum_i \mathbb{E}[X_i^2]\Big)$

$= \dfrac{1}{n^2}\big((n^2-n)\lambda^2 + n(\lambda+\lambda^2)\big)$

$= \dfrac{1}{n^2}(n^2\lambda^2 + n\lambda) = \lambda^2 + \dfrac{\lambda}{n}$.

$\mathbb{V}(\hat\lambda) = \dfrac{\lambda}{n}$ ... 그냥 당연하잖아 ...

$\mathbb{V}(\hat\lambda) = \mathbb{V}(n^{-1}\sum_{i=1}^n X_i) = \dfrac{1}{n^2}\sum \mathbb{V}(X_i) = \dfrac{1}{n^2}\cdot n\lambda$.

MGF.

$\mathbb{E}[e^{\theta X}] = \sum e^{\theta x} e^{-\lambda}\dfrac{\lambda^x}{x!}$

$= e^{-\lambda}\sum \dfrac{(e^\theta\lambda)^x}{x!}$

$= e^{-\lambda}e^{e^\theta\lambda} = e^{(e^\theta-1)\lambda}$

$\mathbb{E}[X] = \lambda e^\theta e^{(e^\theta-1)\lambda}\big|_{\theta=0} = \lambda$

$\mathbb{E}[X^2] = (\lambda e^\theta + \lambda^2 e^{\theta})e^{(e^\theta-1)\lambda}\big|_{\theta=0}$

$= \lambda + \lambda^2$.

$\text{MSE} = \text{bias}^2\hat\lambda + \mathbb{V}(\hat\lambda) = \dfrac{\lambda}{n}$.

6.2 $\mathbb{P}(\hat{\theta} \le x) = \left(\frac{x}{\theta}\right)^n \quad \forall x \in [0, \theta]$

$\hat{\theta}$ has PDF $\frac{n}{\theta^n} x^{n-1}$, $x \in [0, \theta]$

$$\mathbb{E}[\hat{\theta}] = \int_0^\theta \frac{nx}{\theta^n} x^{n-1} dx = \int_0^\theta \frac{n}{\theta^n} x^n dx = \frac{n}{n+1}\theta$$

$$\text{bias}(\hat{\theta}) = \frac{n}{n+1}\theta - \theta = \frac{1}{n+1}\theta.$$

$$\mathbb{V}(\hat{\theta}) = \mathbb{E}[\hat{\theta}^2] - \mathbb{E}[\hat{\theta}]^2$$

$$\mathbb{E}[\hat{\theta}^2] = \int_0^\theta \frac{nx^2}{\theta^n} x^{n-1} dx = \int_0^\theta \frac{n}{\theta^n} x^{n+1} dx = \frac{n}{n+2}\theta^2.$$

$$\mathbb{V}(\hat{\theta}) = \frac{n}{n+2}\theta^2 - \left(\frac{n\theta}{n+1}\right)^2 = \frac{n(n+1)^2 - (n+2)n^2}{(n+2)(n+1)^2}\theta^2$$

$$= \frac{n}{(n+2)(n+1)^2}\theta^2.$$

$$se(\hat{\theta}) = \sqrt{\frac{n}{(n+2)(n+1)^2}}\,\theta.$$

$$MSE = \text{bias}^2 + \mathbb{V}(\hat{\theta}) = \left(\frac{\theta}{n+1}\right)^2 + \frac{n}{(n+2)(n+1)^2}\theta^2$$

$$= \frac{2n+2}{(n+2)(n+1)^2}\theta^2 = \frac{2}{(n+2)(n+1)}\theta^2.$$

6.3 $\mathbb{E}[\hat{\theta}] = \theta$ $\quad$ bias $= 0$.

$$\mathbb{V}(\hat{\theta}) = \mathbb{V}(2\overline{X}_n) = 4\mathbb{V}(\overline{X}_n) = 4 \cdot \frac{1}{n}\mathbb{V}(X_1) = 4 \cdot \frac{1}{n} \times \frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

$\text{MSE} = \frac{\theta^2}{3n}$, $\quad$ $se = \frac{\theta}{\sqrt{3n}}$

Estimating

the CDF and Statistical Functionals.

7.1 Already Proved!

$\hat{F}_n(x) = \frac{1}{n}\sum Y_i$, where $Y_i = I(X_i \leq x)$. $\begin{cases} \mathbb{P}(Y_i = 1) = F(x) \\ \mathbb{P}(Y_i = 1) = 1 - F(x). \end{cases}$

$Y_i \sim \text{Bernoulli}(F(x))$. $\mathbb{E}[Y_i] = F(x)$, $\mathbb{V}(Y_i) = F(x)(1-F(x))$.

$\mathbb{E}[\hat{F}_n(x)] = \frac{1}{n}\sum \mathbb{E}[Y_i] = F(x)$

$\mathbb{V}(\hat{F}_n(x)) = \frac{1}{n}\mathbb{V}(Y_i) = \frac{1}{n}F(x)(1-F(x))$.

bias $= 0 \rightarrow \text{MSE} = \mathbb{V}(\hat{F}_n(x)) = \frac{1}{n}F(x)(1-F(x))$.

MSE $\to 0$ as $n \to \infty$. i.e. $\hat{F}_n(x) \xrightarrow{qm} F(x)$.

$\rightarrow \hat{F}_n(x) \xrightarrow{P} F(x)$.

7.2 $p = \mathbb{E}[X]$, $X \sim \text{Bernoulli}(p)$. $\hat{F}_n(x) = \frac{1}{n}\sum I(X_i \leq x)$.

$\hat{p} = \int x \, d\hat{F}_n = \frac{1}{n}\sum X_i$

$\mathbb{V}(\hat{p}) = \frac{1}{n}p(1-p)$. $\hat{se}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$

$\mathbb{P}(\sup_x |F(x) - \hat{F}_n(x)| > \varepsilon) \leq 2e^{-2n\varepsilon^2}$. $\varepsilon_n^2 = \frac{1}{2n}\log\left(\frac{2}{\alpha}\right)$

$\mathbb{P}(\hat{F}_n(x) - \varepsilon \leq F(x) \leq \hat{F}_n(x) + \varepsilon \;\; \forall x) \geq 1 - \alpha$

$\alpha = 0.1$. $(\bar{X}_n - \varepsilon_n, \bar{X}_n + \varepsilon_n)$ where $\varepsilon_n = \sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)}$.

$$\widehat{p-q} = \hat{p} - \hat{q} = \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{m}\sum_{j=1}^{m} Y_j$$

$$\hat{q} = \frac{1}{m}\sum_{j=1}^{m} Y_j . \quad \hat{G}_m$$

$$\mathbb{V}(\hat{q}) = \frac{1}{m} q(1-q) , \quad \hat{se}(\hat{q}) = \sqrt{\frac{1}{m} q(1-q)}$$

$$\widehat{p-q} = \frac{1}{n}\sum X_i - \frac{1}{m}\sum Y_j .$$

$$\mathbb{V}(\widehat{p-q}) = \mathbb{V}(\hat{q}) + \mathbb{V}(\hat{p}) = \frac{1}{n} p(1-p) + \frac{1}{m} q(1-q)$$

$$\hat{se}(\widehat{p-q}) = \sqrt{\quad ''}$$

$$(\widehat{p-q} - z_{\alpha/2}\hat{se} , \widehat{pq} + z_{\alpha/2}\hat{se})$$

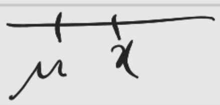
7.3 CLT: $\dfrac{\bar{X}_n - \mu}{\sqrt{\mathbb{V}(\bar{X}_n)}} \rightsquigarrow Z$

$\hat{F}_n(x) = \frac{1}{n}\sum I(X_i \le x) = \frac{1}{n}\sum Y_i = \bar{Y}_n$ where $Y_i \sim$ Bernoulli$(F(x))$!

$$\frac{\bar{Y}_n - \mu}{\sqrt{\frac{1}{n}p(1-p)}} \rightsquigarrow Z . \Rightarrow \bar{Y}_n \rightsquigarrow \mu .$$

$$x < \mu . \Rightarrow \mathbb{P}(\bar{Y}_n \le x) \le \mathbb{P}(|\bar{Y}_n - \mu| \ge \mu - x)$$

$$\le \frac{\mathbb{V}(\bar{Y}_n)}{(\mu - x)^2} = \frac{p(1-p)}{n(\mu - x)^2} \to 0$$



$$x > \mu \Rightarrow \mathbb{P}(\bar{Y}_n > x) \le \mathbb{P}(|\bar{Y}_n - \mu| > x - \mu)$$

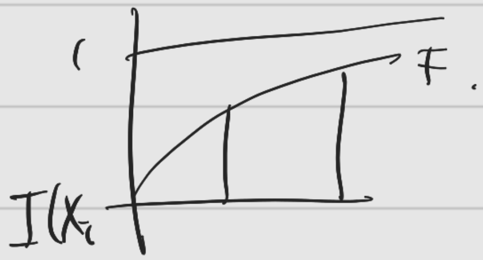
$$\leq \frac{p(1-p)}{n(x-\mu)^2}.$$

$$P(\bar{V}_{1n} \leq x) \geq 1 - \frac{p(1-p)}{n(x-\mu)^2} \to 1.$$

17.5 $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}[X]\mathbb{E}[Y]$

$\mathbb{E}[\hat{F}_n(x)] = F(x)$

$\mathbb{V}(\hat{F}_n(x)) = \frac{1}{n} F(x)(1-F(x)).$



$$\mathbb{E}[\hat{F}_n(x)\hat{F}_n(y)] = \frac{1}{n^2}\mathbb{E}\Big[\sum_i I(X_i \leq x) \sum_j I(X_j \leq y)\Big]$$

$$= \frac{1}{n^2}\Big(\sum_{i \neq j} \mathbb{E}[I(X_i \leq x)]\,\mathbb{E}[I(X_j \leq y)] + \sum_{i=j} \mathbb{E}[I(X_i \leq x) I(X_j \leq y)]\Big)$$

$$= \frac{1}{n^2}\big((n^2-n)F(x)F(y) + nF(\min\{x,y\})\big)$$

Thus, $\mathrm{Cov}(\ ,\ )$

$$= \frac{1}{n} F(\min\{x,y\}) - \frac{1}{n} F(x)F(y).$$

17.6 $\mathbb{E}[\hat{\theta}] = \mathbb{E}[\hat{F}_n(b)] - \mathbb{E}[\hat{F}_n(a)] = F(b) - F(a).$

$\mathbb{V}(\hat{\theta}) = \mathbb{E}[\hat{\theta}^2] - \mathbb{E}[\hat{\theta}]^2.$

$\mathbb{E}[\hat{\theta}^2] = \mathbb{E}[\hat{F}_n(b)^2 - 2\hat{F}_n(a)\hat{F}_n(b) + \hat{F}_n(b)^2]$

$$\mathbb{E}[\hat{F}_n(b)^2] = \mathbb{V}(\hat{F}_n(b)) + \mathbb{E}[\hat{F}_n(b)]^2$$

$$= \frac{1}{n} F(b)(1-F(b)) + F(b)^2$$

$$\mathbb{E}[\hat{F}_n(a)^2] = \frac{1}{n} F(a)(1-F(a)) + F(a)^2$$

$$\mathbb{E}[\hat{F}_n(a)\hat{F}_n(b)] = \frac{1}{n^2} \sum_{i,j} \mathbb{E}[I(X_i \le a) I(X_j \le b)]$$

$$= \frac{1}{n^2} \Big(\sum_{i \ne j}$$

$$\sum_{i=j} \Big)$$

$$= \frac{1}{n^2} \left((n^2-n) F(a)F(b) + n F(\min\{a,b\}) \right)$$

$$\to \mathbb{E}[\hat{\theta}^2] = \frac{1}{n} F(b)(1-F(b)) + F(b)^2$$

$$+ \frac{1}{n} F(a)(1-F(a)) + F(a)^2$$

$$- 2\frac{1}{n}((n-1)F(a)F(b) + F(\min\{a,b\}))$$

$$\mathbb{V}(\hat{\theta})^2 = \frac{1}{n}\Big(F(b)(1-F(b)) + F(a)(1-F(a))$$

$$-2(n-1)F(a)F(b) + F(\min\{a,b\})\Big)$$

$$+2F(a)F(b)$$

$$= \frac{1}{n}(F(b)(1-F(b)) + F(a)(1-F(a)) + 2F(a)F(b)$$

$$+ F(\min\{a,b\}))$$

$$= \frac{1}{n}\left(F(b) + F(a) - (F(b)-F(a))^2 + F(\min\{a,b\}) \right)$$

By Hoeffding, $\mathbb{P}(|\hat{F}_n(b)-F(b)|>\varepsilon) \le 2e^{-n\varepsilon^2}$

$\mathbb{P}(|\hat{F}_n(a)-F(a)|>\varepsilon) \le 2e^{-n\varepsilon^2}$

$\mathbb{P}(|\hat{\theta}-\theta|>\varepsilon) \le \mathbb{P}(| \quad |>\frac{\varepsilon}{2})+\mathbb{P}(| \quad |>\frac{\varepsilon}{2})$

$\le 4e^{-\frac{n\varepsilon^2}{4}}$

7.7

7.8

7.9 $X_1,\cdots,X_{100} \sim \text{Bernoulli}(p_1)$

$Y_1,\cdots,Y_{100} \sim \text{Bernoulli}(p_2)$.

$\theta = p_1-p_2$. $\hat{\theta} = \frac{1}{100}\sum X_i - \frac{1}{100}\sum Y_j$

$\hat{\sigma}_1^2 = \frac{1}{100}\sum(X_i-\overline{X}_{100})^2$, $\sigma_2^2 = \frac{1}{100}\sum(Y_j-\overline{Y}_{100})^2$.

$\hat{\sigma}^2 = \hat{\sigma}_1^2+\hat{\sigma}_2^2$. $\hat{\sigma} = \frac{1}{10}\sqrt{\sum_i(X_i-\overline{X}_{100})^2+\sum_j(Y_j-\overline{Y}_{100})^2}$

Given data, point estimate of $\hat{\theta}$ is $\frac{1}{100}(90-85) = \frac{1}{20}$

$\hat{se}$ is $\frac{1}{10}\sqrt{\left(\frac{9}{10}\right)^2\times 10+\left(\frac{1}{10}\right)^2\times 90 + \left(\frac{17}{20}\right)^2\times 15+\left(\frac{3}{20}\right)^2\times 85}$

$= \frac{1}{10}\sqrt{9+\frac{3\times 17^2}{80}+\frac{3^2\times 17}{80}}$

$= \frac{1}{10}\sqrt{9+\frac{3\times 17}{4}} = \frac{1}{10}\sqrt{\frac{87}{4}}$ $\quad \frac{51}{36}$

$= \frac{\sqrt{87}}{20}$

95 percent conf. int. $(\hat{\theta} - z_{\alpha/2}\,\hat{se},\ \hat{\theta} + z_{\alpha/2}\,\hat{se})$,

$\alpha = 0.1$

Bootstrap.

8.1

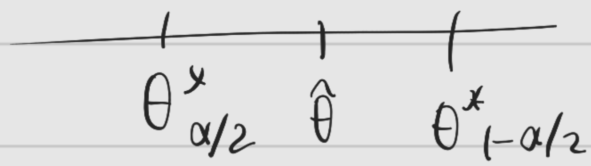
1. Consider the data in Example 8.6. Find the plug-in estimate of the correlation coefficient. Estimate the standard error using the bootstrap. Find a 95 percent confidence interval using the Normal, pivotal, and percentile methods.

$$\hat{\rho} = \frac{\sum (Y_i - \bar{Y})(Z_i - \bar{Z})}{\sqrt{\sum (Y_i - \bar{Y})^2 \sum (Z_i - \bar{Z})^2}}$$

$$\hat{se}_{boot} = \sqrt{v_{boot}} = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(T^*_{n,b} - \frac{1}{B}\sum_{r=1}^{B} T^*_{n,r}\right)^2}.$$

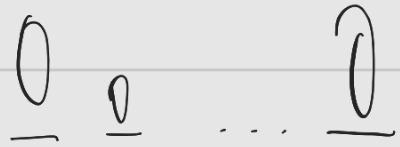
Normal $\hat{\rho} \pm z_{\alpha/2}\, \hat{se}_{boot}$ where $\alpha = 0.05$

Pivotal



Percentile

8.4



$B = n^2$. ${}_nH_n = \binom{n+n-1}{n} = \binom{2n-1}{n}$

8.5 $\mathbb{E}(\bar{X}_n^* \mid X_1, \dots, X_n) = \frac{1}{n}\sum \mathbb{E}(X_i^* \mid X_1, \dots, X_n)$

$= \frac{1}{n}\sum \frac{X_1 + \dots + X_n}{n}$

$= \frac{X_1 + \dots + X_n}{n} = \bar{X}_n$.

$\mathbb{E}(\bar{X}_n^*) = \mathbb{E}(\mathbb{E}(\bar{X}_n^* \mid X_1, \dots, X_n))$

$= \mathbb{E}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n}\sum \mathbb{E}[X_i] = \mathbb{E}(X)$.

$\mathbb{V}(\bar{X}_n^* \mid X_1, \dots, X_n) = \frac{1}{n^2}\sum \mathbb{V}(X_i^* \mid X_1, \dots, X_n)$

$\mathbb{V}(X_i^* \mid X_1, \dots, X_n) = \frac{1}{n}\sum (X_i - \bar{X}_n)^2 = \frac{n-1}{n} S^2$.

$\Rightarrow \mathbb{V}(\bar{X}_n^* \mid \ ) \ \frac{1}{n^2}\sum (X_i - \bar{X}_n)^2$.

$\mathbb{V}(\bar{X}_n^*) = \mathbb{E}(\mathbb{V}(\bar{X}_n^* \mid X_1, \dots, X_n)) + \mathbb{V}(\mathbb{E}(\bar{X}_n^* \mid X_1, \dots, X_n))$

$= \mathbb{E}\ \frac{n-1}{n^2} S^2 + \mathbb{V}(\bar{X}_n)$

$= \frac{n-1}{n^2}\sigma^2 + \frac{\sigma^2}{n} = \frac{\sigma^2}{n}\left[2 - \frac{1}{n}\right]$

$\mathbb{E}(X_i^*) = \frac{1}{n}\mathbb{E}(X_1) + \cdots + \frac{1}{n}\mathbb{E}(X_n) = \mu$

$\mathbb{E}(X_i^{*2}) = \frac{1}{n}\mathbb{E}(X_1^2) + \cdots + \frac{1}{n}\mathbb{E}(X_n^2) = \mu^2 + \sigma^2$.

$\mathbb{V}(X_i^*) = \sigma^2$.

$\mathbb{E}(\overline{X}_n^*) = \mu$

$\mathbb{V}(\overline{X}_n^*) = \frac{1}{n}\sigma^2$.

8.8

8. Let $T_n = \overline{X}_n^2$, $\mu = \mathbb{E}(X_1)$, $\alpha_k = \int |x-\mu|^k dF(x)$ and $\widehat{\alpha}_k = n^{-1}\sum_{i=1}^n |X_i - \overline{X}_n|^k$. Show that

$$v_{\text{boot}} = \frac{4\overline{X}_n^2 \widehat{\alpha}_2}{n} + \frac{4\overline{X}_n \widehat{\alpha}_3}{n^2} + \frac{\widehat{\alpha}_4}{n^3}.$$

$X_1^*, \cdots, X_n^* \sim \widehat{F}_n$.

$T_n = \overline{X}_n^2$.

$T_{n,1}^*, \cdots, T_{n,B}^*$ $\qquad X_1^*, \cdots X_n^*$

$$\mathbb{E}[T_{n,b}^*] = \mathbb{E}[\overline{X}_n^{*2}] = \mathbb{V}(\overline{X}_n^*) + \mathbb{E}[\overline{X}_n^*]^2$$

$$= \frac{1}{n}\sum(X_i - \overline{X}_n)^2 + (\overline{X}_n)^2.$$

$$= \hat{\alpha}_2 + (\overline{X}_n)^2.$$

$$\mathbb{E}[(T_{n,b}^* - \frac{1}{B}\sum_r T_{n,r}^*)^2] = \mathbb{E}[T_{n,b}^{*2}] - \frac{1}{B^2}\left(\sum_r \mathbb{E}[T_{n,r}^*]\right)^2.$$

$$\mathbb{E}[T_{n,b}^{*2}] = \mathbb{E}[\overline{X}_n^{*4}] = \mathbb{E}\left(\frac{X_1^* + \cdots + X_n^*}{n}\right)^4.$$

$$\mathbb{E}[v_{boot}] =$$

$\mathbb{E}(X_i^*) = \overline{X}_n$.

$\mathbb{E}(X_i^{*2}) = \frac{1}{n}\mathbb{E}(X_1^2) + \cdots + \frac{1}{n}\mathbb{E}(X_n^2) = \mu^2 + \sigma^2$

$\mathbb{E}(X_i^{*4}) = \frac{1}{n}\sum X_i^4$

$$\mathbb{E}\left(T_{n,b}^{*} - \frac{1}{B}\sum_{r} T_{n,r}^{*}\right)^2 \qquad \frac{B-1}{B} \quad - \quad \frac{1}{B}\sum_{r\neq b}.$$

$$\left(\frac{B-1}{B}\right)^2 \mathbb{E}\left[T_{n,b}^{*2}\right] - \frac{2(B-1)}{B^2}\sum_{r\neq b}\mathbb{E}[T_{n,b}^{*}]\,\mathbb{E}[T_{n,r}^{*}] + \frac{1}{B^2}\mathbb{E}\left(\sum_{r\neq b} T_{n,r}\right)^2$$

$$T_{n,b}^{*} \sim Y \qquad \underbrace{\frac{1}{B}\sum_{b}\left(Y_b - \frac{1}{B}\sum_{r} Y_r\right)^2}_{\text{sample variance} \times \frac{B-1}{B}}.$$

$$\frac{B-1}{B} \times \frac{1}{B-1}$$

$$\frac{B-1}{B}\,\mathbb{V}(Y) = \frac{B-1}{B}\left(\mathbb{E}(\bar{X}_i^{*4}) - \mathbb{E}(\bar{X}_i^{*2})^2\right)$$

$$= \frac{B-1}{B}\Big($$

$$\mathbb{E}\left((\bar{X}_n + S_n)^2\right)$$