ProbStats L14a

Confidence Intervals

Tue, Mar 28
TA Anency,
7-9 pm (Zoom)

March 28
2023

Wed, March 29

TA Austin Help Hours

3:15 - 4:30 MEB 3115

4:30 - 6:00 MEB 3105

Review Estimation

Assume distribution $f(\theta)$ pdf parameters

lower case $x_1, x_2 \dots x_n$

Assume $X_1, X_2, \dots X_n \overset{iid}{\sim} f(\theta)$

RVs.

Create a statistic $\hat{\theta} = T(X_1, \dots X_n)$

estimator $\hat{\theta}$

unbiased estimator if $E[\hat{\theta}] = \theta$

Flip a coin (2 sides $1, 0$) n times

$$F_1, F_2, \dots, F_n \overset{\sim}{\text{iid}} \text{Ber}(p)$$

Prob of 1

ave-func$(F_1 \dots F_n)$
function$(F_1 \dots F_n)$
return $\frac{1}{n} \sum_{i=1}^{n} F_i$

function$(F_1 \dots F_n)$
min$(F_1 \dots F_n)$ $\overset{a.s.}{=} 0$
max$(F_1 \dots F_n)$ $\overset{a.s.}{=} 1$
return $\left(\frac{max + min}{2}\right)$ $\overset{a.s.}{=} 1/2$

$\hat{p}$ = ave-func$(F_1 \dots F_n)$

$E[\hat{p}] = p$ unbiased.

$E[F_i] = p$

CLT $\bar{F}_n = \hat{p}$
$E[\bar{F}_n] = E[F_i] = p$

Confidence Intervals

assign probabilistic guarantees about parameters

$X_1, X_2, \ldots X_n \overset{iid}{\sim} f(\theta)$

const. parameter. "ground truth"

↳ create two statistics L_n, R_n

$L_n = \text{func}(\underbrace{X_1 \ldots X_n}_{\text{R.V.}})$

↳ R.V. also R_n R.V.

$$\Pr(\underbrace{L_n \leq \theta \leq R_n}_{100(1-\alpha)\%\text{ confidence interval}}) = 1-\alpha$$

e.g. $\alpha = 0.05$

↳ 95% confidence interval.

Confidence Intervals for Normal R.V.

unknown θ ; known (w/ known variance)

$$X_1, \ldots X_n \underset{iid}{\sim} N(\mu, \sigma^2)$$

CLT

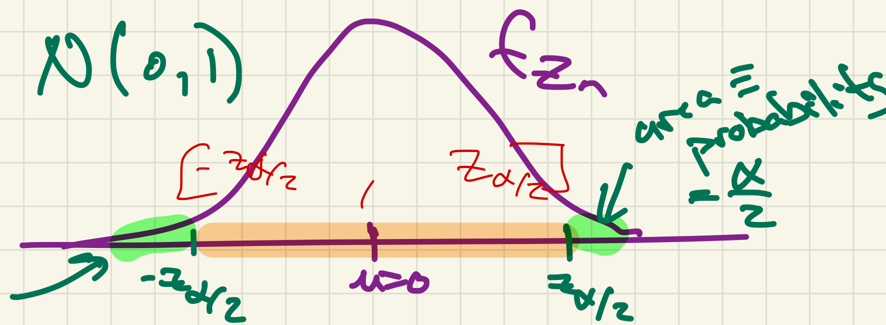
$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \qquad Z_n \sim N(0,1)$$

$$\Pr\left(-z_{\alpha/2} < Z_n \leq z_{\alpha/2}\right) = 1-\alpha$$

$z_{\alpha/2} = (1-\alpha/2)$-quantile of $N(0,1)$

$z_{0.025} \approx 1.96$

$z_{0.005} \approx 2.58$



$X_1, \dots X_n \underset{\text{iid}}{\sim} N(\mu, \sigma^2)$ known

$Var[X_i] = \sigma^2$

$$\boxed{\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i}$$

$Z_n \sim N(0,1)$

$100(1-\alpha)\%$ - confidence interval

$$Pr\left(\boxed{\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \le \mu} \le \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

const.

known const.

parameters

known $\quad Var[\bar{X}_n] = \frac{\sigma^2}{n} \quad$ std-dev$[\bar{X}_n] = \frac{\sigma}{\sqrt{n}}$

$$= Pr\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \le \mu\right) = Pr\left(\bar{X}_n \le \mu + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$= Pr\left(\bar{X}_n - \mu \le z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = Pr\left(\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \le z_{\alpha/2}\right)$$

$z_{\alpha/2}$

Normal $N(\mu, \sigma^2)$ $(1-\alpha)$-conf. inter

$$P_r\left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Interval $\left[\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$

$$\text{length} = \cancel{\bar{X}_n} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - \left(\cancel{\bar{X}_n} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$= z_{\alpha/2}\left(2\cdot\frac{\sigma}{\sqrt{n}}\right)$$