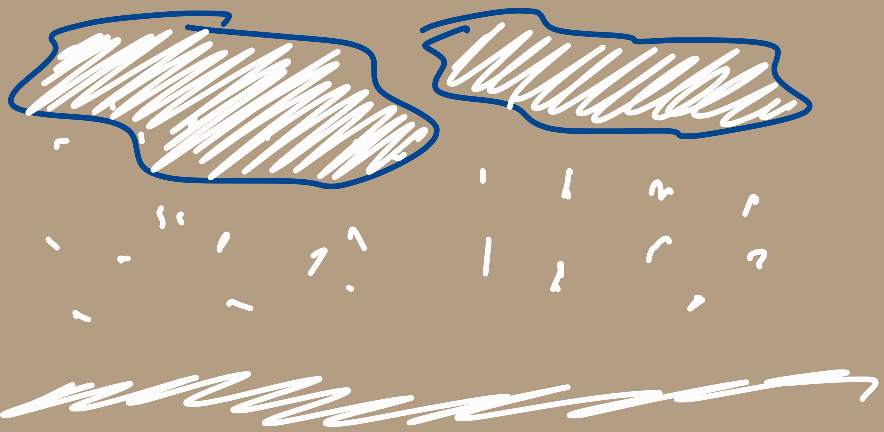
ProbStats L14c

Confidence Intervals

t-Distribution



April 4, 2023

Statistics

RVs. constant

Sample $X_1, X_2, \dots X_n \underset{iid}{\sim} f(\theta)$

What can we say (probabilistically) about $f(\theta)$, ... about θ

a statistic $\underset{\text{R.V.}}{\hat{\theta}} = T(X_1, X_2, \dots X_n)$

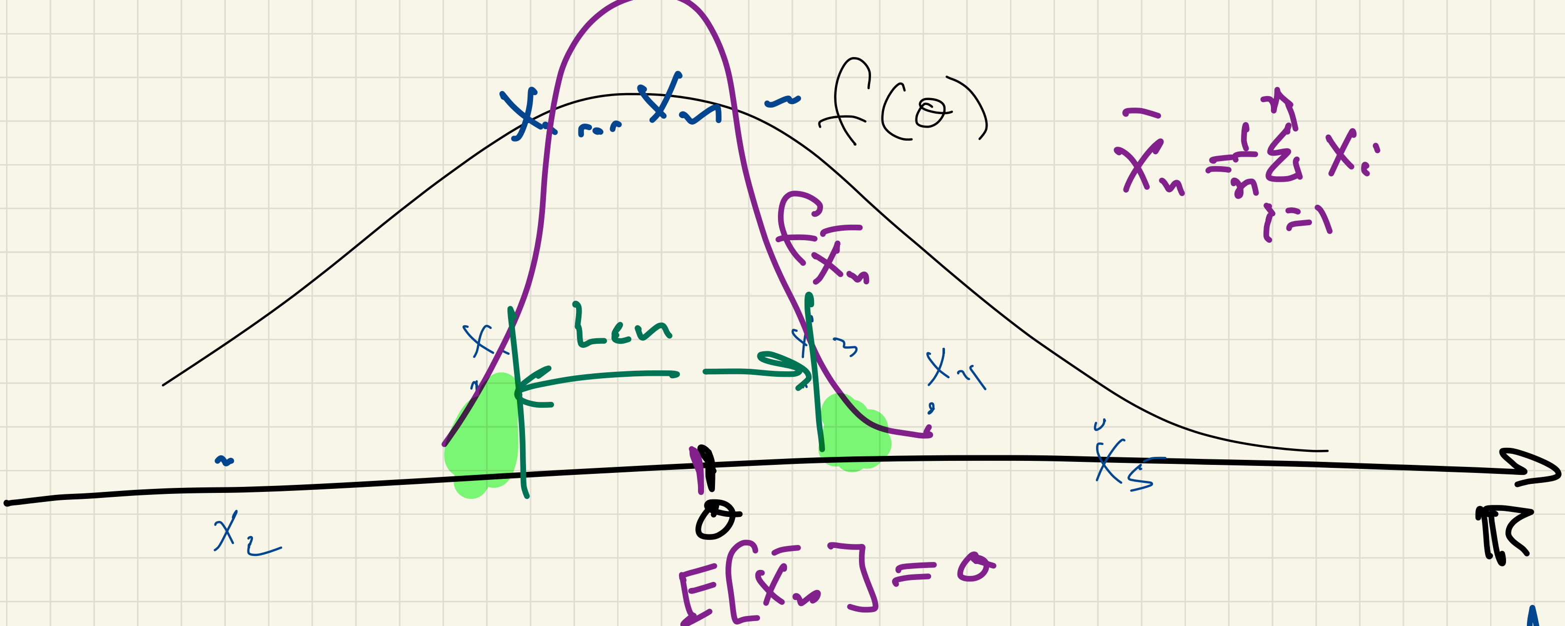
estimator of θ

- $\text{bias}(\hat{\theta}) = E[\hat{\theta}] - \theta$

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

- confidence interval $[L_n(\hat{\theta}), R_n(\hat{\theta})]$

$$\Pr(L_n \le \theta \le R_n) = 1-\alpha$$



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

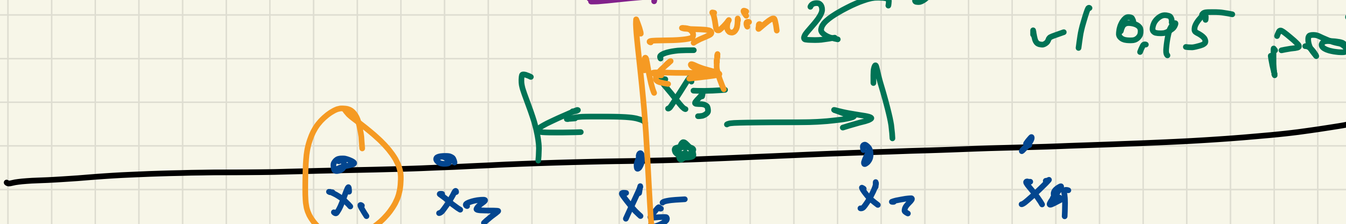
Real Data

5% chance distance θ to $\bar{X}_5$

$\geq \text{Len} = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

also $\theta \in \bar{X}_5 \pm \text{Len}$

w/ 0.95 probability



What if $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

and we do not know σ^2

first estimate variance w/ sample variance

$$S_n = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - E[X_i])^2$$

before if know σ^2

$$\bar{X}_n \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$(1-\alpha)100\%$ confidence interval for $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 unknown

$$\bar{X}_n \pm \underbrace{t_{\alpha/2}}_{} \frac{\sqrt{S_n}}{\sqrt{n}}$$

Student's t-distribution

1908

published

name: student

"William Gossett"

t-distribution

$$T_n \sim t(\underline{n-1})$$

degrees of freedom

Similar $X \sim N(0, 1)$

as $n \to \infty$ approaches $N(0,1)$

Compare

$\bar{X}_n \pm \underbrace{z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}_{\text{half length}}$ Normal

$\bar{X}_n \pm \underbrace{t_{\alpha/2} \frac{\sqrt{S_n}}{\sqrt{n}}}_{t \text{ half length}}$ t-distribution

$t_{\alpha/2} > z_{\alpha/2}$

t-half length is bigger than the z-half length probably

maybe $\sqrt{S_n}$ is too small

$$\boxed{\Pr\left(L_n \leq \mu \leq R_n\right) = 1-\alpha}$$

$$L_n = \bar{X}_n - t_{\alpha/2}\,\frac{\sqrt{S_n}}{\sqrt{n}}$$

$n = 40$ snow stations

$\bar{X}_n = 620$ in (before $\sigma^2 = 36$ in^2)

95% - confidence interval

$\underline{S_n = 34 \text{ in}^2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \underset{620}{\bar{X}_n})^2$

$\bar{X}_n \pm \underbrace{t_{0.025} \frac{\sqrt{S_n}}{\sqrt{40}}}$ $\qquad t_{0.025} = 2.02$

$620 \pm (2.02) \frac{\sqrt{34}}{\sqrt{40}} = 620 \pm 1.86$