STAT 5430

Lec 32, M, Apr 14

- No homework this week!
- Exam 2 is on W, April 16, 6:15-8:15 PM, 3rd floor seminar room
- No class on that W.
- I'll post: study guide (sufficiency/completeness/tests)
  No Bayes tests on exam
    - practice exams
    - bring new 1 page (front/back) formula sheet on exam 2 material (I'll post one to use if you'd like)
    - can bring calculator & previous formula sheet for exam 1
- I'll provide table of distributions / STAT 542 facts on test as before

Interval Estimation I

"Inverting a Test"

Theorem: Let $X_1, \ldots, X_n$ have joint pdf/pmf $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ and let $A(\underset{\sim}{\theta_0})$ denote the acceptance region of a *simple test of size* α for testing $H_0 : \underset{\sim}{\theta} = \underset{\sim}{\theta_0}$ vs. $H_1 : \underset{\sim}{\theta} \neq \underset{\sim}{\theta_0}$ (for $p = 1$, $H_1 : \theta < \theta_0$ or $\theta > \theta_0$ is allowed too). Define sets $C_{\underset{\sim}{x}} \subset \Theta$, $\underset{\sim}{x} \in \mathbb{R}^n$ as

test is 0 or 1

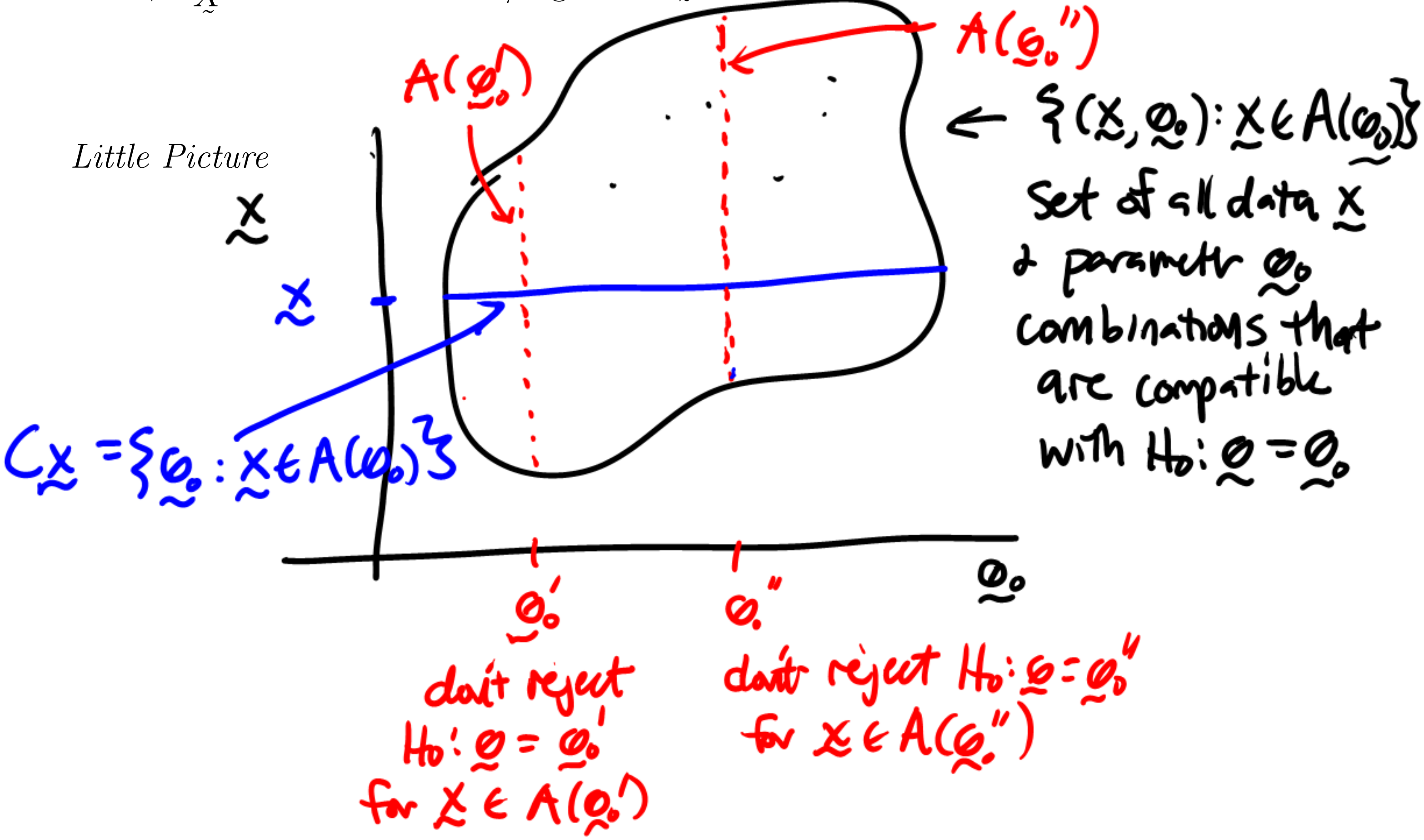
$$C_{\underset{\sim}{x}} = \{\underset{\sim}{\theta_0} : \underset{\sim}{x} \in A(\underset{\sim}{\theta_0})\}$$

given data $\underset{\sim}{x}$

recall $\underset{\sim}{x} \in A(\underset{\sim}{\theta_0})$ means "don't reject $H_0: \underset{\sim}{\theta} = \underset{\sim}{\theta_0}$"

Then, $C_{\underset{\sim}{X}}$ is a confidence set/region for $\underset{\sim}{\theta}$ with confidence coefficient $1 - \alpha$.

Little Picture



Interval Estimation I

Inverting a Test, cont'd

Proof of Theorem: Note that

1.

$$\begin{aligned}
\min_{\underset{\sim}{\theta}_0 \in \Theta} P_{\underset{\sim}{\theta}_0}\big(\underset{\sim}{X} \in A(\underset{\sim}{\theta}_0)\big) &= \min_{\underset{\sim}{\theta}_0 \in \Theta} P_{\underset{\sim}{\theta}_0}(\text{“do not reject } H_0 : \underset{\sim}{\theta} = \underset{\sim}{\theta}_0\text{”}) \\
&= \min_{\underset{\sim}{\theta}_0 \in \Theta} [1 - P_{\underset{\sim}{\theta}_0}(\text{“reject } H_0 : \underset{\sim}{\theta} = \underset{\sim}{\theta}_0\text{”})] \\
&= \min_{\underset{\sim}{\theta}_0 \in \Theta} [1 - \alpha] \\
&= 1 - \alpha,
\end{aligned}$$

and

2. for any $\underset{\sim}{\theta}_0 \in \Theta$, any $\underset{\sim}{x} \in \mathbb{R}^n$, it holds that

$$\underset{\sim}{x} \in A(\underset{\sim}{\theta}_0) \;\Leftrightarrow\; \underset{\sim}{\theta}_0 \in C_{\underset{\sim}{x}}.$$

Hence,

$$\begin{aligned}
\min_{\underset{\sim}{\theta}_0 \in \Theta} P_{\underset{\sim}{\theta}_0}(\underset{\sim}{\theta}_0 \in C_{\underset{\sim}{X}}) &= \min_{\underset{\sim}{\theta}_0 \in \Theta} P_{\underset{\sim}{\theta}_0}(\underset{\sim}{X} \in A(\underset{\sim}{\theta}_0)) \\
&= 1 - \alpha.
\end{aligned}$$

Interval Estimation I

Inverting a Test: Illustration

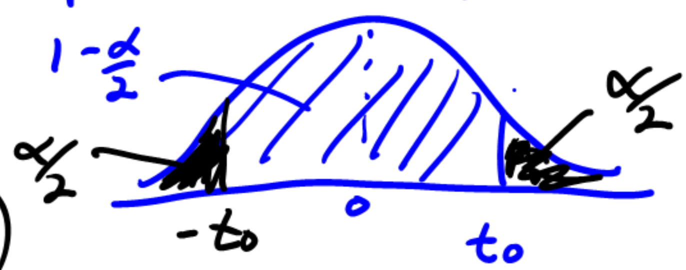
Example: Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$, $-\infty < \mu < \infty$, $\sigma > 0$. Find

1. a C.I. (confidence interval) for μ with C.C. $1-\alpha$ (two-sided)
2. a 1-sided lower confidence bound for μ with C.C. $1-\alpha$, i.e., $(L(\underset{\sim}{X}),\ \infty)$

Solution for 1. Consider a test function

$$\phi_{\mu_0}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } \dfrac{|\bar{X}_n - \mu_0|}{S/\sqrt{n}} > t_0 \\ 0 & \text{o.w.} \end{cases}$$

for testing $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ (given some $\mu_0 \in \mathbb{R}$)
where $t_0 \equiv (1-\frac{\alpha}{2})$ percentile of T_{n-1} distribution.



$E_{\mu_0}\phi_{\mu_0}(\underset{\sim}{X})$
$= P_{\mu_0}\left(|\bar{X}_n - \mu_0| / S/\sqrt{n} > t_0 \right)$
$= \alpha$, i.e. $\phi_{\mu_0}(\underset{\sim}{X})$ is a simple test (0 or 1) of size α for $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$

Note: acceptance region of $\phi_{\mu_0}(\underset{\sim}{x})$ is $A(\mu_0) = \left\{ \underset{\sim}{x} : \dfrac{|\bar{X}_n - \mu_0|}{S/\sqrt{n}} \leq t_0 \right\}$
$= \{ \underset{\sim}{x} : \phi_{\mu_0}(\underset{\sim}{x}) = 0 \}$

Hence, $C_{\underset{\sim}{x}} = \{ \mu_0 : \underset{\sim}{x} \in A(\mu_0) \}$

depends on $\underset{\sim}{x}$

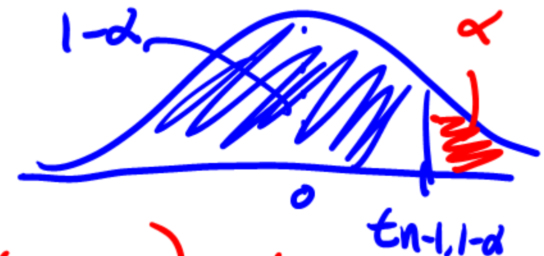
and $C_{\underset{\sim}{x}}$ is a CI with C.C. $1-\alpha$

$$\begin{aligned} C_{\underset{\sim}{x}} &= \left\{ \mu_0 : \frac{|\bar{X}_n - \mu_0|}{S/\sqrt{n}} \leq t_0 \right\} \\ &= \left\{ \mu_0 : -t_0 \leq \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \leq t_0 \right\} \\ &= \left\{ \mu_0 : -t_0 \frac{S}{\sqrt{n}} \leq \bar{X}_n - \mu_0 \leq t_0 S/\sqrt{n} \right\} \\ &= \left\{ \mu_0 : \bar{X}_n - t_0 S/\sqrt{n} \leq \mu_0 \leq \bar{X}_n + t_0 S/\sqrt{n} \right\} \\ &= \left[\bar{X}_n - t_0 S/\sqrt{n},\ \bar{X}_n + t_0 S/\sqrt{n} \right] \end{aligned}$$

Solution to 2. Consider the following test for $H_0: \mu = \mu_0$ vs $H_1: \mu > \mu_0$ given by

$$\phi_{\mu_0}(\underset{\sim}{x}) = \begin{cases} 1 & (\bar{X}_n - \mu_0)/{S/\sqrt{n}} > t_{n-1,1-\alpha} \\ 0 & \text{o.w} \end{cases}$$

where $t_{n-1,1-\alpha} \equiv 1-\alpha$ percentile



Note: $E_{\mu_0} \phi_{\mu_0}(\underset{\sim}{x}) = P_{\mu_0}\left(\frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} > t_{n-1,1-\alpha} \right) = \alpha$
↑ size

Acceptance region for $\phi_{\mu_0}(\underset{\sim}{x})$ is

$$A(\mu_0) \equiv \left\{ \underset{\sim}{x} : \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \le t_{n-1,1-\alpha} \right\}$$

For given $\underset{\sim}{x}$,

$$\begin{aligned} C_{\underset{\sim}{x}} &= \{ \mu_0 : \underset{\sim}{x} \in A(\mu_0) \} \\ &= \left\{ \mu_0 : \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \le t_{n-1,1-\alpha} \right\} \\ &= \left\{ \mu_0 : \bar{X}_n - \mu_0 \le \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right\} \\ &= \left\{ \mu_0 : \bar{X}_n - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \le \mu_0 \right\} \\ &= \left[\bar{X}_n - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha}, \infty \right) \end{aligned}$$

↑ lower confidence interval for μ

Interval Estimation I

Pivotal Quantities

$\swarrow$ vector of parameters $\underset{\sim}{\theta}$

Definition: Let $X_1, \ldots, X_n$ be joint pdf/pmf $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$. Then a random variable $Q(\underset{\sim}{X}, \underset{\sim}{\theta})$ is called a **pivot** or **pivotal quantity** if the distribution of $Q(\underset{\sim}{X}, \underset{\sim}{\theta})$ under $\underset{\sim}{\theta}$ *does not depend* on $\underset{\sim}{\theta}$.

Note: $Q(\underset{\sim}{X}, \underset{\sim}{\theta})$ is NOT a statistic (because can depend on $\underset{\sim}{\theta}$)

$P_{\underset{\sim}{\theta}}(Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \in A) = P(Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \in A)$

Some examples: (pivots, unlike statistics, can be functions of parameters $\underset{\sim}{\theta}$)

$\frac{X_1 - X_2}{X_3 - X_4}$ ancillary statistic

1. Let $X_1 \ldots X_n$ be iid $N(\mu, \sigma^2)$ random variables.

   $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$

   pivot $\left\{ \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \right.$ pivot $\left\{ \frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim T_{n-1} \text{ distribution} \right.$

2. Let f_0 be a pdf on $\mathbb{R}$. Let $X_1 \ldots X_n$ be iid with random variables common pdf $f(x|\underset{\sim}{\theta})$ where

$$\underset{\sim}{\theta} = (\theta_1, \theta_2) \qquad f(x|\underset{\sim}{\theta}) = \frac{1}{\theta_2} f_0\left(\frac{x - \theta_1}{\theta_2}\right), \quad x \in \mathbb{R},$$

for $\underset{\sim}{\theta} = (\theta_1, \theta_2)$, $\theta_1 \in \mathbb{R}$ (location parameter) and $\theta_2 > 0$ (scale parameter).

Then, $Q(\underset{\sim}{X}, \underset{\sim}{\theta}) = \frac{\bar{X}_n - \theta_1}{\theta_2} \leftarrow$ pivot

Why? Note: $Y_i \equiv \frac{X_i - \theta_1}{\theta_2} \sim f_0(y)$ (i.e. $Y_1, \ldots, Y_n$ iid $f_0(y)$)

and $Q(\underset{\sim}{X}, \underset{\sim}{\theta}) = \frac{1}{n}\sum_{i=1}^{n} Y_i \leftarrow$ distribution doesn't depend on $\underset{\sim}{\theta}$

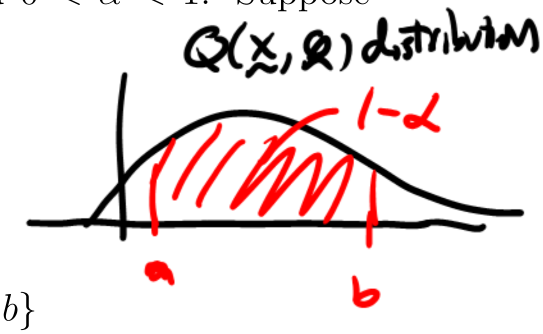
Interval Estimation I

Interval Estimation via Pivotal Quantities

Remarks:

.

1. Let $Q(\underset{\sim}{X}, \underset{\sim}{\theta})$ be a pivotal quantity ($\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$) and $0 < \alpha < 1$. Suppose $-\infty \le a \le b \le \infty$ are such that

   $$P(a \le Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \le b) = P_{\underset{\sim}{\theta}}(a \le Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \le b) = 1 - \alpha.$$

   Then,

   given data $\underset{\sim}{X}$ →
   $$C_{\underset{\sim}{X}} = \{\underset{\sim}{\theta} : \underset{\sim}{\theta} \in \Theta,\ a \le Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \le b\}$$

   is a **confidence region** for $\underset{\sim}{\theta}$ with CC $(1 - \alpha)$

   $$\text{That is, } \min_{\underset{\sim}{\theta} \in \Theta} P_{\underset{\sim}{\theta}}(\underset{\sim}{\theta} \in C_{\underset{\sim}{X}}) = \underbrace{\min_{\underset{\sim}{\theta} \in \Theta} P_{\underset{\sim}{\theta}}(a \le Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \le b)}_{\underbrace{P(a \le Q(\underset{\sim}{X}, \underset{\sim}{\theta}) \le b)}_{1-\alpha}} = 1 - \alpha.$$

2. If $\Theta \subset \mathbb{R}$ and $Q(\underset{\sim}{X}, \underset{\sim}{\theta})$ is monotone in $\underset{\sim}{\theta} \in \mathbb{R}$, then the region $C_{\underset{\sim}{X}}$ will be an interval.

Interval Estimation I

Interval Estimation via Pivotal Quantities

shape is known

Example: Let $X_1 \dots X_n$ be iid Gamma(δ_0, θ) where $\theta > 0$ (δ_0 fixed/known). Using a pivotal quantity based on $\sum_{i=1}^n X_i$, find a CI for θ with C.C. $1-\alpha$.

scale (unknown)

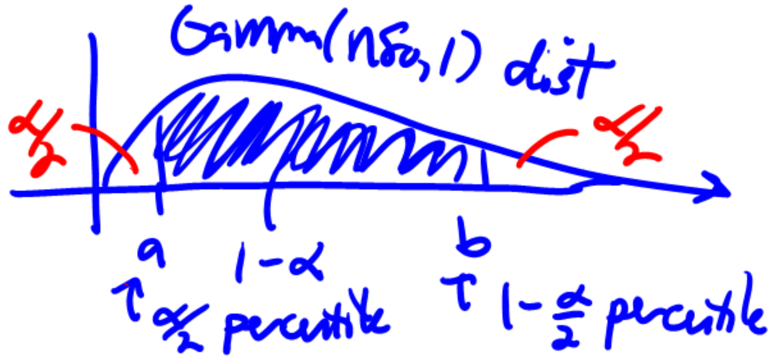
Solution: $\sum_{i=1}^n X_i \sim \text{Gamma}(n\delta_0, \theta)$

$$Q(\underset{\sim}{X}, \theta) = \sum_{i=1}^n X_i \Big/ \theta \sim \text{Gamma}(n\delta_0, 1)$$

pivot

dist. doesn't depend on θ

Find a & b such that



Confidence interval for $\theta > 0$ is $\{\theta : a \le Q(\underset{\sim}{X}, \theta) \le b\} = \{\theta : a \le \frac{\sum X_i}{\theta} \le b\}$

$= \{\theta > 0 : \frac{a}{\sum X_i} \le \theta \le \frac{b}{\sum X_i}\}$

$= \left[\frac{a}{\sum X_i}, \frac{b}{\sum X_i}\right]$