STAT 5430
Lec 36, F, Apr 25

- Homework 8 is assigned & due M, Apr 28
  by midnight
  ↳ 2nd to last homework
  (practice on inverting tests & pivotal quantities
  for making CIs)

- Exam 2 solutions & grading key
  posted

- Final Exam on Tuesday, May 13, 7:30-9:30 AM
  - Comprehensive - but focus on material since
    Exam 2 (interval estimation)
  - Formula sheet for new material/interval & 2 formula
    sheets previous. material
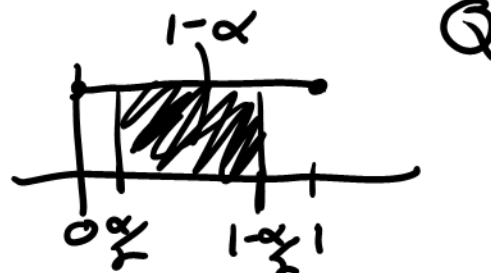   (3 sheets (front/back) total)
  - Practice Exams

see Canvas

ASIDE on pivots

Result: T is statistic with continuous cdf $F(t|\theta)$, $t \in \mathbb{R}$.

Then, $Q(T, \theta) \equiv F(T|\theta) \sim \text{UNIF}(0,1)$.

Make CI for θ:

$$\{\theta : \tfrac{\alpha}{2} \le Q(T, \theta) \equiv F(T|\theta) \le 1 - \tfrac{\alpha}{2}\}$$

Problem 9.11: Test $H_0 : \theta = \theta_0$ using

$$\phi(x) = \begin{cases} 0 & \tfrac{\alpha}{2} \le F(T|\theta) \le 1 - \tfrac{\alpha}{2} \\ 1 & \text{o.w} \end{cases}$$

$$A(\theta_0) = \{T : \tfrac{\alpha}{2} \le F(T|\theta_0) \le 1 - \tfrac{\alpha}{2}\}$$

$$C_T \equiv \{\theta_0 : T \in A(\theta_0)\}$$
$$= \{\theta_0 : \tfrac{\alpha}{2} \le F(T|\theta_0) \le 1 - \tfrac{\alpha}{2}\}$$

$$P_{\theta_0}(\theta_0 \in C_T) = P_{\theta_0}\left(\tfrac{\alpha}{2} \le \underbrace{F(T|\theta_0)}_{\text{UNIF}} \le 1 - \tfrac{\alpha}{2}\right) = 1 - \alpha$$

STAT 5430: Summary to date
Where we have been
& where we are headed

- Completed

  – Introduction to Statistical Inference

  – Point Estimation

    * MME/MLE

  – Criteria for Evaluating Point Estimators

    * bias, variance, UMVUE, MSE

  – Elements of Decision Theory

    * Minimax, finding Bayes estimators

  – Sufficiency and Point Estimation

    * Factorization/Rao-Blackwell/Lehman-Scheffe Theorems

  – Hypothesis Testing

    * MP/UMP, Likelihood Ratio/Bayes Tests

  – Interval Estimation I

    * Inverting Tests/Pivotal Quantities/Asymptotic Pivots/VST

- Next: Interval Estimation II

  – MGB Method (Pivot-based)

  – Bayes Intervals

  – Evaluating Interval Estimators
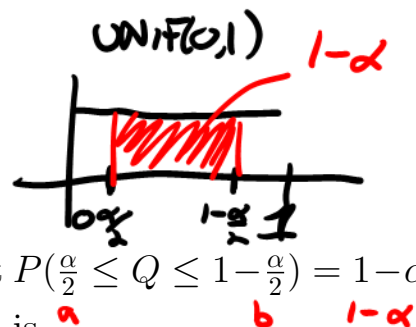
Interval Estimation II

Mood, Graybill & Boes (MGB) Interval Method

uses probability integral transform to make a certain pivot (Sec 9.2.3, C & B)

Main Idea: Suppose a statistic T is a continuous random variable having a cdf $F(t|\theta) = P(T \leq t|\theta)$, $t \in \mathbb{R}$, which depends on a real-valued parameter $\theta \in \Theta \subset \mathbb{R}$. If $\theta \in \Theta$ is the data-generating parameter, then (since T is continuous) by the probability integral transform **(PIT)**

$$F(T|\theta) \sim \text{Uniform}(0,1),$$

and hence $Q(T, \theta) \equiv F(T|\theta)$ is a pivotal quantity.

Then, for $Q \sim \text{Uniform}(0,1)$ and $\alpha \in (0,1)$, it holds that $P(\frac{\alpha}{2} \leq Q \leq 1 - \frac{\alpha}{2}) = 1 - \alpha$ and, given T, a confidence region for θ with C.C. $1 - \alpha$ is

$$C_T \equiv \left\{ \theta \in \Theta : \frac{\alpha}{2} \leq Q(T, \theta) \leq 1 - \frac{\alpha}{2} \right\} = \left\{ \theta \in \Theta : \frac{\alpha}{2} \leq F(T|\theta) \leq 1 - \frac{\alpha}{2} \right\}.$$

For given a value $T = t$, if it turns out that $F(t|\theta)$ is increasing in θ or decreasing in θ, then the above confidence region based on this $T = t$ value will be an **interval**, say, $C_{T=t} = [\theta_L(t), \theta_U(t)]$ with endpoints determined by $\theta_L(t) = \min\{a(t), b(t)\}$ and $\theta_U(t) = \max\{a(t), b(t)\}$ for

$$\frac{\alpha}{2} = F(t|a(t)) = P(T \leq t|a(t)), \qquad \frac{\alpha}{2} = 1 - F(t|b(t)) = P(T \geq t|b(t));$$

$a(t), b(t) \in \Theta$ are points where, given $T = t$, the cdf $F(t|\theta)$ "crosses" $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ as a function of θ.

This is the MGB method with an extension allowing for discrete statistics T in addition to continuous ones.

Interval Estimation II
Mood, Graybill & Boes (MGB) Interval Method

Theorem: Let T be a statistic (a discrete or continuous random variable is allowed) with cdf $F(t|\theta) = P(T \leq t|\theta)$, $\theta \in \Theta \subset \mathbb{R}$. Suppose $\mathcal{T} = $ "the set of all possible values of T".

for each possible value t of T $(t \in \mathcal{T})$

Let $0 < \alpha < 1$. Suppose, for each $t \in \mathcal{T}$, there exist functions $\theta_L(t)$ and $\theta_U(t)$ such that

$\mathcal{T}$ given

(a) if $F(t|\theta)$ is a decreasing function of θ for each t, then *upper tail*

lower tail

$$P\left(T \leq t|\theta_U(t)\right) = \frac{\alpha}{2} \quad \& \quad P\left(T \geq t|\theta_L(t)\right) = \frac{\alpha}{2}$$

given t   *given t*

(b) or if $F(t|\theta)$ is an increasing function of θ for each t, then

lower tail   *upper tail*

$$P\left(T \leq t|\theta_L(t)\right) = \frac{\alpha}{2} \quad \& \quad P\left(T \geq t|\theta_U(t)\right) = \frac{\alpha}{2}$$

given t   *given t*

Then,

$$[\theta_L(T), \theta_U(T)]$$

is a CI for θ with a confidence coefficient (C.C.) satisfying C.C. $\geq 1 - \alpha$.

In particular,

when T is a continuous random variable (namely, $F(t|\theta)$ is continuous in t), then $[\theta_L(T), \theta_U(T)]$ has

$$\text{C.C.} = \min_{\theta \in \Theta} P_\theta \left(\theta \in [\theta_L(T), \theta_U(T)]\right) = 1 - \alpha.$$

↑ T continuous, CC is exactly $1-\alpha$

when T is a discrete random variable (namely, $F(t|\theta)$ is a step function in t), then $[\theta_L(T), \theta_U(T)]$ has

✓ not exactly $1-\alpha$

$$\text{C.C.} = \min_{\theta \in \Theta} P_\theta \left(\theta \in [\theta_L(T), \theta_U(T)]\right) \geq 1 - \alpha.$$

✓ not exactly $1-\alpha$ for discrete T

118

Interval Estimation II

Mood, Graybill & Boes (MGB) Interval Method

Remarks:

1. A confidence interval I (a function of the data) for $\theta \in \Theta \subset \mathbb{R}$ is called a **conservative** $(1 - \alpha)$ **confidence interval** if

$$\min_{\theta \in \Theta} P_\theta(\theta \in I) \geq 1 - \alpha$$

lower bound on coverage

2. The theorem above says that the confidence interval $\left[\theta_L(T), \theta_U(T)\right]$ will have a C.C. of exactly $(1-\alpha)$ when the statistic T (based on the data) is a continuous random variable. But, when T is a discrete random variable, the C.C. of the interval $\left[\theta_L(T), \theta_U(T)\right]$ may not exactly equal $(1 - \alpha)$, but cannot be smaller. Hence, when T is discrete, the interval $\left[\theta_L(T), \theta_U(T)\right]$ will be a conservative $(1 - \alpha)$ confidence interval.

Interval Estimation II

Mood, Graybill & Boes (MGB) Interval Method: Illustration

Example 1. (Continuous random variables) Suppose $X_1, \ldots, X_n$ are iid $N(\theta, \log \theta)$,
$1 < \theta < \infty$. Apply the Mood-Graybill-Boes method to obtain a $\overset{1-\alpha}{\text{CI}}$ for θ based on
$T = \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = (n-1)S^2$.

Solution: first, we need to get cdf of T

Note: $\dfrac{(n-1)S^2}{\underset{\underset{\text{variance}}{\searrow} \log \theta = \frac{T}{\log \theta}}{\to}} \sim \chi_{n-1}^2$.

Then,
$$F(t|\theta) = P(T \le t | \theta) = P\left(\frac{T}{\log \theta} \le \frac{t}{\log \theta} \Big| \theta\right)$$

$$= P\left(\chi_{n-1}^2 \le \frac{t}{\log \theta}\right) \equiv G\left(\frac{t}{\log \theta}\right),$$

χ_{n-1}^2 density

area $G\left(\frac{t}{\log \theta}\right)$

where
$G(\cdot)$ is the cdf
of χ_{n-1}^2

$t/\log \theta$

So, $\mathcal{T} \equiv$ "possible values of T" $= (0, \infty)$.
Pick/fix a possible value $t \in (0, \infty)$ of T.
(Note: given t, $F(t|\theta) = G\left(\frac{t}{\log \theta}\right) \downarrow$ as $\theta \uparrow$)
Find $\theta_u \equiv \theta_u(t)$ & $\theta_L \equiv \theta_L(t)$ so that
$$\frac{\alpha}{2} = P(T \le t | \theta_u) \quad \& \quad \frac{\alpha}{2} = P(T \ge t | \theta_L)$$

$\overset{\text{since } T}{\underset{\text{continuous}}{\to}}$

$$= G\left(\frac{t}{\log \theta_u}\right) \qquad = 1 - P(T \le t | \theta_L)$$

$$= 1 - G\left(t/\log \theta_L\right)$$

$$\Rightarrow \frac{t}{\log \theta_U} = \chi^2_{n-1,\frac{\alpha}{2}} \qquad \& \qquad \frac{t}{\log \theta_L} = \chi^2_{n-1,1-\frac{\alpha}{2}}$$

$$\uparrow \frac{\alpha}{2} \text{ percentile} \qquad\qquad \uparrow 1-\frac{\alpha}{2} \text{ percentile}$$

$$\Rightarrow \text{ solve for } \theta_U \ \& \ \theta_L \ \& \ \text{set } [\theta_L(t), \theta_U(t)]$$