STAT 5430

Lec 30, W, Apr 9

- No homework this week! Homework 7 solutions posted
- Exam 2 is coming up (≈ 1 week away)
  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

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

- Next: Interval Estimation I
  - General Concepts
  - Inverting Tests
  - Pivotal Quantities
  - Asymptotically Pivotal Quantites/Variance Stabilizing Transformation

Interval Estimation I

Introduction

Definition: Let $X_1, \ldots, X_n$ have joint pdf/pmf $f(\underset{\sim}{x}|\theta)$, $\theta \in \Theta \subset \mathbb{R}$ (real-valued θ), and let $L(\underset{\sim}{X})$ and $U(\underset{\sim}{X})$ be two statistics such that $L(\underset{\sim}{X}) \leq U(\underset{\sim}{X})$. Then, $\uparrow$ $\theta \in \mathbb{R}^1$ (one parameter)

1. the random interval $I(\underset{\sim}{X}) = \big[L(\underset{\sim}{X}), U(\underset{\sim}{X})\big]$ is called an **interval estimator** for θ.

   $\uparrow$ "two-sided interval"
   (lower $L(\underset{\sim}{X})$ & upper $U(\underset{\sim}{X})$ bound)

2. $I(\underset{\sim}{X}) = \big(-\infty, U(\underset{\sim}{X})\big]$ is a **one-sided upper interval estimator**.

   $U(\underset{\sim}{X})$ is called "upper bound" for θ

3. $I(\underset{\sim}{X}) = \big[L(\underset{\sim}{X}), \infty\big)$ is a **one-sided lower interval estimator**.

   $L(\underset{\sim}{X})$ is called "lower bound" for θ

4. For $\theta \in \Theta$, the **coverage probability** of an interval estimator is

$$P_\theta\left(\theta \in I(\underset{\sim}{X})\right).$$

   $\uparrow$ θ that generates data $\underset{\sim}{X}$ under $f(\underset{\sim}{x}|\theta)$

5. The **confidence coefficient** (CC) of an interval estimator is given by

$$\min_{\theta \in \Theta} P_\theta\left(\theta \in I(\underset{\sim}{X})\right).$$

   $\uparrow$ minimal coverage probability
   over all possible $\theta \in \Theta$

Interval Estimation I

$a=1,\ b=2$

Example

↙ positive parameter $\quad X_{(n)} \le \theta$

Example: Let $X_1, \ldots, X_n$ be iid Uniform$(0, \theta)$, $\theta > 0$. Consider the intervals $\left[aX_{(n)}, bX_{(n)}\right]$ for fixed constants $1 \le a \le b$, or $\left[c + X_{(n)}, \infty\right)$ where $c > 0$ is fixed. Here $X_{(n)} = \max_{1 \le i \le n} X_i$. Find the coverage probabilities and confidence coefficients of these intervals.

Solution: Coverage prob of $[aX_{(n)}, bX_{(n)}]$

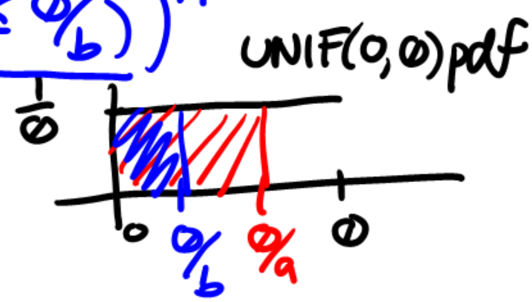
$$= P_\theta(\theta \in [aX_{(n)}, bX_{(n)}])$$
$$= P_\theta(aX_{(n)} \le \theta \le bX_{(n)})$$
$$= P_\theta\left(\frac{\theta}{b} \le X_{(n)} \le \frac{\theta}{a}\right)$$

Recall: $P(X_{(n)} \le x) = [P(X_1 \le x)]^n$

$$= P_\theta(X_{(n)} \le \theta/a) - P_\theta(X_{(n)} \le \theta/b)$$

Note: $X_{(n)}$ is continuous

$$= (P_\theta(X_1 \le \theta/a))^n - (P_\theta(X_1 \le \theta/b))^n$$



$$= \left(\frac{\theta}{a} \cdot \frac{1}{\theta}\right)^n - \left(\frac{\theta}{b} \frac{1}{\theta}\right)^n$$
$$= a^{-n} - b^{-n}$$

So, C.C. of $[aX_{(n)}, bX_{(n)}]$ is

$$\min_{\theta > 0} P_\theta(\theta \in [aX_{(n)}, bX_{(n)}]) = a^{-n} - b^{-n}$$

you can control by a & b choice

Coverage prob of $[c + X_{(n)}, \infty)$

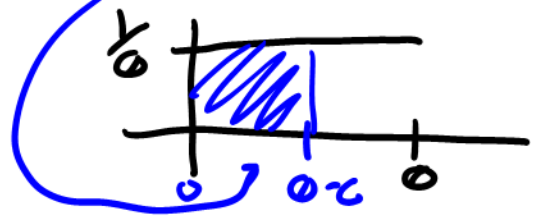
$$= P_\theta(\theta \in [c + X_{(n)}, \infty))$$
$$= P_\theta(c + X_{(n)} \le \theta)$$

$$= P_\theta\left(X_{(n)} \le \theta - c \right)$$

Note: $c > 0$, $\theta > 0$,

$$= \begin{cases} 0 & \text{if } 0 < \theta \le c \\ \left[\underline{P_\theta(X_1 \le \theta - c)} \right]^n & \text{if } \theta > c > 0 \end{cases}$$

UNIF$(0,\theta)$ pdf



$$= \begin{cases} 0 & \text{if } 0 < \theta \le c \\ \left(\frac{\theta - c}{\theta} \right)^n & \text{if } c < \theta \end{cases}$$

So, C.C. of $[c + X_{(n)}, \infty)$ is

$$\min_{\theta > 0} P_\theta\left(\theta \in [c + X_{(n)}, \infty) \right) = 0$$

Interval Estimation I

Overview

Remarks:

1. An interval estimator $I(\underset{\sim}{X}) = \left[L(\underset{\sim}{X}), U(\underset{\sim}{X})\right]$ **together** with its coefficient coefficient is called a **confidence interval** for a real-valued parameter.

   $\theta \in \mathbb{R}^1$

2. If $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ (vector-valued), then the concept of "confidence intervals" is replaced by **confidence regions** (set estimator).

   or confidence sets

General Methods of Interval Estimation

1. Inverting a Test

2. Pivotal Quantities & Asymptotically Pivotal Quantities

3. Bayes "Credible" Intervals