STAT 610: Discussion 5

1 Summary

- Hypothesis test: after observing samples $\mathbf{X} \sim \mathbb{P}_\theta$, decide either to retain $H_0 : \theta \in \Theta_0$ (null hypothesis) or to reject H_0 and conclude $H_1 : \theta \in \Theta_0^c$ (alternative hypothesis) is true.
- Basic definitions:
  - Type I error: reject H_0 if H_0 is true.
  - Type II error: retain H_0 if H_0 is false.
  - Let R denote the rejection region for a test. Then

    $$\beta(\theta) = \mathbb{P}_\theta(\mathbf{X} \in R) = \begin{cases} \mathbb{P}(\text{Type I error}) & \text{if } \theta \in \Theta_0 \\ 1 - \mathbb{P}(\text{Type II error}) & \text{if } \theta \in \Theta_0^c \end{cases}$$

    is called the power function.
  - A test with power function $\beta(\theta)$ is a size (level) α test if $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$ $(\leq \alpha)$.
- The likelihood ratio test statistic for $H_0 : \theta \in \Theta_0$ v.s. $H_1 : \theta \in \Theta_0^c$ is

  $$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta \mid \mathbf{x})}{\sup_{\theta \in \Theta} L(\theta \mid \mathbf{x})}.$$

  A *likelihood ratio test* (LRT) is any test with rejection region $\{\mathbf{x} : \lambda(\mathbf{x}) \leq c\}$, where $0 \leq c \leq 1$. Normally, we choose c such that the test

  $$T(\mathbf{X}) = \begin{cases} 1 & \text{if } \lambda(\mathbf{X}) \leq c \\ 0 & \text{if } \lambda(\mathbf{X}) > c \end{cases}$$

  has size α, that is, $\sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\lambda(\mathbf{X}) \leq c) = \alpha$.
- **Theorem 8.2.4.** If $W(\mathbf{X})$ is a sufficient statistic for θ and $\lambda^*(w)$ and $\lambda(\mathbf{x})$ are the LRT statistic based on W and $\mathbf{X}$, then $\lambda^*(W(\mathbf{x})) = \lambda(\mathbf{x})$.
- A test T^* of size α is a *uniformly most powerful* (UMP) test if and only if $\beta_{T^*}(\theta) \geq \beta_T(\theta)$ for all $\theta \in \Theta_0^c$ and T of level α.
- **Theorem 8.3.12. (Neyman-Pearson Lemma)** Consider $H_0 : \theta = \theta_0$ v.s. $H_1 : \theta = \theta_1$, then

  $$T^*(\mathbf{X}) = \begin{cases} 1 & f(\mathbf{X} \mid \theta_1) > cf(\mathbf{X} \mid \theta_0) \\ 0 & f(\mathbf{X} \mid \theta_1) < cf(\mathbf{X} \mid \theta_0) \end{cases},$$

  is a UMP test of size $\alpha_c := E(T^* \mid \theta_0)$. Note that $f(\mathbf{X} \mid \theta_1) = cf(\mathbf{X} \mid \theta_0)$ can be arbitraty.

2 Questions

1. Suppose that we observe m i.i.d Ber(θ) random variables, denoted by $Y_1, \ldots, Y_m$. Show that the LRT of $H_0 : \theta \le \theta_0$ v.s. $H_1 : \theta > \theta_0$ will reject H_0 if $\sum_{i=1}^m Y_i > b$.

$$\{x \mid \lambda(x) \le c\} \equiv \{\Sigma Y_i > b\}$$

2. A random sample $X_1, \ldots, X_n$ is drawn from a Pareto distribution with pdf

$$f(x \mid \theta, \nu) = \frac{\theta \nu^\theta}{x^{\theta+1}} I_{[\nu,\infty)}(x), \quad \theta > 0, \nu > 0.$$

   (a) Find the MLEs of θ and ν.

   (b) Show that the LRT of

$$H_0 : \theta = 1, \quad \text{v.s.} \quad H_1 : \theta \ne 1, \ \nu \text{ unknown},$$

   has reject region of the form $\{\mathbf{x} : T(\mathbf{x}) \le c_1 \text{ or } T(\mathbf{x}) \ge c_2\}$, where $0 < c_1 < c_2$ and

$$T = \log\left[\frac{\prod_{i=1}^n X_i}{X_{(1)}^n}\right].$$

$$\{\lambda(\underset{\sim}{X}) < c\} \equiv \{g(T) < c\}$$

$$\equiv \{T < c_1 \text{ or } T > c_2\}$$

3. Suppose X is one observation from a population with distribution Beta(θ, 1).

   (a) For testing $H_0 : \theta \le 1$ v.s. $H_1 : \theta > 1$, find the size and the power function of the test that rejects H_0 if $X > 0.5$.

   (b) Find the most powerful level α test of $H_0 : \theta = 1$ v.s. $H_1 : \theta = 2$.

Q_1: $Y_1, \ldots, Y_m \sim Ber(\theta)$

$$\begin{cases} H_0: & \theta \le \theta_0 \\ H_0: & \theta > \theta_0 \end{cases}$$

$$\Lambda(\underset{\sim}{Y}) = \frac{\sup\limits_{\boxed{\theta \le \theta_0}} L(\theta \mid \underset{\sim}{Y})}{\boxed{\sup\limits_{\theta \in \Theta} L(\theta \mid \underset{\sim}{Y})}}$$

$\hat{\theta} = \bar{X}$

Find the solution for numerator:

$$L(\theta \mid \underset{\sim}{Y}) = \binom{m}{\sum Y_i} \theta^{\sum Y_i} (1-\theta)^{m - \sum Y_i}$$

$$l(\theta \mid \underset{\sim}{Y}) = \log \binom{m}{\sum Y_i} + \sum Y_i \log \theta + \left(m - \sum Y_i\right) \log(1-\theta)$$

$$l'(\theta \mid \underset{\sim}{Y}) = 0 + \frac{\sum Y_i}{\theta} - \frac{m - \sum Y_i}{1-\theta}$$

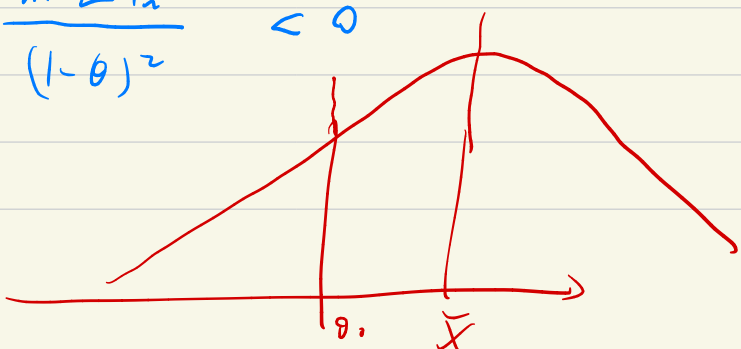
Solve $l'(\theta \mid \underset{\sim}{Y}) = 0 \Rightarrow \hat{\theta} = \bar{X}$

But, $\bar{X}$ may not in $(-\infty, \theta_0)$

$$l''(\theta \mid \underset{\sim}{Y}) = -\frac{\sum Y_i}{\theta^2} - \frac{m - \sum Y_i}{(1-\theta)^2} < 0$$

$\Rightarrow$ l is concave

So if $\bar{X} > \theta_0$



$\Rightarrow$ θ_0 is the maximizier.

So in the numerator, $\hat{\theta} = \min\{\bar{X}, \theta_0\}$.

$$\Lambda(\underset{\sim}{Y}) = \frac{L(\hat{\theta} \mid \underset{\sim}{Y})}{L(\bar{X} \mid \underset{\sim}{Y})} = \frac{\cancel{\binom{m}{\sum Y_i}} \hat{\theta}^{\sum Y_i} (1-\hat{\theta})^{m-\sum Y_i}}{\cancel{\binom{m}{\sum Y_i}} \bar{X}^{\sum Y_i} (1-\bar{X})^{m-\sum Y_i}}$$

$$= \begin{cases} \boxed{1} & \text{if } \underline{\bar{X} < \theta_0} \\ \left(\frac{\theta_0}{\bar{X}}\right)^{\sum Y_i} \left(\frac{1-\theta_0}{1-\bar{X}}\right)^{m-\sum Y_i} & \boxed{\bar{X} > \theta_0} \end{cases}$$

Rejection region,

$$\Bigg\{ \underbrace{\left(\frac{\theta_0}{\bar{X}}\right)^{m\bar{X}} \left(\frac{1-\theta_0}{1-\bar{X}}\right)^{m(1-\bar{X})}}_{\substack{\| \\ g(\bar{X})}} < c \Bigg\}$$

If $g(\bar{x}) \downarrow$

we have equivalent RR.

$\Rightarrow$ $\{ \bar{X} > b \}$.

$$g(t) = \exp\left[mt\{\log\theta_0 - \log t\} + m(1-t)\{\log(1-\theta_0) - \log(1-t)\} \right]$$

$$g'(t) = g(t) \times \left[m(\log\theta_0 - \log t) + \cancel{mt}\left[-\frac{1}{t}\right] - m(\log(1-\theta_0) - \log(1-t)) + \cancel{m(1-t)}\frac{1}{1-t} \right]$$

$$= \underbrace{g(t)}_{>0} \left[m\underbrace{\left\{\log\frac{\theta_0}{t} - \log\frac{1-\theta_0}{1-t}\right\}}_{\leq 0} \right]$$

$t > \theta_0$

$$\Rightarrow \frac{\theta_0}{t} < 1 \Rightarrow \log\frac{\theta_0}{t} < 0$$

$$\Rightarrow \frac{1-\theta_0}{1-t} > 1 \Rightarrow \log\frac{1-\theta}{1-t} > 0$$

So $g'(t) < 0$

$\Rightarrow \{\bar{X} > b\}$ is the RR for LRT #

Q2: $f(x|\theta, \nu) = \frac{\theta \nu^\theta}{x^{\theta+1}} \mathbb{1}\{x > \nu\}$

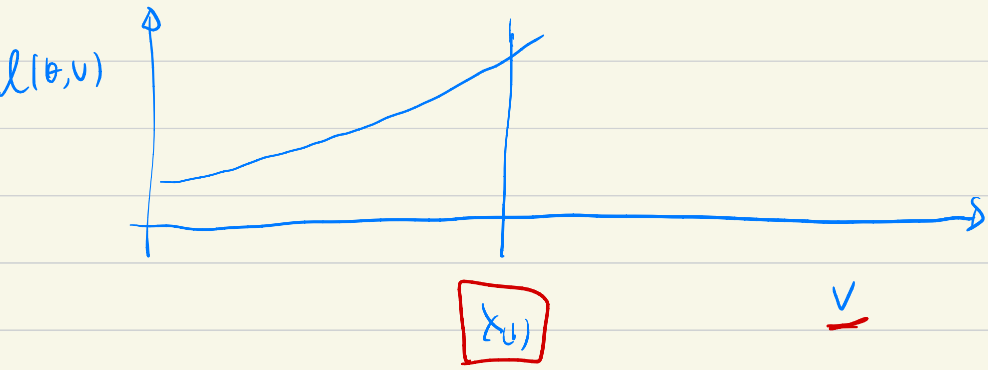
a) MLE

$$L(\theta,\nu) = \theta^n \nu^{n\theta} \left(\prod x_i\right)^{-\theta-1} \mathbb{1}\{X_{(1)} \geq \nu\}$$

Consider $\boxed{\nu < X_{(1)}}$

$$\ell(\theta,\nu) = n\log\theta + n\theta\log\nu - (\theta+1)\log\prod X_i$$

$$\partial_\nu \ell(\theta,\nu) = 0 + \frac{n\theta}{\nu} + 0 > 0$$



So $\hat{\nu} = X_{(1)}$

$$\partial_\theta \ell(\theta, X_{(1)}) = \frac{n}{\theta} + n\log X_{(1)} - \log\prod X_i$$

$$= \frac{n}{\theta} - \boxed{\log\prod\frac{X_i}{X_{(1)}}} = T$$

Solve $\partial_\theta \ell(\theta, X_{(1)}) = 0$

$$\hat{\theta} = \frac{n}{T}$$

Check $\frac{d}{d\theta^2} l(\theta, X_{(1)}) = -\frac{n}{\theta^2} < 0$

So we have the MLE $\left(X_{(1)}, \frac{n}{T}\right)$ #

b)

$$\Lambda(\underset{\sim}{X}) = \frac{L(\theta=1, \nu = X_{(1)})}{L(\theta=\frac{n}{T}, \nu = X_{(1)})}$$

$$= \frac{X_{(1)}^{n} \left(\prod X_i\right)^{-2}}{\left(\frac{n}{T}\right)^{n} X_{(1)}^{\frac{n^2}{T}} \left(\prod X_i\right)^{-\frac{n}{T}-1}}$$

$$= \frac{T^n}{n^n} X_{(1)}^{n\left(1-\frac{n}{T}\right)} \left(\prod X_i\right)^{-\left(1-\frac{n}{T}\right)}$$

$$= \frac{T^n}{n^n} \left\{ \underbrace{\boxed{\frac{\prod X_i}{X_{(1)}^{n}}}}_{= e^T} \right\}^{-\left(1-\frac{n}{T}\right)}$$

$$= \frac{T^n}{n^n} e^{-T+n}$$

The RR for LRT

$$\left\{ \frac{T^n}{n^n} e^{-T+n} < c \right\}$$

$$= \left\{ \underbrace{T^n e^{-T}}_{= g(T)} < c' \right\}$$

$$\left[\begin{array}{l} g(0) = 0 \\ \lim_{t \to \infty} g(t) = 0 \end{array} \right]$$

$$g'(T) = n T^{n-1} e^{-T} - T^n e^{-T}$$

$$g'(T) = 0 \quad \Rightarrow \quad n T^{n-1} \cancel{e^{-T}} = T^n \cancel{e^{-T}}$$

$$\Rightarrow \quad n T^{n-1} = T^n$$

$$\Rightarrow \quad \boxed{n} = T$$

$\because$ we only have one solution.

$\Rightarrow$

So, the RR for LLT

$$\{ T < c_1 , \quad T > c_2 \}$$

$\#$

Q3: $X \sim Beta(\theta, 1)$ $f_X(x) = \frac{1}{Beta(\theta,1)} x^{\theta-1}$ $x \in (0,1)$

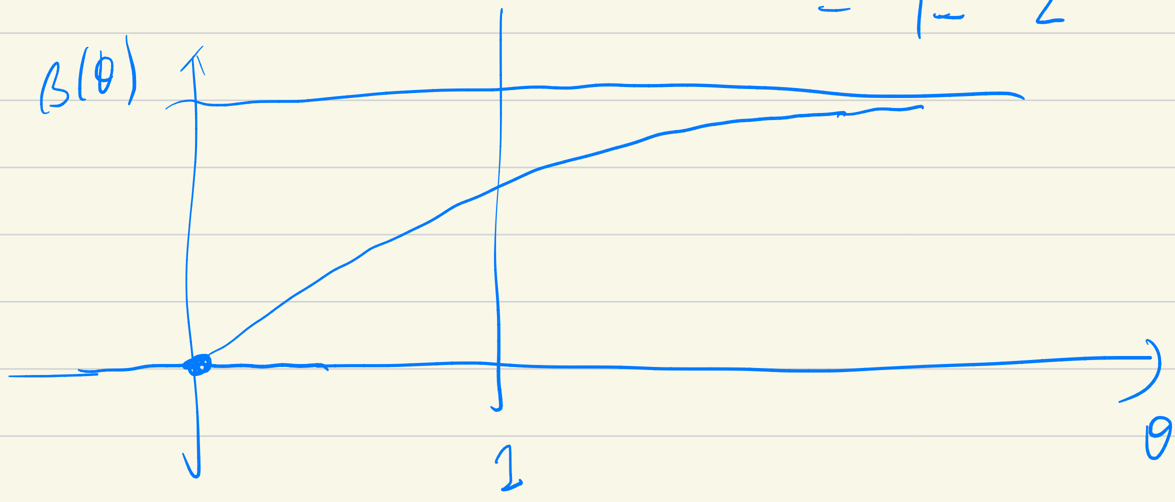
$= \theta x^{\theta-1}$ $x \in (0,1)$

a) $\{X > 0.5\}$

$\beta(\theta) = P(X > 0.5 \mid \theta)$

$= \int_{0.5}^{1} \theta x^{\theta-1} dx = x^{\theta} \Big|_{0.5}^{1}$

$= 1 - 2^{-\theta}$



$\Rightarrow$ Size is $\beta(1) = 1 - \frac{1}{2} = \frac{1}{2}$

power is $1 - 2^{-\theta}$ #

b) Apply N-P Lemma.

$\because$ X is continuous

$\Rightarrow$ don't need to discuss $f_1(x) = c\, f_0(x)$

$f_1(x) = 2x$

$f_0(x) = 1$

RR $\{ f_1(x) > c\, f_0(x) \}$

$= \{ 2x > c \}$

$\Rightarrow \{ x > c \}$

By NP this is the RR for UMP-test.

$\alpha = P(X > c \mid \theta = 1)$

$= \int_c^1 1\, dx = 1 - c$

So $c = 1 - \alpha$

So the UMP test for α

is $\{ X > 1-\alpha \}$ #