Lecture 14: Most Powerful Tests

Mathematical Statistics II, MATH 60062/70062

Thursday March 10, 2022

Reference: Casella & Berger, 8.3.2

Recap: Neyman-Pearson Lemma

Consider testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1,$$

where H_0 and H_1 are both simple hypotheses. The PDFs/PMFs of $\boldsymbol{X} = (X_1, \ldots, X_n)$ corresponding to θ_0 and θ_1 are $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0)$ and $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1)$, respectively. Consider the test function

$$\phi(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_0)} > k \\ 0 & \text{if } \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_0)} < k, \end{cases}$$

rejection region

for $k \geq 0$, where

\alpha and rejection region changes with respect to k

$$\alpha = P_{\theta_0}(\boldsymbol{X} \in R) = E_{\theta_0}[\phi(\boldsymbol{X})].$$

Any test satisfying the above definition of $\phi(\boldsymbol{x})$ is a **(uniformly) most power level α test**.

Most powerful Poisson test

Suppose that $X_1, \ldots, X_n$ are iid $\mathrm{Pois}(\theta)$, where $\theta > 0$. Find the MP level α test for

$$H_0 : \theta = 10 \quad \text{versus} \quad H_1 : \theta = 12.$$

Form the ratio

$$\frac{f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0)} = \frac{f_{\boldsymbol{X}}(\boldsymbol{x} \mid 12)}{f_{\boldsymbol{X}}(\boldsymbol{x} \mid 10)} = \frac{\prod_{i=1}^n \frac{e^{-12} 12^{x_i}}{x_i!}}{\prod_{i=1}^n \frac{e^{-10} 10^{x_i}}{x_i!}} = e^{-2n} 1.2^{\sum_{i=1}^n x_i}.$$

By the Neyman-Pearson Lemma, the MP level α test uses the rejection region

$$R = \left\{ \boldsymbol{x} \in \mathcal{X} : e^{-2n} 1.2^{\sum_{i=1}^n x_i} > k \right\}$$

where k satisfies

$$\alpha = P_{\theta=10}(\boldsymbol{X} \in R) = P\left(e^{-2n} 1.2^{\sum_{i=1}^n X_i} > k \mid \theta = 10 \right).$$

Note that

$$e^{-2n}1.2^{\sum_{i=1}^{n} x_i} > k \iff \sum_{i=1}^{n} x_i > \log_{1.2}(ke^{2n}) = k'.$$

Thus, an equivalent rejection region can be defined through the sufficient statistic $T = T(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$:

$$R = \{\boldsymbol{x} \in \mathcal{X} : T(\boldsymbol{x}) > k'\},$$

where k' is chosen such that

$$\alpha = P_{\theta=10}(\boldsymbol{X} \in R) = P(T(\boldsymbol{X}) > k' \mid \theta = 10),$$

where $T(\boldsymbol{X}) \sim \mathrm{Pois}(n)$.

Note that

$$e^{-2n}1.2^{\sum_{i=1}^n x_i} > k \iff \sum_{i=1}^n x_i > \log_{1.2}(ke^{2n}) = k'.$$

Thus, an equivalent rejection region can be defined through the sufficient statistic $T = T(\boldsymbol{X}) = \sum_{i=1}^n X_i$:

$$R = \{\boldsymbol{x} \in \mathcal{X} : T(\boldsymbol{x}) > k'\},$$

where k' is chosen such that

$$\alpha = P_{\theta=10}(\boldsymbol{X} \in R) = P(T(\boldsymbol{X}) > k' \mid \theta = 10),$$

where $T(\boldsymbol{X}) \sim \mathrm{Pois}(\boldsymbol{n})$. 10n

When $n = 5$,

- $57 \leq k' < 58$, $\alpha = 0.116$; $58 \leq k' < 59$, $\alpha = 0.092$
- $60 \leq k' < 61$, $\alpha = 0.056$; $61 \leq k' < 62$, $\alpha = 0.042$

Proof of Neyman-Pearson Lemma

We will prove the sufficiency part only.

Define the test function

$$\phi(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_0)} > k \\ 0 & \text{if } \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\theta_0)} < k, \end{cases}$$

where $k \geq 0$, and $\alpha = P_{\theta_0}(\boldsymbol{X} \in R) = E_{\theta_0}[\phi(\boldsymbol{X})]$. Note that $\phi(\boldsymbol{x})$ is a size α test (and of course, a level α test).

We want to show that $\phi(\boldsymbol{x})$ is a MP level α test. That is, for any other level α test with test function $0 \leq \phi^*(\boldsymbol{x}) \leq 1$ that satisfies $E_{\theta_0}[\phi^*(\boldsymbol{X})] \leq \alpha$,

$$E_{\theta_1}[\phi(\boldsymbol{X})] \geq E_{\theta_1}[\phi^*(\boldsymbol{X})].$$

\beta(\theta) = P_{\theta}(x \in R) = E_{\theta}[\phi(x)]

Since $E_{\theta_0}[\phi(\boldsymbol{X})] = \alpha$ and $E_{\theta_0}[\phi^*(\boldsymbol{X})] \le \alpha$,

$$E_{\theta_0}[\phi(\boldsymbol{X}) - \phi^*(\boldsymbol{X})] = E_{\theta_0}[\phi(\boldsymbol{X})] - E_{\theta_0}[\phi^*(\boldsymbol{X})] \ge 0.$$

Consider the function

$$b(\boldsymbol{x}) = [\phi(\boldsymbol{x}) - \phi^*(\boldsymbol{x})][f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0)].$$

For all $\boldsymbol{x} \in \mathcal{X}$, $b(\boldsymbol{x}) \ge 0$, as in the following conditions.

- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) > 0$, $\phi(\boldsymbol{x}) = 1$ and $b(\boldsymbol{x}) \ge 0$.
- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) < 0$, $\phi(\boldsymbol{x}) = 0$ and $b(\boldsymbol{x}) \ge 0$.
- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) < 0$, $b(\boldsymbol{x}) = 0$.

Since $E_{\theta_0}[\phi(\boldsymbol{X})] = \alpha$ and $E_{\theta_0}[\phi^*(\boldsymbol{X})] \leq \alpha$,

$$E_{\theta_0}[\phi(\boldsymbol{X}) - \phi^*(\boldsymbol{X})] = E_{\theta_0}[\phi(\boldsymbol{X})] - E_{\theta_0}[\phi^*(\boldsymbol{X})] \geq 0.$$

Consider the function

$$b(\boldsymbol{x}) = [\phi(\boldsymbol{x}) - \phi^*(\boldsymbol{x})][f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0)].$$

For all $\boldsymbol{x} \in \mathcal{X}$, $b(\boldsymbol{x}) \geq 0$, as in the following conditions.

- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) > 0$, $\phi(\boldsymbol{x}) = 1$ and $b(\boldsymbol{x}) \geq 0$.
- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) < 0$, $\phi(\boldsymbol{x}) = 0$ and $b(\boldsymbol{x}) \geq 0$.
- When $f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) - k f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0) < 0$, $b(\boldsymbol{x}) = 0$.

Therefore,

$$[\phi(\boldsymbol{x}) - \phi^*(\boldsymbol{x})] f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_1) \geq k[\phi(\boldsymbol{x}) - \phi^*(\boldsymbol{x})] f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta_0)$$

Integrating both sides, we get

$$E_{\theta_1}[\phi(\boldsymbol{X}) - \phi^*(\boldsymbol{X})] \geq k E_{\theta_0}[\phi(\boldsymbol{X}) - \phi^*(\boldsymbol{X})] \geq 0$$

$$\implies E_{\theta_1}[\phi(\boldsymbol{X})] \geq E_{\theta_1}[\phi^*(\boldsymbol{X})].$$

Neyman-Pearson Lemma with a sufficient statistic

Consider testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1,$$

and suppose that $T = T(\boldsymbol{X})$ is a sufficient statistic. The PDFs/PMFs of T corresponding to θ_0 and θ_1 are $g_T(t \mid \theta_0)$ and $g_T(t \mid \theta_1)$, respectively. Consider the test function

$$\phi(t) = \begin{cases} 1 & \text{if } \frac{g_T(t|\theta_1)}{g_T(t|\theta_0)} > k \\ 0 & \text{if } \frac{g_T(t|\theta_1)}{g_T(t|\theta_0)} < k, \end{cases}$$

for $k \geq 0$, where, with rejection region $S \subset \mathcal{T}$

$$\alpha = P_{\theta_0}(T \in S) = E_{\theta_0}[\phi(T)].$$

Any test satisfying the definition is a most power level α test.

Most powerful Normal test

Suppose that $X_1, \ldots, X_n$ are iid $\mathcal{N}(\mu, \sigma_0^2)$, where $-\infty < \mu < \infty$ and σ_0^2 is known. Find the MP level α test for

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu = \mu_1,$$

where $\mu_0 > \mu_1$.

The sample mean $T = T(\boldsymbol{X}) = \bar{X}$ is a sufficient statistic for $\mathcal{N}(\mu, \sigma_0^2)$, and is distributed as $T \sim \mathcal{N}(\mu, \sigma_0^2/n)$,

$$g_T(t \mid \mu) = \frac{1}{\sqrt{2\pi\sigma_0^2/n}} e^{-\frac{n}{2\sigma_0^2}(t-\mu)^2}.$$

Consider the ratio

$$\frac{g_T(t \mid \mu_1)}{g_T(t \mid \mu_0)} = \frac{\frac{1}{\sqrt{2\pi\sigma_0^2/n}} e^{-\frac{n}{2\sigma_0^2}(t-\mu_1)^2}}{\frac{1}{\sqrt{2\pi\sigma_0^2/n}} e^{-\frac{n}{2\sigma_0^2}(t-\mu_0)^2}} = e^{-\frac{n}{2\sigma_0^2}[(t-\mu_1)^2-(t-\mu_0)^2]}.$$

The MP level α test rejects H_0 when

$$e^{-\frac{n}{2\sigma_0^2}[(t-\mu_1)^2-(t-\mu_0)^2]} > k \iff t < \frac{(2\sigma_0^2 \log k)/n + \mu_1^2 - \mu_0^2}{2(\mu_1 - \mu_0)} = k',$$

where k' satisfies

$$\alpha = P_{\mu_0}(T < k') = P\left(Z < \frac{k' - \mu_0}{\sigma_0/\sqrt{n}}\right)$$

$$\implies \frac{k' - \mu_0}{\sigma_0/\sqrt{n}} = -z_\alpha \implies t' = \mu_0 - z_\alpha \sigma_0/\sqrt{n}.$$

Thus, the MP level α test rejects H_0 when $\bar{X} < \mu_0 - z_\alpha \sigma_0/\sqrt{n}$.

Uniformly most powerful tests

Let $\mathcal{C}$ be a class of tests for testing

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta > \Theta_0^c.$$

A test in class $\mathcal{C}$, with power function $\beta(\theta)$, is a **uniformly most powerful (UMP) class $\mathcal{C}$ test** if

$$\beta(\theta) \geq \beta^*(\theta)$$

for all $\theta \in \Theta_0^c$, where $\beta^*(\theta)$ is the power function of any other test in $\mathcal{C}$.

The **Neyman-Pearson Lemma** is only applicable to test *simple-versus-simple* hypotheses, not to problems involving *composite* hypotheses. E.g.,

- $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$
- $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

Monotone likelihood ratio

A family of PDFs/PMFs $\{g_T(t \mid \theta) : \theta \in \Theta\}$ for a univariate random variable T has a **monotone likelihood ratio (MLR)** if for all $\theta_2 > \theta_1$, the ratio

$$\frac{g_T(t \mid \theta_2)}{g_T(t \mid \theta_1)}$$

is a monotone (nonincreasing or nondecreasing) function of t on $\{t : g_T(t \mid \theta_1) > 0 \text{ or } g_T(t \mid \theta_2) > 0\}$.

Monotone likelihood ratio

A family of PDFs/PMFs $\{g_T(t \mid \theta) : \theta \in \Theta\}$ for a univariate random variable T has a **monotone likelihood ratio (MLR)** if for all $\theta_2 > \theta_1$, the ratio

$$\frac{g_T(t \mid \theta_2)}{g_T(t \mid \theta_1)}$$

is a monotone (nonincreasing or nondecreasing) function of t on $\{t : g_T(t \mid \theta_1) > 0 \text{ or } g_T(t \mid \theta_2) > 0\}$.

E.g., the family of $T \sim \mathrm{Bin}(n, \theta)$ has an MLR, since

$$\frac{g_T(t \mid \theta_2)}{g_T(t \mid \theta_1)} = \frac{\binom{n}{t}\theta_2^t(1-\theta_2)^{n-t}}{\binom{n}{t}\theta_1^t(1-\theta_1)^{n-t}} = \left(\frac{1-\theta_2}{1-\theta_1}\right)^n \left[\frac{\theta_2(1-\theta_1)}{\theta_1(1-\theta_2)}\right]^t > 0$$

is an increasing function of t for all $\theta_2 > \theta_1$.

Monotone likelihood ratio

A family of PDFs/PMFs $\{g_T(t \mid \theta) : \theta \in \Theta\}$ for a univariate random variable T has a **monotone likelihood ratio (MLR)** if for all $\theta_2 > \theta_1$, the ratio

$$\frac{g_T(t \mid \theta_2)}{g_T(t \mid \theta_1)}$$

is a monotone (nonincreasing or nondecreasing) function of t on $\{t : g_T(t \mid \theta_1) > 0 \text{ or } g_T(t \mid \theta_2) > 0\}$.

E.g., the family of $T \sim \mathrm{Bin}(n, \theta)$ has an MLR, since

$$\frac{g_T(t \mid \theta_2)}{g_T(t \mid \theta_1)} = \frac{\binom{n}{t}\theta_2^t(1-\theta_2)^{n-t}}{\binom{n}{t}\theta_1^t(1-\theta_1)^{n-t}} = \left(\frac{1-\theta_2}{1-\theta_1}\right)^n \left[\frac{\theta_2(1-\theta_1)}{\theta_1(1-\theta_2)}\right]^t > 0$$

is an increasing function of t for all $\theta_2 > \theta_1$.

Note: If $T \sim g_T(t \mid \theta) = h(t)c(\theta)e^{w(\theta)t}$, then $\{g_T(t \mid \theta) : \theta \in \Theta\}$ has an MLR.

Karlin-Rubin Theorem

Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

Suppose that T is a sufficient statistic for θ and the family $\{g_T(t \mid \theta) : \theta \in \Theta\}$ has an MLR. Then the test that rejects H_0 if and only if $T > t_0$ is a UMP level α test, where

$$\alpha = P_{\theta_0}(T > t_0).$$

Karlin-Rubin Theorem

Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

Suppose that T is a sufficient statistic for θ and the family $\{g_T(t \mid \theta) : \theta \in \Theta\}$ has an MLR. Then the test that rejects H_0 if and only if $T > t_0$ is a UMP level α test, where

$$\alpha = P_{\theta_0}(T > t_0).$$

Similarly, when testing

$$H_0 : \theta \geq \theta_0 \quad \text{versus} \quad H_1 : \theta < \theta_0,$$

the test that rejects H_0 if and only if $T < t_0$ is a UMP level α test, where $\alpha = P_{\theta_0}(T < t_0)$.