STA 522, Spring 2022
Introduction to Theoretical Statistics II

Lecture 11

Department of Biostatistics
University at Buffalo

AGENDA

- Extending Neyman-Pearson Lemma, MLR family
- non-existence of UMP tests
- Interval Estimation
- Method of Finding Interval Estimates

Review: Neyman Pearson Lemma & Most Powerful Tests

- Consider testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, where (1) the pdf or pmf corresponding to θ_i is $f(\underline{x} \mid \theta_i)$ for $i = 0, 1$; (2) the test has a rejection region R that satisfies $\underline{x} \in R \quad \text{if} \quad f(\underline{x} \mid \theta_1) > kf(\underline{x} \mid \theta_0)$ and $\underline{x} \in R^c \quad \text{if} \quad f(\underline{x} \mid \theta_1) < kf(\underline{x} \mid \theta_0)$ for some $k \geq 0$; and (3) $\alpha = P_{\theta_0}(\underline{X} \in R)$.
- Then (a) **(Sufficiency)** any test that satisfies (2) and (3) above is a UMP level α test; and (b) **(Necessity)** if there exists a test satisfying (2) and (3) above with $k > 0$, then every UMP level α test is a size α test (satisfies (3) above), and every UMP level α test satisfies (2) above, except perhaps on a set A satisfying $P_{\theta_0}(\underline{X} \in A) = P_{\theta_1}(\underline{X} \in A) = 0$.
- Suppose $T(\underline{X})$ is a sufficient statistic for θ, and let $g(t \mid \theta_i)$ be the pdf or pmf of T corresponding to θ_i for $i = 0, 1$. Then any test based on T with rejection region S (a subset of the sample space of T) is a UMP level α test if it satisfies (1) for some $k \geq 0$, $t \in S$ if $g(t \mid \theta_1) > kg(t \mid \theta_0)$ and $t \in S^c$ if $g(t \mid \theta_1) < kg(t \mid \theta_0)$

Extending the Neyman-Pearson Lemma

- Can we extend the Neyman-Pearson Lemma to composite hypotheses (hypotheses that specify more than one possible distribution for the sample)?

– Yes, but only for one-sided hypotheses ($H : \theta \geq \theta_0$ or $H : \theta < \theta_0$).

– A UMP level α test must be UMP for all values in the alternative hypothesis.

Monotone Likelihood Ratio (MLR)

Definition: A family of pdfs or pmfs $\{g(t \mid \theta) : \theta \in \Theta\}$ for a univariate random variable T with real-valued parameter θ has a **monotone likelihood ratio (MLR)** if, for every $\theta_2 > \theta_1$,

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

is a monotone (non-increasing or non-decreasing) function of t on

$$\{t : g(t \mid \theta_1) > 0 \text{ or } g(t \mid \theta_2) > 0\}.$$

Comments About MLR

- MLR is a property of a family of distributions.
- $\mathrm{N}(\theta, \sigma^2)$ (with σ^2 known), poisson(θ), and binomial(n, θ) all have an MLR.
- In general, any regular exponential family

Karlin-Rubin Theorem

Theorem

Consider testing $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$. Suppose that T is a sufficient statistic for θ and the family of pdfs or pmfs $\{g(t \,|\, \theta) \,:\, \theta \in \Theta\}$ of T has an MLR. Then for any t_0, the test that rejects H_0 if and only if $T > t_0$ is a UMP level α test, where $\alpha = P_{\theta_0}(T > t_0)$.

Example (Contd.): Let $X_1, X_2, \ldots, X_n \sim$ iid $N(\theta, \sigma^2)$ population, σ^2 known. Consider testing $H_0' : \theta \geq \theta_0$ vs. $H_1' : \theta < \theta_0$.

Consider the test that rejects H_0' if $\overline{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$. $\overline{X}$ is sufficient.

We'll show that the distribution of $T = \overline{X}$ has an MLR, and apply the Karlin-Rubin theorem.

For $\theta_2 > \theta_1$:

$$\begin{aligned}\frac{g(t \mid \theta_1)}{g(t \mid \theta_2)} &= \frac{\exp\left(-\frac{n}{2\sigma^2}(t-\theta_2)^2\right)}{\exp\left(-\frac{n}{2\sigma^2}(t-\theta_1)^2\right)} \\ &= \exp\left[\frac{n}{\sigma^2}t(\theta_2-\theta_1)\right]\exp\left[-\frac{n}{2\sigma^2}(\theta_2^2-\theta_1^2)\right]\end{aligned}$$

which is non-decreasing in t as $\theta_2 - \theta_1 > 0$.

Thus the distribution of $T = \overline{X}$ has an MLR.

Therefore, from Karlin-Rubin theorem it follows that this test is UMP level α for this problem.

Nonexistence of UMP Test

Example: Let $X_1, X_2, \ldots, X_n \sim$ iid $\mathrm{N}(\theta, \sigma^2)$, with σ^2 known. Consider testing

$$H_0 : \theta = \theta_0$$
$$\text{vs. } H_1 : \theta \neq \theta_0.$$

We'll show that there does not any UMP test at any level $0 < \alpha < 1$.

For a specified value of α, a level α test in this problem is any test that satisfies

$$P_{\theta_0}(\text{reject } H_0) \leq \alpha.$$

Suppose $\theta_1 < \theta_0$. By Corollary to the NP Lemma with sufficient statistic, the test with rejection region

$$R = \left\{ \underline{x} \, : \, \overline{x} < \theta_0 - \frac{\sigma z_\alpha}{\sqrt{n}} \right\}$$

has the highest possible power at θ_1; call this Test 1.

By part (b) of the NP Lemma, any other level α test that has the same power as Test 1 at θ_1 must have the same rejection region, except possibly for a set A with measure zero.

So if a UMP level α test exists, it must be Test 1, since no other level α test has as high a power as Test 1 at θ_1.

Now consider Test 2, which has rejection region

$$R = \left\{ \underline{x} \,:\, \overline{x} > \theta_0 + \frac{\sigma z_\alpha}{\sqrt{n}} \right\}.$$

This is also a level α test.

We can show that for any $\theta_2 > \theta_0$, $\beta_2(\theta_2) > \beta_1(\theta_2)$.

So Test 1 cannot be a UMP level α test, since Test 2 has a higher power than Test 1 at θ_2.

Therefore, no UMP level α test exists in this problem.

Since a global UMP test does not exist, we can restrict to the class of unbiased tests. (Recall that for an unbiased test the power function at each $\theta \in \Theta_0^c$ is $\geq$ the level of the test.)

Consider Test 3, which rejects $H_0 : \theta = \theta_0$ in favor of $H_1 : \theta \neq \theta_0$ if and only if

$$\overline{X} > \theta_0 + \sigma z_{\alpha/2}/\sqrt{n} \ \text{ or } \ \overline{X} < \theta_0 - \sigma z_{\alpha/2}/\sqrt{n}$$

is actually a UMP unbiased level α test; i.e., it is UMP in the class of unbiased tests.

p-Values

Definition: A p-**value**, $p(\underline{X})$, is a test statistic satisfying $0 \leq p(\underline{x}) \leq 1$ for every sample point $\underline{x}$. Small values of $p(\underline{X})$ give evidence that H_1 is true. A p-value is **valid** if, for every $\theta \in \Theta_0$ and every $0 \leq \alpha \leq 1$,

$$P_\theta(p(\underline{X}) \leq \alpha) \leq \alpha.$$

If $p(\underline{X})$ is a valid p-value, then the test that rejects H_0 if and only if $p(\underline{X}) \leq \alpha$ is a level α test.

Theorem (8.3.27; Determining Valid p-Values)

Let $W(\underline{X})$ be a test statistic such that large values of W give evidence that H_1 is true. For each sample point $\underline{x}$, define

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} P_\theta \left[W(\underline{X}) \geq W(\underline{x}) \right].$$

Then $p(\underline{X})$ is a valid p-value.

Interval Estimation

Defintion: An **interval estimate** of a real-valued parameter θ is any pair of functions, $L(\underline{x})$ and $U(\underline{x})$, of a sample that satisfy $L(\underline{x}) \leq U(\underline{x})$ for all $\underline{x} \in \mathcal{X}$. If $\underline{X} = \underline{x}$ is observed, the inference $L(\underline{x}) \leq \theta \leq U(\underline{x})$ is made. The random interval $[L(\underline{X}), U(\underline{X})]$ is called an **interval estimator**.

Example: Suppose $X_1, X_2, X_3, X_4 \sim$ iid $\mathrm{N}(\mu, 1)$. Then $[\overline{X} - 1, \overline{X} + 1]$ is an interval estimator for the population mean μ. What is $P\left(\mu \in [\overline{X} - 1, \overline{X} + 1]\right)$?

Note that interval estimators are less precise than point estimators, but are more likely to be correct.

Recall that $P(\overline{X} = \mu) = 0$, for instance, i.e., there is no chance we are correct if we estimate μ using $\overline{X}$.

Coverage Probability & Confidence coefficient

Definition: For an interval estimator $[L(\underline{X}), U(\underline{X})]$ of a parameter θ, the **coverage probability** of $[L(\underline{X}), U(\underline{X})]$ is the probability that the random interval $[L(\underline{X}), U(\underline{X})]$ covers the true parameter θ.

Symbolically, it is denoted by either $P_\theta\left(\theta \in [L(\underline{X}), U(\underline{X})]\right)$ or $P\left(\theta \in [L(\underline{X}), U(\underline{X})] \,|\, \theta\right)$.

Note that the coverage probability is usually a function of θ.

Definition: For an interval estimator $[L(\underline{X}), U(\underline{X})]$ of a parameter θ, the **confidence coefficient** of $[L(\underline{X}), U(\underline{X})]$ is the infimum of the coverage probabilities,

$$\inf_\theta P_\theta\left(\theta \in [L(\underline{X}), U(\underline{X})]\right).$$

We use **confidence interval** to mean the interval estimator along with its corresponding confidence coefficient.

Note that since θ is fixed, but unknown, the probability statements above refer to $\underline{X}$, not θ. We can think of such a probability as $P_\theta\left(L(\underline{X}) \le \theta, U(\underline{X}) \ge \theta\right)$.

Example (Scale Uniform Interval Estimator): Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform$(0, \theta)$ population, and let $Y = X_{(n)}$ be the nth order statistic.

We are interested in an interval estimator of θ.

We consider two candidate estimators:

(a) $[aY, bY]$, where $1 \leq a < b$; and

(b) $[Y + c, Y + d]$, where $0 \leq c < d$.

Note that θ is necessarily larger than y.

Determine the coverage probability and confidence coefficient for each estimator.

(a) We have

$$\begin{aligned}
P_\theta(\theta \in [aY, bY]) &= P_\theta(aY \le \theta \le bY) \\
&= P_\theta\left(\frac{1}{b} \le \frac{Y}{\theta} \le \frac{1}{a}\right) \\
&= P_\theta\left(\frac{1}{b} \le T \le \frac{1}{a}\right) \qquad \left(T = \frac{Y}{\theta}\right)
\end{aligned}$$

The pdf of T is $f_T(t) = nt^{n-1}$, $0 \le t \le 1$. Therefore,

$$P_\theta(\theta \in [aY, bY]) = P_\theta\left(\frac{1}{b} \le T \le \frac{1}{a}\right) = \int_{1/b}^{1/a} nt^{n-1}\, dt = \left(\frac{1}{a}\right)^n - \left(\frac{1}{b}\right)^n$$

which is free of θ. So, confidence coefficient $=$

$$\inf_\theta P_\theta(\theta \in [aY, bY]) = \left(\frac{1}{a}\right)^n - \left(\frac{1}{b}\right)^n$$

(b) Here

$$P_\theta(\theta \in [Y+c, Y+d]) = \left(1-\frac{c}{\theta}\right)^n - \left(1-\frac{d}{\theta}\right)^n$$

which depends on θ. Note that

$$\lim_{\theta\to\infty} P_\theta(\theta \in [Y+c, Y+d]) = \lim_{\theta\to\infty}\left[\left(1-\frac{c}{\theta}\right)^n - \left(1-\frac{d}{\theta}\right)^n\right] = 0$$

So, confidence coefficient $= \inf_\theta P_\theta(\theta \in [Y+c, Y+d]) = 0$.

Methods of Finding Interval Estimators

(a) Invert a test statistic.

(b) Use pivotal quantities.

Correspondence Between Confidence Intervals and Hypothesis Testing

- There is a very strong correspondence between hypothesis testing and interval estimation.
- In general, every confidence interval corresponds to a test, and vice versa.

Example (Inverting a Normal Test): Let $X_1, X_2, \ldots, X_n \sim$ iid $\mathrm{N}(\mu, \sigma^2)$. Suppose we are testing $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$. Consider the rejection region

$$R = \left\{ \underline{x} \,:\, |\overline{x} - \mu_0| > z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\}.$$

The acceptance region of the hypothesis test (the subset of the sample space for which $H_0 : \mu = \mu_0$ is accepted) is

$$\begin{aligned} A(\mu_0) &= \left\{ \underline{x} \,:\, \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \overline{x} \leq \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\} \\ &= \left\{ \underline{x} \,:\, \overline{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \overline{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\}. \end{aligned}$$

Note that since $P(\underline{x} \in R \,|\, \mu = \mu_0) = \alpha$, we can deduce that

$$P(\underline{x} \in A(\mu_0) \,|\, \mu = \mu_0) = 1 - \alpha$$

for every μ_0.

So $P_\mu(\underline{x} \in A(\mu)) = 1 - \alpha$.

The $1 - \alpha$ confidence interval (the subset of the parameter space containing plausible values of μ) is

$$C(\underline{x}) = \left\{ \mu \,:\, \overline{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\}.$$

So we see that

$$\underline{x} \in A(\mu_0) \iff \mu_0 \in C(\underline{x}),$$

i.e., $\underline{x}$ is in the acceptance region for $H_0 : \mu = \mu_0$ if and only if μ_0 is a plausible value for the parameter μ.

Correspondence Between Confidence Intervals and Hypothesis Testing

- Both hypothesis tests and confidence intervals look for consistency between sample statistics and population parameters.
- The hypothesis test fixes the parameter and asks what sample values are consistent with that fixed value (the acceptance region).
- The confidence interval fixes the sample value and asks what parameter values make this sample value most plausible (the confidence interval).