STAT 713 sp 2023 Lec 11 slides

Unbiased tests, uniformly most powerful tests, Neyman-Pearson and Karlin-Rubin

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

As we asked what makes a good estimator, we now ask: What makes a good test?

Every "sensible" test should have the following property.

Unbiasedness of a test of hypotheses

A test of H_0: $\theta \in \Theta_0$ vs H_1: $\theta \in \Theta_1$ with power function γ is called *unbiased* if

$$\gamma(\theta_1) \geq \gamma(\theta_0) \text{ for all } \theta_0 \in \Theta_0 \text{ and } \theta_1 \in \Theta_1.$$

An unbiased test is more likely to reject H_0 when $\theta \in \Theta_1$ than when $\theta \in \Theta_0$.

Exercise: For $Y \sim \text{Binom}(3, p)$, H_0: $p \leq 1/2$ v H_1: $p > 1/2$, check unbiasedness:

1. Reject H_0 if $Y = 1$.
2. Reject H_0 if $Y = 3$.
3. Flip a coin and reject H_0 if it is "heads".

Consider comparing tests belonging to a class $\mathcal{C}$, for example

$$\begin{aligned} \mathcal{C}_\alpha &= \{ \text{ all tests with level } \alpha \text{ } \} \\ \mathcal{C}_{\alpha,U} &= \{ \text{ all tests with level } \alpha \text{ that are unbiased } \}. \end{aligned}$$

Uniformly most powerful tests

Given a class $\mathcal{C}$ of tests of H_0: $\theta \in \Theta_0$ vs H_1: $\theta \in \Theta_1$, a test in $\mathcal{C}$ with power curve γ is the *uniformly most powerful (UMP)* test in $\mathcal{C}$ if

$$\gamma(\theta) \geq \tilde{\gamma}(\theta) \quad \text{for all } \theta \in \Theta_1,$$

where $\tilde{\gamma}$ is the power curve of any other test in $\mathcal{C}$.

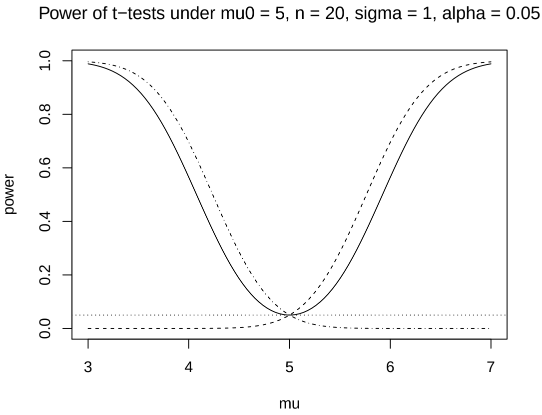
From now on we will refer by

1. *UMP test* to a test which is UMP in $\mathcal{C}_\alpha$.
2. *UMP unbiased (UMPU) test* to a test which is UMP in $\mathcal{C}_{\alpha,U}$.

Example: Let $\mathbf{X} \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$. Consider testing $H_0\colon \mu = \mu_0$ v $H_1\colon \mu \neq \mu_0$ with

$$\phi_2(\mathbf{X}) = \mathbf{1}(|T| > t_{n-1,\alpha/2}), \quad \phi_R(\mathbf{X}) = \mathbf{1}(T > t_{n-1,\alpha}), \quad \phi_L(\mathbf{X}) = \mathbf{1}(T < -t_{n-1,\alpha}),$$

where $T = \sqrt{n}(\bar{X}_n - \mu_0)/S_n$.



Discuss unbiasedness, UMP, UMPU property of ϕ_2, ϕ_R, ϕ_L.

One-sided and two-sided sets of hypotheses

Let $\tau = \tau(\theta)$ be one-dimensional.

1. Hypotheses are called *one-sided* if they are of the form

$$H_0\colon \tau \leq \tau_0 \text{ vs } H_1\colon \tau > \tau_0 \quad (\text{or } H_0\colon \tau \geq \tau_0 \text{ vs } H_1\colon \tau < \tau_0)$$

2. Hypotheses are called *two-sided* if they are of the form

$$H_0\colon \tau = \tau_0 \text{ vs } H_1\colon \tau \neq \tau_0$$

Neyman-Pearson Lemma

Let $\mathbf{X}$ have the likelihood $\mathcal{L}(\theta; \mathbf{X})$, and suppose we wish to test

$$H_0\colon \theta = \theta_0 \text{ versus } H_1\colon \theta = \theta_1.$$

Then for any $k > 0$ the test

$$\phi(\mathbf{X}) = 1 \iff \frac{\mathcal{L}(\theta_0; \mathbf{X})}{\mathcal{L}(\theta_1; \mathbf{X})} < k$$

is at least as powerful as any other test with the same or smaller size.

Moreover, it is the unique such test (ignoring events occurring w/prob 0).

N-P Lemma points us toward ratios of likelihoods in our search for UMP tests.

Exercise: Prove first part of result.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Exponential}(\beta)$, β unknown, and consider

$$H_0\colon \beta = \beta_0 \text{ versus } H_1\colon \beta = \beta_1.$$

Give the form of the most powerful test.

Corollary (UMP test a function of a sufficient statistic)

The UMP test (N-P) of H_0: $\theta = \theta_0$ vs H_1: $\theta = \theta_1$ is a function of a minimal sufficient statistic, and if $T = T(\mathbf{X})$ is sufficient, the test is equivalent to

$$\phi(\mathbf{X}) = 1 \iff \frac{f_T(T(\mathbf{X}); \theta_0)}{f_T(T(\mathbf{X}); \theta_1)} < k.$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Poisson}(\lambda)$, λ unknown, and consider

$$H_0\colon \lambda = \lambda_0 \text{ versus } H_1\colon \lambda = \lambda_1.$$

Give the form of the most powerful test.

Theorem (Karlin-Rubin: Monotone LR gives UMP for one-sided tests)

Let $\Theta \subset \mathbb{R}$ and let T be a 1-dimensional sufficient statistic with pdf/pmf $f_T(t;\theta)$.

1. *If testing H_0: $\theta \leq \theta_0$ vs H_1: $\theta > \theta_0$, choose any $\theta_1 > \theta_0$...*
2. *If testing H_0: $\theta \geq \theta_0$ vs H_1: $\theta < \theta_0$, choose any $\theta_1 < \theta_0$...*

In either case the unique UMP test is found as:

If $\dfrac{f_T(t;\theta_0)}{f_T(t;\theta_1)}$ is non-increasing (non-decr) in t, reject H_0 when $T > c$ ($T < c$).

If $\dfrac{f_T(t;\theta_0)}{f_T(t;\theta_1)}$ is monotone in t, we say T has a *monotone likelihood ratio (MLR)*.

Simpler: Just check if $\dfrac{\mathcal{L}(\theta_0;\mathbf{X})}{\mathcal{L}(\theta_1;\mathbf{X})}$ is monotone in a suff. stat when $\theta_0 \neq \theta_1$.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim}$ Exponential(β). Give the UMP size-α test for

1. H_0: $\beta \leq \beta_0$ vs H_1: $\beta > \beta_0$.
2. H_0: $\beta \geq \beta_0$ vs H_1: $\beta < \beta_0$.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, σ^2 known. Give UMP size-α test for

1. $H_0\colon \mu \leq \mu_0$ vs $H_1\colon \mu > \mu_0$.
2. $H_0\colon \mu \geq \mu_0$ vs $H_1\colon \mu < \mu_0$.

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(\theta)$. Give the UMP size-$\alpha$ test for

1. H_0: $\theta \leq \theta_0$ vs H_1: $\theta > \theta_0$.
2. H_0: $\theta \geq \theta_0$ vs H_1: $\theta < \theta_0$.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Geometric}(p)$. Give a level-$\alpha$ test that is UMP for

1. $H_0\colon p \leq p_0$ vs $H_1\colon p > p_0$.
2. $H_0\colon p \geq p_0$ vs $H_1\colon p < p_0$.

The Karlin-Rubin result says we can find UMP tests for:

- One-sided tests of hypotheses.
- When the parameter is one-dimensional.
- When there is a one-dimensional sufficient statistic.

This does not cover all situations!