Mathematical Statistics Recitation 7

TA: Sonia Reilly

March 13, 2026

Lecture Review

Lecture 13: Simple vs. Simple Testing

H_0: null hypothesis, H_1: alternative hypothesis

Test: function of observations that returns H_0 or H_1. Often of the form H_0 if $T(X) > c$, and H_1 if $T(X) \leq c$, for some *test statistic* T.

Rejection region = set of x's for which we reject null H_0, e.g. $\{x \mid T(x) \leq c\}$.

Simple hypotheses fully describe distribution, composite include multiple distributions.

Example: for $X \sim N(0, \sigma^2)$,

- Simple vs. simple: $H_0: \ \sigma^2 = 1$ vs. $H_1: \ \sigma^2 = 10$
- Simple vs. composite (two-sided): $H_0: \ \sigma^2 = 1$ vs. $H_1: \ \sigma^2 \neq 1$
- Simple vs. composite (one-sided): $H_0: \ \sigma^2 = 1$ vs. $H_1: \ \sigma^2 > 1$

Likelihood ratio (LR) test: for some "critical value" c,

$$X \to \begin{cases} H_0 & \text{if } \dfrac{p_0(X)}{p_1(X)} > c \\ H_1 & \text{if } \dfrac{p_0(X)}{p_1(X)} \leq c \end{cases}$$

where $p_0(x)$ = likelihood if H_0, $p_1(x)$ = likelihood if H_1

Two types of errors

- Type I = false positive = H_0 is true but we conclude H_1, probability α
- Type II = false negative = H_1 is true but we conclude H_0, probability β

Truth \ Test	H_0	H_1
H_0	Correct, $1-\alpha$	Type I error, α "level"
H_1	Type II error, β	Correct, $1-\beta$ "power"

We want to choose a c that balances the tradeoff between the two errors.

Neyman–Pearson Lemma

Likelihood ratio test is optimal for simple hypotheses, i.e., if LR with some c has level α and power $1-\beta$, any other test with level $\leq \alpha$ has power $\leq 1-\beta$ (can't get more power with same or lower level).

Lecture 14: Simple vs. Composite Testing

If H_1 is not simple, power $1-\beta = \min_{H_1} P(\text{output } H_1 \mid H_1)$. The minimization is often hard or impossible, so we typically control only α, not power. In this case we often say we reject H_0 instead of accepting H_1.

Good test statistic: distinguishes between H_0 and H_1, and has known distribution under H_0.

- E.g., observe X_i's from $N(\mu, 1)$. Hypotheses: $H_0 : \mu = 0$, $H_1 : \mu \neq 0$. Test: reject if $|\bar{X}_n| \geq c$ for some c.

p-value

- Measure of how unlikely our data is to occur under H_0. Specifically, probability of the same or more extreme (less likely under H_0) data if H_0 was true.
- For a test with rejection region $T(x) \leq c$, p-value of data X is $P(T \leq T(X) \mid H_0)$. (If rejection region is $T(x) \geq c$, p-value is $P(T \geq T(X) \mid H_0)$.)
- Alternate interpretation: Smallest significance level at which we would reject H_0.

  Because p-value $= P(T \leq T(X) \mid H_0)$ is the significance level α of a test with $c = T(X)$. For smaller α, $c \leq T(X)$, so you would not reject given the same X.

Converting between confidence intervals and tests

- CI $\to$ test: accept H_0 if θ_0 is in the CI.

  If $C_n(X_1, \ldots, X_n)$, e.g.,
  $$\left(\bar{X}_n \pm \frac{z_{\alpha/2}}{\sqrt{n}} \cdot S_n\right)$$
  is a $1-\alpha$ CI for θ, then
  $$A(\theta_0) = \{X_1, \ldots, X_n \mid \theta_0 \in C_n(X_1, \ldots, X_n)\}$$
  is the acceptance region of a level α test of $H_0 : \theta = \theta_0$.

- Test $\to$ CI: form CI out of any hypotheses we would be willing to accept.

  If $A(\theta_0)$ is the acceptance region of a level α test of $H_0 : \theta = \theta_0$, then a $1-\alpha$ CI is
  $$C_n(X_1, \ldots, X_n) = \{\theta : (X_1, \ldots, X_n) \in A(\theta)\}.$$

Problems

1. **[9.5 and 9.23]** True or False? Give a justification for your answer.

   (a) If the significance level of a test is decreased, the power will increase.
   False. No, typically the power will also decrease — the tradeoff is between α and β, not α and $1-\beta$.

   (b) If a test is rejected at the significance level α, the probability that the null hypothesis is true equals α.
   False. Cannot say anything about $P(H_0)$, only that seeing our current data $\to$ any stronger evidence against H_0 has prob. $\leq \alpha$ if H_0 true.

   (c) The probability that the null hypothesis is falsely rejected is equal to the power of the test.
   False. That's the level.

   (d) The power of a test is determined by the null distribution of the test statistic.
   False. Power is the probability that we get it right if H_1 is true.

   (e) The likelihood ratio is a random variable.
   True. Depends on X, so it is random.

   (f) If the p-value is 0.03, the corresponding test will reject at the significance level 0.02.
   False. p-value is the smallest significance level at which you would reject.

   (g) Suppose that a 99% CI for the mean μ of a normal distribution is found to be $(-2.0, 3.0)$. A test of $H_0 : \mu = -3$ versus $H_A : \mu \neq -3$ would reject H_0 at the 0.01 significance level.
   True. A $1 - 0.99 = 0.01$ significance level test of $H_0 : \mu = -3$ accepts if -3 is in a 0.99 CI of μ. Since it is not, it will reject.

2. **[9.19]** Under H_0 a random variable has the CDF $F_0(x) = x^2, 0 \leq x \leq 1$, and under H_1 it has the CDF $F_1(x) = x^3, 0 \leq x \leq 1$.

   (a) What is the form of the likelihood ratio test of H_0 versus H_1?

   (b) What is the rejection region of a level α test?

   (c) What is the power of the test?

   Solution:

   (a) First find pdfs:
   $$p_0(x) = \frac{d}{dx}(x^2) = 2x \quad \text{on } [0,1]$$
   $$p_1(x) = \frac{d}{dx}(x^3) = 3x^2 \quad \text{on } [0,1]$$
   Likelihood ratio test has the form
   $$\boxed{\begin{cases} H_0 & \text{if } \frac{2}{3x} > c \\ H_1 & \text{if } \frac{2}{3x} \leq c \end{cases}}$$
   since
   $$\frac{p_0(x)}{p_1(x)} = \frac{2x}{3x^2} = \frac{2}{3x}.$$

(b) If the test has level α,

$$P\left(\frac{2}{3x} \le c \mid H_0\right) = P\left(x \ge \frac{2}{3c} \mid H_0\right) = 1 - \left(\frac{2}{3c}\right)^2 = 1 - \frac{4}{9c^2} = \alpha.$$

$$\Rightarrow c = \frac{2}{3\sqrt{1-\alpha}},$$

so the rejection region is

$$\boxed{x \ge \sqrt{1-\alpha}}.$$

(c) The power is

$$1 - P(x \le \sqrt{1-\alpha} \mid H_1) = 1 - (\sqrt{1-\alpha})^3 = \boxed{1 - (1-\alpha)^{3/2}}.$$

3. **[9.20]** Consider two probability density functions on $[0, 1]$: $f_0(x) = 1$, and $f_1(x) = 2x$. Among all the tests of the null hypothesis $H_0 : X \sim f_0(x)$ versus the alternative $H_1 : X \sim f_1(x)$, with significance level $\alpha = 0.10$, how large can the power possibly be?

Solution:

Since these are simple hypotheses, the Neyman–Pearson lemma tells us that for a given level α, the likelihood ratio test will have the highest possible power.

$$\begin{cases} H_0 & \text{if } \frac{1}{2x} > c' \\ H_1 & \text{if } \frac{1}{2x} \le c' \end{cases} \iff \begin{cases} H_0 & \text{if } x < c \\ H_1 & \text{if } x \ge c \end{cases}$$

$$P(X \ge c \mid H_0) = 1 - c = \alpha \Rightarrow c = 1 - \alpha,$$

so the level-α LR test is

$$\begin{cases} H_0 & \text{if } x < 1-\alpha \\ H_1 & \text{if } x \ge 1-\alpha \end{cases}$$

The power of this test is

$$1 - P(x < 1-\alpha \mid H_1) = 1 - (1-\alpha)^2 = 2\alpha - \alpha^2 = 0.20 - 0.01 = \boxed{0.19}.$$

4. **[9.28]** Suppose that a test statistic T has a standard normal null distribution.

(a) If the test rejects for large values of $|T|$, what is the p-value corresponding to $T = 1.50$?

(b) Answer the same question if the test rejects for large T.

Solution:

(a) For a two-tailed test (rejecting for large $|T|$), the p-value is

$$p = P(|T| \ge 1.50) = 2P(T \ge 1.50) = 2(1 - \Phi(1.50)).$$

Using $\Phi(1.50) \approx 0.9332$ from the given table:

$$p = 2(1 - 0.9332) = 2(0.0668) = \boxed{0.1336}.$$

(b) For a one-tailed test, rejecting for large T,

$$p = P(T \ge 1.50) = 1 - \Phi(1.50) = \boxed{0.0668}.$$