STATS 300A: Theory of Statistics I	Autumn 2016/17

Lecture 12: November 8

Lecturer: Joseph Romano *Scribes: Kenneth Tay*

12.1 Hypothesis Testing

Let $\mathcal{S}$ be the sample space for X, and let $X \sim P_\theta$, $\theta \in \Omega$. Assume that Ω is the disjoint union of Ω_0 and Ω_1. We can formulate null and alternative hypotheses:

$$H_0 : \theta \in \Omega_0,$$
$$H_1 : \theta \in \Omega_1.$$

Definition 12.1 *A hypothesis $H_i : \theta \in \Omega_i$ is **simple** if Ω_i contains a single value of θ, i.e. $\Omega_i = \{\theta_0\}$ for some θ_0.*

*A hypothesis which is not simple is called a **complex hypothesis**.*

Definition 12.2 *A **test function** φ is a function $\varphi : \mathcal{S} \to [0, 1]$ with the interpretation that for $x \in \mathcal{S}$, we reject H_0 with probability $\varphi(x)$.*

- In this definition of test functions, we are allowing for randomized tests, i.e. $0 < \varphi(x) < 1$ for some values of x.
- If $\varphi(x) = 0$ or 1 for all x, φ is a non-randomized test. In this case, we define the **rejection region** as the set $\{x : \phi(x) = 1\}$.

Recall that we have the following possible scenarios in hypothesis testing:

	H_0 is true	H_1 is true
Reject H_0	Type 1 error	Good decision
Don't reject H_0	Good decision	Type 2 error

We want to construct a test that has "small probabilities" of Type 1 and Type 2 error. In general, we are more concerned with Type 1 error, so we only consider tests whose probability of Type 1 error does not exceed some pre-specified α, i.e.

$$\mathbb{E}_\theta \varphi(X) \leq \alpha \text{ for all } \theta \in \Omega_0.$$

α is called the **level of significance**.

Definition 12.3 *For a test φ, its **size** is $\sup_{\theta \in \Omega_0} \mathbb{E}_\theta \varphi(X)$.*

Definition 12.4 *A test φ is **uniformly most powerful (UMP)** at level α if its size is $\leq \alpha$ and for any other test φ' with size $\leq \alpha$,*

$$\mathbb{E}_\theta \varphi(X) \geq \mathbb{E}_\theta \varphi'(X) \text{ for all } \theta \in \Omega_1.$$

*If the alternative hypothesis is simple, we say that φ is a **most powerful (MP)** test at level α.*

12.1.1 Simple H_0 vs. Simple H_1

In this case, we can write $H_0 : X \sim P_0$, $H_1 : X \sim P_1$ for fixed distributions P_0 and P_1. We may also assume that P_i has density $p_i = \frac{dP_i}{d\mu}$ w.r.t. some dominating measure μ.

The Neyman-Pearson Lemma gives us a way of constructing a most powerful test for this setting:

Lemma 12.5 (Neyman-Pearson Lemma)

(i) *For testing P_0 vs. P_1, there exists a (possibly randomized) test φ and a constant k such that:*

  (a) *$\mathbb{E}_0 \varphi(X) = \alpha$, and*

  (b) *$\varphi(X) = 1$ if $\dfrac{p_1(X)}{p_0(X)} > k$, and $\varphi(X) = 0$ if $\dfrac{p_1(X)}{p_0(X)} < k$.*

(ii) *(Sufficiency to get a MP level α test) A sufficient condition for a test φ to be MP level α is it satisfies (a) and (b) in (i).*

(iii) *(Necessity) If φ is MP level α, then for some k, it satisfies (b), and it also satisfies (a) unless there exists a test φ' whose size is $< \alpha$ with power = 1.*

Proof:

(i) For each $c \in \mathbb{R}$, define

$$\begin{aligned}\alpha(c) &= P_0\{p_1(X) > cp_0(X)\} \\ &= 1 - F_0(c),\end{aligned}$$

where F_0 is the cdf of $\dfrac{p_1(X)}{p_0(X)}$ under P_0. Note that $\alpha(\cdot)$ is non-increasing and right-continuous. Also, for any c, if we define $\alpha(c^-) = \lim_{c' \uparrow c} \alpha(c')$, we have

$$P_0 \left\{ \frac{p_1(X)}{p_0(X)} = c \right\} = \alpha(c^-) - \alpha(c).$$

In general, there exists some c^* such that $\alpha(c^{*-}) \leq \alpha \leq \alpha(c^*)$. Define a test function

$$\varphi(X) = \begin{cases} 1 & \text{if } \frac{p_1(X)}{p_0(X)} > c^*, \\ \frac{\alpha - \alpha(c^*)}{\alpha(c^{*-}) - \alpha(c^*)} & \text{if } \frac{p_1(X)}{p_0(X)} = c^*, \\ 0 & \text{if } \frac{p_1(X)}{p_0(X)} < c^*. \end{cases}$$

Then

$$\begin{aligned}
\mathbb{E}_0\varphi(X) &= P_0\left\{\frac{p_1(X)}{p_0(X)} > c^*\right\} + \frac{\alpha-\alpha(c^*)}{\alpha(c^{*-})-\alpha(c^*)}P_0\left\{\frac{p_1(X)}{p_0(X)} = c^*\right\} \\
&= \alpha(c^*) + \frac{\alpha-\alpha(c^*)}{\alpha(c^{*-})-\alpha(c^*)}[\alpha(c^{*-})-\alpha(c^*)] \\
&= \alpha,
\end{aligned}$$

as required.

(ii) Suppose φ satisfies (a) and (b) from (i). Let φ' be any other level α test. Consider the value of the integral

$$\int [\varphi(x)-\varphi'(x)][p_1(x)-kp_0(x)]d\mu(x).$$

We break the integral up over 3 regions:

$$\begin{aligned}
S^+ &= \{x : \varphi(x) > \varphi'(x)\}, \\
S^- &= \{x : \varphi(x) < \varphi'(x)\}, \text{ and} \\
S^0 &= \{x : \varphi(x) = \varphi'(x)\}.
\end{aligned}$$

- Over S^0, the integral is clearly equal to 0.
- For $x \in S^+$, $\varphi(x) > \varphi'(x) \geq 0$, which means (by φ's definition) that $p_1(x) \geq kp_0(x)$. Hence, the integrand is non-negative, so the integral over S^+ is non-negative.
- $x \in S^-$, $\varphi(x) < \varphi'(x) \leq 1$, which means that $p_1(x) \leq kp_0(x)$. Hence, the integrand is again non-negative, so the integral over S^+ is non-negative.

Summarizing, we have

$$\begin{aligned}
\int [\varphi(x)-\varphi'(x)][p_1(x)-kp_0(x)]d\mu(x) &\geq 0, \\
\int [\varphi(x)-\varphi'(x)]p_1(x)d\mu(x) &\geq k\int [\varphi(x)-\varphi'(x)]p_0(x)d\mu(x), \\
\mathbb{E}_1[\varphi-\varphi'] &\geq k(\mathbb{E}_0\varphi - \mathbb{E}_0\varphi') \\
&= k(\alpha - \mathbb{E}_0\varphi') \\
&\geq 0, \\
\mathbb{E}_1\varphi &\geq \mathbb{E}_1\varphi',
\end{aligned}$$

as required.

(iii) The proof is in the book. It is essentially the proof of (ii) backwards.

■

Corollary 12.6 *For testing P_0 vs. P_1, the power of a MP level α test is $> \alpha$ (unless $P_0 = P_1$).*

Proof: Let $\varphi_0(X) = \alpha$, i.e. always reject with probability α. This test has power α, but it is not a likelihood ratio test. Hence, by Lemma 12.5(iii), it can't be a MP test. So any MP test will have larger power than φ_0, i.e. power $> \alpha$. ■

Note: We can always restrict attention to tests based on a sufficient statistic T. This is because we can always define a test $\psi(T) = \mathbb{E}[\varphi(X) \mid T]$. For this test, $\mathbb{E}\psi(T) = \mathbb{E}\varphi(X)$.

12.1.1.1 Example: Normal setting

Let $X_1, \dots, X_n$ be i.i.d., $X_i \sim \mathcal{N}(\theta, \sigma^2)$ with σ known. $H_0 : \theta = 0$, $H_1 : \theta = \theta_1$ for some specified θ_1.

To use the Neyman-Perason Lemma (ii), we need to find a constant k such that (i)(a) and (i)(b) are satisfied. For a given x, the likelihood ratio at x is given by

$$
\begin{aligned}
LR(x) &= \frac{\prod \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x_i - \theta_1)^2\right]}{\prod \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x_i - 0)^2\right]} \\
&= \exp\left\{-\frac{1}{2\sigma^2}\sum(x_i - \theta_1)^2 + \frac{1}{2\sigma^2}\sum x_i^2\right\} \\
&= \exp\left\{\frac{\theta_1}{\sigma^2}\sum x_i - \frac{n\theta_1^2}{2\sigma^2}\right\}.
\end{aligned}
$$

Now,

$$
\begin{aligned}
& & LR(X) &> k & &\text{for some constant } k \\
&\Leftrightarrow & \frac{\theta_1}{\sigma^2}\sum X_i - \frac{n\theta_1^2}{2\sigma^2} &> k_1 & &\text{for some constant } k_1 \\
&\Leftrightarrow & \theta_1 \sum X_i &> k_2 & &\text{for some constant } k_2 \\
&\Leftrightarrow & \theta_1 \bar{X}_n &> k_3 & &\text{for some constant } k_3 \\
&\Leftrightarrow & \theta_1\left(\frac{\bar{X}_n\sqrt{n}}{\sigma}\right) &> \tilde{k} & &\text{for some constant } \tilde{k}.
\end{aligned}
$$

Note that under the null, $\frac{\bar{X}_n\sqrt{n}}{\sigma} \sim N(0,1)$.

Case 1: $\theta_1 > 0$.

Our likelihood ratio test becomes $\frac{\bar{X}_n\sqrt{n}}{\sigma} > \tilde{\tilde{k}}$ for some constant $\tilde{\tilde{k}}$. In order for (i)(a) to hold, we need $\tilde{\tilde{k}} = Z_{1-\alpha} = \Phi^{-1}(1-\alpha)$.

Thus, the MP level α test is Reject if $\frac{\bar{X}_n\sqrt{n}}{\sigma} > Z_{1-\alpha} = \Phi^{-1}(1-\alpha)$.

Case 2: $\theta_1 < 0$.

Our likelihood ratio test becomes $\frac{\bar{X}_n\sqrt{n}}{\sigma} < \tilde{\tilde{k}}$ for some constant $\tilde{\tilde{k}}$. In order for (i)(a) to hold, we need $\tilde{\tilde{k}} = Z_{\alpha} = \Phi^{-1}(\alpha)$.

Thus, the MP level α test is Reject if $\frac{\bar{X}_n\sqrt{n}}{\sigma} < Z_{\alpha} = \Phi^{-1}(\alpha)$.

Some comments are in order:

1. Note that for $\theta_1 > 0$, the MP level α test does not depend on θ_1! Hence, in the normal model, for testing $H_0 : \theta = 0$ vs. $H_1 : \theta > 0$, the test we derived above is UMP level α. (The same is true for $H_0 : \theta = 0$ vs. $H_1 : \theta < 0$.)

2. However, for testing $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$, no UMP test exists!

12.1.1.2 Example: Discrete setting

Let X take on only values in $\{1, 2, 3, 4, 5\}$, and define probability distributions P_0 and P_1 by the following table:

x	1	2	3	4	5
$p_0(x)$	$^1/_4$	$^1/_{100}$	$^1/_{100}$	$^3/_{100}$	$^7/_{10}$
$p_1(x)$	$^1/_2$	$^1/_{10}$	$^2/_{100}$	$^5/_{100}$	$^{33}/_{100}$

We are looking at $H_0 : X \sim P_0$ vs. $H_1 : X \sim P_1$, with $\alpha = 0.05$. As a first step, we compute the likelihood ratios:

x	1	2	3	4	5
$\frac{p_1(x)}{p_0(x)}$	2	10	2	$^5/_3$	$^{33}/_{70}$

We look at the likelihood ratios from biggest to smallest. If $X = 2$ is in the rejection region, then $\mathbb{E}_0\varphi(X) = \frac{1}{100}$ is still within our "budget".

Next, we look at the likelihood ratio 2. We can't include both $X = 1$ and $X = 3$ in the rejection region as that would push $\mathbb{E}_0\varphi(X)$ far above α. We could have a randomized test to take care of this.

If we wanted a non-randomized test, the best test would be to reject if $X = 2, 3,$ or 4, even though $X = 1$ has a higher likelihood ratio than $X = 4$. The Neyman-Pearson Lemma tells us that we can find a (randomized) test with better power than this.

12.1.2 p-values

Definition 12.7 *For non-randomized tests in general, let S_α be the rejection region for a level α test. Assume that $S_\alpha \subseteq S_{\alpha'}$ if $\alpha' > \alpha$.*

Then the p-value is given by

$$\hat{p} := \int_\alpha \{X \in S_\alpha\}.$$

Intuitively, a p-value is the smallest value of α leading to a rejection of H_0 if X is observed.