STA 522, Spring 2022
Introduction to Theoretical Statistics II

Lecture 9

Department of Biostatistics
University at Buffalo

AGENDA

- Hypothesis testing, LRT
- Properties of tests, finding c in LRT
- Methods of evaluating tests

Review: likelihood ratio test

- Recall the **likelihood function** $L(\theta \mid \underline{x}) = f(\underline{x} \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$. The **likelihood ratio test (LRT) statistic** for testing $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_0^c$ is $\lambda(\underline{x}) = \dfrac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})}$.

- Note that the likelihood ratio test statistic can be viewed as

$$\lambda(\underline{x}) = \frac{L(\hat{\theta}_0 \mid \underline{x})}{L(\hat{\theta} \mid \underline{x})} = \frac{\text{restricted maximization}}{\text{unrestricted maximization}},$$

  where $\hat{\theta}$ is the MLE obtained by maximizing $L(\theta \mid \underline{x})$ over the entire parameter space Θ, and $\hat{\theta}_0$ is the MLE obtained by maximizing over the restricted parameter space Θ_0.

- A **likelihood ratio test (LRT)** is any test that has a rejection region of the form

$$\{\underline{x} : \lambda(\underline{x}) \leq c\},$$

  where $c \in [0, 1]$.

Example: Let $X_1, X_2, \ldots, X_n \sim$ iid from a (location) exponential population with pdf $f(x \mid \theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x)$, where $\theta \in \Theta = \mathbb{R}$. Suppose we wish to test $H_0 : \theta \leq a$ vs. $H_1 : \theta > a$ where a is a known value (e.g. 0) supplied by the experimenter. Find the LRT rejection region.

The likelihood function for θ is:

$$L(\theta \mid \underline{x}) = \prod_{i=1}^{n} e^{-(x_i-\theta)} \; I(x_i \geq \theta) = e^{-n(\overline{x}-\theta)} \; I(x_{(1)} \geq \theta)$$

$L(\theta \mid \underline{x})$ is an increasing function in θ for $\theta \in (-\infty, x_{(1)}]$. So unrestricted MLE is $\hat{\theta} = x_{(1)}$ so that $\sup_{\theta \in \Theta} L(\theta \mid \underline{x}) = L(x_{(1)} \mid \underline{x}) = e^{-n\left(\overline{x}-x_{(1)}\right)}$.

Under H_0, the restricted range $\theta \in \Theta_0 = (-\infty, a]$ MLE of θ is

$$\hat{\theta}_0 = \begin{cases} x_{(1)} & \text{if } \; x_{(1)} \leq a \\ a & \text{if } \; x_{(1)} > a \end{cases}$$

Therefore, LRT is:

$$\lambda(\underline{x}) = \begin{cases} 1 & \text{if } x_{(1)} \leq a \\ e^{-n\left(x_{(1)}-a\right)} & \text{if } x_{(1)} > a \end{cases}$$

Therefore the rejection region for the LRT is:

$$\{\underline{x} : \lambda(\underline{x}) \leq c\} = \left\{\underline{x} : x_{(1)} \geq a - \frac{\log c}{n}\right\}$$

for some $0 < c < 1$.

NOTE: The LRT rejection region depends on the data only through $X_{(1)}$. In the normal example discussed last week, the LRT rejection region depends on data only through $\overline{X}$.

LRT and sufficiency

Note: Sufficiency means that all the information about θ in $\underline{x}$ is contained in a sufficient statistic $T(\underline{x})$. Intuitively, a test based on T should be as good as the test based on the complete sample $\underline{X}$. The following theorem formalizes this.

Theorem (8.2.4)

If $T(\underline{X})$ is a sufficient statistic for θ and $\lambda^*(t)$ and $\lambda(\underline{x})$ are the LRT statistics based on T and $\underline{X}$, respectively, then

$$\lambda^*(T(\underline{x})) = \lambda(\underline{x})$$

for every $\underline{x}$ in the sample space.

Proof: Since $T(\underline{X})$ is a sufficient statistics, therefore by the Factorization theorem, we have

$$f(\underline{x} \mid \theta) = g(T(\underline{x}) \mid \theta)\ h(\underline{x})$$

Therefore

$$\begin{aligned}
\lambda(\underline{x}) &= \frac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})} \\
&= \frac{\sup_{\Theta_0} f(\underline{x} \mid \theta)}{\sup_{\Theta} f(\underline{x} \mid \theta)} \\
&= \frac{\sup_{\Theta_0} g(T(\underline{x}) \mid \theta)\ h(\underline{x})}{\sup_{\Theta} g(T(\underline{x}) \mid \theta)\ h(\underline{x})} \\
&= \frac{\sup_{\Theta_0} g(T(\underline{x}) \mid \theta)}{\sup_{\Theta} g(T(\underline{x}) \mid \theta)} \\
&= \frac{\sup_{\Theta_0} L^*(\theta \mid T(\underline{x}))}{\sup_{\Theta} L^*(\theta \mid T(\underline{x}))} \\
&= \lambda^*(T(\underline{x}))
\end{aligned}$$

This completes the proof.

Example: Let $X_1, X_2, \ldots, X_n \sim$ iid from a population with pdf $f(x \mid \theta) = \theta x^{\theta-1} I_{(0,1)}(x)$, $\theta > 0$. Suppose we wish to test $H_0 : \theta = 1$ vs. $H_1 : \theta \neq 1$. Find the LRT rejection region.

Note at the outset that the restricted MLE is simply $\hat{\theta}_0 = 1$.

For $\theta \in \Theta = (0, \infty)$ the likelihood function is given by

$$L(\theta \mid \underline{x}) = \theta^n \left(\prod_{i=1}^{n} x_i \right)^{(\theta-1)} \implies \log L(\theta \mid \underline{x}) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i$$

therefore

$$\frac{\partial \ L(\theta \mid \underline{x})}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i \gtreqless 0 \iff \theta \lesseqgtr -\frac{n}{\sum_{i=1}^{n} \log x_i}$$

Therefore, the MLE of θ is $\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log X_i}$.

Therefore the LRT statistic is

$$
\begin{aligned}
\lambda(\underline{x}) &= \frac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})} \\
&= \exp\left[n \log \theta_0 + (\theta_0 - 1) \sum_{i=1}^{n} \log x_i - n \log \hat{\theta} - (\hat{\theta} - 1) \sum_{i=1}^{n} \log x_i \right] \\
&= \exp\left[n \log\left(\frac{\theta_0}{\hat{\theta}}\right) + (\theta_0 - \hat{\theta}) \sum_{i=1}^{n} \log x_i \right]
\end{aligned}
$$

Note that $\lambda(\underline{x})$ depends on $\underline{x}$ only through $\sum_{i=1} \log x_i$.

The rejection region of the LR test is given by:

$$
\left\{ \underline{x} : \exp\left[n \log\left(\frac{\theta_0}{\hat{\theta}}\right) + (\theta_0 - \hat{\theta}) \sum_{i=1}^{n} \log x_i \right] \leq c \right\}
$$

Example (LRT under the presence of nuisance parameters): Let $X_1, X_2, \ldots, X_n \sim$ iid $\mathrm{N}(\mu, \sigma^2)$ (both parameters unknown). Suppose we wish to test $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$. Find the LRT rejection region.

Note that here σ^2 is a nuisance parameter.

The unrestricted MLEs of μ and σ^2 are $\hat{\mu} = \overline{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Under H_0, the restricted MLE for μ is

$$\hat{\mu}_0 = \begin{cases} \overline{X} & \text{if } \overline{X} \leq \mu_0 \\ \mu_0 & \text{if } \overline{X} > \mu_0 \end{cases}$$

The corresponding MLE of σ^2 is

$$\hat{\sigma}_0^2 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 & \text{if } \overline{X} \leq \mu_0 \\ \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2 & \text{if } \overline{X} > \mu_0 \end{cases}$$

The LRT statistic is given by:

$$\lambda(\underline{x}) = \begin{cases} 1 & \text{if } \overline{X} \le \mu_0 \\ \dfrac{L(\mu_0, \sigma_0^2)}{L(\hat{\mu}, \hat{\sigma}^2)} & \text{if } \overline{X} > \mu_0 \end{cases}$$

The rejection region is given by

$$\{\underline{x} : \lambda(\underline{x}) \le c\}$$

It can be shown that (HW, exercise 8.37) the above rejection region can be equivalently expressed as (t-test)

$$\overline{X} > \mu_0 + c' \sqrt{\frac{S^2}{n}}$$

Errors in Hypothesis Testing

Definition: Suppose we are testing

$$H_0 : \theta \in \Theta_0$$
$$\text{vs. } H_1 : \theta \in \Theta_0^c.$$

If $\theta \in \Theta_0$, but the test incorrectly rejects H_0, then the test has made a **Type I error**.

If, on the other hand, $\theta \in \Theta_0^c$, but the test decides to accept H_0, then the test has made a **Type II error**.

		Decision	
		Accept H_0	Reject H_0
Truth	H_0	Correct decision	Type I Error
	H_1	Type II Error	Correct decision

Computing Error Probabilities

Definition: Let R denote the rejection region of a hypothesis test.

If $\theta \in \Theta_0$, then the probability of a Type I error is

$$P_\theta(\underline{X} \in R).$$

If $\theta \in \Theta_0^c$, then the probability of a Type II error is

$$P_\theta(\underline{X} \notin R) = 1 - P_\theta(\underline{X} \in R).$$

Power Function

Definition: The **power function** of a hypothesis test with rejection region R is the function of θ defined by

$$\beta(\theta) = P_\theta(\underline{X} \in R)$$

$$= \begin{cases} \text{probability of a Type I error} & \text{if } \theta \in \Theta_0 \\ 1 - \text{ probability of a Type II error} & \text{if } \theta \in \Theta_0^c. \end{cases}$$

Comments on the Power function:

(a) Ideally, we want $\beta(\theta) = 0$ for all $\theta \in \Theta_0$ and $\beta(\theta) = 1$ for all $\theta \in \Theta_0^c$.

(b) Depends on the hypothesis test (what are we testing?).

(c) Depends on the rejection region (value of c).

(d) It's a function of θ, not the data.

(e) Since it's a probability, $0 \leq \beta(\theta) \leq 1$ for all θ.

Example: Suppose $X \sim \text{binomial}(5, \theta)$, and we are testing $H_0 : \theta \leq \frac{1}{2}$ vs. $H_1 : \theta > \frac{1}{2}$. Consider the two rejection regions

$$R_1 = \{x : x = 5\}$$

$$R_2 = \{x : x = 3, 4, 5\}.$$

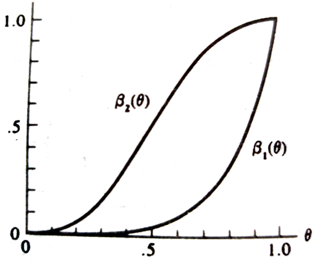
Note that with R_1, we reject H_0 if and only if we observe all successes, whereas with R_2, we reject H_0 if and only if we observe at least 3 successes. Determine the power function for each test.

Here

$$\beta_1(\theta) = P_\theta(X \in R_1) = P_\theta(X = 5) = \binom{5}{5}\theta^5(1-\theta)^{5-5} = \theta^5$$

$$\beta_2(\theta) = P_\theta(X \in R_2) = \sum_{j=3}^{5} P_\theta(X = j) = \sum_{j=3}^{5}\binom{5}{j}\theta^j(1-\theta)^{5-j}$$

Comments about the two power functions



(a) $\beta_2(\theta)$ has higher Type I error and lower Type II error.

(b) $\beta_1(\theta)$ has lower Type I error and higher Type II error.

(c) Ideally, what we will do is try to maximize power while controlling Type I error.

(d) This is how we will choose c in our previous calculations of rejection regions.

Size and Level

Definition: For $0 \le \alpha \le 1$, a test with power function $\beta(\theta)$ is a **size** α **test** if

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha.$$

For $0 \le \alpha \le 1$, a test with power function $\beta(\theta)$ is a **level** α **test** if

$$\sup_{\theta \in \Theta_0} \beta(\theta) \le \alpha.$$

Notes: the set of size α tests is a subset of the set of level α tests.

By specifying the level of a test, we are only controlling the Type I error, not the Type II error.

Choosing c For LRTs

- Restricting to size α tests allows us to determine the value of c to use in the LRT.

- We can build a size α LRT by choosing c so that

$$\sup_{\theta\in\Theta_0} P_\theta(\underline{X} \in R) = \alpha, \quad \text{i.e.,} \quad \sup_{\theta\in\Theta_0} P_\theta(\lambda(\underline{X}) \leq c) = \alpha.$$

Example (contd.): Let $X_1, X_2, \ldots, X_n \sim$ iid $\mathrm{N}(\theta, 1)$. Suppose we wish to test $H_0 : \theta = \theta_0$vs. $H_1 : \theta \neq \theta_0$. We saw that the LRT rejection region is given by

$$R = \{\underline{x} \, : \, |\overline{x} - \theta_0| \geq k\},$$

where $k = \sqrt{\frac{-2 \log c}{n}}$. Find the value of c so that we have a size α test.

Since $\Theta_0 = \{\theta_0\}$ is singleton, hence

$$\text{size} = \sup_{\Theta_0} P_\theta \left(|\overline{X} - \theta_0| \geq k\right) = P_{\theta_0} \left(|\overline{X} - \theta_0| \geq k\right)$$

Now, under H_0, $\overline{X} \sim N(\theta_0, 1/n)$ so that $Z = \sqrt{n}(\overline{X} - \theta_0) \sim N(0, 1)$. Therefore the size of the LRT being α implies

$$\begin{aligned}
\alpha &= P_{\theta_0} \left(|\sqrt{n}(\overline{X} - \theta_0)| \geq \sqrt{n}\ k\right) \\
&= P_{\theta_0}(|Z| \geq \sqrt{n}\ k) \\
&= P(Z \geq \sqrt{n}\ k) + P(Z \leq -\sqrt{n}\ k) \\
&= P(Z \geq \sqrt{n}\ k) + P(-Z \geq -\sqrt{n}\ k) = 2\ P(Z \geq \sqrt{n}\ k)
\end{aligned}$$

Let z_α be the upper α-th quantile of Z such that $P(Z \geq z_\alpha) = \alpha$.

Here $\alpha/2 = P(Z \geq \sqrt{n}\ k)$, which implies

$$\sqrt{n}\ k = z_{\alpha/2} \implies k = \frac{1}{\sqrt{n}} z_{\alpha/2} \implies c = \exp\left(-z^2_{\alpha/2}/2\right)$$

Example (contd.): Let $X_1, X_2, \ldots, X_n \sim$ iid from a location exponential population with pdf

$$f(x \mid \theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x).$$

Suppose we wish to test $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$. We showed that the LRT rejection region is given by

$$R = \left\{ \underline{x} \,:\, x_{(1)} \geq \theta_0 - \frac{\log c}{n} \right\}.$$

Find the value of c so that we have a size α test.

HW. See p. 386 in the textbook.

Evaluating Tests

Definition: A test with power function $\beta(\theta)$ is **unbiased** if

$$\beta(\theta') \geq \beta(\theta'')$$

for every $\theta' \in \Theta_0^c$ and $\theta'' \in \Theta_0$.

Definition: Let $\mathcal{C}$ be a class of tests for testing $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_0^c$. 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 every $\theta \in \Theta_0^c$ and every $\beta'(\theta)$ that is a power function of a test in class $\mathcal{C}$.

Note: if we take $\mathcal{C}$ to be the class of all level α tests, the test described in the above definition is called a **UMP level α test**.

Homework

- Read p. 374 – 379, 382 – 387.
- Exercises: TBA.