STA261 Lecture 10 — 2017-08-09

Neil Montgomery

Last edited: 2017-08-14 19:06

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.
4. Decide on a test statistic (not obvious how to do this!)

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.
4. Decide on a test statistic (not obvious how to do this!)
5. Decide on α and determine the rejection region (not obvious how to do this!) of this size.

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.
4. Decide on a test statistic (not obvious how to do this!)
5. Decide on α and determine the rejection region (not obvious how to do this!) of this size.
6. Observe data and see it if lands in rejection region.

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.
4. Decide on a test statistic (not obvious how to do this!)
5. Decide on α and determine the rejection region (not obvious how to do this!) of this size.
6. Observe data and see it if lands in rejection region.
7. ???

(classical) hypothesis testing procedure

1. Propose a statistical model, i.e. a family of densities (or pmfs) indexed by a parameter.
2. Specify a null hypothesis and an alternative hypothesis.
3. Plan to collect a sample $X_1, \ldots, X_n$.
4. Decide on a test statistic (not obvious how to do this!)
5. Decide on α and determine the rejection region (not obvious how to do this!) of this size.
6. Observe data and see it if lands in rejection region.
7. ???
8. Profit!

likelihood ratio | Neyman-Pearson

(This theorem has been renumbered to be a Lecture 10 theorem!)

Theorem 10.0: Suppose $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$ (each contain one value only.) The test of size α that rejects H_0 when:

$$\frac{L(\theta_0)}{L(\theta_1)} < c$$

for some non-negative constant c is at least as powerful as any test of size less than or equal to α.

(proof to be given next class)

examples

Example 10.1: $N(\mu, \sigma^2)$ (σ^2 known) with $H_0 : \mu = \mu_0$ and $H_1 : \mu = \mu_1$ with $mu_1 > \mu_0$.

Example 10.2: Bernoulli(p) with $H_0 : p = p_0$ and $H_1 : p = p_1$ with $p_1 > p_0$.

Example 10.3: $N(\mu, \sigma^2)$ (μ known) with $H_0 : \sigma^2 = \sigma_0^2$ and $H_1 : \sigma^2 = \sigma_1^2$ with $\sigma_1^2 > \sigma_0^2$.

implications of Neyman-Pearson

It turns out the likelihood ratio result also applies to hypotheses of these forms:

$$H_0 : \theta \leqslant \theta_0 \qquad \text{versus} \qquad H_1 : \theta > \theta_0$$

implications of Neyman-Pearson

It turns out the likelihood ratio result also applies to hypotheses of these forms:

$$H_0 : \theta \leqslant \theta_0 \qquad \text{versus} \qquad H_1 : \theta > \theta_0$$

However, they don't directly apply to the most common type of test for:

$$H_0 : \theta = \theta_0 \qquad \text{versus} \qquad H_1 : \theta \neq \theta_0$$

likelihood ratio test in in general

Denote by Θ the set of all possible paramater values ("the parameter space") and denote by Θ_0 and Θ_1 the subsets corresponding to H_0 and H_1.

likelihood ratio test in in general

Denote by Θ the set of all possible paramater values ("the parameter space") and denote by Θ_0 and Θ_1 the subsets corresponding to H_0 and H_1.

Define:

$$\Lambda = \frac{\max\limits_{\theta \in \Theta_0} L(\theta)}{\max\limits_{\theta \in \Theta} L(\theta)}$$

likelihood ratio test in in general

Denote by Θ the set of all possible paramater values ("the parameter space") and denote by Θ_0 and Θ_1 the subsets corresponding to H_0 and H_1.

Define:

$$\Lambda = \frac{\max\limits_{\theta \in \Theta_0} L(\theta)}{\max\limits_{\theta \in \Theta} L(\theta)}$$

The *generalized likelihood ratio test*, or LRT, rejects when $\Lambda < c$ for some constant.

Example 10.4: $N(\mu, \sigma^2)$ population (σ^2 known) with $H_0 : \mu = \mu_0$ and $H_1 : \mu \neq \mu_0$

equivalence of LRT and confidence interval

The rejection region just derived could be written as:

$$\overline{X} \leqslant \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ or } \overline{X} \geqslant \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

equivalence of LRT and confidence interval

The rejection region just derived could be written as:

$$\overline{X} \leqslant \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ or } \overline{X} \geqslant \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Recall the $(1-\alpha)\cdot 100\%$ confidence interval for μ:

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

equivalence of LRT and confidence interval

The rejection region just derived could be written as:

$$\overline{X} \leqslant \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ or } \overline{X} \geqslant \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Recall the $(1-\alpha)\cdot 100\%$ confidence interval for μ:

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

H_0 is rejected if and only if μ_0 is outside the C.I.

numerical example $N(\mu, 1)$

Example 10.5: Fix $n = 20$. Testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. Set $\alpha = 0.05$.

1. What is the rejection region?

numerical example $N(\mu, 1)$

Example 10.5: Fix $n = 20$. Testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. Set $\alpha = 0.05$.

1. What is the rejection region?
2. I simulated data from a $N(0.5, 1)$ distribution and $\overline{x} = 0.439$. What is the conclusion?

numerical example $N(\mu, 1)$

Example 10.5: Fix $n = 20$. Testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. Set $\alpha = 0.05$.

1. What is the rejection region?
2. I simulated data from a $N(0.5, 1)$ distribution and $\overline{x} = 0.439$. What is the conclusion?
3. (A question for those with experience. . .) Does this example suggest any criticisms of "classical" hypothesis testing?

p-values

A p-value is defined as *the smallest α at which H_0 would be rejected.*

p-values

A p-value is defined as *the smallest* α *at which* H_0 *would be rejected.*

A less formal but more useful “definition” is *the probability of observing a more extreme value of the test statistic.*

p-values

A p-value is defined as *the smallest* α *at which* H_0 *would be rejected.*

A less formal but more useful "definition" is *the probability of observing a more extreme value of the test statistic.*

Numerical example 10.5 continued. . .

p-values

A p-value is defined as *the smallest* α *at which* H_0 *would be rejected.*

A less formal but more useful "definition" is *the probability of observing a more extreme value of the test statistic.*

Numerical example 10.5 continued...

I could simulate from a $N(2,1)$ distribution to get a new $\overline{x} = 1.2451801$ What is the p-value this time?