Chapter 9: Hypothesis Testing

Sections

- 9.1 Problems of Testing Hypotheses - we are still here
- Skip: 9.2 Testing Simple Hypotheses - reorganized
- Skip: 9.3 Uniformly Most Powerful Tests - reorganized
- Skip: 9.4 Two-Sided Alternatives - reorganized
- 9.5 The t Test
- 9.6 Comparing the Means of Two Normal Distributions
- 9.7 The F Distributions
- 9.8 Bayes Test Procedures
- 9.9 Foundational Issues

After-School example

- $X_1, \ldots, X_n$ i.i.d. $N(\mu, 6^2)$, where $n = 220$ and we observe $\overline{X}_n = 21.1$
- Test $H_0 : \mu \geq 22$ versus $H_1 : \mu < 22$
- Testing procedure: Reject H_0 iff $\overline{X}_n \leq 22 - c$
- Given significance level $\alpha = 0.05$, we obtained that $c = 0.665$ thus $22 - 0.665 = 21.335$
- Do we reject H_0?
- If we use significance level $\alpha = 0.1$, will we reject H_0?
- If we use significance level $\alpha = 0.01$, will we reject H_0?
- Find all the signifiance levels α such that we reject H_0
- Find the smallest significance level α_0 such that we reject H_0

p-values

- Hypothesis testing end in either “reject” or “not reject”.
- Seems inefficient use of data. How close were we to making the other decision? What if we want to use a different level?

Def: p-value

The *p-value* is the smallest level α_0 such that we would reject the null hypothesis at level α_0 after seeing the data

- We reject H_0 if and only if the p-value we get is smaller than the pre-determined level of significance α_0
- We can also say that the observed test statistic is *just significant* at level equal to the p-value

Calculating p-values

- Suppose the test is of the form “reject H_0 if $T \geq c$"
- Let t be the observed value of T
- Then p-value = $\sup_{\theta \in \Omega_0} P(T \geq t|\theta)$
- The maximum is often obtained on the boundary of Ω_0
- Tail area under H_0
- For tests of the form “reject H_0 if $T \leq c$", p-value = $\sup_{\theta \in \Omega_0} P(T \leq t|\theta)$

p-value for disease example

- $X_1, \ldots, X_{80}$ i.i.d. $\mathrm{Bernoulli}(p)$
- Hypotheses: $H_0 : p \leq 0.02$ and $H_1 : p > 0.02$
- Test: Reject H_0 if $Y = \sum_{i=1}^{80} X_i > c$.
- Suppose we observe $Y = 6$. Find the p-value for the observed data.

c	1	2	3	4	5	6
$P(Y > c \mid p = 0.02)$	0.477	0.216	0.077	0.022	0.005	0.001

Tests and Confidence intervals

There is a relationship between a confidence interval for θ and a hypothesis of the form

$$H_0 : \theta = \theta_0 \quad \text{and} \quad H_1 : \theta \neq \theta_0$$

- We can obtain a $\gamma = 1 - \alpha_0$ confidence set from an α_0 level test.
- We can obtain an $\alpha_0 = 1 - \gamma$ level test from a $100\gamma\%$ confidence set for θ

For one-sided test, such as

$$H_0 : \theta \leq \theta_0 \quad \text{and} \quad H_1 : \theta > \theta_0$$

we only get one direction (in general):

- We can obtain a $\gamma = 1 - \alpha_0$ confidence set from an α_0 level test.
- Only in special cases can we obtain a $\alpha_0 = 1 - \gamma$ level test from a one-sided confidence interval

Tests and Confidence intervals

Theorem 9.1.1: Test $\longrightarrow$ Confidence Set

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from a distribution that is indexed by a parameter θ. Let $g(\theta)$ be the parameter of interest and δ_{g_0} be a level α_0 test of the hypothesis

$$H_{0,g_0} : g(\theta) = g_0 \quad \text{and} \quad H_{1,g_0} : g(\theta) \neq g_0$$

Define $\omega(\mathbf{x}) = \{g_0 : \delta_{g_0} \text{ does not reject } H_{0,g_0} \text{ if } \mathbf{X} = \mathbf{x} \text{ is observed }\}$
Then the random set $\omega(\mathbf{X})$ satisfies

$$P(g(\theta) \in \omega(\mathbf{X}) | \theta = \theta_0) \geq \gamma$$

for all $\theta_0 \in \Omega$, i.e. $\omega(\mathbf{X})$ is a $100\gamma\%$ *confidence set for* $g(\theta)$.

Also works for one-sided tests (Theorem 9.1.3)

Tests and Confidence intervals

Theorem 9.1.2: Confidence Set $\longrightarrow$ Test

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from a distribution that is indexed by a parameter θ. Let $g(\theta)$ be the parameter of interest and let $\omega(\mathbf{X})$ be a $100\gamma\%$ confidence set for $g(\theta)$. Let δ_{g_0} be a test of the hypothesis

$$H_{0,g_0} : g(\theta) = g_0 \quad \text{and} \quad H_{1,g_0} : g(\theta) \neq g_0$$

where δ_{g_0} rejects H_{0,g_0} iff $g_0 \notin \omega(\mathbf{X})$.
Then δ_{g_0} is a level $\alpha_0 = 1 - \gamma$ test of the above hypothesis.

Example

Let $X_1, \dots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ and

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \sigma' = \left(\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}\right)^{1/2}$$

We know that

$$\left(\overline{X}_n - T_{n-1}^{-1}\left(\frac{\gamma+1}{2}\right)\frac{\sigma'}{\sqrt{n}},\ \overline{X}_n + T_{n-1}^{-1}\left(\frac{\gamma+1}{2}\right)\frac{\sigma'}{\sqrt{n}}\right)$$

is a $100\gamma\%$ confidence interval for μ.

- Construct a level $\alpha_0 = 1 - \gamma$ test of the hypothesis

$$H_0 : \mu = \mu_0 \qquad \text{and} \qquad H_1 : \mu \neq \mu_1$$

Constructing tests: Likelihood ratio tests

Def: Likelihood Ratio Test (LRT)

The statistic

$$\Lambda(\mathbf{x}) = \frac{\sup_{\theta\in\Omega_0} f_n(\mathbf{x}|\theta)}{\sup_{\theta\in\Omega} f_n(\mathbf{x}|\theta)}$$

is called the *likelihood ratio statistic*. The *likelihood ratio test (LRT)* of

$$H_0 : \theta \in \Omega_0 \quad \text{vs} \quad H_1 : \theta \in \Omega_1$$

is to reject H_0 if $\Lambda(\mathbf{x}) \leq k$ for some constant k

- Note: If $\hat{\theta}$ is the MLE of θ then

$$\sup_{\theta\in\Omega} f_n(\mathbf{x}|\theta) = f_n(\mathbf{x}|\hat{\theta})$$

Example: Z-test as a LRT

- Let $X_1, \ldots, X_2$ be i.i.d. $N(\mu, \sigma^2)$, where σ^2 is known
- Consider the hypotheses

$$H_0 : \mu = \mu_0 \qquad \text{vs} \qquad H_1 : \mu \neq \mu_0$$

- Find the likelihood ratio test of these hypotheses

Example: Two-sided Z-test

Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, 1)$, $n = 25$, and suppose we want to test the hypotheses

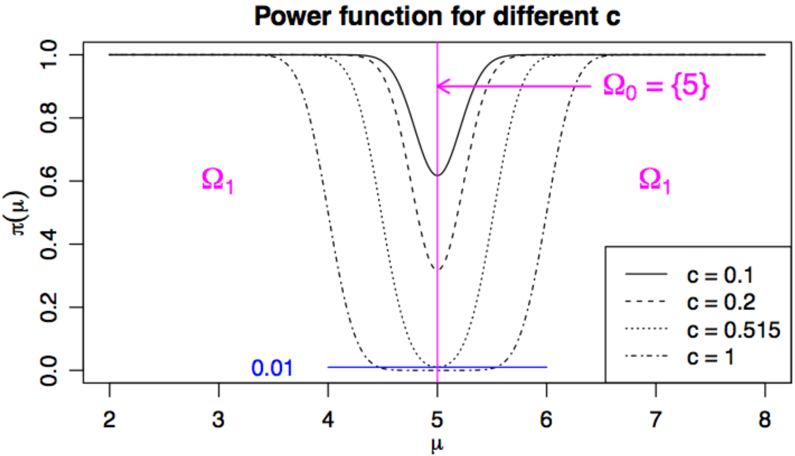
$$H_0 : \mu = \mu_0 \qquad \text{and} \qquad H_1 : \mu \neq \mu_0$$

Let δ_c be the test that rejects H_0 iff $|\overline{X}_n - \mu_0| \geq c$

- Find the power function $\pi(\mu|\delta_c)$
- Find the value of c so that δ_c is of size 0.01
- Find the value of c so that δ_c is of size α where $\alpha \in (0, 1)$

Example: Two-sided Z-test

Power function for $\mu_0 = 5$



Notes on hypothesis testing

- Decisions are expressed in terms of H_0
- “Do not reject H_0" does not mean that we should accept H_0 as true. Some use the phrase “There is no evidence that H_0 is not true”.
- “critical regions vs. “rejection regions"; “