

Topic 15: Simple Hypotheses

November 10, 2009

In the simplest set-up for a **statistical hypothesis**, we consider two values θ_0, θ_1 in the parameter space. We write the test as

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1.$$

H_0 is called the **null hypothesis**. H_1 is called the **alternative hypothesis**.

The possible actions are:

- **Reject the hypothesis.** Rejecting the hypothesis when it is true is called a **type I error** or a **false positive**. Its probability α is called the **size of the test** or the **significance level**.
- **Fail to reject the hypothesis.** Failing to reject the hypothesis when it is false is called a **type II error** or a **false negative**, has probability β . The **power of the test** $1 - \beta$.

	H_0 is true	H_1 is true
reject H_0	type I error	OK
fail to reject H_0	OK	type II error

Given observations X , the rejection of the hypothesis is based on whether or not the data X lands in a **critical region** C . Thus,

$$\text{reject } H_0 \quad \text{if and only if} \quad X \in C.$$

Given a choice α for the size of the test, the choice of a critical region C is called **best or most powerful** if for any choice of critical region C^* for a size α test, and

$$\beta = P_{\theta_1}\{X \notin C\}, \quad \beta^* = P_{\theta_1}\{X \notin C^*\}. \quad (1)$$

we have the lowest probability of a type II error, $\beta \leq \beta^*$.

1 The Neyman-Pearson Lemma

The Neyman-Pearson lemma tell us that the best test for a simple hypothesis is a **likelihood ratio test**.

Theorem 1 (Neyman-Pearson Lemma). *Let $L(\theta|\mathbf{x})$ denote the likelihood function for the random variable X corresponding to the probability measure $P_\theta, \theta \in \Theta$. If there exists a critical region C of size α and a nonnegative constant k such that*

$$\frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} \geq k \quad \text{for } \mathbf{x} \in C$$

and

$$\frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} \leq k \quad \text{for } \mathbf{x} \notin C,$$

then C is the most powerful critical region of size α .

2 Examples

Example 2. Let $X = (X_1, \dots, X_n)$ be independent normal observations with unknown mean and known variance σ^2 . The hypothesis is

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu = \mu_1.$$

The likelihood ratio

$$\begin{aligned} \frac{\mathbf{L}(\mu_1|\mathbf{x})}{\mathbf{L}(\mu_0|\mathbf{x})} &= \frac{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_1-\mu_1)^2}{2\sigma^2}\right) \cdots \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_n-\mu_1)^2}{2\sigma^2}\right)}{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_1-\mu_0)^2}{2\sigma^2}\right) \cdots \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_n-\mu_0)^2}{2\sigma^2}\right)} \\ &= \frac{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_1)^2\right)}{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2\right)} \\ &= \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n ((x_i - \mu_1)^2 - (x_i - \mu_0)^2)\right) \\ &= \exp\left(-\frac{\mu_0 - \mu_1}{2\sigma^2} \sum_{i=1}^n (2x_i - \mu_1 - \mu_0)\right) \end{aligned}$$

The likelihood test is equivalent to

$$(\mu_1 - \mu_0) \sum_{i=1}^n x_i \geq k_1,$$

or for some k_α , known as the **critical value** for the test statistic \bar{x}

$$\bar{x} \geq k_\alpha \quad \text{when } \mu_0 < \mu_1 \quad \text{or} \quad \bar{x} \leq k_\alpha \quad \text{when } \mu_0 > \mu_1.$$

To determine k_α , note that under the null hypothesis, \bar{X} is $N(\mu_0, \sigma^2/n)$ and

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

is a standard normal. Set z_α so that $P\{Z \geq z_\alpha\} = \alpha$. Then, for $\mu_0 < \mu_1$,

$$\bar{X} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha = k_\alpha.$$

For $\mu_1 < \mu_0$, we have $\bar{X} \leq -k_\alpha$.

Equivalently, we can use the standardized score Z as our test statistic and $-z_\alpha$ as the critical value.

The power should

- increase as a function of $|\mu_1 - \mu_0|$,
- decrease as a function of σ^2 , and
- increase as a function of n .

In this situation, the type II error probability,

$$\begin{aligned} \beta &= P_{\mu_1}\{X \notin C\} = P_{\mu_1}\{\bar{X} < \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_\alpha\} \\ &= P_{\mu_1}\left\{\frac{\bar{X} - \mu_1}{\sigma_0/\sqrt{n}} < z_\alpha - \frac{|\mu_1 - \mu_0|}{\sigma_0/\sqrt{n}}\right\} = \Phi\left(z_\alpha - \frac{|\mu_1 - \mu_0|}{\sigma_0/\sqrt{n}}\right) \end{aligned}$$

For $\mu_0 = 10$ and $\mu_1 = 5$ and $\sigma = 3$. Consider the 16 observations and choose a level $\alpha = 0.05$ test, then

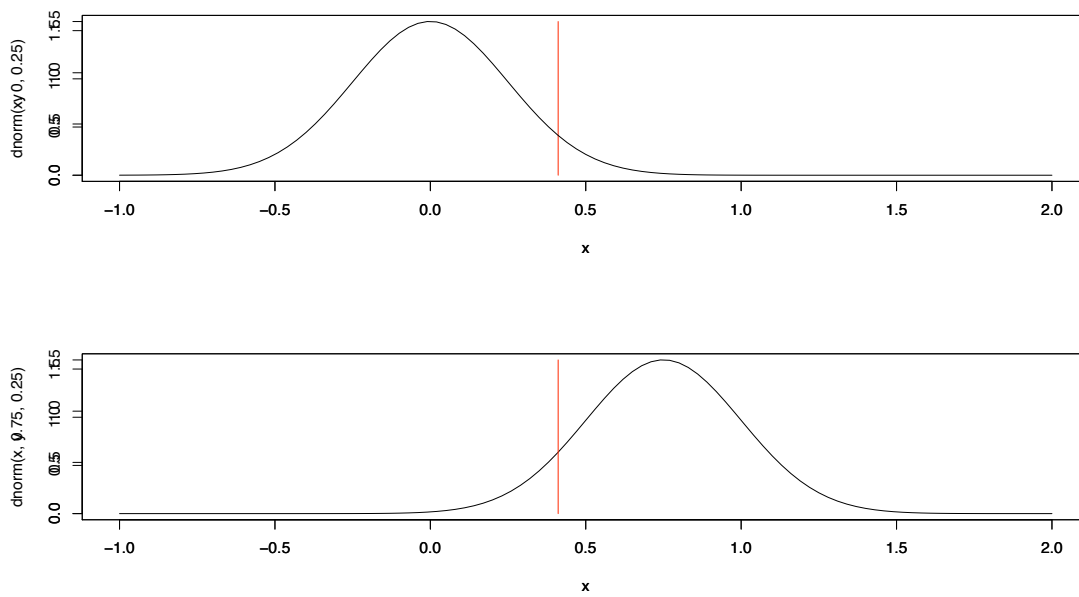


Figure 1: Top graph: Density of \bar{X} for normal data under the null hypothesis - $\mu = 0$ and $\sigma/\sqrt{n} = 0.25$. With an $\alpha = 0.05$ level test, the critical value $k_\alpha = 0.4112$. The area to the right of the red line and below the density function is α . **Bottom graph:** Density of \bar{X} for normal data under the alternative hypothesis. $\mu = 3/4$ and $\sigma/\sqrt{n} = 0.25$. The probability of a type II error is $\beta = 0.0877$. The area to the left of the red line and below the density function is β .

```
> qnorm(0.05)
[1] -1.644854
```

Thus the critical value is $z_\alpha = -1.645$ for the test statistic Z . Not let's look at the data.

```
> x
[1] 8.887753 2.353184 12.123175 10.020566 9.247956 3.711350
[7] 13.907150 9.079790 8.826202 6.288765 12.120783 10.994228
[13] 12.522522 4.529421 8.191806 10.195854
> mean(x)
[1] 8.937532
```

Then

$$Z = \frac{8.937 - 10}{3/\sqrt{16}} = -1.417.$$

$z_\alpha = -1.645 > -1.417$ and we fail to reject the null hypothesis.

To compute the probability of a type II error, note that for $\alpha = 0.05$,

$$z_\alpha - \frac{|\mu_1 - \mu_0|}{\sigma_0/\sqrt{n}} = 1.645 - \frac{5}{3/\sqrt{16}} = -5.022$$

```
>> pnorm(-5.022)
[1] 2.556809e-07
```

This is called the z -test. If n is sufficiently large, the even if the data are not normally distributed, \bar{X} is well approximated by a normal distribution and, as long as the variance σ^2 is known, the z -test is used in this case. In addition, the z -test can be used when $g(\bar{X}_1, \dots, \bar{X}_n)$ can be approximated by a normal distribution using the delta method.

Example 3 (Bernoulli trials). Here $X = (X_1, \dots, X_n)$ is a sequence of Bernoulli trials with unknown success probability θ , the likelihood

$$\mathbf{L}(\theta|\mathbf{x}) = (1 - \theta)^n \left(\frac{\theta}{1 - \theta} \right)^{x_1 + \dots + x_n}.$$

For the test

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1$$

the likelihood ratio

$$\frac{\mathbf{L}(\theta_1|\mathbf{x})}{\mathbf{L}(\theta_0|\mathbf{x})} = \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^n \left(\left(\frac{\theta_1}{1 - \theta_1} \right) / \left(\frac{\theta_0}{1 - \theta_0} \right) \right)^{x_1 + \dots + x_n}$$

Consequently, the test is to reject H_0 whenever

$$\sum_{i=1}^n x_i \geq k_\alpha \quad \text{when } \theta_0 < \theta_1 \quad \text{or} \quad \sum_{i=1}^n x_i \leq k_\alpha \quad \text{when } \theta_0 > \theta_1.$$

Note that under H_0 , $\sum_{i=1}^n X_i$ has a $\text{Bin}(n, \theta)$ distribution. Thus, in the case $\theta_0 < \theta_1$, we choose k_α so that

$$\sum_{k=k_\alpha}^n \binom{n}{k} \theta_0^k (1 - \theta_0)^{n-k} \leq \alpha. \quad (6)$$

In general, we cannot choose k_α to obtain the sum α . Thus, we take the minimum value of k_α to achieve the inequality in (6).

If $n\theta_0$ is sufficiently large, then, by the central limit theorem, $\sum_{i=1}^n X_i$ has a normal distribution. If we standardize

$$Z = \frac{\bar{X} - \theta_0}{\sqrt{\theta_0(1 - \theta_0)/n}}$$

is approximately a standard normal random variable and we perform the z-test as in the previous exercise.

For example, if we take $\theta_0 = 1/2$ and $\theta_1 > 1/2$ and $\alpha = 0.05$, then with 60 heads in 100 coin tosses

$$Z = \frac{0.60 - 0.50}{0.05} = 2.$$

```
> qnorm(0.95)
```

```
[1] 1.644854
```

Thus, $z_{0.05} = 1.645 < 2$ and we reject the null hypothesis.

Example 4. Now consider the hypotheses

$$H_0 : p = 0.7 \quad \text{versus} \quad H_1 : p = 0.8.$$

for a test of the probability that a feral bee hive survives a winter. The test statistic is

$$Z = \frac{\bar{X} - p}{\sqrt{p(1 - p)/n}}$$

and for an α level test, the critical value is z_α where α is the probability that a standard normal is at least z_α

For this study, 112 colonies have been chosen and 88 survive. Thus $\bar{X} = 0.7875$ and $Z = 1.979$. If the significance level is $\alpha = 0.05$, then we will reject H_0 .

If, instead, we collect our data over two years, and use $g(\bar{X}) = \sqrt{\bar{X}}$ as our test statistic, then using the delta method, we have

$$Z = \frac{g(\bar{X}) - g(q)}{|g'(q)|\sqrt{q(1 - q)/n}} = \frac{\sqrt{\bar{X}} - \sqrt{q}}{1/(2\sqrt{q})\sqrt{q(1 - q)/n}} = \frac{\sqrt{\bar{X}} - \sqrt{q}}{\sqrt{(1 - q)/4n}}$$

with $q = p^2 = 0.49$ under the null hypothesis. Thus, if 61 hives survive two years, $Z = 1.1262$ and we fail to reject H_0 .