LECTURE: HYPOTHESIS TESTING (III)

1. Sample Size Selection

Note: Here we mainly focus on the mean μ although something this also works with the proportion p

Setting: Parameter of interest is μ

(1) Alternative hypothesis, $H_a : \mu > \mu_0$

(2) Null hypothesis, $H_0 : \mu = \mu_0$

(3) Test statistic, $\bar{Y}$

(4) Rejection region (RR), $\{\bar{Y} \geq k\}$

Recall:

$$\alpha = \text{Type I error } = P(\text{ reject Null when Null is true })$$
$$\beta = \text{Type II error } = P(\text{ accept Null when Null is false })$$

Last time: Given α we found the threshold k and calculated β

Problem: In real life you don't want to just run the hypothesis test and calculate β. If you did that, β might be too large, and your test could be meaningless!

Goal: Fix α and β and calculate k (and n)

Here you need to fix three more parameters:

(5) μ_a = the true/alternative value of μ

(6) α = the maximum type I error you are willing to accept

(7) β = the maximum type II error you are willing to accept

STEP 1: Using the definition of α and converting to Z:

$$\begin{aligned}\alpha &= P(\bar{Y} \text{ is in RR when Null is true }) \\ &= P(\bar{Y} \geq k) \text{ when } \mu = \mu_0 \\ &= P\left(\frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{k - \mu_0}{\sigma/\sqrt{n}}\right) \\ &= P(Z \geq z_\alpha)\end{aligned}$$

Here z_α is chosen so that $P(Z \geq z_\alpha) = \alpha$

STEP 2: Using the definition of β and converting to Z:

$$\begin{aligned}\beta &= P(\bar{Y} \text{ is outside of RR when Alt is true }) \\ &= P(\bar{Y} \leq k) \text{ when } \mu = \mu_a \\ &= P\left(\frac{\bar{Y} - \mu_a}{\sigma/\sqrt{n}} \leq \frac{k - \mu_a}{\sigma/\sqrt{n}}\right) \\ &= P(Z \leq -z_\beta)\end{aligned}$$

Here z_β is chosen so that $P(Z \leq -z_\beta) = \beta$. We use the negative sign for convenience, since the $z-$value we are looking for will always fall to the left of the graph of the standard normal distribution.

STEP 3: We then have two equations we can solve simultaneously:

$$\frac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha \text{ and } \frac{k-\mu_a}{\sigma/\sqrt{n}} = -z_\beta$$

If we solve both equations for k, we get:

$$k = \mu_0 + z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right) \text{ and } k = \mu_a - z_\beta\left(\frac{\sigma}{\sqrt{n}}\right)$$

Note: The first equation is the same we got in the large sample section

STEP 4: Solving for n we get

$$\begin{aligned}
\mu_0 + z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right) &= \mu_a - z_\beta\left(\frac{\sigma}{\sqrt{n}}\right) \\
(z_\alpha + z_\beta)\left(\frac{\sigma}{\sqrt{n}}\right) &= \mu_a - \mu_0 \\
\sqrt{n} &= \frac{(z_\alpha + z_\beta)\sigma}{\mu_a - \mu_0} \\
n &= \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu_a - \mu_0)^2}
\end{aligned}$$

Summary:

$$k = \mu_0 + z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right) \text{ and } n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu_a - \mu_0)^2}$$

2. P-VALUES

Another way of reporting the probability of a type I error is the p-value. This is what is most often given in the scientific literature.

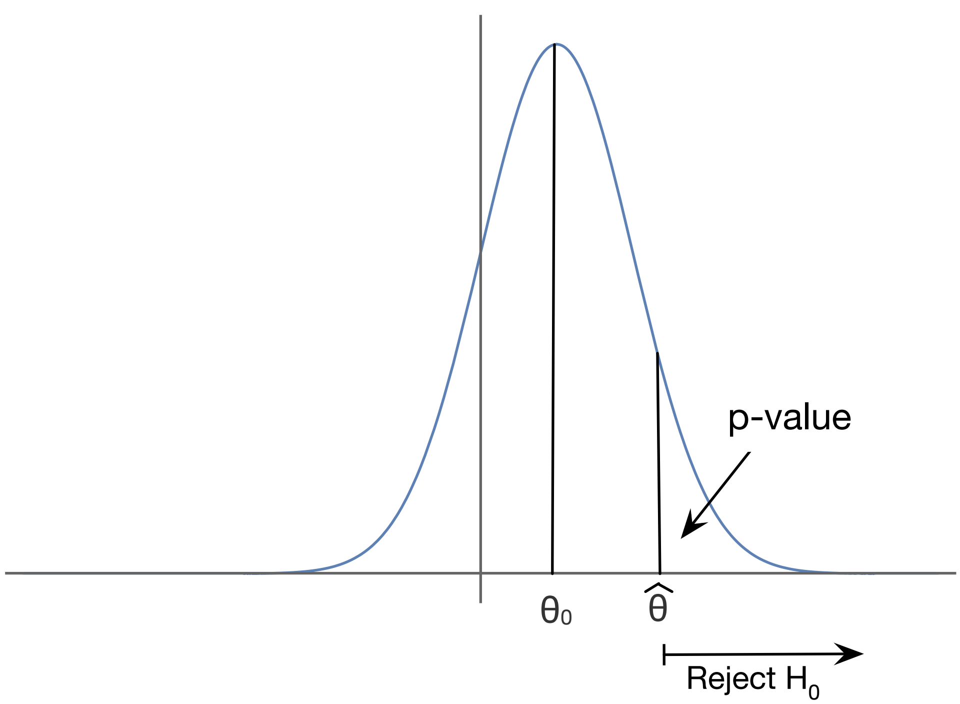
Definition:

If t is the observed value of $\hat{\theta}$ then the $p-$**value** of a test is

$$p = P(\hat{\theta} \geq t \text{ given Null is true })$$

Interpretation: The $p-$value is the smallest α error for which we reject Null. For any α greater than the $p-$value, we reject Null, and for any α lower than the $p-$value, we don't reject Null.

This provides more info than just saying "The null hypothesis was rejected for $\alpha = 0.05$" It's the *smallest* value for which we reject Null.



Example 1:

Ice cream preferences in the US

$p =$ proportion of people who prefer chocolate ice cream

You claim that more than 50% of people prefer chocolate.

Suppose you sample 100 people and 60 of them prefer chocolate

What is the p-value for this test?

The Alt hypothesis is $p > 0.5$ and the null hypothesis is $p = 0.5$

Here $\hat{p} = 0.6$ and we previously estimated $\hat{\sigma} = 0.05$ therefore

$$\begin{aligned}
p\text{-value} &= P(\hat{p} \geq 0.6 \text{ given Null is true}) \\
&= P(\hat{p} \geq 0.6 \text{ given } p = 0.5) \\
&= P\left(\frac{\hat{p} - 0.5}{\hat{\sigma}} \geq \frac{0.6 - 0.5}{\hat{\sigma}}\right) \\
&= P\left(Z \geq \frac{0.1}{0.05}\right) \\
&= P(Z \geq 2) \\
&= 0.0228
\end{aligned}$$

Note: In a previous example, we saw that if $\alpha = 0.05$ we can reject the null hypothesis. This is consistent with our p-value, since 0.05 is greater than the p-value.

3. Small Sample Hypothesis Testing

Question: What happens when the sample size is not large?

Setting: As in the section on confidence intervals, assume Y_i are iid $N(\mu, \sigma^2)$ where σ is unknown

The parameter of interest is μ

(1) Alternative Hypothesis $\mu > \mu_0$

(2) Null Hypothesis $\mu = \mu_0$

(3) Test statistic: $\overline{Y}$

(4) RR $\left\{\overline{Y} > k\right\}$

Goal: Find k given α

What saves us here is (once again) the Student's $t-$distribution

Recall:

$$T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$$

Has a Student's $t-$distribution with $n-1$ df

Therefore we can repeat everything that we've done so far but replace the standard normal Z with T

Example: If we are doing an upper-tail test with desired level α, then the rejection region is given by:

$$k = \mu_0 + t_\alpha \hat{\sigma} = \mu_0 + t_\alpha \left(\frac{S}{\sqrt{n}}\right)$$

Example 2:

You are a rocket scientist, and you conduct an experiment which involves measuring the launch velocity of a model rocket.

You claim that the launch velocity of the model rocket is more than 29 m/s.

Suppose 8 measurements are taken. The sample mean is 29.59 m/s, and the sample standard deviation is 0.391 m/s.

Is the claim supported at the 0.025 level of significance?

Note: We assume that the launch velocities are normally distributed, which is essential to use to use the t-distribution

The parameter of interest is μ

(1) Alternative hypothesis, $H_a : \mu > 29$

(2) Null hypothesis, $H_0 : \mu = 29$

(3) Test statistic, $\bar{Y}$

(4) Rejection region (RR), $\{\bar{Y} \geq k\}$

For the rejection region, we have:

$$\begin{aligned} k &= \mu_0 + t_\alpha \hat{\sigma} \\ &= 29 + t_\alpha \left(\frac{S}{\sqrt{n}} \right) \\ &= 29 + t_{0.025} \left(\frac{0.391}{\sqrt{8}} \right) \\ &= 29 + 2.365(0.138) \\ &= 29.326 \end{aligned}$$

Hence the rejection region is $\{\bar{Y} \geq 29.326\}$

Answer: Since our measurement $\overline{Y} = 29.59$ lies in the rejection region, we reject the null hypothesis with a level of 0.025. In other words, the claim is supported with $\alpha = 0.025$