

ILRST/STSCI 2100 Discussion 9: Ch 10 One Sample Hypothesis Testing

Hypothesis Testing: Comparing experimental data to a claim, or population value.

Steps:

1. State the null hypothesis, H_0

This is the status quo or "claim" $H_0 : \mu = \mu_0$

Ex. $H_0 : \mu = 5$

Note, must have equal sign and must be related to a population parameter (ie. μ or p)

2. State alternative hypothesis, H_a

Ex. $H_a : \mu > 5$ or $H_a : \mu < 5$ (more general, $H_a : \mu > \mu_0$ or $H_a : \mu < \mu_0$) these are one-sided

$H_a : \mu \neq 5$ (more general, $H_a : \mu \neq \mu_0$) this is two-sided

Note, also written in terms of parameters

3. Calculate test statistic

Depends on whether you know σ

If you know σ : $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Here, μ_0 is the value associated with the null hypothesis.

It is more likely that you don't know σ , so you'll have to use the t-distribution, $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

4. Finding a p-value

Depends on alternative hypothesis.

When $H_a : \mu > \mu_0$, p-value = $P(T > \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}})$

When $H_a : \mu < \mu_0$, p-value = $P(T < \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}})$

When $H_a : \mu \neq \mu_0$, p-value = $2 * P(T > | \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} |)$

5. Making a decision about null hypothesis

If p-value $< \alpha$, then reject null hypothesis.

If p-value $> \alpha$, then fail to reject null hypothesis. (don't write "accept null hypothesis")

Similar for proportions:

Calculate the z-score with $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

Be sure to check that $n * p > 10$ and $n * (1 - p) > 10$

Errors:

Type I Error: "a false positive"

When you reject a true null hypothesis

The probability of a Type I error is the α level.

Type II Error: "a false negative"

When you fail to reject a false null hypothesis

The probability of a type II error is denoted by β .

Check lecture notes for example of calculating probability of Type II error.

Practice Problems:

A scientist is interested in testing the claim that the average salinity of the nearby lake is 35 ppm. However, some are concerned that the salinity is above average. 41 samples of lake water are taken, and the mean is found to be 37 ppm, with a standard deviation of 3 ppm. Do the data support the claim?

In general, with a two sided alternative hypothesis, would you be more likely to reject a null hypothesis?

True or false: The p-value is the probability of accepting the null hypothesis. If false, provide the definition of a p-value.

Answers:

$$H_0 : \mu = 35 \text{ ppm}$$

$$H_a : \mu > 35 \text{ ppm}$$

$$T = \frac{37 - 35}{3/\sqrt{41}} = 4.27, \text{ with } df = 41 - 1 = 40$$

Using table 3, we see that the central area is greater than 99.9%, which means that the area under the upper tail must be less than .05%.

Using an $\alpha = 0.05$, the p-value is less than alpha, we we can reject the null hypothesis and conclude that the salinity in the lake is greater than average.

No, since p-value for two-sided alternative hypothesis is larger; two-sided tests are more conservative (more stringent to reject null).

False, p-value is the probability of obtaining an observation equally or more extreme than the one you have observed, assuming the null hypothesis is true.