Section 7: HYPOTHESIS TESTING

Inferential statistics and hypotheses

Hypothesis testing process:

1. State the null and alternative hypotheses.
2. Determine the level of significance.
3. Calculate the test statistic.
4. Find critical value(s) and determine the regions of acceptance and rejection.
5. State the conclusion.

Inferential statistics: Using information we have about the sample to make inferences about the population

Hypothesis: A statement of expectation about a population parameter that we develop for the purpose of testing it

Alternative hypothesis: The abnormality we're looking for in the data; it's the significance we're hoping to find

Null hypothesis: The opposite claim of the alternative hypothesis

Significance level and type I and II errors

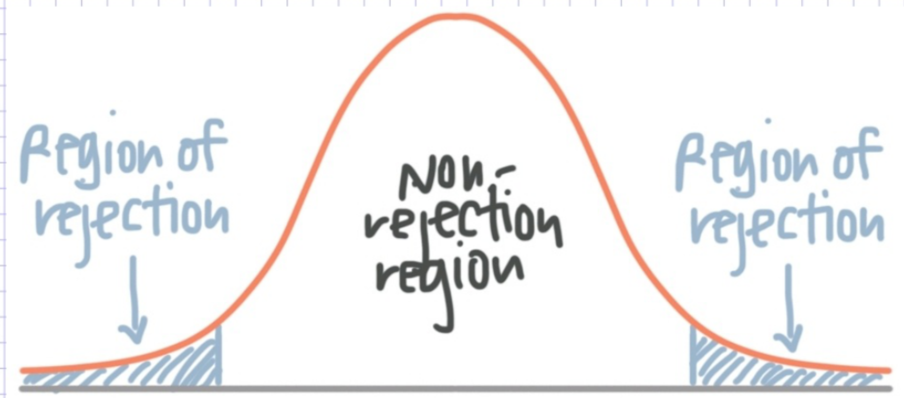
Type I error: When we mistakenly reject a null hypothesis that's actually true. The probability of making a Type I error is given by α, which is the same as the level of significance for the hypothesis test.

Type II error: When we mistakenly accept a null hypothesis that's actually false. The probability of making a Type II error is given by β.

Power: The probability that we'll reject the null hypothesis when it's false (which is the correct thing to do). This is what we want, so we want our test to have a high power.

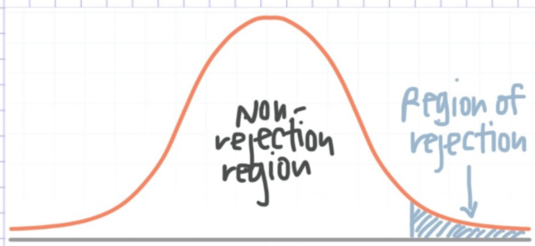
Test statistics for one- and two-tailed tests

Two-tailed test, two-sided test, non-directional test: The alternative hypothesis states that one value is unequal to another, while the null hypothesis states that one value is equal to the other

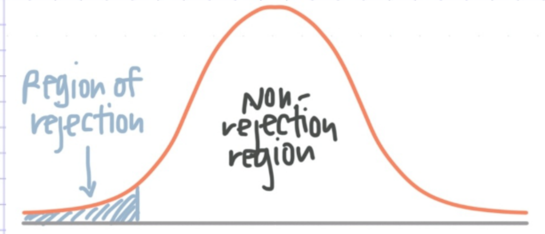


One-tailed test, one-sided test, directional test: Either an upper-tailed test or lower-tailed test

Upper-tailed test, right-tailed test: The alternative hypothesis states that one value is greater than another, while the null hypothesis states that one value is less than or equal to the other



Lower-tailed test, left-tailed test: The alternative hypothesis states that one value is less than another, while the null hypothesis states that one value is greater than or equal to the other



Test statistics:

test statistic = $\frac{\text{observed} - \text{expected}}{\text{standard deviation}}$

When σ is known; $z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

When σ is unknown and/or we have small samples; $t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

For the proportion; $z = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

The p-value and rejecting the null

p-value, observed level of significance: The smallest level of significance at which we can reject the null hypothesis, assuming the null hypothesis is true, or the total area of the region of rejection

If $p \leq \alpha$, reject the null hypothesis

If $p > \alpha$, do not reject the null hypothesis

p-value approach when σ is known:

Lower-tailed test: Reject H_0 when $p \leq \alpha$

Upper-tailed test: Reject H_0 when $p \leq \alpha$

Two-tailed test: Reject H_0 when $p \leq \alpha$

p-value approach when σ is unknown and/or sample size is small:

Lower-tailed test: Reject H_0 when $p \leq \alpha$

Upper-tailed test: Reject H_0 when $p \leq \alpha$

Two-tailed test: Reject H_0 when $p \leq \alpha$

Critical value approach:

Lower-tailed test: Reject H_0 when $z \leq -z_{\alpha}$

Upper-tailed test: Reject H_0 when $z \geq z_{\alpha}$

Two-tailed test: Reject H_0 when $z \leq -z_{\frac{\alpha}{2}}$ or $z \geq z_{\frac{\alpha}{2}}$

Significance, statistical significance: The probability of obtaining the result by chance

Confidence interval for the difference of means

Sampling distribution of the difference of means (SDDM): The probability distribution of every possible difference of means for certain sample sizes n_1 and n_2 from populations 1 and 2, respectively

Mean and standard deviation of the SDDM:

Mean: $\mu_{\bar{x}_1-\bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$

Standard deviation (standard error): $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Confidence interval around the difference of means:

With known σ_1 and σ_2: $(a,b) = (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

With unknown σ_1 and σ_2 and/or small s_1 and s_2:

- Unequal population variances: $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$$\text{with } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2} \quad \text{(round down)}$$

- Equal population variances: $(a,b) = (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \times s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

with $df = n_1 + n_2 - 2$

with pooled standard deviation $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

Hypothesis testing for the difference of means

The statistic with large samples and unequal population variances:

$$z = \frac{(\bar{x}_1-\bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The statistic with large samples and equal population variances:

$$z = \frac{(\bar{x}_1-\bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

The statistic with small samples and unequal population variances:

$$t = \frac{(\bar{x}_1-\bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{with } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2}$$

The statistic with small samples and equal population variances:

$$t = \frac{(\bar{x}_1-\bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{with } df = n_1 + n_2 - 2$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

Matched-pair hypothesis testing

Independent samples: Samples for which the observations from one sample are not related to the observations from the other sample

Dependent samples: Samples for which the observations from one sample are related to the observations from the other sample

Matched-pair test: A hypothesis test with dependent samples

Confidence interval for a matched-pair test:

With σ_d known: $(a,b) = \bar{d} \pm z_{\alpha/2}\frac{\sigma_d}{\sqrt{n}}$

With σ_d unknown and/or $n < 30$: $(a,b) = \bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}}$ with $df = n-1$

Confidence interval for the difference of proportions

Confidence interval around the difference of proportions :

$$(a,b) = (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Negative values in the confidence interval: When the confidence interval contains 0, it means there is likely no difference in proportions. On the other hand, when the confidence interval doesn't contain 0, it means there's likely a difference in proportions.

Hypothesis testing for the difference of proportions

Test statistic for the difference of proportions:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

with $\hat{p} = \dfrac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$