

Hypothesis Testing pt. 3

More Types of Hypothesis Tests

Grinnell College

Hypothesis Test – Single Proportion

$$H_0: p = p_0$$

Under the null hypothesis we have:

$$Z := \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathbf{N}(0,1)$$

- ▶ use `pnorm()` with value of Z

Conditions:

- ▶ Random Sample
- ▶ $n \times p_0 \geq 10$
- ▶ $n \times (1 - p_0) \geq 10$

Hypothesis Test – Difference of Proportions

$$H_0: p_1 - p_2 = 0$$

If $p_1 = p_2$, then both are estimating the same thing.

$$\text{Let } \hat{p}_{pool} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$$

If we are simulating what the null hypothesis looks like, then

$$Z := \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_{pool}(1 - \hat{p}_{pool})}{n_1} + \frac{\hat{p}_{pool}(1 - \hat{p}_{pool})}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}_{pool}(1 - \hat{p}_{pool})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim \mathbf{N(0,1)}$$

► use `pnorm()` with value of Z

Conditions:

- Random Samples
- $n_1 \times \hat{p}_1 \geq 10$ and $n_1 \times (1 - \hat{p}_1) \geq 10$
- $n_1 \times \hat{p}_2 \geq 10$ and $n_1 \times (1 - \hat{p}_2) \geq 10$

Hypothesis Test – Single Mean

$$H_0: \mu = \mu_0$$

The CLT says that $\bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$

If we are simulating what the null hypothesis looks like $\rightarrow \bar{x} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$

Think back to standardizing. If we standardize a Normal distribution it becomes a Standard Normal distribution $N(0,1)$. So...

$$Z := \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathbf{N(0,1)}$$

If we define Z in this way, then we know it follows a Standard Normal distribution and we have a way to calculate p-values.

- ▶ use `pnorm()` function with value of Z

Hypothesis Test – Single Mean

$$H_0: \mu = \mu_0$$

Issue: We probably don't know σ

$$T := \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim \mathbf{t(df = n-1)}$$

- ▶ use `pt()` function with value of T and df

Hypothesis Test – Single Mean

$$H_0: \mu = \mu_0$$

Conditions:

- ▶ Random Sample
- ▶ Normal population **OR** $n \geq 30$

If σ is known:

$$Z := \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathbf{N(0,1)}$$

- ▶ use `pnorm()` function with value of Z

If σ is not known:

$$T := \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim \mathbf{t(df = n-1)}$$

- ▶ use `pt()` function with value of T and df

Hypothesis Test – Difference of Means

$$H_0: \mu_1 - \mu_2 = \mu_0 = 0$$

Conditions:

- ▶ Random Sample
- ▶ Normal population **OR** $n_1 \geq 30$ and $n_2 \geq 30$

If σ is known:

$$Z := \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \sim \mathbf{N(0,1)}$$

- ▶ use `pnorm()` function with value of Z

If σ is not known:

$$T := \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim \mathbf{t(df = \min(n_1, n_2) - 1)}$$

- ▶ use `pt()` function with value of T and df