Hypothesis Testing 4

More Types of Hypothesis Tests

Grinnell College

Hypothesis Test – Single Proportion

H_0: $p = p_0$

Under the null hypothesis we have:

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

- p-val: use Normal chart with value of Z test-stat

Conditions:

- Representative 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.

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

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

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

- p-val: use Normal chart with value of Z test-stat

Conditions:

- Independent Representative samples
- $n_1 \times \widehat{p}_1 \geq 10$ and $n_1 \times (1 - \widehat{p}_1) \geq 10$ (Success conditions)
- $n_2 \times \widehat{p}_2 \geq 10$ and $n_2 \times (1 - \widehat{p}_2) \geq 10$ (Failure conditions)

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...

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

Issue: we probably don't know σ...

Hypothesis Test – Single Mean

H_0: $\mu = \mu_0$

Conditions:

- Representative sample
- Normal population **OR** n large enough (CLT, same rule as before)

Under the null hypothesis we have:

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

- p-val: use t-distribution with df = n-1 and value of T

Hypothesis Test – Difference of Means

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

Conditions:

- Representative samples
- Normal populations **OR** large enough samples (CLT, same as before)

Under the null hypothesis of no difference:

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

- p-val: use t-distribution with df = $\min(n_1, n_2)$ - 1 and value of T