Comparison without σ

Hypothesis Testing with Two Means and Unknown Variance

Prof Hans Georg Schaathun

Høgskolen i Ålesund

24th March 2014

Example Problem

Claim An average Engineering student drink more beer on Saturdays than an average Business student.

- We can form hypotheses
  - $H_0 : \mu_E - \mu_B = 0$
  - $H_1 : \mu_E - \mu_B > 0$
- Poll two populations
  - $E_1, E_2, E_3, \ldots, E_n$
  - $B_1, B_2, B_3, \ldots, B_m$
- Two sample means to use: $\bar{E}$ and $\bar{B}$

Recall

$$Z = \frac{\bar{E} - \bar{B} - \delta_0}{\sqrt{\sigma_E^2/n_E + \sigma_B^2/n_B}}$$

- Null hypothesis: $H_0 : \mu_E - \mu_Y = \delta_0$
- (one-sided) reject if $Z > z_\alpha$
- (two-sided) reject if $|Z| > z_{\alpha/2}$

What do we do when σ_E and σ_B are unknown.

HØGSKOLEN I ÅLESUND
Aalesund University College

Using Sample Standard Deviation

$$Z = \frac{\bar{E} - \bar{B} - \delta_0}{\sqrt{\sigma_E^2/n_E + \sigma_B^2/n_B}}$$

What happens if we replace $\sigma \mapsto s$?

$$T = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

This is t-distributed — not z-distributed

Using the t-distribution

$$T = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

$$T \sim T(\nu)$$

$$\nu = \frac{(s_X^2/n_X + s_Y^2/n_Y)^2}{\frac{(s_X^2/n_X)^2}{n_X - 1} + \frac{(s_Y^2/n_Y)^2}{n_Y - 1}}$$

- If n_X and n_Y are large, then ν is large
- If ν is large, then we use normal distribution

Summary

- Comparing two treatments is a common problem
- Null hypothesis: $H_0 : \mu_X - \mu_Y = 0$
  - more generally $H_0 : \mu_X - \mu_Y = \delta$
- Observable statistic: $D = \bar{X} - \bar{Y}$

$$T = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}} \tag{1}$$

$$\nu = \frac{(s_X^2/n_X + s_Y^2/n_Y)^2}{\frac{(s_X^2/n_X)^2}{n_X-1} + \frac{(s_Y^2/n_Y)^2}{n_Y-1}} \tag{2}$$

- (one-sided) reject if $T > t_\alpha^\nu$
- (two-sided) reject if $|T| > t_{\alpha/2}^\nu$