Small-Sample Tests/CI's for $\mu_1 - \mu_2$

Engineering Statistics
Section 9.2

Josh Engwer

TTU

27 April 2016

Small-Sample t-Test for $\mu_1 - \mu_2$ (Test Statistic)

Proposition

Given any two normal populations with means μ_1 and μ_2.
Let $\mathbf{X} := (X_1, X_2, \ldots, X_{n_1})$ be a random sample from the 1^{st} population.
Let $\mathbf{Y} := (Y_1, Y_2, \ldots, Y_{n_2})$ be a random sample from the 2^{nd} population.
Moreover, let random samples $\mathbf{X}$ & $\mathbf{Y}$ be independent of one another.

Suppose an α-level hypothesis test for $\mu_1 - \mu_2$ is desired with one of the forms:

$$\begin{array}{ll} H_0: & \mu_1 - \mu_2 = \delta_0 \\ H_A: & \mu_1 - \mu_2 > \delta_0 \end{array} \quad \text{OR} \quad \begin{array}{ll} H_0: & \mu_1 - \mu_2 = \delta_0 \\ H_A: & \mu_1 - \mu_2 < \delta_0 \end{array} \quad \text{OR} \quad \begin{array}{ll} H_0: & \mu_1 - \mu_2 = \delta_0 \\ H_A: & \mu_1 - \mu_2 \neq \delta_0 \end{array}$$

Then, the corresponding test statistic $W(\mathbf{X}, \mathbf{Y}; \delta_0)$ is

$$Z = \frac{(\overline{X} - \overline{Y}) - \delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \Longrightarrow \quad Z \sim t_\nu \quad \text{where} \quad \nu = \left\lfloor \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \right\rfloor$$

PROOF: (Beyond the scope of this course.)

Small-Sample t-Test for $\mu_1 - \mu_2$ $(\sigma_1, \sigma_2$ unknown$)$

Proposition

Populations:	*Two Normal Populations with σ_1, σ_2 unknown*
Realized Samples:	$\mathbf{x} := (x_1, x_2, \ldots, x_{n_1})$ *with mean* $\bar{x}$ *and std dev* s_1 $\mathbf{y} := (y_1, y_2, \ldots, y_{n_2})$ *with mean* $\bar{y}$ *and std dev* s_2 *Samples* $\mathbf{x}$ & $\mathbf{y}$ *are independent of one another*
Test Statistic Value $W(\mathbf{x}, \mathbf{y}; \delta_0)$	$t = \dfrac{(\bar{x} - \bar{y}) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $\qquad \nu^* = \left\lfloor \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \right\rfloor$

HYPOTHESIS TEST:	***P-VALUE DETERMINATION:***
$H_0: \mu_1 - \mu_2 = \delta_0$ *vs.* $H_A: \mu_1 - \mu_2 > \delta_0$	*P-value* $= 1 - \Phi_t(t; \nu^*)$
$H_0: \mu_1 - \mu_2 = \delta_0$ *vs.* $H_A: \mu_1 - \mu_2 < \delta_0$	*P-value* $= \Phi_t(t; \nu^*)$
$H_0: \mu_1 - \mu_2 = \delta_0$ *vs.* $H_A: \mu_1 - \mu_2 \neq \delta_0$	*P-value* $= 2 \cdot [1 - \Phi_t(\lvert t\rvert; \nu^*)]$

Decision Rule:

If P-value $\leq \alpha$ *then reject* H_0 *in favor of* H_A

If P-value $> \alpha$ *then accept* H_0 *(i.e. fail to reject* H_0*)*

Small-Sample t-CI for $\mu_1 - \mu_2$ $\quad(\sigma_1, \sigma_2$ unknown)

Proposition

Given two normal populations with means μ_1 and μ_2.
Let $x_1, x_2, \ldots, x_{n_1}$ be a sample taken from the 1^{st} population.
Let $y_1, y_2, \ldots, y_{n_2}$ be a sample taken from the 2^{nd} population.

*Then the $100(1-\alpha)\%$ **small-sample CI for** $\mu_1 - \mu_2$ is approximately*

$$\left((\bar{x}-\bar{y}) - t^*_{\nu^*,\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \quad (\bar{x}-\bar{y}) + t^*_{\nu^*,\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

— *OR WRITTEN MORE COMPACTLY* —

$$(\bar{x}-\bar{y}) \pm t^*_{\nu^*,\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where

$$\nu^* = \left\lfloor \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} \right\rfloor$$

Textbook Logistics for Section 9.2

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Probability of Event	$P(A)$	$\mathbb{P}(A)$
Alternative Hypothesis	H_a	H_A
Sample Sizes	m, n	n_1, n_2
Hypothesized Mean Difference	Δ_0	δ_0

- Ignore "Pooled t Procedures" section (pg 377-378)
  - A **pooled** procedure assumes the pop. variances σ_1^2, σ_2^2 are **equal**.
  - A **pooled estimator of** σ^2 is a **weighted avg** of sample variances S_1^2, S_2^2:

$$S_p^2 := \frac{n_1-1}{n_1+n_2-2} \cdot S_1^2 + \frac{n_2-1}{n_1+n_2-2} \cdot S_2^2$$

  - Pooled t tests are not robust to violations of equal variance assumption.
  - In practice, population variances σ_1^2, σ_2^2 will not be equal.
- Ignore "Type II Error Probabilities" section (pg 378 - 379)
  - Turns out computing β or power of a 2-sample t test is complicated.
- Ignore any mention of **one-sided CI's**.

Fin.