STAT 260 Lecture Notes
Sets 28 and 29 - Hypothesis Tests on the Mean With Two Samples

Suppose we want to investigate the difference (or lack of difference) between two samples.

- Sample 1 has size n_1, mean μ_1, and sample standard deviation s_1 (or standard deviation σ_1).
- Sample 2 has size n_2, mean μ_2, and sample standard deviation s_2 (or standard deviation σ_2).

parameter we are testing

To compare if samples are similar or not, we look at $\underline{\mu_1 - \mu_2}$.

Z or T?

σ_1, σ_2: ns big → Z; ns small → Z

s_1, s_2: ns big → Z; ns small → T: pooled, unpooled

There are four situations to consider:

(a) We know σ_1 and σ_2, and both populations follow the normal distribution.
Z

(b) Both sample sizes are large, that is $n_1 \geq 30$ and $n_2 \geq 30$.
Z

(c) At least one n is small, equal variances. That is, $\dfrac{\text{larger standard deviation}}{\text{smaller standard deviation}} \leq 1.4$.
T pooled

(d) At least one n is small, unequal variances. That is, $\dfrac{\text{larger standard deviation}}{\text{smaller standard deviation}} > 1.4$.
T unpooled

Case 1:

Use the normal distribution (so either we know σ_1 and σ_2 and both populations are normal, or both distributions are anything and the sample sizes are large, i.e. $n_1 \geq 30$ and $n_2 \geq 30$.)

test statistic:

$$Z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}} \qquad \text{or} \qquad Z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Use the confidence interval: $(\overline{x_1} - \overline{x_2}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ or $(\overline{x_1} - \overline{x_2}) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.

Case 2:

At least one n is small, equal variances. So $\dfrac{\text{larger standard deviation}}{\text{smaller standard deviation}} \leq 1.4$

Use the **Pooled t-test.**

test statistic:

$$T = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

degrees of freedom: $\nu = n_1 + n_2 - 2$.

Use the confidence interval: $(\overline{x_1} - \overline{x_2}) \pm t_{\alpha/2}\sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$, again with degrees of freedom $\nu = n_1 + n_2 - 2$.

The value $\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ is sometimes called the **pooled variance** and the notation $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ is used.

Extra assumption needed: Here we need to assume that the data from each of the two populations follow a normal (or near-normal) distribution.

Case 3:

At least one n is small, unequal variances. So $\dfrac{\text{larger standard deviation}}{\text{smaller standard deviation}} > 1.4$

Use the **Unpooled t-test.**

test statistic:

$$T = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

degrees of freedom

$$\nu = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Since degrees of freedom is an integer value, we only use the integer part of the formula for ν above. That is, we always round down with the formula above.

Use the confidence interval: $(\overline{x_1} - \overline{x_2}) \pm t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$, again with degrees of freedom ν using the large formula above.

Extra assumption needed: Here we need to assume that the data from each of the two populations follow a normal (or near-normal) distribution.

Example 1

Consider two independent normal populations and suppose we know σ_1 from population 1 and σ_2 from population 2. (From this we know that the r.v. $\overline{X_1}$ is independent of the r.v. $\overline{X_2}$.)

$$V(\overline{X_1} - \overline{X_2}) = V(\bar{x}_1) + (-1)^2 V(\bar{x}_2) = V(\bar{x}_1) + V(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

The standard error for $\overline{X_1} - \overline{X_2}$ is $= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Example 2

Do grad students experience greater stress levels than undergrad students? We will use the Stress Level Scale to measure. In a study, we rated 97 grad students and 148 undergrad students. The grad students gave a mean of 10.4 on the Stress Level Scale with a standard deviation of 4.83 and the undergrads gave a mean of 9.26 on the Stress Level Scale with a standard deviation of 4.68. Does this data support the hypothesis? Use a significance level of $\alpha = 0.05$.

Group 1: Grads $n_1 = 97$ $\bar{x}_1 = 10.4$ $s_1 = 4.83$

Group 2: Undergrads $n_2 = 148$ $\bar{x}_2 = 9.26$ $s_2 = 4.68$

n_1 and n_2 are large $\Rightarrow$ use z

testing $\mu_1 - \mu_2$ = true difference in means between stress levels for grads (μ_1) and undergrads (μ_2)

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 > \mu_2$
$\Rightarrow$
$H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 > 0$

← rearrange to form $\mu_1 - \mu_2$ because we need this when we standardize.

$\mu_1 - \mu_2$ from H_0

$$z_{obs} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(10.4 - 9.26) - (0)}{\sqrt{\frac{(4.83)^2}{97} + \frac{(4.68)^2}{148}}} = 1.83$$

$p\text{-value} = P(Z > 1.83) = 1 - P(Z \le 1.83)$

same as in H_1

$= 1 - 0.9664$

$= 0.0336$

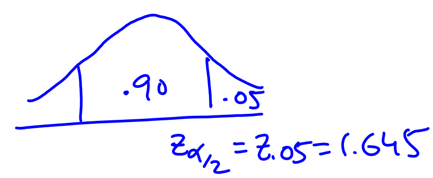
$p\text{-value} = 0.0336 \le \alpha = 0.05 \Rightarrow$ p-value small $\Rightarrow$ reject H_0

We conclude that there is enough evidence to say $\mu_1 > \mu_2$, so grad students do experience more stress than undergrads.

Example 3

Using the info from Example 2, create a 90% confidence interval for $\mu_1 - \mu_2$. What does this confidence interval say about the conclusion we made in the hypothesis test? Does it agree with our earlier findings?

n_1 & n_2 large $\Rightarrow$ use Z.



$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (10.4 - 9.26) \pm 1.645\sqrt{\frac{4.83^2}{97} + \frac{4.68^2}{148}} = [0.1147, 2.1653]$$

The CI estimates $\mu_1 - \mu_2$. Here we can see that all values in the CI range are positive. Thus all reasonable estimates of $\mu_1 - \mu_2$ are positive (e.g. $\mu_1 - \mu_2 > 0$), and so we can say that μ_1 is greater than μ_2.

This does agree with the hypothesis test conclusion in Example 2.

Example 4

The baby food brand Tastee claims it is better than its competitor because it helps babies gain more weight in the first days of life. To test this claim the following observations were made on baby weight gain in the first days of life.

Tastee:	$\overline{x_1} = 36.93$ g	$s_1 = 4.23$ g	$n_1 = 15$
Competitor:	$\overline{x_2} = 31.36$ g	$s_2 = 3.35$ g	$n_2 = 25$

n_1 & n_2 are small $\Rightarrow$ use T (need to check if it's pooled or unpooled)

Use a significance level of $\alpha = 0.05$.

$$\frac{\text{large } s}{\text{small } s} = \frac{4.23}{3.35} = 1.26 \leq 1.4 \implies \text{use pooled } T.$$

testing $\mu_1 - \mu_2$ = the difference in mean weight gain between babies fed Tastee (μ_1) and babies fed competitor (μ_2).

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$ ← trying to prove

$\Longrightarrow$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$

df $\nu = n_1 + n_2 - 2 = 15 + 25 - 2 = 38$

$$T_{obs} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(36.93 - 31.36) - 0}{\sqrt{\frac{14(4.23)^2 + 24(3.35)^2}{15+25-2}\left(\frac{1}{15} + \frac{1}{25}\right)}} = \frac{5.57}{\sqrt{13.68\left(\frac{8}{75}\right)}} = 4.611$$

← same as in H_0

same as in H_1

$$p\text{-value} = P(T_{38} > 4.611)$$

⊛ Notice that $\nu = 38$ is not in the T table. Use closest value $\nu = 40$ instead.

$$P(T_{40} > 4.611) < 0.0005$$

so p-value < 0.0005

p-value < 0.0005 $\qquad \alpha = 0.05$

p-value $\le \alpha \Rightarrow$ p-value is small $\Rightarrow$ reject H_0.

Conclude there is enough evidence to say that $\mu_1 - \mu_2 > 0$ (so $\mu_1 > \mu_2$), so babies fed Tastee brand on average gain more weight than babies fed competitor brand.

Example 5

Using the info from Example 4, create a 95% confidence interval for $\mu_1 - \mu_2$.

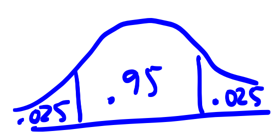
still have n_1 & n_2 small and $\frac{\text{large } s}{\text{small } s} \le 1.4$ so still using pooled T.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\nu, \alpha/2} \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

df $\nu = n_1 + n_2 - 2$
$= 15 + 25 - 2$
$= 38$
so use $\nu = 40$ on table

$$= (36.93 - 31.36) \pm 2.021 \sqrt{\frac{14(4.23^2) + 24(3.35)^2}{15 + 25 - 2}\left(\frac{1}{15} + \frac{1}{25}\right)}$$

$$= [3.129, 8.011]$$



$t_{40, .025} = 2.021$

Example 6

To do a statistical analysis, we can use either Software A or Software B. We want to know if there is a difference in the time it takes to complete the task with the two programs. To test this claim, the following observations are made:

Software A:	$\overline{x_1} = 70.42$ seconds	$s_1 = 20.54$ seconds	$n_1 = 24$
Software B:	$\overline{x_2} = 56.44$ seconds	$s_2 = 9.03$ seconds	$n_2 = 16$

n_1 & n_2 are small $\Rightarrow$ use T (need to check if it's pooled or unpooled)

Use a significance level of $\alpha = 0.05$.

$$\frac{\text{large } s}{\text{small } s} = \frac{20.54}{9.03} = 2.27 > 1.4 \Rightarrow \text{use unpooled } T$$

testing $\mu_1 - \mu_2$ = true difference in mean time for Software A (μ_1) and mean time for Software B (μ_2)

$H_0: \mu_1 = \mu_2 \quad \Longrightarrow \quad H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 \neq \mu_2 \quad\quad\quad\quad H_1: \mu_1 - \mu_2 \neq 0$

trying to prove

since $\neq$ in H_1, this is a two-tailed test

$$T_{obs} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(70.42 - 56.44) - 0}{\sqrt{\frac{20.54^2}{24} + \frac{9.03^2}{16}}} = 2.936$$

$$\text{df} \quad \nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(20.54^2/24 + 9.03^2/16\right)^2}{\frac{(20.54^2/24)^2}{23} + \frac{(9.03^2/16)^2}{15}}$$

$$= \frac{514.1611993}{13.43542589 + 1.73148916} = \underbrace{33}.90018324$$

The integer part. Use $\nu = 33$.

But $\nu = 33$ is not on the table so use $\nu = 30$.

$$p\text{-value} = P(T_{30} < -2.936) + P(T_{30} > 2.936)$$
$$= 2 \cdot P(T_{30} > 2.936)$$

$$0.0025 < P(T_{30} > 2.936) < 0.005$$
$$2(0.0025) < p\text{-value} < 2(0.005)$$
$$0.005 < p\text{-value} < 0.01$$

$0.005 < p\text{-value} < 0.01$ $\quad \alpha = 0.05$

$p\text{-value} \leq \alpha \Rightarrow$ p-value is small $\Rightarrow$ reject H_0

Conclude there is enough evidence to say that $\mu_1 - \mu_2 \neq 0$, so $\mu_1 \neq \mu_2$. There is a difference in average time for Software A and Software B.

Example 7

Using the info from Example 6, create a 95% confidence interval for $\mu_1 - \mu_2$.

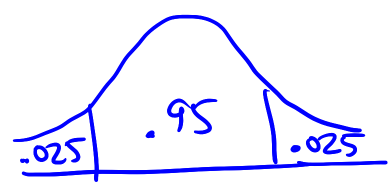
still have n_1 & n_2 small and $\frac{\text{large } s}{\text{small } s} > 1.4 \Rightarrow$ use unpooled T.

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

still have $\nu = 33$ so use $\nu = 30$ on the table.

$$= (70.42 - 56.44) \pm 2.042 \sqrt{\frac{20.54^2}{24} + \frac{9.03^2}{16}}$$

$$= [4.2563, 23.7037]$$



$t_{30, .025} = 2.042$