The *t* test for two independent samples

Chapter 10

Class outline 7-24-08

- review single sample t-test
- Independent samples t-test (chapter 10)
- Questions about problem set?
- REMEMBER TO MEET IN TILLETT 119 TOMORROW.

t-test for independent samples

- Most research compares two (or more) sets of data (two samples as opposed to just one sample)
  - The research question concerns a mean difference between the two sets of data.
- Two sets of data could come from separate samples (independent-measures) or from the same sample (between-subjects)
- Today we will be looking at independent measures only

t-test for independent samples

- Independent-measures (also called between-subjects) design uses two separate samples to answer a question about two populations
- Subscripts in notation
  - n_1 n_2, etc.
- Goal is to evaluate the mean difference between two populations (or two treatment conditions)

t-test for independent samples

- The t statistic for comparing two independent samples is called the **independent-measures t statistic** (the t statistic we learned about earlier is referred to as the **single-sample t statistic)**
- Now, hypotheses are:
  - H_0: $\mu_1 - \mu_2 = 0$
  - H_1: $\mu_1 - \mu_2 \neq 0$

t-test for independent samples

The basic structure of the t is still the same:

$$\frac{\text{sample statistic} - \text{hypothesized population parameter}}{\text{estimated standard error}}$$

- but for the independent samples t:
  - Sample statistic = the difference between sample 1 and sample2
  - Hypothesized population parameter = the hypothesized difference between population 1 and population 2 (based on the null)
  - Estimated standard error = the amount of error that is expected when you use a sample mean difference ($M_1 - M_2$) to represent a population mean difference ($\mu_1 - \mu_2$)

t-test for independent samples

- In symbols:
  - $t = \dfrac{(M_1 - M_2) - (\mu_1 - \mu_2)}{s_{(M_1 - M_2)}}$
- But frequently the null hypothesis sets the expected population mean difference equal to zero
- Simplifies to:
  - $t = \dfrac{(M_1 - M_2)}{s_{(M_1 - M_2)}}$

Estimated standard error of $M_1 - M_2$

- What does standard error mean?
- Each of the two sample means represents its own population mean, but in each case there is some error.
  - M_1 approximates μ_1 2with some error
  - M_2 approximates μ_2 2with some error
- Need to combine (or pool) these errors.
  - $S_{(M_1 - M_2)} = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$

But this only works if $n_1 = n_2$

Pooled variance and estimated standard error of $M_1 - M_2$

- So we need to compute the pooled variance (average of the two sample variances) to adjust for unequal ns (gives more weight to the larger sample).
- Remember: $s^2 = SS / df$
- $s^2_p = \dfrac{SS_1 + SS_2}{df_1 + df_2}$
- We then use the pooled variance in place of the individual sample variances to obtain an unbiased estimate of the standard error for a sample mean difference.
- $s_{(M_1 - M_2)} = \sqrt{\dfrac{s^2_p}{n_1} + \dfrac{s^2_p}{n_2}}$

t-test for independent samples

- Is it significant?
- Need to calculate *df*
  - Independent samples t: $df_{total} = df_1 + df_2$
- Use to look up critical t value in table

One-sample vs. Independent Samples

	Sample data	Hypothesized population parameter	Estimated standard error	Sample variance
Single-sample t statistic	M	μ	$\sqrt{\frac{s^2}{n}}$	$s^2 = \frac{SS}{df}$
Independent-measures t statistic	$(M_1 - M_2)$	$(\mu_1 - \mu_2)$	$\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$	$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$

Table 10-1 (p. 265)
The basic elements of a t statistic for the single-sample t and the independent-measures t.

Example – learning styles

- Educational researcher wants to compare kids in a cooperative learning classroom to kids in a regular classroom (control group)
- Is cooperative learning more effective?
- Example on board
- APA style

Effect Size - *d*

- Cohen's $d = \dfrac{\text{mean difference}}{\text{standard deviation}}$
- best estimate of standard deviation = the pooled standard deviation
- $d = \dfrac{M_1 - M_2}{\sqrt{s^2_p}}$

Effect Size – r^2

- measures how much of the variability in scores can be explained by the treatment effect.
- calculated same as before
- $r^2 = \dfrac{t^2}{t^2 + df}$

One-tailed test

- If have a directional hypothesis, can do a one-tailed test
- Only difference is use different t_{crit}
- Generally prefer 2-tailed tests

Assumptions of Independent Samples *t* tests

- Observations within each sample must be independent
- Two populations from which the samples were drawn must be normally distributed
- Two populations from which the samples were drawn must have equal variances (homogeneity of variance)

More examples…

- z-test example
- one sample t-test example
- independent samples t-test example

Questions?

REMEMBER TO MEET IN TILLETT 119 TOMORROW!!