**Psy 230
Independent Samples t-test**

I. Independent Samples





**Figure 10-3 (p. 314)**Two population distributions. The scores in population I vary from 50 to 70 (a 20-point spread). The scores in population II range from 20 to 30 (a 10-point spread). If you select one score from each of these two populations, the closest two values are *X*1 = 50 and *X*2 = 30. The two values that are farthest apart are *X*1 = 70 and *X*2 = 20.

A. Two Sources of Variability (error)

- each of the two samples will have some error as *M* represents 

- would be nice to simply add and then average the estimated standard errors from each sample

- can't (unless samples are the same size)

 - Pooled Variance --allows the bigger sample to carry more weight in determining the final value



B. The formula

$t= \frac{(M\_{1-M\_{2}})-(μ\_{1-μ\_{2}})}{S(M\_{1}-M\_{2})}$ where

$S\left(M\_{1}-M\_{2}\right)= \sqrt{\frac{s\_{p}^{2}}{n\_{1}}+\frac{s\_{p}^{2}}{n\_{2}}}$

C. Degrees of Freedom

df = df1 + df2

= (n1 - 1) + (n2 - 1)

D. Assumptions

1. observations in each sample are independent

2. underlying populations are normal

            3. the two populations being compared have equal variances (homogeneity of variance)

II. Single Sample vs. Two Independent Samples



A developmental psychologist would like to examine the difference in verbal skills for 10-year-old boys vs. 10-year-old girls. A sample of 10 boys and 10 girls is obtained and each child is given a standardized verbal abilities test. Do these data indicate a significant difference in verbal skills for boys compared to girls? Use two-tailed test and set alpha = .05.

Girls
*M* = 37
*SS* = 150
*n* = 10

Boys
*M* = 31
*SS* = 210
*n* = 10

III. Variability and Effect Size





**D.  Effect Size**

Important limitation of the hypothesis testing procedure:

 It makes a relative comparison: the size of the treatment effect relative to the difference expected by chance. If the standard error is very small, then the treatment effect can also be very small and still be bigger than chance.

Therefore, a significant effect does not necessarily mean a big effect.

 Also, if the sample size is large enough, any treatment effect, no matter how small, can be enough for us to reject the null hypothesis.



**Figure 8-11** The appearance of a 15-point treatment effect in two different situations. In part (a), the standard deviation is σ = 100 and the 15-point effect is relatively small. In part (b), the standard deviation is σ = 15 and the 15-point effect is relatively large. Cohen’s *d* uses the standard deviation to help measure effect size.

**Calculating effect size:**

*Cohen's d = mean difference / standard deviation*

***d* Evaluation**

 0.2 Small effect

 0.5 Medium effect

       0.8 Large effect

       1.10                    Very Large

       1.40                     Extremely Large

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Alternative effect size for t-tests: *r2***

*r2 = t2 / (t2 + df)*

Advantage to this one is that people are familiar with it.

 Values range from 0.00 to 1.00.

 What proportion of the total variability in the scores is accounted for by the treatment?

**Magnitude of *r2* Evaluation**

.09 and below Small effect

between .09&.25 Medium effect

over .25 Large effect

 **The Relationship between Power and Effect Size**



 

**Think about the following:**
Suppose that a researcher normally uses an alpha level of .01 for hypothesis tests, but this time uses an alpha level of .05.

a) What does this change in alpha level do to the amount of power?

b) What does this change in alpha do to the risk of a Type I error?