Week 5 Confidence Interval

4. Comparing two proportions/means

Stat 140 - 04

Mount Holyoke College

Dr. Shan Shan

Slides posted at *http://sshanshans.github.io/stat140*

Parameter	Distribution	Standard Error
Proportion	Normal	$\sqrt{\frac{p(1-p)}{n}}$
Difference in Proportions	Normal	$\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}$
Mean	t, df = n – 1	$\sqrt{\frac{\sigma^2}{n}}$
Difference in Means	t, df = min(n_1, n_2) – 1	$\sqrt{\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2}}$
Correlation	t, df = n – 2	$\sqrt{\frac{1-\rho^2}{n-2}}$

1. Examples
    1. Sleep versus Caffeine
    2. A horror example

- Students were given words to memorize, then randomly assigned to take either a 90 min nap, or a caffeine pill. 2 $1/2$ hours later, they were tested on their recall ability.
- Variables
  - Explanatory variable: sleep or caffeine
  - Response variable: number of words recalled
- Is sleep or caffeine better for memory?

Mednick, Cai, Kanady, and Drummond (2008). "Comparing the benefits of caffeine, naps and placebo on verbal, motor and perceptual memory," Behavioral Brain Research, 193, 79-86.

Poll question

What is the parameter of interest in the sleep versus caffeine experiment?

a. Proportion
b. Difference in proportions
c. Mean
d. Difference in means

- Let μ_1 and μ_2 be the mean number of words recalled after sleeping and after caffeine.
- Is there a difference in average word recall between sleep and caffeine?

Is $\mu_1 = \mu_2$?
Compute a 95% confidence interval for $\mu_1 - \mu_2$.

$$(\bar{x}_1 - \bar{x}_2) \pm t^*_{\min(n_1,n_2)-1}\left(\sqrt{\frac{s_1^2}{n_1}} + \sqrt{\frac{s_2^2}{n_2}}\right)$$

The summaries are given as follows

$\bar{x}_1 = 45$, $\bar{x}_2 = 39$
$s_1 = 4.36$, $s_2 = 2.57$
$n_1 = 20$, $n_2 = 30$
$t^*_{19} = 2.09$

The confidence interval is therefore computed as

$$(45 - 39) \pm 2.09 \left(\sqrt{\frac{4.36^2}{20}} + \sqrt{\frac{2.57^2}{30}} \right) = (3, 9)$$

We are 95% confident that the mean number of words recalled after sleeping is between 3 to 9 higher than the mean number of words recalled after caffeine.

- Randomization condition: do data come from a random sample or suitably randomized experiment?
- Nearly normal condition: do data come from a distribution that is unimodal and symmetric.
- **New:** Independent Groups condition: are the two groups independent?

The slasher horror film has been deplored based on claims that it depicts eroticized violence against predominately female characters as punishment for sexual activities.

Is **survival** for female characters in slasher films associated with **sexual activity**?

Gender	Sexual activity	Outcome of physical aggression		n
		Survival	Death	
Female				
	Present	13.3% (n=11)	86.7% (n=72)	83
	Absent	28.1% (n=39)	71.9% (n=100)	139
Male				
	Present	9.5% (n=7)	90.5% (n=67)	74
	Absent	14.8% (n=28)	85.2% (n=161)	189

Let

- p_1 denote the survival rate when there is sexual activity present in the movie
- p_2 denote the survival rate when there is no sexual activity present in the movie

Is $p_1 = p_2$?
Compute a 95% confidence interval for $p_1 - p_2$.

$$\hat{p}_1 - \hat{p}_2 \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

We found,

$\hat{p_1} = 0.133$, $\hat{p_2} = 0.281$
$n_1 = 83$, $n_2 = 139$
$z^* = 2$

The confidence interval is therefore

$$(0.133 - 0.281) \pm 2 \left(\sqrt{\frac{0.133(1-0.133)}{83} + \frac{0.281(1-0.281)}{139}} \right)$$

This is simplified to be

$$(-0.255, -0.041)$$

We are 95% confident that the survival rate for female characters when there is sexual activity present in the movie is between 25.5% to 4.1% lower than when there is no sexual activity present.

- Randomization condition: do data come from a random sample or suitably randomized experiment?
- The 10% Condition: If the data are sampled without replacement, the sample should not exceed 10% of the population without an adjustment to the standard deviation.
- Independent Groups condition: are the two groups independent?
- **New:** Success/Failure Condition: Both groups are big enough that at least 10 successes and at least 10 failures have been observed in each.