MAT 135, Discussion Questions 4.27

1. How can we tell the difference between a situation where two-samples are dependent versus when they are independent?

dependent samples are related, usually coded by # or by the same person at different times

independent if collected separately, can have different sample sizes

2. What is the formula for the test statistic for a 2-sample proportion test? What do all the variables in the equation stand for?

$$z = \frac{p_1 - p_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

(if difference is assumed to be zero, or add $-\Delta_p$)

$$\bar{p} = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$ (weighted mean)

p_1, p_2 are sample proportions

n_1, n_2 are corresponding sample sizes

3. What is the formula for the confidence interval for the difference between two population proportions?

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

4. Under the assumption that the sample size for both groups is the same, what is the formula for determining the minimum sample size for a confidence interval with a given margin of error E?

$$n = \frac{z_{\alpha/2}\sqrt{2\bar{p}(1-\bar{p})} + z_\beta\sqrt{p_1(1-p_1)+p_2(1-p_2)}}{p_1 - p_2}$$

z_β is a power test

5. The Harris Poll conducted a survey in which they asked "How many tattoos do you currently have on your body?" Of the 1205 males surveyed, 181 said that they had at least one tattoo. Of the 1097 females surveyed, 143 responded that they had at least one tattoo. Conduct a hypothesis test to determine if this evidence supports a claim that men have more tattoos than women at the $\alpha = 0.01$ level of significance.

2 prop Z test

$H_0: p_1 = p_2$

$H_a: p_1 \neq p_2$

$X_1 = 181$

$N_1 = 1205$

$X_2 = 143$

$n_2 = 1097$

$p_1 \neq p_2$

$Z = 1.3679$

$p = .1713$ fail to reject null

There is not enough evidence to think men have more tattoos than women

6. In clinical trials for the treatment of a skin disorder, 642 of 2105 patients receiving the current standard treatment were cured of the disorder, and 697 of 2115 patients receiving a new proposed treatment were cured of the disorder. Does the new procedure have a higher cure rate than the old method? Construct a 95% confidence interval to test the claim.

2 prop Z test

$H_0: p_1 = p_2$

$H_a: p_1 < p_2$

$X_1 = 642$ $X_2 = 697$

$n_1 = 2105$ $n_2 = 2115$

$p_1 < p_2$

$Z = -1.71$

$p = .043 < .05$

reject null

2 Prop Z Int

$X_1 = 642$

$n_1 = 2105$

$X_2 = 697$

$n_2 = 2115$

C level = .95

(-.0526, .00351)

Confidence interval includes zero (barely) so fail to reject null

7. Redo the above problem using a hypothesis test. Does your answer change? If it does, why might that happen?

difference is because of slightly different assumptions about hypothesis tests vs. confidence intervals.

8. In many colleges and universities around the country, educators are changing their approach to instruction from a "teacher/lecture-centered model" to a "student-centered model" where students learn in a laboratory environment in which students can proceed at a pace suitable to their learning needs and lecture is de-emphasized. In one school where this model was introduced, of the 743 students who enrolled in the traditional lecture model, 364 passed; of the 567 in the student-centered model, 335 passed. Is this good evidence that the new approach has a higher pass rate?

2 Prop Z test

$H_0: p_1 = p_2$

$H_a: p_1 < p_2$

$X_1 = 364$

$n_1 = 743$

$X_2 = 335$

$n_2 = 567$

$p_1 < p_2$

$Z = 3.62788$

$p = 1.429 \times 10^{-4} < .05$

reject null

Yes, There is good evidence this model works

9. If we wanted to test that hypothesis at AACC at the 5% significance level, what kind of sample size would be needed for both groups if we wanted to reject H_0 if the passing rate of the student-centered model was 3% different than the traditional lecture-model? For a conservative estimate, assume the passing rate of the course is only 50% as that will give us the most conservative estimate of the sample size needed.

$$n = 1.64^2[(.5^2) + .5^2]/.03^2$$

$$= 1494.22$$

$\longrightarrow$ 1495 for both groups