```
Stat 145, Wed 27-Oct-2021 -- Wed 27-Oct-2021
Biostatistics
Spring 2021


------------------------------
Wednesday, October 27th 2021
------------------------------
Due))   WW ch06Part2 due at 11 pm

------------------------------
Wednesday, October 27th 2021
------------------------------
Wk 9, We
Topic:: Two-proportion hypothesis tests
Read::  Lock5 6.9
HW((   WW ch06Part3 due Sat.
```

Hypothesis testing for the difference of two proportions

For problems in this vein, we have some binary categorical variable (like for 1-proportion problems) of interest, but wish to compare proportions of "successes" in two distinct groups/populations. We have samples of size n_1 and n_2 respectively from these groups, yielding sample proportions $\hat{p}_1, \hat{p}_2$.

If we can assume *both* $\hat{p}_1$, $\hat{p}_2$ have normal distributions (i.e., rules of thumb are met), then the prototype standard error for the difference $\hat{p}_1 - \hat{p}_2$ is

$$\text{SE}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

What we used for building a confidence interval was

$$\text{SE}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.$$

— modified for when
• don't have p_1, p_2
• constructing a CI

When conducting an hypothesis test, where the null hypothesis is that the groups are not really different with regards to the binary categorical explanatory variable (Which-group? variable), we have

$$\mathbf{H}_0 : \; p_1 - p_2 = 0$$

This leads us to replace both $\hat{p}_1, \hat{p}_2$ with a **pooled proportion**

$$\hat{p} \;=\; \frac{(\text{number of successes in sample 1}) + (\text{number of successes in sample 2})}{n_1 + n_2},$$

in which case the approximate standard error is

$$\text{SE}_{\hat{p}_1 - \hat{p}_2} \;=\; \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \;=\; \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}.$$

same

It is this standard error we use to obtain a standardized z-statistic

$$z = \frac{\widehat{p_1} - \widehat{p_2}}{SE_{\widehat{p_1} - \widehat{p_2}}}.$$

(test statistic) = (point estimate) / SE

(unstandardized) test statistic

1. 379 of 460 females support tougher gun-control laws, 318 of 520 males
2. 10 of 24 cocaine addicts treated with desipramine had relapses, compared with 20 of 24 who received placebo
3. Kidsfeet
   domhand vs. beggerfoot

1. $\hat{p}_f = \frac{379}{460}$, $\hat{p}_m = \frac{318}{520}$, and $\hat{p} = \frac{379+318}{460+520} = \frac{697}{980}$

$$SE_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{697}{980}\left(1 - \frac{697}{980}\right)\left(\frac{1}{460} + \frac{1}{520}\right)} = 0.029$$

Standardized test statistic

$$z = \frac{\hat{p}_f - \hat{p}_m}{SE} = \frac{379/460 - 318/520}{0.029} = 7.32$$