JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Section E

Fisher's Exact Test

Recall: HIV Transmission/AZT Example

- Recall 2X2 (contingency) table

		Drug Group		
		AZT	Placebo	
HIV Transmission	Yes	13	40	53
	No	167	143	310
		180	183	363

Hypothesis Testing Problem: AZT and HIV Transmission

- Testing equality of two population proportions using data from two samples

  - H_o: $p_1 = p_2$ $\quad\quad$ H_o: $p_1 - p_2 = 0$
  - H_a: $p_1 \neq p_2$ $\quad\quad$ H_A: $p_1 - p_2 \neq 0$

  - In the context of the 2x2 table, this is testing whether there is a relationship between the rows (HIV status) and columns (treatment type)

Statistical Test Procedures

- (Pearson's) Chi-Square Test (x^2)/Two-sample z-test
  - Both based on central limit theorem "kicking in"
  - Both results are "approximate," but are excellent approximations if sample sizes are large
  - These do not perform so well in smaller samples

Statistical Test Procedures

- Fisher's Exact Test
  - Calculations are difficult
  - Always appropriate to test equality of two proportions
  - Computers are usually used
  - Exact p-value (no approximations)—no minimum sample size requirements

Fisher's Exact Test: HIV Transmission/AZT

- Rationale
  - Suppose H_o is true—no association between AZT and maternal/infant HIV transmission
  - Imagine putting 53 red balls (the infected) and 310 blue balls (non-infected) in a jar
  - Shake it up

Fisher's Exact Test

- Now choose 180 balls (that's AZT group)
  - The remaining balls are the placebo group
- We calculate the probability you get 13 or fewer red balls among the 180
  - That is the one-sided p-value
- The two-sided p-value is just about (but not exactly) twice the one-sided
- p-value
  - It accounts for the probability of getting either extremely few red balls or a lot of red balls in the AZT group
  - The p-value is the probability of obtaining a result as or more extreme (more imbalance) than you did by chance alone

Using Stata: AZT/HIV Example

- Results from *csi* command, with *exact* option

```
. csi 13 40 167 143, exact

                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+----------
           Cases |        13          40  |         53
        Noncases |       167         143  |        310
-----------------+------------------------+----------
           Total |       180         183  |        363
                 |                        |
            Risk |  .0722222    .2185792  |   .1460055
                 |                        |
                 |      Point estimate    |   [95% Conf. Interval]
                 |------------------------+----------------------
 Risk difference |         -.146357       |  -.2171766   -.0755374
      Risk ratio |         .3304167       |   .1829884    .5966235
 Prev. frac. ex. |         .6695833       |   .4033765    .8170116
 Prev. frac. pop |         .3320248       |
                 +-----------------------------------------------
                               1-sided Fisher's exact P = 0.0001
                               2-sided Fisher's exact P = 0.0001
```

Small Sample Application

- Sixty-five pregnant women, all who were classified as having a high risk of pregnancy induced hypertension, were recruited to participate in a study of the effects of aspirin on hypertension*

- The women were randomized to receive either 100 mg of aspirin daily, or a placebo during the third trimester of pregnancy

Notes: *Schiff, E. et al. The use of aspirin to prevent pregnancy-induced hypertension and lower the ratio of thromboxane A2 to prostacyclin in relatively high risk pregnancies. *New England Journal of Medicine*, 321, 6.

Display Data in a 2x2 Table

- Results

		Group		
		Aspirin	Placebo	
Hypertension	Yes	4	11	15
	No	30	20	50
		34	31	65

Display Data in a 2x2 Table

- Sample proportion of subjects with hypertension

$$\hat{p}_{aspirin} = \frac{4}{34} = .12$$

$$\hat{p}_{placebo} = \frac{11}{31} = .35$$

Smaller Sample

- In this example . . . (just FYI)

$$n_{aspirin} * \hat{p}_{aspirin} * (1 - \hat{p}_{aspirin}) = 34 * .12 * .88 = 3.6$$

$$n_{placebo} * \hat{p}_{placebo} * (1 - \hat{p}_{placebo}) = 31 * .35 * 65 = 7.1$$

Fishers Exact

- Results from *csi* command, with *exact* option

```
. csi 4 11  30 20, exact

                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+----------
           Cases |         4          11  |         15
        Noncases |        30          20  |         50
-----------------+------------------------+----------
           Total |        34          31  |         65
                 |                        |
            Risk |  .1176471    .3548387  |   .2307692
                 |                        |
                 |      Point estimate    |   [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |        -.2371917       | -.4374335   -.0369498
      Risk ratio |         .3315508       |  .1176925    .9340096
 Prev. frac. ex. |         .6684492       |  .0659904    .8823075
 Prev. frac. pop |         .3496503       |
                 +-------------------------------------------------
                                  1-sided Fisher's exact P = 0.0236
                                  2-sided Fisher's exact P = 0.0378
```

Chi Square

- Results from *csi* command, without *exact* option

```
. csi 4 11  30 20

                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
           Cases |         4          11  |         15
        Noncases |        30          20  |         50
-----------------+------------------------+------------
           Total |        34          31  |         65
                 |                        |
            Risk |  .1176471    .3548387  |   .2307692
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |        -.2371917       | -.4374335   -.0369498
      Risk ratio |         .3315508       |  .1176925    .9340096
 Prev. frac. ex. |         .6684492       |  .0659904    .8823075
 Prev. frac. pop |         .3496503       |
                 +-------------------------------------------------
                           chi2(1) =     5.14  Pr>chi2 = 0.0234
```

Fishers Exact

- 95% CI: not quite correct in smaller samples, but “good enough”

```
. csi 4 11  30 20, exact

                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+----------
           Cases |         4          11  |         15
        Noncases |        30          20  |         50
-----------------+------------------------+----------
           Total |        34          31  |         65
                 |                        |
            Risk |  .1176471    .3548387  |   .2307692
                 |                        |
                 |      Point estimate    |   [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |        -.2371917       | -.4374335   -.0369498
      Risk ratio |         .3315508       |  .1176925    .9340096
 Prev. frac. ex. |         .6684492       |  .0659904    .8823075
 Prev. frac. pop |         .3496503       |
                 +-------------------------------------------------
                              1-sided Fisher's exact P = 0.0236
                              2-sided Fisher's exact P = 0.0378
```

Comparing Proportions between Independent Populations

- To get a p-value for testing:
  - H_o: $p_1 = p_2$
  - H_A: $p_1 = p_2$

- Two sample z-test or chi-squared test (give same p-value): work better in "bigger" samples and will match results of Fishers Exact Test

- Fisher's exact test: always appropriate

Comparing Proportions between Independent Populations

- To create a 95% confidence interval for the difference in two proportions:

$$\hat{p}_1 - \hat{p}_2 \pm 2SE(\hat{p}_1 - \hat{p}_2)$$

  - Fine for "bigger samples," can be used as a "guideline" in smaller samples

- Not quite correct in "smaller samples" but will give you a good sense of width/range of CI