BIOS 665 Homework 1
Joyce Choe
9/12/2024

1.

a) Contingency table of Treatment & Outcome

	Outcome		
Treatment	Favorable	Unfavorable	Total
Placebo	34	61	95
Test	51	39	90
Total	85	100	185

b) $\hat{p}_{placebo}$: The proportion of placebo participants with favorable response is 0.3579 (34/95) with 95% CI: (0.2562, 0.4596).

$\hat{p}_{test}$: The proportion of test participants with favorable response is 0.5667 (51/90) with 95% CI: (0.4587, 0.6746).

c) $\hat{p}_{diff}$: The estimated difference in proportions of favorable response between placebo and test participants is -0.2088 from $\hat{p}_{placebo} - \hat{p}_{test}$ with 95% CI: (-0.3602, -0.0573).

d) Test whether any association exists between treatment group and outcome using a two-sample binomial proportion large-sample chi-squared test. Conditions for this statistical test include: sufficiently large expected cell counts (> 10) , two independent groups, random samples from discrete binomial distribution, and a binary outcome (response).

$$H_0 : p_{placebo} = p_{test}$$

The null hypothesis is: there is no association between treatment group and response. The randomization chi-squared test statistic (Q_p) is 8.11 with a p-value of 0.0044. At $\alpha = 0.05$, we reject H_0 that the treatment group and response are independent.

e) Since we reject H_0, we conclude that an association likely exists between treatment group and outcome, with the test group having a higher estimated proportion of favorable outcome (response) than placebo group.

2.

a) The odds of satisfactory outcome in high dose group is 4.094 times the odds of satisfactory outcome in low dose group at Center A.
(This is the odds ratio estimate).

b) The 95% CI for the asymptotic odds ratio is (1.919, 8.733).

c) Using a chi-squared test due to sufficient expected cell counts (> 10), the odds ratio appears to be a valid estimate for measuring significant association between dose group and response. The OR 95% CI does not contain the null value (1), an indicator of no association. The chi-squared test statistic (Q_p) is 13.97 with a p-value of 0.0002. At $\alpha = 0.05$, we reject H_0: dose group and response are not associated. We conclude that there dose group and response are likely associated at Center A.

d) The odds of satisfactory outcome in high dose group is 7.00 times the odds of satisfactory outcome in low dose group at Center B with exact 95% CI: (0.62, 99.87), which contains the null value (1), indicating no association between treatment and response.

e) We choose Fisher's Exact test to test for association between treatment and response due to insufficient expected cell counts (< 10). With a two-sided p-value of 0.1534 and $\alpha = 0.05$, we fail to reject H_0: dose group and response are not associated. We conclude that dose group and response are not associated at Center B.

3.

a) The McNemar test is a one-sample test of association, where the two groups (placebo vs. new treatment) are dependent or paired from a single random sample (same patient).

$$H_0 : p_{placebo} = p_{newtreatment}$$

Since the marginal cell counts are sufficiently large (> 30), a large-sample approximation chi-squared test statistic with 1 d.f. (Q_m) will be 14.143. The p-value is 0.0002, and at $\alpha = 0.05$, we reject H_0 and conclude that the probability of clearing protein outcome is not the same between placebo and treatment groups.

b) Since the probability of clearing protein outcome is significantly different between placebo and treatment, we want to know whether placebo or new treatment has a higher probability of clearing protein. Based on the contingency table, we can calculate the proportion difference estimate from $\hat{p}_{treatment} - \hat{p}_{placebo} = (129 + 55)/250 - (129 + 22)/250 = 0.736 - 0.604 = 0.132$. A positive proportion difference means that the result is more favorable for the new treatment at clearing proteins than placebo.

c) A two-sided 95% CI for the difference in proportions (Mcnemar's test) is (0.065, 0.199).

4.

a) The sensitivity estimate for the screening test is 0.7541. The specificity estimate for the screening test is 0.7895.

b) A two-sided 90% CI for sensitivity is (0.6900, 0.8182).

c) A two-sided 95% CI for specificity is (0.7075, 0.8715).

d) The estimated proportion of non-diseased patients among those who test negative is 0.23685 / 0.40898 = 0.579 (negative predictive value).

5.

a) 89 samples per group is required for 0.9 power 2-sided $\alpha = 0.05$ using Pearson chi-square test and expected proportions (0.66, 0.42). The total sample size required is 178.

b) For 1:2 control:test ratio, 50 samples are required for control group and 100 samples are required for test group for 0.80 power 2-sided $\alpha = 0.05$ Pearson chi-square test expected proportions (0.66, 0.42).

c) For 70 samples per group (balanced allocation), the power of the study is 0.82, $\alpha = 0.05$, using Pearson chi-square test and expected proportions (0.66, 0.42).

```
* BIOS 665 hw 1;

* 1a-e);
data one;
    input treat $ outcome $ count;
    cards;
placebo fav 34
placebo un 61
test fav 51
test un 39
;

/* equivalent except one-sided p-value is reversed;
data one_;
    input treat $ outcome $ count @@;
    cards;
test fav 51 test un 39
placebo fav 34 placebo un 61
;

/*
proc print data=one;
sum count;
run;
*/

proc freq data=one order=data;
    weight count;
    table treat*outcome / expected chisq nocol riskdiff (correct) measures;
run;

* 2a-c;
data twoA;
    input dose $ response $ count;
    cards;
high yes 58
high no 17
low yes 25
low no 30
;

*chi square test;
proc freq data=twoA order=data;
    weight count;
    table dose*response / expected chisq nocol nopct measures;
run;

* 2d-e;
data twoB;
    input dose $ response $ count;
    cards;
high yes 6
high no 2
low yes 3
low no 7
;

* fisher's exact test;
proc freq data=twoB order=data;
    weight count;
    table dose*response / expected nocol nopct chisq;
    exact or;
run;

* 3a-c;
data three;
    input placebo $ treat $ count;
    cards;
clear clear 129
clear not 22
not clear 55
not not 44
;

* preview table;
proc freq data=three order=data;
```

```
    weight count;
    table placebo*treat / nopercent nocol;
run;

* Mcnemar's matched pairs association test;
ods select McnemarsTest;
proc freq data=three order=data;
    weight count;
    table placebo*treat / agree;
    *exact mcnem;
run;

* 4;
data four;
    input disease $ test $ count @@;
    cards;
present + 92 present - 30
absent + 20 absent - 75
;

* Sensitivity/specificity screening test;
proc freq data=four order=data;
    weight count;
    tables disease*test / riskdiff alpha=0.05;
run;

proc freq data=four order=data;
    weight count;
    tables disease*test / riskdiff alpha=0.10;
run;

* 5;
/*In designing a randomized clinical trial for the evaluation of a treatment for cardiovascular disease in terms of a favorab

* a)
Using a two-sided 0.05 significance level with balanced allocation to these two groups, determine the sample size that would
proc power;
    twosamplefreq test=pchi
    grouppproportions= (0.66 0.42)
    power = 0.9
    ntotal=.;
run;

* b)
With twice as many patients for the treatment as for the appropriate control, determine the sample sizes needed to provide ab
proc power;
    twosamplefreq test=pchi
    grouppproportions = (0.66 0.42)
    power = 0.8
    groupweights = (2 1)
    ntotal=.;
run;

* c)
If at the end of the study there were 70 patients enrolled in each group, what is the power of the study if the expected prop
proc power;
    twosamplefreq test=pchi
    grouppproportions = (0.66 0.42)
    power = .
    groupns = ( 70 70);
run;
```