Introduction to Biostatistics PHP 2500

Final Exam
Friday, December 14, 2007

(Closed Book)

There are 30 questions, each worth 3.33 points.

Write out your solutions and circle your final answers. Show all your work on these pages. If you need more space write on the back of the page.

You have three hours to complete this exam.

Name: Solutions

1. Zinc density was measured for thirteen newborn babies (group 1) and for thirteen babies that where were born prematurely (group 2). Some summary statistics are:

$$\overline{X}_1 = 1511.9 \qquad S_1 = 318.036$$
$$\overline{X}_2 = 1339.4 \qquad S_2 = 272.804$$
$$\overline{X}_1 - \overline{X}_2 = 172.5 \qquad S_{diff} = 171.802$$

$$\sqrt{\frac{S_1^2}{13} + \frac{S_2^2}{13}} = 116.212 \quad \text{and} \quad \sqrt{S_p^2\left(\frac{1}{13} + \frac{1}{13}\right)} = 116.212$$

$$\frac{[(S_1^2/13) + (S_2^2/13)]^2}{[(S_1^2/13)^2/(12) + (S_2^2/13)^2/(12)]} = 23.46$$

Give a 99% confidence interval for the difference in the expected values of Zinc density in the two groups under each of the following assumptions:

(a) Assume that the observations are all independent and normally distributed, and that all have the same variance.

$\overline{X}_1 - \overline{X}_2 \pm t_{\alpha/2}^{n_1+n_2-2} \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $\qquad t_{.005}^{24} = 2.797$

$172.5 \pm 2.797\,(116.212)$

172.5 ± 325.05

$-152.55,\ 497.55$

(b) Assume that the observations are all independent and normally distributed, but that the groups might have different variances.

$\overline{X}_1 - \overline{X}_2 \pm t_{\alpha/2} \sqrt{\frac{S_1^2}{13} + \frac{S_2^2}{13}}$ $\qquad$ w/ $df = \begin{cases} 13-1 = 12 & \min(n_1-1, n_2-1) \\ 23.46 \approx 24 & \text{Welch-Satterthwaite } df \end{cases}$

either 13 ok $\begin{cases} t_{.005}^{12} = 3.055 \\ t_{.005}^{24} = 2.797 \end{cases}$

$172.5 \pm 2.797(116.212)$

$-152.55,\ 497.55$

* notice that (a) + (b) give the same answer. This is b/c $n_1 = n_2$.

Problem 1 continued

(c) Assume that babies were matched on gestation time, gender, and weight. That is, the babies are paired. (The quantity s_{diff} is the sample standard deviation of the thirteen paired differences.)

$\bar{d} \pm t^{12}_{\alpha/2} \frac{s_{diff}}{\sqrt{13}}$ $\quad t^{12}_{\alpha/2} = 3.055$

$172.5 \pm 3.055 \frac{171.802}{\sqrt{13}}$

172.5 ± 145.56

26.94 , 318.07

Now suppose that the data are not normally distributed. That is, assume that these data really are independent observations from some probability distribution with unknown mean and variance, but that the distribution is not normal.

(e) What characteristic of the distribution is most important for determining whether the confidence interval in parts (a)-(c) will be approximately correct?

Symmetry (lack of skewness)

(f) What theorem is used to justify the application of a large sample confidence interval in this case?

the central Limit theorem.

why b/c even if $X \sim ?$ still $\bar{X} \sim N(E(X), \frac{Var(X)}{n})$

2. A colleague comes to you looking for statistical advice about a study she is planning. She wants to show that the probability of an allergic reaction to a new immunization shot is less than 10%. Suppose there are two valid scientific hypotheses: H_0: the probability is 10% and H_1: the probability is 20%. Her current plan is to enroll 49 participants and reject H_0 if the observed allergic proportion is greater than 15%.

→ this sentence is irrelevant.

(a) Approximate the Type I error of her plan.

$H_0: \theta = .1$
$H_1: \theta = .2$
Reject H_0 when $\hat{\theta} > .15$

So, $\alpha = P(\text{Reject } H_0 \mid H_0 \text{ true})$

$= P(\hat{\theta} > .15 \mid \theta_0 = .1)$

$= P\left(\frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} > \frac{.15 - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \mid \theta_0 = .1\right)$

$= P(Z > 1.1667) = \boxed{.1210}$

(b) Approximate the Power of her plan.

Same set-up as above

$1 - \beta = P(\text{Reject } H_0 \mid \theta_a = .2)$

$= P(\hat{\theta} > .15 \mid \theta_a = .2)$

$= P\left(Z > \frac{.15 - .2}{\sqrt{\frac{.2(.8)}{49}}}\right)$

$= P(Z > -0.875)$

$= 1 - .1922$

$= \boxed{.8078}$

Problem 2 continued

(c) If she would like to keep the Type I error at 5%, how should she modify her plan?

$$P(\hat{\theta} > ? \mid \theta = \theta_0) \leq 0.05$$

$$P\left(Z > \frac{? - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \mid \theta_0 = .1\right) = 0.05$$

So $1.645 = \frac{? - .1}{\sqrt{\frac{.1(.9)}{49}}}$

$? = \boxed{.1705}$

So using a cutoff of 0.1705 does the trick.

the question is a little too open-ended, so it was ✗ to find the larger sample size that would keep the same power when $\alpha \to 5\%$

For example w/ 98 participants

$P(\hat{\theta} > .15 \mid \theta = .1) = .05$ ←

$P\left(Z > \frac{.15 - .1}{\sqrt{\frac{.1(.9)}{98}}} \mid \theta = .1\right) = P(Z > 1.645)$

$= 0.05$ ✓

(d) After conducting the study, she finds 11.2% is the observed allergic proportion. What is the approximate p-value?

$$P(\hat{\theta} > .112 \mid \theta = \theta_0)$$

$$= P\left(Z > \frac{.112 - .1}{\sqrt{\frac{.1(.9)}{49}}} \mid \theta = .1\right)$$

$$= P(Z > 0.28 \mid \theta = .1)$$

$$= 0.3897$$

$\boxed{\text{p-value} = 0.3897}$

3. A study of ~~350~~ 84 participants was conducted to compare the responsiveness of talk therapy between men and women. The data are provided below.

	Men	Women	Total
Responder	18 a	36 b	54
Not Responder	4 c	26 d	30
Total	22	62	84

Fisher's exact = 0.069
1-sided Fisher's exact = 0.038

OR = 3.25

(a) Estimate the Odds Ratio of responding to talk therapy for men versus women and provide a 95% CI for that OR.

$\log\left(\frac{ad}{bc}\right) \pm 1.96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$ is a CI for log OR OR = 3.25

1.18 ± 1.195

$[-0.0151, 2.375] \Rightarrow [e^{-0.0151}, e^{2.375}] = [0.985, 10.75]$

(b) Based on your answer in part (a), would you reject the null hypothesis that the OR=1 at the 5% level? (choose one)

☐ Yes
☒ No b/c 1 is in the 95% CI

(c) Use a Chi-square test to test the null hypothesis that gender and response are independent at the 5% level. (Calculate the test statistic, state the critical value and p-value, and state your conclusions.)

$\chi^2 = \sum\frac{(O-E)^2}{E} = \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$

$= \frac{84[18(26)-4(36)]^2}{22(62)(54)(30)} = 3.99$ ← Test statistic

$\chi^2_{0.05} = 3.84$ ← critical value

So $3.99 > 3.84$ so $p < 0.05$ so Reject H_0.

Problem 3 continued

(d) Is the null hypothesis you tested in part (b) equivalent to the null hypothesis you tested in part (c)? (choose one)

☒ Yes $OR = 1 \Leftrightarrow P(R|M) = P(R|W)$

☐ No

(e) Are your answers in part (b) and part (c) consistent in rejecting the null hypothesis? Answer yes or no, and explain why this is the case.

NO, Because the χ^2-Test and OR Test are based on different metrics so they have different standard errors. Hence they will not always agree. Notice also, the OR-Test/CI give the same 'result' as Fisher's Exact Test.

(f) Under the null hypothesis that gender and response are independent, what is the expected number of female non-responders?

$E[\text{women non-responders}] = P(\text{woman and non-responder}) \cdot N$

$= P(\text{woman}) P(\text{non-responder}) \cdot N$ b/c of independence

$= \frac{62}{84} \cdot \frac{30}{84} \cdot 84$

$= \frac{62(30)}{84}$ → Formula is on crib sheet also

$= \boxed{22.14}$

4. A large study is planned to investigate the sensitivity of digital mammography. Previous studies suggest that the sensitivity is 85%.

(a) What sample size would be required to estimate the sensitivity to within a margin of error of 5% when using a 90% confidence interval?

$$N = \frac{4\theta(1-\theta)(Z_{\alpha/2})^2}{L^2} = \frac{4(.85)(.15)(1.645)^2}{.1^2}$$

$$= 138.00$$

$$= \boxed{138}$$

Here $L = 2 \cdot MOE = 2(.05) = .1$

If you used $L = .05$ you got 552 people.

(b) With only 100 participants, what is the margin of error for a 95% CI for the sensitivity?

$$L = \sqrt{\frac{4\theta(1-\theta)Z_{\alpha/2}^2}{N}} = \sqrt{\frac{4(.85)(.15)1.96^2}{100}} = .14$$

So $L = .14$ so $MOE = .07$ or $\boxed{7\%}$

(c) What sample size would be needed to test the null hypotheses that the sensitivity is 80%, with 80% power and two-sided Type I error of 5%?

$$N = \frac{\left[Z_{\alpha/2}\sqrt{\theta_0(1-\theta_0)} + Z_\beta\sqrt{\theta_a(1-\theta_a)}\right]^2}{(\theta_0 - \theta_a)^2}$$

$\theta_0 = .8$

$\theta_a = .85$

$$= \frac{\left[1.96\sqrt{.8(.2)} + 0.84\sqrt{.85(.15)}\right]^2}{(.05)^2}$$

$$= 469.97$$

$\boxed{N = 470 \text{ people}}$

*If you reverse θ_0 and θ_a then you got $N = 430$ people

5. The table below displays data on resting energy expenditure (REE) for *twelve* different cystic fibrosis patients. Six patients received placebo and six received the drug. The null hypothesis is that the drug does not work (i.e., there is no affect of drug on REE measurements). However the investigator has forgotten if participants were matched. Look at the output below and decide if these data should be analyzed as matched or as two separate groups.

	Resting Energy Expenditure (REE)	
Id	Placebo	Drug
1	224	213
2	80	89
3	75	33
4	541	440
5	74	-32
6	85	-29

. ttest placebo=drug

```
Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 placebo |       6    179.8333    76.05762    186.3024     -15.679    375.3457
    drug |       6         119    74.13456    181.5919   -71.56896     309.569
---------+--------------------------------------------------------------------
    diff |       6    60.83333    21.75227    53.28195    4.917351    116.7493
------------------------------------------------------------------------------
     mean(diff) = mean(placebo - drug)                             t =   2.7966
 Ho: mean(diff) = 0                         degrees of freedom =        5

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.9809         Pr(|T| > |t|) = 0.0381          Pr(T > t) = 0.0
```

. ttest placebo=drug, unpaired

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 placebo |       6    179.8333    76.05762    186.3024     -15.679    375.3457
    drug |       6         119    74.13456    181.5919   -71.56896     309.569
---------+--------------------------------------------------------------------
combined |      12    149.4167    51.45776    178.2549     36.1589    262.6744
---------+--------------------------------------------------------------------
    diff |            60.83333    106.2106               -175.8187    297.4853
------------------------------------------------------------------------------
    diff = mean(placebo) - mean(drug)                             t =   0.5728
Ho: diff = 0                               degrees of freedom =       10

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.7103         Pr(|T| > |t|) = 0.5795          Pr(T > t) = 0.2897
```

Note: the ttest output is the same as above with the 'unpaired unequal' option.

Problem 5 continued

(a) Based on the t-tests, we should treat these data as (pick one):

- ☒ matched and reject the null hypothesis at the 5% level
- ☐ matched and fail to reject the null hypothesis at the 5% level
- ☐ unmatched and reject the null hypothesis at the 5% level
- ☐ unmatched and fail to reject the null hypothesis at the 5% level

(b) Justify your choice of the matched or unmatched analysis and cite statistics from the output to support your choice.

the matched SE of the diff is much smaller than the pooled SE of diff.

Matched $SE(\bar{x}_1-\bar{x}_2)=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}-cov}=21.75$

unmatched $SE(\bar{x}_1-\bar{x}_2)=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}=106.21$

the reduction is due to the correlation/cov from matching.

If the matching was not helpful the estimated SE would be very close.

After examining the t-tests you decide to examine several non-parametric tests.

```
. signrank placebo=drug

Wilcoxon signed-rank test

        sign |      obs    sum ranks    expected
-------------+---------------------------------
    positive |        5           20        10.5
    negative |        1            1        10.5
        zero |        0            0           0
-------------+---------------------------------
         all |        6           21          21

Ho: placebo = drug
             z =   1.992
    Prob > |z| =   0.0464
```

Problem 5 continued

```
. ranksum ree, by(group)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       group |      obs     rank sum    expected
-------------+---------------------------------
 Placebo (0) |        6           44          39
    Drug (1) |        6           34          39
-------------+---------------------------------
    combined |       12           78          78

Ho: ree(group==0) = ree(group==1)
              z =   0.801
    Prob > |z| =   0.4233
```

(c) Based on the t-tests and non-parametric tests, we should treat these data as (pick one):

Here the extra tests don't add or take away much.

- [x] matched and reject the null hypothesis at the 5% level
- [ ] matched and fail to reject the null hypothesis at the 5% level
- [ ] unmatched and reject the null hypothesis at the 5% level
- [ ] unmatched and fail to reject the null hypothesis at the 5% level

(d) If a sign test were preformed on these data, what would be the p-value?

the data are 5 minuses and one plus.

So p-value = $P(k \geq 5 \mid \theta = 1/2) = \binom{6}{6}(1/2)^6 + \binom{6}{5}(1/2)^6 = 0.1094$

(one-sided) $k \sim Bin(6, 1/2)$

need to add other side $P(k \geq 5 \mid \theta = 1/2) + P(k \leq 1 \mid \theta = 1/2) = 2(.1094) = \boxed{0.2188}$ → p-value

(e) Explain why the sign test might sometimes disagree with the sign rank test.

The sign test ignores the magnitude of the ranks, while the Sign Rank uses the Rank information. So they can be quite different as they are here.

6. In a study of two treatments, A and B, patients are to be paired on age, gender, and a measure of disease severity. One member of each pair will be selected (randomly) to receive treatment A, and the other will receive treatment B. After treatment we will determine for each pair whether the patient on treatment A did better or the patient on B did better. We will observe n patient pairs. We will reject the hypothesis that the two treatments are equivalent if either treatment gives better results in all n pairs.

(a) Is this a one-sided or a two-sided test?

two-sided

(b) If our study consists of n = 5 patient pairs, what is the size (Type I error probability) of our test?

$k \sim Bin(n, \theta)$ $n=5$ $\theta = 1/2$

$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ true})$

$= P(k=0 \mid \theta=1/2) + P(k=5 \mid \theta=1/2)$

$= \binom{5}{0}(1/2)^5 + \binom{5}{5}(1/2)^5 = 2(.0313) = \boxed{0.0625}$

(c) If our study consists of n = 5 patient pairs, what is the power of our test if in fact treatment A gives better results with probability 0.6?

$\text{Power} = P(\text{Reject } H_0 \mid H_a)$

$= P(k=0 \mid \theta=.6) + P(k=5 \mid \theta=.6)$

$= \binom{5}{0}(.6)^0(.4)^5 + \binom{5}{5}(.6)^5(.4)^0$

$= .4^5 + .6^5 = \boxed{0.0880}$

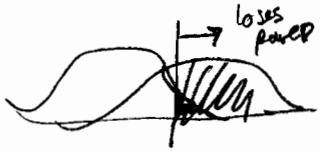
Problem 6 continued

(d) The calculations described here are the exact version of what test? (choose one)

- ☐ Sign Rank test
- ☒ McNemar's test (Exact is Sign test).
- ☐ Fisher's exact test
- ☐ Chi-square test

7. A study is planned to detect manufacturing differences in the weight of Tylenol pills. With a two-sided Type I error of 10%, the study can detect a 2.5 microgram differential from the target weight while accounting for a variance of 36 micrograms with 84% power. $\sigma^2 = 36$ $(\mu_a - \mu_0)^2 = 6.25$ (2.5^2) $Z_{.05} = 1.645$ ($Z_{\alpha/2}$) $Z_\beta = 0.99$ ($Z_{.16}$)

(a) If the Type I error is set at 5%, the power of the study will?



- ☐ increase
- ☒ decrease b/c $Z_{.025}$ ↑ from $Z_{.05}$ Here Z_β must come down
- ☐ remain the same
- ☐ not enough information is provided to answer

$$n = \frac{(Z_{\alpha/2} + Z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2}$$

(b) If the variance increases to 51.84 and the differential increases to 3 micrograms, the power of the study will?

- ☐ increase
- ☐ decrease
- ☒ remain the same — why? b/c $\frac{\sigma^2}{(\mu_a - \mu_0)^2} = \frac{36}{6.25} = 5.76$ and now $\frac{51.84}{9} = 5.76$ } Ratio stays constant, so formula does not change
- ☐ not enough information is provided to answer

note: one quick way to get the answer was to solve for n in the beginning ($n = 50$) and then plug in for each problem.

(c) If the test is changed to a one-sided test, the power of the study will?

- ☐ increase
- ☐ decrease
- ☐ remain the same
- ☒ not enough information is provided to answer

Partial credit b/c if in same direction the tail area effect goes up in that direction.

Need to know if alternative and Reject are in the same direction OR opposite directions