Math 316 Spring 2022 **Exam 3** **March 29, 2022**

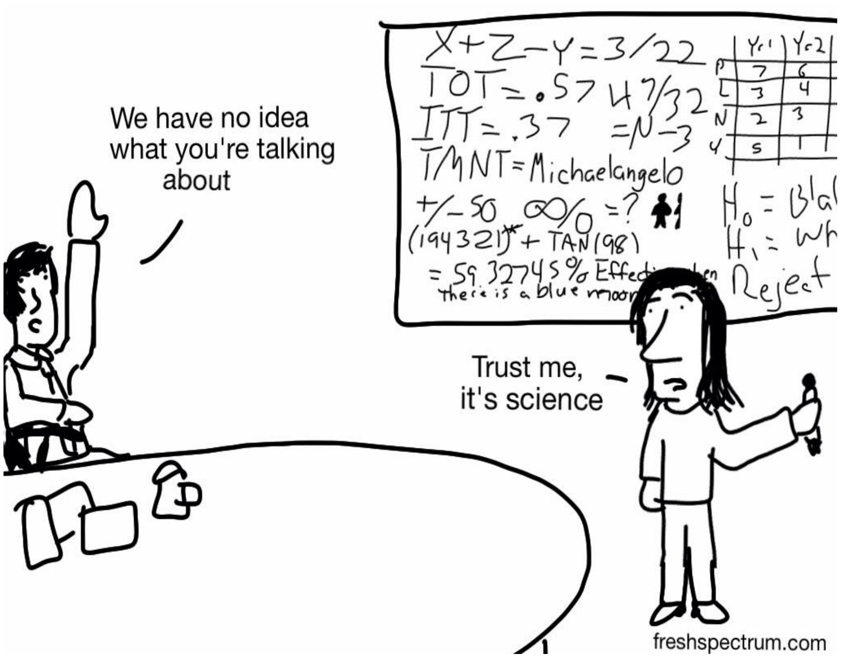
Name: Solutions

Problem	1	2	3	4 / 5	Total
Possible	48	22	18	12	100
Received					

DO NOT OPEN YOUR EXAM UNTIL TOLD TO DO SO.

You may use a page of handwritten notes (both sides) and a calculator, and the supplied *Statistical Tables* and *Summary of Formulas and Tests*.

FOR FULL CREDIT, SHOW ALL WORK RELATED TO FINDING EACH SOLUTION.



48 points 1. Exercise vs. diet weight loss. **Data for this problem are on the final page of the exam.**

/6 First, assume the five persons are paired and perform a (non-directional) t test to determine whether we would reject $H_0: \mu_{Diet} = \mu_{Exercise}$. Use $\alpha = 0.10$.

$t_s = \frac{2.4}{\frac{2.97}{\sqrt{5}}} = \frac{2.4}{1.33} = 1.81$ Table 4, df $= 5-1=4$

$.05 <$ one tail $< .10$

$\Rightarrow .10 < P < .20$

Do not reject H_0.

/3 If the test were directional (and we suspected before doing this study that weight loss due to Exercise would be greater than weight loss due to Diet), how would that change our results?

$.05 < P .10$ so $P < \alpha$, reject H_0.

/3 If the test were directional (and we suspected before doing this study that weight loss due to Exercise would be less than weight loss due to Diet), how would that change our results?

Wrong direction!
So do not reject H_0.

/6 Next, still assuming the five persons are paired, find a (two-sided) 90% confidence interval for the difference $\mu_{Diet} - \mu_{Exercise}$ in the two weight loss strategies.

$2.4 \pm 2.132\left(\frac{2.97}{\sqrt{5}}\right) = 2.4 \pm 2.83$
$= (-.43, 5.23)$

Same conclusion as above: don't reject H_0, since 0 is in CI.

/5 For the confidence interval that you just found, approximately how large would the sample size have needed to be so that the margin of error $t_{\alpha/2} SE_{\bar{D}}$ would have been ≤ 0.1?

We'll need a much larger n to go from 2.83 to .1, so use $t = 1.8$ or 1.7 or even 1.645.

$t \cdot \frac{2.97}{\sqrt{n}} \leq 0.1 \Rightarrow n \geq \left[\frac{(1.8)(2.97)}{0.1}\right]^2$
$= 2858$

(Problem 1 continued)

/12 Now treat the data as unpaired but approximately normally distributed. Determine whether we should reject $H_0: \mu_{Diet} = \mu_{Exercise}$ by:

- Performing a (non-directional) t test. Use $\alpha = 0.10$.
- Finding a 90% confidence interval for the difference of the two weight loss strategies.

You can use $df = 7$.

$$t_s = \frac{7 - 4.6}{\sqrt{\frac{3.16^2}{5} + \frac{0.89^2}{5}}} = \frac{2.4}{1.4682} = 1.63$$

Table 4, $df = 7$

$.05 <$ one tail $< .10$

$\Rightarrow .10 < P < .20$

Cannot reject H_0

CI $2.4 \pm 1.895(1.4682)$

$-0.382 < \mu_1 - \mu_2 < 5.182$

/6 Assuming once again that the data are paired, perform a Sign Test on the given data to determine whether or not we should reject $H_0: \mu_{Diet} = \mu_{Exercise}$.

$N_+ = 4$ $\quad B_s = 4$ $\quad$ Table 7, $n_d = 5$

$N_- = 1$

$P > .20$ $(\alpha = .10)$

so do not reject H_0.

/7 Still assuming the data are paired, perform a Wilcoxon Signed-Rank Test on the given data to determine whether we should reject $H_0: \mu_{Diet} = \mu_{Exercise}$.

Signed ranks

Diff	Rank	SR
5	4.5	4.5
3	3	3
5	4.5	4.5
−2	2	−2
1	1	1

$W_+ = 4.5 + 3 + 4.5 + 1 = 13$

$W_- = 2$

$W_s = 13$

Table 8, $n = 5$

$P = .188 > \alpha$

so do not reject H_0

22 points **2.** We are interested in what proportion of younger (4 to 6 years old) children and what proportion of older (7 to 9 years old) children are iron deficient. Samples of 98 younger children and 98 older children were taken: 6 of the younger children and 15 of the older children were found to be iron deficient.

		Younger children	Older children
Iron deficient?	Yes	6	15
	No	92	83
	Total	98	98

/7 Use these data to construct a 95% confidence interval for the proportion of all younger children who have iron deficiency.

$$\tilde{p} = \frac{6+2}{98+4} = \frac{8}{102} \qquad p = \frac{8}{102} \pm 1.96\sqrt{\frac{\frac{8}{102}\left(1-\frac{8}{102}\right)}{98+4}} = .078 \pm .052$$

$$= (.026, .131)$$

/9 Use the above data to construct a 95% confidence interval for the difference in proportions of younger and older children that are iron deficient.

$$\tilde{p}_1 = \frac{6+1}{98+2} = .07 \qquad p_1 - p_2 = .07 - .16 \pm 1.96\sqrt{\frac{(.07)(1-.07)}{98+2} + \frac{(.16)(1-.16)}{98+2}}$$

$$\tilde{p}_2 = \frac{15+1}{98+2} = .16 \qquad = -.09 \pm .088$$

$$= (-.178, -.002)$$

/6 Suppose that we are interested in how iron deficiency varies amongst younger and oldest siblings in the same family, and suppose that the above data came from 98 families, each with a younger and older sibling, so the two samples are actually paired. With this in mind, we can display the above data in the following more useful format.

		Younger sibling iron deficient? Yes	Younger sibling iron deficient? No
Older sibling iron deficient?	Yes	6	**9**
	No	**3**	81

We are not interested in the case when both siblings are iron deficient (the 6 in the table) or the case when both are not iron deficient (81). We are only interested in the two cases in which one sibling is iron deficient and the other is not (3 and 9). Test the hypothesis H_0: the probability of iron deficiency is the same for older siblings as it is for younger siblings (so naturally we have the alternative H_A: the probabilities are different).

$$\chi_s^2 = \frac{(3-9)^2}{3+9} = 3 \qquad \text{Table 9, } df = 1$$

$$.05 < p < .10$$

18 points **3.** In a classic study of peptic ulcer, blood types were determined for 100 ulcer patients. The accompanying table shows data for these patients as well as for an independently chosen group of 200 healthy controls from the same city.

Blood type	Ulcer patients	Controls	Total
O	30 (40)	90 (80)	120
A or B	50 (50)	100 (100)	150
AB	20 (10)	10 (20)	30
Total	100	200	300

/9 Find the missing expected values ( ) and use a χ^2 test to determine whether there is a difference in percentage of each blood type in ulcer patients vs. non-ulcer patients. Use $\alpha = .01$.

$$X_s^2 = \frac{(30-40)^2}{40} + \cdots + \frac{(10-20)^2}{20}$$

$$= \frac{100}{40} + \frac{100}{80} + 0 + 0 + \frac{100}{10} + \frac{100}{20}$$

$$= 18.75$$

Table 9

$df = (3-1)(2-1)$

$= 2$

$P < .0001$

If both sample sizes were quadrupled (4 × larger), but the percentage of each blood type for ulcer and for non-ulcer patients remained the same:

/3 What would the test statistic X_s^2 be? 4 X larger, so 75

/2 Would the P-value be smaller or larger? Smaller

/4 According to this sample, how much more likely is someone with Type O blood to have an ulcer than someone with Type AB blood? That is, what is $\frac{\hat{P}r\{\, Ulcer \mid Type\ O\ Blood \,\}}{\hat{P}r\{\, Ulcer \mid Type\ AB\ Blood \,\}}$?

$$\frac{30/120}{20/30} = .375$$

7 points **4.** A cross between white and yellow summer squash gave progeny of the following colors:

Color	White	Yellow	Green	Total
Number of progeny	145	35	20	200
Expected	150	37.5	12.5	200

To determine whether or not these data consistent with the 12:3:1 ratio predicted by a certain genetic model, use a chi-square test at $\alpha = .01$.

$$\chi_s^2 = \frac{(145-150)^2}{150} + \frac{(35-37.5)^2}{37.5} + \frac{(20-12.5)^2}{12.5} = 4.83$$

$df = 3 - 1 = 2$

Table 9 $\quad .05 < P < .10$

$P > \alpha$ so do not reject H_0

5 points **5.** Suppose that we computed a 90% confidence interval for a population proportion of $(0.2, 0.4)$. What sample size was used in computing this?

$\tilde{p}$ must be the center of the interval 0.3

and it must be that $z_{\alpha/2}\sqrt{\dfrac{\tilde{p}(1-\tilde{p})}{n+z_{\alpha/2}}} = 0.1$

where $z_{\alpha/2} = 1.645$, $\tilde{p} = .3$:

$$1.645\sqrt{\frac{(.3)(1-.3)}{n+1.645^2}} = .1$$

$$\Rightarrow n = \frac{(1.645)^2(.3)(.7)}{(.1)^2} - 1.645^2$$

$\approx 54.$

For Problem 1 we will consider the following data showing weight loss for five persons due to Exercise and for five other persons due to Diet.

Person	**Exercise**	**Diet**	**Difference**
1	11	6	5
2	7	4	3
3	9	4	5
4	3	5	-2
5	5	4	1
Mean	**7.0**	**4.6**	**2.4**
SD	**3.16**	**0.89**	**2.97**

We will analyze these data in several ways. We want to test the hypothesis that exercise and diet result in different amounts of weight loss. That is, where $\mu_{Exercise}$ is the average weight loss due to Exercise and μ_{Diet} is the average weight loss due to Diet, we are testing the hypotheses

$$H_0: \mu_{Diet} = \mu_{Exercise}$$

$$H_A: \mu_{Diet} \neq \mu_{Exercise}$$

If you like, you can tear this page out of your exam to more easily refer to it when working Problem 1.