Pharmacy 2012, Winter 2020, Stats Assignment 2 Solution

(Total 50 points)

1. (17 points) In a study comparing men that survive heart attacks to those that do not survive, emergency room records were obtained for a sample of 65 men who ultimately died and 65 men who survived. The arterial blood pressure at the time of admittance to the emergency room was determined. The mean arterial blood pressure for the sample of surviving patients is 91.8 and the sample standard deviation is 12.71. For the patients who did not survive, the mean and sample standard deviation are 87.31 and 13.60 respectively.

   (a) Is there a significant difference between the mean arterial blood pressures of the two groups?

(2) i. State the hypotheses.

- $H_o : \mu_1 - \mu_2 = 0$
- $H_a : \mu_1 - \mu_2 \neq 0$
  OR
- $H_o : \mu_1 = \mu_2$
- $H_a : \mu_1 \neq \mu_2$

(2) ii. Calculate the pooled estimate of the standard deviation.

- $s_p^2 = \frac{64(12.71)^2+64(13.60)^2}{65+65-2} \approx 173.25$
- $s_p = \sqrt{173.25} \approx 13.16$

(Subtract 1 point for each error.)

(2) iii. Calculate the test statistic.

- $t = \frac{91.8-87.31}{13.16\sqrt{\frac{1}{65}+\frac{1}{65}}} = \frac{4.49}{2.3089} \approx 1.95$
  (Subtract 1 point for each error. If the difference was taken in the other direction, the observed test statistic would be ≈ -1.95, which is OK, as the alternative is 2 sided.)

(1) iv. What are the degrees of freedom of the test statistic?

- 65 + 65 - 2 = 128

(2) v. Bound the P value as accurately as possible using the t-table on the class web site.

- $P = 2P(T > |t_{obs}|)$ using 128 degrees of freedom.
- The degrees of freedom do not appear in the table. Then **we can use the next smallest value**, 120, **which leads to a conservative test**.
- Using 120 degrees of freedom, as t_{obs} is between 1.658 and 1.980, $P(T > |t_{obs}|)$ is between .025 and .05, so $.05 < p - value < .10$.
- One can also use an ∞ degrees of freedom, as the degrees of freedom are large. Then t_{obs} is between 1.645 and 1.960, $P(T > |t_{obs}|)$ is between .025 and .05, so $.05 < p - value < .10$.

(1) vi. Are the results statistically significant at the $\alpha = .01$ level of significance?

- NO. (Because $P > .01$ we conclude that the results are not statistically significant at the $\alpha = .01$ level.)

(5) (b) Construct a 99% confidence interval for the mean difference in arterial blood pressure.

- $\bar{X} - \bar{Y} \pm t_{\alpha/2,n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
- Using 120 degrees of freedom, because 130 - 2 = 128 is not in the tables.
  $t_{.005,120} = 2.617$
- $91.8 - 87.31 \pm 2.617(13.16)\sqrt{\frac{1}{65} + \frac{1}{65}}$
- $4.49 \pm 34.44(.1754)$
- 4.49 ± 6.041

- $(-1.55, 10.53)$
(Because the directionality of the difference was not specified in the statement of the problem, the interval $(-10.53, 1.55)$, is also OK.)
(For this assignment it's also OK to use an infinite number of degrees of freedom, in which case use $t_{.005,\infty} = 2.576$, leading to interval 4.49 ± 5.95 - $(-1.46, 10.44)$, or the interval $(-10.44, 1.46)$)
(Subtract 1 point for each error.)

(2) (c) Explain how this confidence interval can be used to test the null hypothesis, and confirm that it gives the same answer as does the test procedure.

- In general, a $100(1-\alpha)\%$ confidence interval contains those values which are not significant at level α when used in H_0 for a two-sided test.
- The 99% interval contains zero so we DO NOT reject the null hypothesis that the mean difference is zero, using $\alpha = .01$.
- This agrees with the conculusion of the formal hypothesis test at level $\alpha = .01$.

2. (18 points) Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under conditions of stress. A researcher conducted a study to investigate whether a program of regular exercise might lower the resting (unstressed) concentration of HBE in the blood. He measured blood HBE levels in January and again in May in 10 participants in a physical fitness program. The results are shown in the table below

	HBE Level (pg/ml)		
Participant	January	May	Difference
1	42	22	20
2	47	29	18
3	37	9	28
4	9	9	0
5	33	26	7
6	70	36	34
7	54	38	16
8	27	32	-5
9	41	33	8
10	18	14	4
Mean	37.8	24.8	13.0
SD	17.6	10.9	12.4

(2) (a) State the null and alternative hypotheses.

- $H_0 : \mu_D = 0$, $H_a : \mu_D > 0$

(2) (b) Evaluate the test statistic.

- $T = \frac{13.0-0}{12.4/\sqrt{10}} \approx 3.32$

(1) (c) What are the degrees of freedom?

- df = 10 - 1 = 9

(2) (d) Bound the P value as accurately as possible using the t-table on the class web site.

- The t-table gives $P(T > 3.250) = .005$ and $P(T > 3.690) = .0025$ so $.0025 < P < .005$.
If a different t-table or a computer was used to calculate the p-value, leading to a different answer, deduct 1 point. Students, please use the class t-tables, as this is what will be used on exams.

(2) (e) What assumption are you making about the population from which the data are sampled?

- You are assuming that the differences are a sample from a normal population. Equivalently, you could state that both the January and May measurements were samples from normal populations, in which case the normality of the differences follows.

(2) (f) Using a type I error probability of $\alpha = .05$, conclude whether or not the results are statistically significant.

- We would conclude that the results are statistically significant at the $\alpha = .05$ level because $p - value < .05$.

(2) (g) Describe what would constitute a type II error in this problem.

- Concluding that the mean difference was 0, when in fact, the mean difference was greater than 0. (Anything reasonable along this line.)

(5) (h) Construct a 95% confidence interval for the mean difference in HBE level.

- Using 9 df, $t_{.025,9} = 2.262$ from the t table.
- $13 \pm t_{.025,9} 12.4/\sqrt{10}$
- $13 \pm 2.262(12.4)/\sqrt{10}$
- 13 ± 8.87
- $(4.13, 21.87)$
  (Answers to 1 or more decimal places of accuracy are OK. Deduct 1 point for each error.)

3. (Total 15 points) Sixty-five pregnant women at a high risk of preganancy-induced hypertension participated in a randomized controlled trial comparing 100mg of aspirin daily and a placebo during the third trimester of pregnancy. The observed rates of hypertension are shown in the following table

	Aspirin treated	Placebo treated	Total
Hypertension	4	11	15
No Hypertension	30	20	50
Total	34	31	65

(a) Assess whether aspirin is effective in reducing the risk of hypertension.

(2) i. State the hypotheses.

- If p_1 is the probability of hypertension in the asprin group and p_2 is the probability in the placebo group, the hypotheses are $H_0 : p_1 = p_2$ and $H_A : p_1 < p_2$.

(2) ii. Calculate the observed value of the test statistic.

- The form of the test statistic is

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p(1 - \hat{p}_p)(\frac{1}{n_1} + \frac{1}{n_2})}}$$

where $\hat{p}_1 = \frac{4}{34} \approx .118$, $\hat{p}_2 = \frac{11}{31} \approx .355$, $\hat{p}_p = \frac{4+11}{34+31} \approx .231$, $n_1 = 34$, and $n_2 = 31$, giving $Z \approx -2.27$ (or $Z \approx 2.27$, depending on the definition of p_1 and p_2).
(Subtract 1 point for each error.)

(2) iii. Calculate the p-value using the normal table.

- $P = P(Z < Z_{obs}) \approx P(Z < -2.27) \approx .0116$ from the standard normal table. (Note: if p_1 and p_2 were defined as the probability of NO hypertension, the direction of the inequality in the p-value calculation would be reversed, but at the same time, the sign of Z_{obs} is reversed, so the p-value will remain the same.)

(1) iv. Would you reject the null hypothesis when testing at level $\alpha = .05$?

- Yes.

(6) (b) Calculate a 95% confidence interval for the difference in hypertension rates in the aspirin and placebo groups.

- The form of the confidence interval is

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$
$$= \frac{4}{34} - \frac{11}{31} \pm 1.96\sqrt{\frac{\frac{4}{34}(1-\frac{4}{34})}{34} + \frac{\frac{11}{31}(1-\frac{11}{31})}{31}}$$
$$\approx -.237 \pm 1.96(.102) \approx -.237 \pm .200$$

or $(-.437, -.037)$. If the roles of p_1 and p_2 are reversed, the confidence interval will be $(.037, .437)$. **(Substract 1 point for each error.)**

(1) (c) Based on the 95% confidence interval, conclude whether or not the results are significant at the $\alpha = .05$ level.

- Because the 95% confidence interval does not contain 0, we conclude that the results are statistically significant at the $\alpha = .05$ level.