Dawson College: **Probability and Statistics**: 201-BZS-05

Last Name: SOLUTIONS

First Name: _____

Student ID: _____

Assignment 2

Please answer all of the following questions in the space provided. Write clearly and make sure to use correct notation. Please state any variables, distributions and assumptions that you are using/making.

Question 1. *(5 marks)* A sample of 7 measurements, randomly selected from a normally distributed population, resulted in a sample mean $\bar{x} = 7.8$ and sample standard deviation $s = 1.03$. Using $\alpha = 0.05$, is this sufficient evidence to conclude that he mean of the population is less than 7.9? (Use the p-value approach).

NORMALLY DISTRIBUTED POPULATION - CAN USE Z-TABLE

$df = 7 - 1 = 6$, $\alpha = 0.05$

$H_0: \mu \geq 7.9$

$H_a: \mu < 7.9$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7.8 - 7.9}{1.03/\sqrt{7}} = -0.2567$$

$$P(t < -0.2567) = \frac{0.424 + 0.387}{2} = 0.4055 > \alpha = 0.05$$

FAIL TO REJECT H_0

WITH 5% SIGNIFICANCE THE MEAN OF THE POPULATION IS NOT LESS THAN 7.9.

Question 2. In a random sample of 336 BZS students, 297 said that it was the best class they had ever taken in their entire lives.

(a) *(3 marks)* Construct a 95% C.I.E. (confidence interval estimate) for the proportion of students having taken BZS that consider it the best class ever.

$p' = \frac{297}{336} = 0.884$ $\quad np' = 336(0.884) = 297 > 5$

$n(1-p') = (0.116)(336) = 39 > 5$ } WE CAN USE Z-TABLE

$\alpha = 0.05 \Rightarrow z(\alpha/2) = 1.96$

$E = z(\alpha/2)\sqrt{\frac{p'(1-p')}{n}} = 1.96\sqrt{\frac{0.884(0.116)}{336}} = 0.034 = 3.4\%$

$\therefore$ 95% C.I.E. $85\% < p < 91.8\%$

(b) *(3 marks)* The stats from part (a) sound a bit sketchy to you and so you want to do another experiment to get a smaller confidence interval. What can you do to decrease the error for the confidence interval estimate to 2%? You should use any information gained from the first experiment.

WE CAN USE $p = 0.884$

$n = (z(\alpha/2))^2 \frac{p(1-p)}{E^2} = (1.96)^2 \frac{(0.884)(0.116)}{(0.02)^2}$

$= 984.83$

USE SAMPLE SIZE 985

Question 3. *(5 marks)*A study was designed to compare the attitudes of two groups of nursing students toward computers. Group 1 had previously taken a statistical methods course that involved significant computer interaction through the use of statistical packages. Group 2 had taken a statistical methods course that did not use computers. The students' attitudes were measured by administering the Computer Anxiety Index (CAIN). The results were as follows:

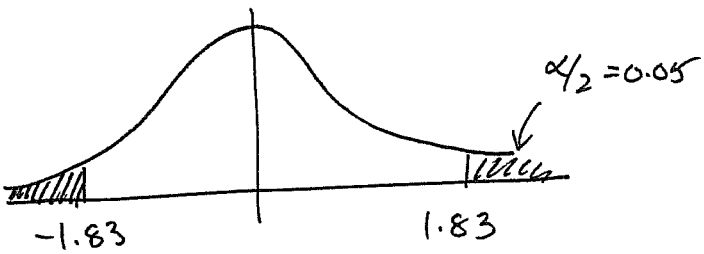
Group 1 (with computers): $n = 10$ $\bar{x} = 60.3$ $s = 7.5$
Group 2 (without computers): $n = 15$ $\bar{x} = 67.2$ $s = 2.1$

Does the data show that the mean score for those with computer experience is different from the mean score for those without computer experience? Use $\alpha = 0.10$. (Determine if the samples are dependent or independent. Use the appropriate method from class and use the classical approach. Assume normal distributions).

$H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 \neq 0$

TWO TAIL TEST



$$df = \frac{\left(\frac{7.5^2}{10} + \frac{2.1^2}{15}\right)^2}{\frac{7.5^4}{10^2(9)} + \frac{2.1^2}{15^2(14)}} = 9.961 \rightarrow 9 \quad \text{so} \quad t(0.05) = 1.83$$

TEST STATISTIC

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(60.3 - 67.2) - 0}{\sqrt{\frac{7.5^2}{10} + \frac{2.1^2}{15}}} = -2.83$$

IN REJECTION REGION, REJECT H_0.

∴ AT 5% SIGNIFICANCE, THE MEANS ARE DIFFERENT.

Question 4. *(5 marks)* The July 28, 1990, issue of Science News reported that smoking boosts death risk for diabetics. The death risk is increased more for women than men. Suppose as a follow-up study we investigated the smoking rates for male and female diabetics and obtained the following data.

Gender	n	Number Who Smoke
Male	500	215
Female	500	170

At the 0.05 level of significance, test the research hypothesis that the smoking rate (proportion of smokers) is higher for males than for females. (Use the p-value approach)

$$\bar{p}_p = \frac{215+170}{500+500} = 0.385$$

$n_1 \bar{p}_p = n_2 \bar{p}_p = 500(0.385) = 192.5 > 5$

$n_1(1-\bar{p}_p) = n_2(1-\bar{p}_p) = 500(0.615) = 307.5 > 5$

} WE CAN USE Z-TABLE

$H_0: p_m - p_f \leq 0$

$H_a: p_m - p_f > 0$

$$\bar{p}_m = \frac{215}{500} = 0.43$$

$$\bar{p}_f = \frac{170}{500} = 0.34$$

TEST STATISTIC

$$z = \frac{\bar{p}_m - \bar{p}_f}{\sqrt{\bar{p}_p(1-\bar{p}_p)\left(\frac{1}{n_m}+\frac{1}{n_f}\right)}} = \frac{0.43-0.34}{\sqrt{0.385(1-0.385)\left(\frac{1}{500}+\frac{1}{500}\right)}}$$

$= 2.92$

RIGHT TAIL TEST

$P(z > 2.95) = 0.0018 < 0.05 = \alpha$ $\therefore$ REJECT H_0.

AT 5% SIGNIFICANCE SMOKING RATE IS HIGHER FOR MALES THAN FOR FEMALES.

Question 5. *(5 marks)* Does test failure reduce academic aspirations and thereby contribute to the decision to drop out of school? These were the concerns of a study titled "Standards and School Dropouts: A National Study of Tests Required for High School Graduation." The table reports the responses of 283 students selected from schools with low graduation rates to the question "Do tests required for graduation discourage some students from staying in school?"

	① *Urban*	② *Suburban*	③ *Rural*	*Total*
Yes	57	27	47	131
No	23	16	12	51
Total	80	43	59	182

Does there appear to be a relationship at the 0.05 level of significance between a student's response and the school's location?(Use the classical approach).
Hint: To see if there is a relationship between a student's response and the school's location, your null hypothesis can be that the response proportions from all the schools are the same as the response proportion that you observed from the urban school. If you reject this hypothesis then there is a relationship between a student's response and the school's location.

FROM URBAN SCHOOL, PROPORTION OF STUDENTS THAT RESPONDED YES:

$$p = \frac{57}{80} = 0.7125$$

$H_0: p_1 = p_2 = p_3 = 0.7125$

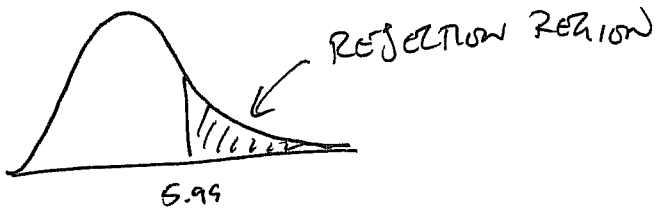
H_a: ONE OF $p_i \neq 0.7125$ WITH $i \in \{1,2,3\}$

$e_1 = 0.7125(80) = 57$, $e_2 = 0.7125(43) = 30.64$, $e_3 = 0.7125(59) = 42.04$

$e_i \geq 5$ FOR $i = 1,2,3$ SO WE CAN USE χ^2-DISTRIBUTION

$df = 3 - 1 = 2$

$\chi^2(0.05) = 5.99$



$$\chi^2 = \frac{(57-57)^2}{57} + \frac{(27-30.64)^2}{30.64} + \frac{(47-42.04)^2}{42.04} = 1.018$$

NOT IN REJECTION REGION, FAIL TO REJECT H_0.

AT 5% SIGNIFICANCE, NO RELATIONSHIP BETWEEN RESPONSE AND LOCATION.