NAME: SOLUTIONS **MARK: /50**

TEST 4
201-BZS-05
Probability & Statistics
Instructor: Emilie Richer
Date: May 1st 2009

Question 1 (8 marks)
Of the 500 managers and professionals polled, 27.8% of them work late 5 days a week on average.

A. Find a 99% confidence interval for the true proportion of managers and professionals who work late 5 days a week.
B. What is the minimum sample size required in order to have a margin of error of 2% with 99% confidence?

ONE proportion
Sample proportion $p' = 0.278$
$n = 500$

$np' = 139 > 5$
$n(1-p') = 361 > 5$ we can use z-table

A. $E = z_{\alpha/2}\sqrt{\frac{p'(1-p')}{n}} = 2.56\sqrt{\frac{0.278(0.722)}{500}} = 0.05129$

99% C.I. is $\boxed{0.227 < p < 0.329}$

B. $E = 0.02$ $z_{\alpha/2} = 2.575$

$0.02 = 2.575\sqrt{\frac{0.278(0.722)}{n}}$

$0.000060326 = \frac{0.2007}{n}$

$n = 3326.92$ We would need $\boxed{3327}$ people

Question 2 (7 marks)

A student group maintains that each day the average student must travel for at least 25 minutes one way to reach college. The college admissions office obtained a random sample of 31 one-way travel times from students. The sample had a mean of 19.4 minutes and a sample standard deviation of 9.6 minutes. Does the admissions office have sufficient evidence to reject the students' claim? Use $\alpha = 0.1$.

$H_0: \mu \geq 25$

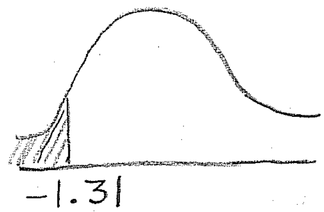
$H_a: \mu < 25$

$n = 31$ $n \geq 30$, we can use t-table

$\bar{x} = 19.4$

$s = 9.6$ $t_{\alpha/2} = t_{0.1} = -1.310$

(with 30 degrees of freedom)



test statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{19.4 - 25}{9.6/\sqrt{31}} = -3.25$$

$\boxed{\text{Reject } H_0.}$

There is enough evidence to reject students' claim.

Question 3 (10 marks)

The length of major league baseball games are approximately normally distributed and last on average 2 hours and 50.1 minutes with a standard deviation of 21.0 minutes. It has been claimed that New York Yankees baseball games last, on average, longer than the games of the other major league teams. To test the truth of this statement a sample of eight Yankee games was randomly identified and the length of each game obtained (in minutes):

199 196 202 213 187 169 169 188

A. At the 0.05 significance level, does this data show sufficient evidence to conclude that the mean time of Yankee baseball games is longer than that of other major league baseball teams?

B. Find β (the probability of a type II error) if the mean length of a Yankee game is actually 180 minutes.

A. MAJOR LEAGUE GAMES $\mu_0 = 170.1$

$\sigma = 21$ MINUTES

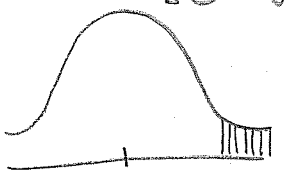
$\bar{x} = 1523/8 = 190.375$

$H_0 : \mu \le 170$

$H_a : \mu > 170.1$

Z-test because σ is known

games are normally distributed so we can use z-table



TEST STAT : $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$z_{0.05} = 1.645$ $= \dfrac{190.375 - 170.1}{21/\sqrt{8}}$

$\boxed{\text{Reject } H_0}$ $= 2.73$

YANKEES' GAMES ARE LONGER

B. Acceptance region $z \le 1.96$ which corresponds

TO $1.645 = \dfrac{\bar{x} - 170}{21/\sqrt{8}}$ $\bar{x} = 182.21$

$P(\bar{x} \le 182.21) = P\left(z \le \dfrac{182.21 - 180}{21/\sqrt{8}}\right) = P(z \le 0.298)$

IF $\mu = 180$ $= 0.5 + 0.1179$

Question 4 (5 marks)

Of 120 Marxist-Leninists polled, 110 supported a new government bill legalizing cannabis while 35 of 200 Conservatives supported the bill. Construct a 95% confidence interval estimate for the difference in levels of support between the two political parties.

TWO POPULATIONS

P_1

$n_1 = 120$

$p_1' = \frac{110}{120} = 0.92$

P_2

$n_2 = 200$

$p_2' = \frac{35}{200} = 0.175$

We can use z-table because

$n_1 p_1' = 110 \geq 5$

$n_1(1-p_1') = 10 \geq 5$

$n_2 p_2' = 35 \geq 5$

$n_2(1-p_2') = 165 \geq 5$

$$E = z_{0.025}\sqrt{\frac{p_1'(1-p_1')}{n_1} + \frac{p_2'(1-p_2')}{n_2}}$$

$$= 1.96\sqrt{\frac{(0.92)(0.08)}{120} + \frac{(0.175)(0.825)}{200}}$$

$$= 0.0716$$

$$(p_1'-p_2') - E < P_1 - P_2 < (p_1'-p_2') + E$$

$$\boxed{0.6734 < P_1 - P_2 < 0.8166}$$

Question 5 (8 marks)

Two different statistics classes were tested with a common exam. A sample of students from each class was chosen and their exam scores collected yielding the data below. Evaluate whether Class A did significantly better than Class B at 0.01 level of significance.

	Class A	Class B
Sample size	12	15
Average test score	72.3	69.5
Sample standard deviation	6.1	5.6

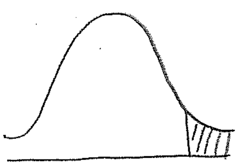
2 SAMPLE t-test

$H_0: \mu_A - \mu_B \le 0$

$H_a: \mu_A - \mu_B > 0$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}} = \frac{\left(\frac{6.1^2}{12} + \frac{5.6^2}{15}\right)^2}{\frac{6.1^4}{12^2 \cdot 11} + \frac{5.6^4}{15^2 \cdot 14}} = \frac{26.95}{1.186}$$

$= 22.72$ ROUND down to 22



$t_{0.01} = 2.508$

test stat

$$t = \frac{(72.3 - 69.5) - 0}{\sqrt{\frac{6.1^2}{12} + \frac{5.6^2}{15}}} = 1.23$$

DO NOT REJECT H_0.

CLASS A did NOT do significantly BETTER than CLASS B.

* NOTE Here, THere WAS INSUFFICIENT dATA to JUSTIFY USe OF t-tABle. We would

Question 6 (7 marks)

Boxes of Smarties are packaged at a factory at two different times of the day. There have recently been complaints about the inconsistencies of packaging speed at the two different times. Ten employees are selected at random and their speed of Smarties packaging is measured, yielding the results below. Is there significant evidence to support the claim that the Smarties packaging speed is different at the two times of day? Conduct a hypothesis test at 0.05 level of significance.

Employee	1	2	3	4	5	6	7	8	9	10
Speed Before Lunch (Smarties/min)	53	72	61	45	39	62	58	55	66	67
Speed After Lunch Smarties/min)	50	45	63	42	40	56	55	58	65	65

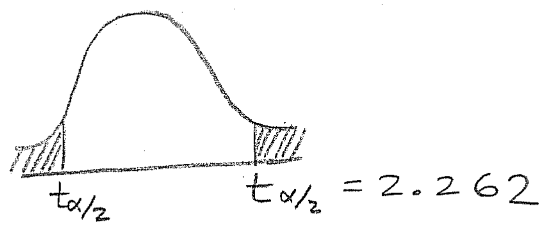
difference (d) 3 27 -2 3 -1 6 3 -3 1 2

$\bar{d} = 3.9$ $S_d = \sqrt{\dfrac{811 - \frac{(39)^2}{10}}{9}} = 8.556$

σ UNKNOWN, POPULATION NORMAL, we use t-table

$H_0 : \mu_d = 0$

$H_a : \mu_d \neq 0$



df = 9

test stat:

$$t = \frac{3.9 - 0}{8.556/\sqrt{10}}$$

$= 1.44$

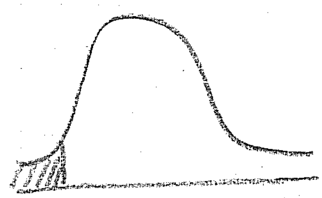
FAIL TO REJECT H_0

There is NOT ENOUGH evidence TO support the claim.

Question 7 (5 marks)
A random sample of 200 Dawson students has a current average of 62% in their English class (with standard deviation of σ = 7%). Can we conclude that the average Dawson student English mark is a passing one (60%)? Conduct a test at 5% level of significance.

$\sigma = 7$
$\bar{x} = 62$
$n = 200$

$H_0 : \mu \geq 60$
$H_a : \mu < 60$



$z_{0.05} = -1.645$

$n = 200 \geq 30$
so we can use z-table

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{62 - 60}{7/\sqrt{200}}$$
$$= 4.04$$

DO NOT REJECT H_0
THE Average English MARK is 60% or over at 5% significance.