Elvan Ceyhan December 24, 2010

Math 202: Statistics for Social Sciences

Fall 2010 Midterm 2

Calculator allowed, duration 90 minutes.

Instructions: There are six problems in this exam. Please inspect the exam and make sure you have all 6 pages (5 pages of questions and 1 cover page). Do all your work on these pages. If you use the back of a page, make sure to indicate it.

Remember: ***You must show your work to get proper credit.***

Academic Honesty Code: Koç University Academic Honesty Code stipulates that "copying from others or providing answers or information, written or oral, to others is cheating." By taking this exam, you are assuming full responsibility for observing the Academic Honesty Code.

NAME and SURNAME: KEY

LECTURE TIME: 9:30_____ 12:30_____

Prob 1:	/15
Prob 2:	/15
Prob 3:	/20
Prob 4:	/20
Prob 5:	/15
Prob 6:	/15
Total:	/100

Formulas:

Confidence interval (CI) : $\bar{X} \mp z \times SE$ or $\bar{X} \mp t \times SE$ or $\hat{p} \mp z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$(\hat{p}_1 - \hat{p}_2) \mp z \times SE_{difference}$ or $(\bar{X}_1 - \bar{X}_2) \mp z \times SE_{difference}$ or $(\bar{X}_1 - \bar{X}_2) \mp t \times SE_{difference}$

Standard Error (SE) : $SE_{difference} = \sqrt{(SE_1)^2 + (SE_2)^2}$, $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ for percentages,

and $SE = \frac{SD}{\sqrt{n}}$ for averages.

Standard Deviation (SD) :

sum all (entry-average)^2 values,
then divide by n-1, and then take the square root
Here, entry: each observation in the data set, n: sample size

Problem 1: (15 points) Ministry of Labor wants to know how a new government policy is perceived (algılanmak) among the workers and whether the perception differs with union (sendika) membership. For this purpose, a group of blue-collar (mavi yakalı) workers are surveyed about their views on this policy. Below is the breakdown of their proportions in regard to their views and union membership status.

	Views on government policy			
Membership status	Supportive	Indifferent	Opposed	Row totals
Union	0.07	0.09	0.28	0.44
Non-union	0.18	0.17	0.21	0.56
Column totals	0.25	0.26	0.49	1.00

If a worker is selected randomly, what is the chance that s/he is
(a) a union member and opposed to the government policy?
(b) not a member of any union?
(c) opposed to the government policy given that s/he is a union member?
(d) a union member given that s/he is opposed?
(e) Are the two events "being a union member" and "being opposed" independent? Explain why or why not.

(a) 0.28 or 28%

(b) 0.56 or 56%

(c) $\frac{0.28}{0.44} \cong 0.64$ or 64%

(d) $\frac{0.28}{0.49} \cong 0.57$ or 57%

(e) chance of being a union member is 0.44
" " " opposed is 0.49
" " " a union member and opposed is 0.28

Checking Independence:

$(0.44) \times (0.49) \stackrel{?}{=} 0.28$

$0.2156 \neq 0.28$

Hence the two events are not independent

Problem 2: (15 points)

(a) (5 points) A fair die is rolled 10 times. What is the chance that all the rolls show 5 spots or less?

(b) (5 points) A fair coin is tossed 10 times. Find the chance that there will be exactly 3 heads among the first 5 tosses, and exactly 4 heads among the last 5 tosses.

(c) (5 points) A fair coin is tossed 10 times. Find the chance that there will be exactly 7 heads.

(a) chance of 5 or less on all rolls?

$n = 10$
$p = 5/6$
$k = 10$

$$\binom{10}{10}\left(\frac{5}{6}\right)^{10}\left(\frac{1}{6}\right)^{0} = \left(\frac{5}{6}\right)^{10} \cong 0.1615$$

or 16.15%

(b) 10 tosses — first 5 tosses: 3H — last 5 tosses: 4H

$$\binom{5}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{2} \times \binom{5}{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{1}$$

$$= 10 \times \left(\frac{1}{2}\right)^{5} \times 5 \times \left(\frac{1}{2}\right)^{5} = 50 \times \left(\frac{1}{2}\right)^{10} = 0.0488 \cong 5\%$$

(c) $$\binom{10}{7}\left(\frac{1}{2}\right)^{7}\left(\frac{1}{2}\right)^{3} = \frac{10 \times 9 \times 8 \times 7!}{3! \times 7!}\left(\frac{1}{2}\right)^{10} = 120 \times \left(\frac{1}{2}\right)^{10}$$

$$\cong 0.12 \text{ or } 12\%$$

Problem 3 (20 points)
There are 200 subjects in a small clinical trial on vitamin C. Half the subjects are assigned at random to treatment (2 g of vitamin C daily) and half to control (2 g of placebo) group. Over the period of the experiment, the treatment group had on average 1.3 colds and the SD was 3.1. The controls did a little worse: they averaged 2.6 colds and the SD was 2.9.

(a) (12 points) Is the difference in averages statistically significant (that is, is vitamin C effective to avoid cold)? (Use α=5%).

(b) (8 points) Also, construct a 95% confidence interval for the difference of the average number of colds for vitamin C and control groups.

(a)

Vit C	Placebo
$n_1 = 100$	$n_2 = 100$
$\bar{X}_1 = 1.3$	$\bar{X}_2 = 2.6$
$SD_1 = 3.1$	$SD_2 = 2.9$

$SE_1 = \frac{3.1}{\sqrt{100}} = 0.31$

$SE_2 = \frac{2.9}{\sqrt{100}} = 0.29$

$SE_{diff} = \sqrt{0.31^2 + 0.29^2}$

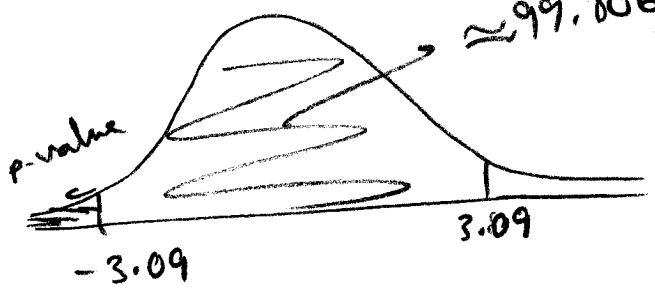
$= \sqrt{0.1802} = 0.42$

$H_0: \mu_V = \mu_P$ or $\mu_V - \mu_P = 0$

$H_a: \mu_V < \mu_P$ or $\mu_V - \mu_P < 0$

$n_1, n_2 > 30$, both are large, so use z-test (i.e., normal approx.)

$z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{SE_{diff}} = \frac{1.3 - 2.6}{0.42} = -3.09$



p-value $= \frac{100 - 99.806}{2}$

$= 0.097\%$

since p-value $< 5\%$, we reject H_0, and conclude that vit. C is effective in avoiding cold.

(b) 95% CI for $\mu_V - \mu_P$ is

$(\bar{X}_1 - \bar{X}_2) \mp z_{CI} \times SE_{diff} = (1.3 - 2.6) \mp 2 \times 0.42$

$= -1.3 \mp 0.84 = (-2.14, -0.46)$

Problem 4: (20 points)
A newspaper article says that on the average, college freshmen in the entire country spend 7.5 hours a week watching TV. One administrator does not believe that this applies at his college, Plower University, which nearly has 3000 freshmen. He takes a random sample of 20 freshmen and interviews them. On average, they report 9.6 hours a week watching TV and the SD is 9 hours.

(a) (12 points) Is the difference between 9.6 and 7.5 real (that is, do the students at Plower University spend more time watching TV than the nationwide average)? Test the hypotheses at 5% level.

(b) (8 points) Construct a 90% confidence interval for average number of hours spent watching TV for the students at Plower University.

(a) $n = 20$ $\quad H_0: \mu = 7.5$ $\quad SE = \frac{9}{\sqrt{20}} = \frac{9}{4.47} = 2.01$

$\bar{x} = 9.6$ $\quad H_a: \mu > 7.5$

$SD = 9$

$n = 20 < 30$, so it is a small sample

$$t = \frac{\bar{x} - \mu_0}{SE} = \frac{9.6 - 7.5}{2.01} = \frac{2.1}{2.01} = 1.04$$

$df = 19$, $t_{crit} = 1.73$

Since 1.04 is not larger than 1.73, we do not reject H_0.

(b) 90% CI for μ is

$$\bar{x} \mp t_{CI} \times SE = 9.6 \mp 1.73 \times 2.01$$

$$= 9.6 \mp 3.52 = (6.08, 13.12)$$

Problem 5: (15 points)
A box contains 20 red and 20 white balls (indistinguishable except for color). Mr. Foreknows claims he has psychic abilities so that he can pick the red balls from the box (when drawing with his eyes covered). To test his claim, Dr. Doubtful makes him draw 400 balls with replacement, and observes that 228 are red. Is that too many reds? Or chance variation? Test at 5% level.

20R & 20W. $n = 400$, 228 are R.

$H_0: p = 1/2$ or $p = 50\%$

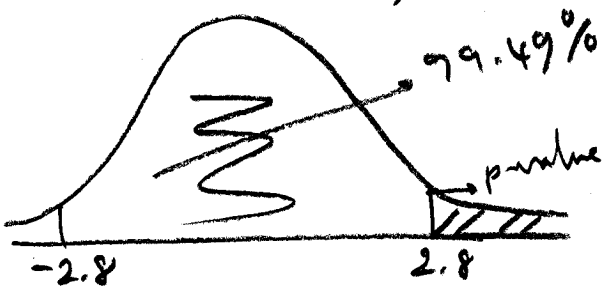
$H_a: p > 1/2$ or $p > 50\%$

sample percentage $= \frac{228}{400} = 0.57$ or 57%

$SE = \sqrt{\frac{0.5(1-0.5)}{400}} = \frac{0.5}{20} = 0.025$ or 2.5%

Since n is large, use z-test.

$$z = \frac{57\% - 50\%}{2.5\%} = \frac{7}{2.5} \cong 2.8$$



p-value $= \frac{100 - 99.49}{2} = 0.255\%$

$p < 5\%$, so reject H_0, so he can pick red balls better than chance.

Problem 6: (15 points)
In the town of Ramsville, there are 25000 eligible voters in total. A random sample of 1600 voters is taken to estimate the percentage of supporters of the candidate for Mayor, Mr. Eager. It turns out that 917 voters support Mr. Eager. Find a 95% confidence interval for the percentage of supporters of Mr. Eager among all 25000 eligible voters. Based on this confidence interval, would you recommend Mr. Eager to run for the Mayor's office?

$n = 1600$, 917 supports, sample percentage $= \frac{917}{1600} = 0.573$ or 57.3%

$SE = \sqrt{\frac{0.573(1-0.573)}{1600}} = \frac{\sqrt{0.24}}{40} = \frac{0.49}{40} \cong 0.0124$ or 1.24%

n is large, so 95% CI is

$57.3 \mp 2 \times 1.24 = 57.3 \mp 2.48$

$= (54.82, 59.78)$

Yes, I would recommend him run for Mayor's office, because he gets more than 50% of the votes with 95% confidence.