Mine Çağlar

January 2, 2018

**Math 202: Statistics for Social Sciences**

**Fall 2017 FINAL EXAM**

**Calculator OK, 2 hrs. 15 minutes**

Instructions: There are six parts to this exam I-VI. Please inspect the exam and make sure you have all six pages of questions. Do all your work on these pages. If you use the back of a page, make sure to indicate that.

Remember: ***You must show your work to get proper credit. In hypothesis testing questions, show all steps of the test and state your conclusion in plain English.***

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:** SOLUTION

| | |
|---|---|
| Part I: | /15 |
| Part II: | /25 |
| Part III: | /20 |
| Part IV: | /15 |
| Part V: | /15 |
| Part VI: | /20 |
| Total: | /110 |

Formulas:

1) Confidence interval (CI) : $\overline{X} \mp z\, SE$ with $SE = \dfrac{SD}{\sqrt{n}}$

or $\overline{X} \mp t\, SE$ or $\hat{p} \mp z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ or

$(\hat{p}_1 - \hat{p}_2) \mp z\, SE_{2-sample}$ or

$(\overline{X}_1 - \overline{X}_2) \mp z\, SE_{2-sample}$

or $(\overline{X}_1 - \overline{X}_2) \mp t\, SE_{2-sample}$

where $SE_{2-sample} = \sqrt{(SE_1)^2 + (SE_2)^2}$,

and $SE = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ for percentages.

2) Binomial formula:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0,1,2,\dots,n$$

3) Chi-squared: $$\chi^2 = sum\,of \frac{(observed\ frequency - expected\ frequency)^2}{expected\ frequency}$$

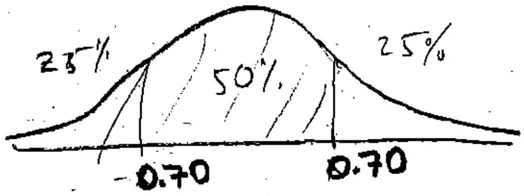
4) ANOVA: Between group variance: $$s_B^2 = \frac{\sum n_i(\overline{x}_i - \overline{x}_{GM})^2}{k-1}$$

Within group variance: $$s_W^2 = \frac{\sum (n_i - 1)s_i^2}{\sum (n_i - 1)}$$ and $$F = \frac{s_B^2}{s_W^2}$$

Here $s_i$ is $SD$ of $i^{th}$ sample and $s_i^2$ is the variance, namely $SD^2$ (the square of $SD$).

**Part I. (15 points)** The histogram for the scores of a survey (on a scale 1 to 9), can be approximated well with a bell-shaped curve with mean of 5 and a standard deviation of 2.

1. (5 points) What is the 75th percentile of this distribution?



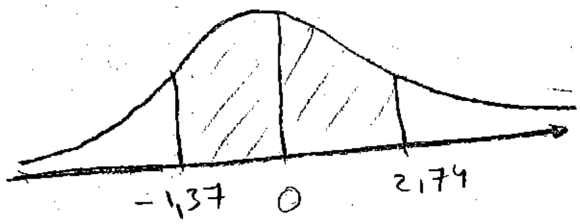

| Area | z |
|---|---|
| 50% | 0,70 |

$$\frac{X-5}{2} = 0{,}70$$

$$X = 5 + 1{,}4 = 6{,}4$$

2. (6 points) What is the chance that the sample mean of 30 students will be between 4.5 and 6?

$$SE = \frac{SD}{\sqrt{N}} = \frac{2}{\sqrt{30}} = 0{,}365$$




$$X_1 = \frac{4{,}5-5}{0{,}365} = \frac{-0{,}5}{0{,}365} = -1{,}37$$

$$X_2 = \frac{6-5}{0{,}365} = \frac{1}{0{,}365} = 2{,}74$$

$$\text{Area} = \frac{82{,}30 + 99{,}40}{2} = 90{,}85\ \%$$

There is 90.85 chance that the sample mean will be between 4.5 and 6.

3. (4 points) Find the 40th percentile of the following data set:

4 3 8 7 6 4 9 7 6 5 6 5 9

3 4 4 5 5 6 6 6 7 7 8 9 9

$$\frac{13 \cdot 40}{100} = 5{,}2 \longrightarrow 6\text{-th number} \longrightarrow \boxed{6}$$

**Part II.** **(25 points)** A psychologist investigates certain learning ability by experimenting with rats in the lab. It is believed that 1 out of 5 rats have this ability in the population of rats.

1. (7 points) Find the probability that in a sample of 13 rats at most 2 have the learning ability.

By the Binomial Formula: $p = \frac{1}{5} = 0.2$

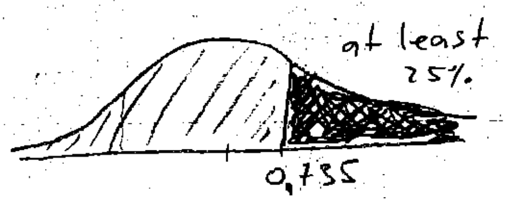
$$P(\text{At most } 2) = \binom{13}{0}(0.8)^{13} + \binom{13}{1}(0.2)(0.8)^{12} + \binom{13}{2}(0.2)^2(0.8)^{11}$$

$$= 0.5016$$

2. (8 points) Find the probability that in a ~~sample of 35~~ rats at least 25% have the learning ability, by approximating with a normal distribution.

$N = 35$

$$SE = \sqrt{\frac{(0.2)(0.8)}{35}} = \sqrt{\frac{0.16}{35}} = 0.068$$




$$\frac{0.25 - 0.20}{0.068} = 0.735$$

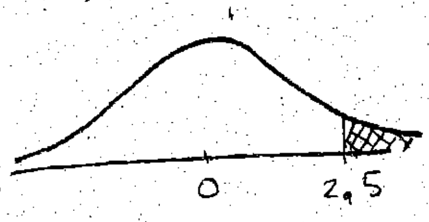
Area: $\frac{100 - 54.67}{2} = \underline{22.66\%}$

3. (10 points) The psychologist wants to see if rewarding helps to increase the learning ability. Her experiment on 35 rats demonstrates that 13 of them show this ability if they are rewarded on a regular basis previously. Does rewarding help to gain ~~the learning~~ ability?

$H_0: p = 0.2$ $\qquad \hat{p} = 0.37$

$H_a: p > 0.2$ $\qquad SE = \sqrt{\frac{(0.2)(0.8)}{35}} = 0.068$

$$z = \frac{0.37 - 0.2}{0.068} = 2.5$$




| Area | z |
| --- | --- |
| 98.76 | 2.5 |

$$P = \frac{100 - 98.76}{2} = 0.62\% < 1\%$$

we reject Ho!

Rewarding helps mice gain the learning ability.

**Part III. (20 points)** Researchers tested the placebo effect as follows. 24 volunteers were randomly divided into two groups. Each volunteer was put inside a magnetic resonance imaging (MRI) machine. In the first group, electric shocks were applied to their arms and the blood oxygen level-dependent (BOLD) signal (a measure related to neural activity in the brain) was recorded during the pain. The treatment in the second group was identical to that in the first group, except that, prior to applying the electric shocks, the researchers put a cream on the volunteer's arms. The volunteers were informed that the cream would block the pain experience, when in fact, it was just a regular skin lotion (that is, a placebo). If the placebo is effective in reducing the pain experience, the BOLD measurements should be higher, on average, in the first group than the second group.

The average BOLD signal measurements were obtained as 0.64 and 0.43, respectively in the first and second groups. The standard deviations were reported, again respectively, as 0.27 and 0.19

1. (14 points) Is the placebo significantly effective in reducing the pain experience?

$N_1 = N_2 = 12$

$SE_1 = \frac{SD_1}{\sqrt{N}} = \frac{0.27}{\sqrt{12}} = 0.078$

$df = 12 + 12 - 2 = 22$

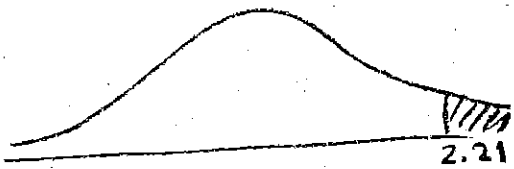
$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

$SE_2 = \frac{SD_2}{\sqrt{N}} = \frac{0.19}{\sqrt{12}} = 0.055$

$SE = \sqrt{SE_1^2 + SE_2^2} = 0.095$

$t = \frac{0.64 - 0.43}{0.095} = 2.21$




the placebo is significantly effective in reducing the pain

$2.07 < 2.11 < 2.51$

$\rightarrow$ $2.5\% > P > 1\%$

$\boxed{P < 5\%}$ - Reject $H_0$.

2. (6 points) Construct a 99% confidence interval for the difference in the pain experience between the first group and the second group.

$t = 2.82$

$(\bar{x}_1 - \bar{x}_2) \pm t \cdot SE$

$(0.64 - 0.43) \pm (2.82) \cdot (0.095)$

$0.21 \pm 0.2679$

$(-0.058, 0.4779)$

**Part IV. (15 points)** A history professor in a university polled a random sample of 57 colleagues about the number of conferences that the professors attended in the past two years and the number of papers submitted by those professors to refereed journals during the same period. In order to test whether attending professional meetings would result in publishing significantly more papers, he categorized the data in the following frequency table.

| *Number of papers submitted* | *Number of conferences attended* | | | |
|---|---|---|---|---|
| | 0-1 | 2-4 | At least 5 | |
| Less than or equal to 1 | 8 | 9 | 3 | 20 |
| Greater than or equal to 2 | 7 | 24 | 6 | 37 |
| | 15 | 33 | 9 | 57 |

Perform a hypothesis test of independence.

Ho: Number of meetings attended and numb. of papers published are independent.

Ha: They are dependent.

Expected Values

| | | |
|---|---|---|
| $\frac{20\cdot15}{57}=5.26$ | $\frac{20\cdot33}{57}=11.57$ | $\frac{20\cdot9}{57}=3.15$ |
| $\frac{37\cdot15}{57}=9.73$ | $\frac{37\cdot33}{57}=21.42$ | $\frac{37\cdot9}{57}=5.84$ |

$$X^2 = \sum \frac{(obs. - exp)^2}{exp} =$$

$$= \frac{(8-5.26)^2}{5.26} + \frac{(9-11.57)^2}{11.57} + \frac{(3-3.15)^2}{3.15} + \frac{(7-9.73)^2}{9.73} + \frac{(21.42-24)^2}{21.42} + \frac{(6-5.84)^2}{5.84}$$

$$= 1.427 + 0.571 + 0.0071 + 0.766 + 0.311 + 0.0044 =$$

$$= 3.0865$$

d.f $= (3-1)(2-1) = 2$

$10\% < P < 30\%$

Don't Reject Ho.

Number of papers published and number of meetings attended are independent.

**Part V.** **(15 points)** A market research company has investigated the customer preference for precious stones. The following table displays the frequency distribution obtained from 58 random customers shopping for jewellery in Kapalıçarşı.

| | Ruby | Sapphire | Emerald | Diamond |
|---|---|---|---|---|
| Frequency | 7 | 15 | 10 | 26 |

1. (6 points) An experienced jeweler estimates that diamond is preferred by 40% of all customers, and the other precious stones are demanded equally likely by the remaining 60%. Find the frequencies expected by the jeweler in a random sample of 58.

| | Ruby | Sapphire | Emerald | Diamond |
|---|---|---|---|---|
| Exp. Freq. | 11.6 | 11.6 | 11.6 | 23.2 |

$(40\%)(58) = 23.2$

$(60\%)58 = 34.8$

$34.8/3 \cong 11.6$

2. (9 points) Do the data support the distribution estimated by the jeweler, at $\alpha = 1\%$?

$$\chi^2 = \frac{(7-11.6)^2}{11.6} + \frac{(15-11.6)^2}{11.6} + \frac{(10-11.6)^2}{11.6} + \frac{(26-23.2)^2}{23.2}$$

$$= 3.38$$

d.f. $= 4-1 = 3$

$\alpha = 1\%$

$\chi^2_{0.01,3} = 11.34$

Since $3.38 < 11.34$ do not reject $H_0$.

The data support the distribution estimated by the jeweler.

**Part VI. (20 points)** The length of a random sample of songs from three different kinds of CD's is listed below, in minutes.

| | Jazz | Classical | Pop |
|---|---|---|---|
| | 7 | 3 | 6 |
| | 7 | 5 | 3 |
| | 6 | 4 | 4 |
| | 3 | 3 | 5 |
| | 4 | 6 | 2 |
| | 2 | | |
| Mean | 4.8 | 4.2 | ? 4 |
| Variance ($s^2$) | 4.6 | 1.7 | ? |

1. (5 points) Find the missing mean and variance.

$$\text{mean} = \frac{6+3+4+5+2}{5} = 4$$

$$s^2 = \frac{(6-4)^2 + \cdots + (2-4)^2}{5-1} = 2.5$$

2. (15 points) Can one conclude that there is a difference in mean lengths for the three types of music?

$H_0: \mu_1 = \mu_2 = \mu_3$

$H_a$: at least one mean is different

$k = 3$ $\quad$ d.f.N $= k-1 = 2$ $\quad$ $F_{0.05} = 3.81$

$N = 16$ $\quad$ d.f.D $= N-k = 13$

$$\bar{X}_{GM} = \frac{(4.8)(6)+(4.2)(5)+(4)(5)}{16} \simeq 4.36$$

$$S_B^2 = \frac{6(4.8-4.36)^2 + 5(4.2-4.36)^2 + 5(4-4.36)^2}{3-1}$$

$$= 0.97$$

$$S_W^2 = \frac{(6-1)(4.6)+(5-1)(1.7)+(5-1)(2.5)}{16-3} = 3.06$$

$$F = \frac{S_B^2}{S_W^2} = \frac{0.97}{3.06} = 0.32$$

Since $0.32 < 3.81$, do not reject $H_0$.

There is no significant difference between the mean length of different kinds of songs.