Spring 2011 Midterm #1

Closed book & notes; only a single-sided and handwritten A4 formula sheet and a calculator allowed; 90 minutes. No questions accepted!

Instructions: There are seven pages (one cover and six pages with questions) in this exam. Please inspect the exam and make sure you have all 6 pages. You may only use your calculator and your formula sheet. Do all your work on these pages. If you use the back of a page, make sure to indicate that. **You may not exchange any kind of material with another student.**

Remember: ***You must show all 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 **SIGNATURE:** ____________________

INSTRUCTOR : ____________________ **LECTURE TIME :** ____________________

1	/20
2	/20
3	/20
4	/20
5	/20
Total:	/100

Show all your work to be eligible for partial credit. Round your **final** answers to **two decimal** places.

Question 1

The following table shows the number of days it took for a number of employees to complete a certain job (For example, each of the first 10 people completed the job in 2 days, each of the next 17 people completed the job in 4 days, and so on) :

Number of Days	Frequency
2	10
4	17
5	18
7	12
9	40
10	3

(a) Calculate the mean, median, Q_1 (i.e., the first quartile), Q_3 (i.e., the third quartile), mode, IQR (interquartile range), variance and standard deviation.

- $\bar{X} = \dfrac{\sum x_i f_i}{\sum f_i} = \dfrac{2(10)+4(17)+\dots+10(3)}{10+17+\dots+3}$

$= \dfrac{652}{100} = 6.52$

- Here $n=100$ is even, so median is the average of 50^{th} & 51^{th} values: Median $= \frac{7+7}{2} = 7$
- For Q_1, $np = 100(0.25) = 25$ is an integer, so Q_1 is average of 25^{th} & 26^{th} values $\Rightarrow Q_1 = \frac{4+4}{2} = 4$
- For Q_3, $np = 100(0.75) = 75$, so we need 75^{th} & 76^{th} values $\rightarrow Q_3 = \frac{9+9}{2} = 9$
- mode $= 9$ (occurs 40 times)
- $IQR = Q_3 - Q_1 = 9-4 = 5$
- $\sum x_i^2 = 2^2(10) + 4^2(17) + \dots + 10^2(3) = 4890$

so $Var = \dfrac{\sum x_i^2 - n\bar{x}^2}{n-1} = \dfrac{4890 - 100(6.52)^2}{99} = 6.45$

- $SD = \sqrt{6.45} = 2.54$

(b) Is the data distribution bell-shaped? Explain your answer.

No, the data is not distributed in bell-shaped form, Because average is 6.52, but the most frequent value is a 9. For bell-shaped distributions, mean & mode should be similar.

Question 2:

A regression equation for Y = grade point average (GPA) and X = number of classes skipped in a typical week was found based on self-reported data for a sample of n = 1673 students at Sarıyer University. The resulting equation is:

GPA = 3.25 – 0.06 × (# of Skipped classes per week)

(a) Give an interpretation of the estimated slope.

slope = –0.06. We expect the GPA to decrease by 0.06 for each additional skipped class per week.

(b) Give an interpretation of the estimated intercept.

intercept = 3.25. We expect the GPA to be 3.25 if the student skips no classes (per week).

(c) Predict GPA for someone who typically skips 3 classes a week. Show your work.

For $X = 3$, $GPA = 3.25 - 0.06(3)$
$= 3.07$

(d) If average GPA was 3.18, then what is the average number of skipped classes per week?

$\bar{Y} = 3.18$, $\bar{X} = ?$ $\quad 3.18 = 3.25 - 0.06(\bar{X})$

$\Rightarrow 0.06\bar{X} = 3.25 - 3.18 = 0.07$

$\Rightarrow \bar{X} = \frac{0.07}{0.06} = 1.17$

(e) Three students at this university have (X,Y) pairs as (1,3.10), (2,3.20), and (2,3.30), find the sum of (vertical) error squares.

X	Y	estimated Y	error	error²
1	3.10	3.19	–0.09	0.0081
2	3.20	3.13	0.07	0.0049
2	3.30	3.13	0.17	0.0289
				+ $\boxed{\text{sum} = 0.0419}$

For X = 1, estimated Y is $GPA = 3.25 - 0.06(1) = 3.19$

" " 2, " "" $GPA = 3.25 - 0.06(2) = 3.13$

(f) Suppose that for this data, $SS_{YY} = 0.0044 \times SS_{XX}$. Then calculate the correlation coefficient, r, and interpret it in the context of the problem.

Since $\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$ and $r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$, we have

$SS_{xy} = \hat{\beta}_1 SS_{xx} = r\sqrt{SS_{xx} SS_{yy}} = r\sqrt{(SS_{xx})^2\, 0.0044}$

$\Rightarrow -0.06\, SS_{xx} = r\, SS_{xx}(0.0632) \Rightarrow r = \frac{-0.06}{0.0632} = -0.90$

So, there is a strong negative linear relationship between # of skipped classes per week and GPA.

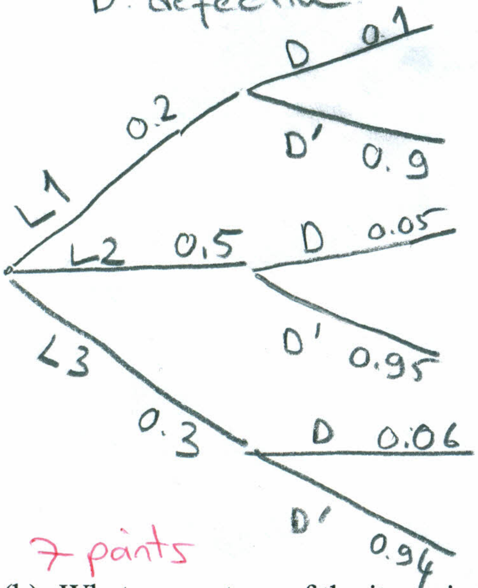
Question 3:

Three production lines contribute to the total pool of a company's product. Line 1 provides 20% to the pool and 10% of its products are defective (hasarlı, defolu); Line 2 provides 50% to the pool and 5% of its products are defective; Line 3 contributes 30% to the pool and 6% of its products are defective.

7 points

(a) Draw a probability tree, which shows all possible outcomes for a product (i.e., which line does the product come from and whether if the product is defective or not).

D: defective



$P(L1 \cap D) = P(L1)\,P(D \mid L1) = 0.02$

$P(L1 \cap D') = 0.18$

$P(L2 \cap D) = 0.025$

$P(L2 \cap D') = 0.475$

$P(L3 \cap D) = 0.018$

$P(L3 \cap D') = 0.282$

$+$ ———

1.00

7 points

(b) What percentage of the items in the pool is defective?

$$P(D) = P(L1)\,P(D \mid L1) + P(L2)\,P(D \mid L2) + P(L3)\,P(D \mid L3)$$
$$= 0.02 + 0.025 + 0.018$$
$$= 0.063$$

6 points

(c) Suppose an item is randomly selected from the pool and found to be defective. What is the probability that it came from Line 1?

$$P(L1 \mid D) = \frac{P(L1 \cap D)}{P(D)} = \frac{0.02}{0.063} = 0.317$$

Question 4:
The mean height of adult males in Andorra is 173 cm with a standard deviation of 7 cm. The z-scores of the National Basketball Team of Andorra are

$$\{2.18, 2.31, 1.94, 3.23, 4.20, 4.12, 3.88\}$$

(a) How tall is the shortest player on the team?

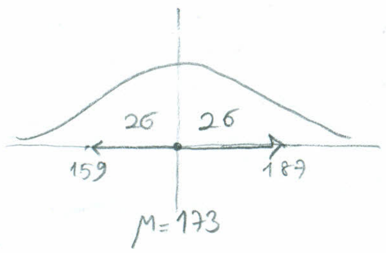
$$z = \frac{x - \mu}{\sigma} \qquad z = 1.94 = \frac{x - 173}{7} \Rightarrow x = 186.58 \text{ cm}$$

(b) What is the average height of the National Team?

$$\bar{z} = \frac{\sum z_i}{n} = \frac{(2.18 + 2.31 + \dots + 3.88)}{7} = 3.123$$

$$\bar{z} = 3.123 = \frac{\bar{x} - \mu}{\sigma} \Rightarrow \bar{x} = 194.86 \text{ cm}$$

(c) What proportion of the adult male population in Andorra is between 1.59m and 1.87m if you know that height is distributed in a symmetric bell-shaped form? What is the proportion if nothing is known about the distribution of height?



$(159, 187) \equiv \mu \pm 2\sigma$

By Emprical Rule we can say that 95% of the adult male population in Andorra is between 1.59 m and 1.87 m

Nothing is known about the distribution $\Rightarrow$ By Chebyshev's Rule we can say that $\underbrace{\left(1 - \frac{1}{2^2}\right)}\%$

at least 75% of the adult male population in Andorra is between 1.59 m and 1.87 m

(d) What is the mean of the z-scores for the whole population? What is the standard deviation?

For each value, x, in the population, we find $z = \frac{x - \mu}{\sigma}$

So, average of all z's is $\frac{\mu - \mu}{\sigma} = 0$ and

standard deviation of all z's is $\frac{\sigma}{\sigma} = 1$.

Question 5:

(a) Three coins are tossed. The event A is defined as "at most one head", and the event B is defined as "all heads or all tails". Are the events A and B independent?

See next page

(b) Dr. Yellowcreek is trying to construct a problem involving two events A and B so that $A \cup B = \mathbf{S}$, where **S** is the sample space, and $A \cap B = \emptyset$. A third event C is also defined. He has already decided that $P(A) = 0.4$. How should he choose $P(B|C)$ so that B and C are independent?

See next page

(c) There are 10 identical chop sticks. 4 are painted red. Three chop sticks are selected randomly, in succession, without replacement. What is the probability that the last two are red, if the first one is not? Do this problem using combinations and/or permutations.

See next page

Question 5

a.) $S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

each simple event is assigned prob. $\frac{1}{8}$

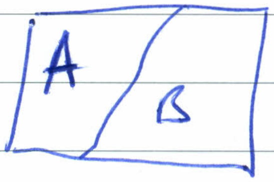
$A = \{TTT, HTT, THT, TTH\} \rightarrow P(A) = \frac{4}{8} = \frac{1}{2}$

$B = \{TTT, HHH\} \rightarrow P(B) = \frac{2}{8} = \frac{1}{4}$

$A \cap B = \{TTT\} \rightarrow P(A \cap B) = \frac{1}{8}$

Since $\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, they are independent

b.)



$P(B) = 1 - P(A) = 0.6$

$P(B|C) = P(B) = \underline{0.6}$

c.) $\{$The first one is not red!$\} = A$

$B = \{$1st one not red, 2nd and 3rd are $\}$

$S = \{$all groups of 3 chopsticks$\}$

$N(S) = P_{10,3} \quad N(A) = 6 \cdot P_{9,2} \quad N(B) = 6 \cdot P(4,2)$

$\boxed{\text{See Note}}$ $\frac{P(B)}{P(A)} = \frac{N(B)}{N(A)} = \frac{6 \cdot 4 \cdot 3}{6 \cdot 9 \cdot 8} = \frac{12}{72} = \frac{1}{6}$

KOÇ ÜNİVERSİTESİ

Note to (c)

$$P_r\left(\text{last two red} \,/\, \text{1st one is not red}\right)$$

$$= \cancel{P_r(} \frac{P_r[\text{last two red and 1st one is not}]}{P_r[\text{1st one is not red}]}$$

$$= \frac{P(B)}{P(A)}$$

Also:

Using combinations:

1st one is not red. How many ways can we choose the last two?

$$\rightarrow \binom{9}{2}$$

How many of those pairs are red?

$$\rightarrow \binom{4}{2}$$

$$\Rightarrow \frac{\binom{4}{2}}{\binom{9}{2}} = \frac{\frac{4!}{2!2!}}{\frac{9!}{7!2!}} = \frac{4.3}{9.8} = \frac{1}{6}$$