ISEG - Lisbon School of Economics and Management

Statistics I

1st Semester of 2019/2020

Repeat Exam

3rd February 2020 Duration: 120 minutes

Name: ______________________ Number: ________

Justify your answers carefully and present all the calculations you consider necessary.

Q	1a	1b	1c	1d	2a	2b	3a	3b	3c	3d	4a	4b	4c	5	6a	6b
Val	0.5	0.5	0.5	0.5	1.5	1.5	1	1	1	2.5	1	1	1	2.5	2	2

1. Let A and B be two events of a sample space S such that $P(A) = \frac{1}{3}$ and $P(B) = \frac{1}{2}$. Classify the following statements as true or false, justifying carefully your answer.

   (a) $P(A \cup B) > 5/6$.

   $$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 5/6 - P(A \cap B) \leq 5/6$$

   Therefore, the statement is false.

   (b) If A and B are independent events, then $P(A \cup B) = 2/3$.

   True ~~False~~, because

   $$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1/3 + 1/2 - 1/6 = 5/6 - 1/6 = 4/6 = 2/3$$

   (c) $P(A|B) \leq \frac{2}{3}$.

   True

   $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{1/2} \leq \frac{\min(P(A), P(B))}{1/2}$$

$$= \frac{1/3}{1/2} = 2/3$$

(d) A and B are a partition of the sample space S.

False , because

$P(A \cup B) < 1$

2. A firm has 3 assembly lines, A_1, A_2, A_3, to produce a certain product. It is known that 1% of products from line A_1 are defective, 5% of products from line A_2 are defective and 10% of products from line A_3 are defective.

(a) Knowing that line A_1 produces 50% and line A_2 produces 30% of the total production, what is the probability that a random chosen product is defective?

D: the product is defective

A_i: the product chosen was produced in line A_i with $i = 1, 2, 3$

From the question, we know that

$P(D|A_1) = 0.01$ $\quad P(D|A_2) = 0.05$ $\quad P(D|A_3) = 0.1$

$P(A_1) = 0.5$ $\quad P(A_2) = 0.3$ $\quad P(A_3) = 0.2$

$P(D) = P(D|A_1) \times P(A_1) + P(D|A_2)P(A_2) + P(D|A_3)P(A_3)$

$= 0.04$

(b) If a random chosen product is defective, what is the probability that it comes from assembly line A_2?

$$P(A_2 | D) = \frac{P(D|A_2) \times P(A_2)}{P(D)}$$

$$= \frac{0.3 \times 0.05}{0.04} = 0.375$$

3. Assume that X is a continuous random variable, such that its density function is given by

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & a < x < b \\ 0, & \text{otherwise} \end{cases},$$

with $a, b > 0$.

(a) Find the admissible set of values for a and b.

f_X is a density function if $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$ and $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$.

From the second condition we get

$$\int_{-\infty}^{+\infty} f_X(x)\,dx = 1 \Leftrightarrow \int_a^b \frac{1}{x^2}\,dx = 1 \Leftrightarrow \left[-x^{-1}\right]_a^b = 1$$

$$\Leftrightarrow \frac{1}{a} - \frac{1}{b} = 1 \Leftrightarrow b - a = ab$$

$(a,b) \in \{(x,y) \in \mathbb{R}^2 : y - x = x\dots$

From now on, assume that $a = \frac{1}{2}$ and $b = 1$.

(b) Compute the cumulative density function.

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du = \begin{cases} 0, & x < 1/2 \\ \int_{1/2}^{x} \frac{1}{u^2}\,du, & 1/2 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

$$= \begin{cases} 0, & x < 1/2 \\ 2 - 1/x, & 1/2 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

(b) Compute the quantile of order 0.25, $q_{0.25}$.

$$q_{0.25} = \min\{x \in \mathbb{R} : F_X(x) \geq 0.25\}$$

$$= \min\{x \in \mathbb{R} : 2 - 1/x \geq 1/4\}$$

$$= \min\{x \in \mathbb{R} : 7/4 \geq 1/x\} = \min\{x \in \mathbb{R} : x \geq 4/7\}$$

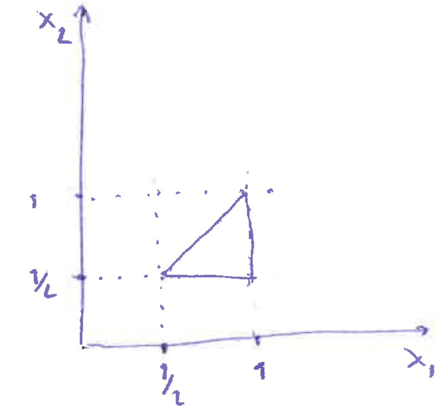
$$= 4/7$$

(c) Assume that the random variables X_1 and X_2 are independent and identically distributed to X. Compute $P(X_2 \leq X_1)$.

X_1 and X_2 are independent and identically distributed to X then

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) = \begin{cases} \frac{1}{(x_1 x_2)^2}, & 1/2 < x_1 < 1 \\ & 1/2 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$P(X_2 \leq X_1) = \int_{1/2}^{1} \int_{1/2}^{x_1} \frac{1}{(x_1 x_2)^2}\, dx_2\, dx_1$$



$$= \int_{1/2}^{1} \frac{1}{x_1^2}\left(2 - \frac{1}{x_1}\right) dx_1 =$$

$$= 2\int_{1/2}^{1} \frac{1}{x_1^2}\, dx_1 - \int_{1/2}^{1} \frac{1}{x_1^3}\, dx_1$$

$$= 2 - \left.\frac{x_1^{-2}}{-2}\right]_{1/2}^{1} = 2 + (1/2 - 2)$$

$$= 1/2$$

4. Assume that Z_1 and Z_2 are random variables that represent, respectively, the number of computers sold from brand 1 and brand 2 in a random day. The probability function of (Z_1, Z_2) is given by

Z_1/Z_2	0	1	2
0	1/30	1/10	1/15
1	2/15	1/6	7/30
2	1/15	1/10	1/10

(a) What is the probability that, in a random day, brand 1 sells more computers than brand 2?

$$P(Z_1 > Z_2) = P(Z_1 = 1, Z_2 = 0)$$

$$+ P(Z_1 = 2, Z_2 = 0) + P(Z_1 = 2, Z_2 = 1)$$

$$= 2/15 + 1/15 + 1/10$$

$$= 3/10$$

(b) What is the probability that, in a random day, brand 1 sells at least one computer given that brand 2 does not sell any computer?

$P(Z_2=0) = \sum_{z_1=0}^{2} P(Z_2=0, Z_1=z_1) = 7/30$

$P(Z_1 \geq 1 \mid Z_2 = 0) = P(Z_1 = 1 \mid Z_2 = 0) + P(Z_1 = 2 \mid Z_2 = 0)$

$= \frac{P(Z_1=1, Z_2=0)}{P(Z_2=0)} + \frac{P(Z_1=2, Z_2=0)}{P(Z_2=0)} = \frac{2/15}{7/30} + \frac{1/15}{7/30}$

$= 6/7$

(c) Are Z_1 and Z_2 independent random variables?

No, they are not independent because

$P(Z_1=0, Z_2=0) = 1/30 \neq P(Z_1=0)P(Z_2=0) = 7/450$

$P(Z_1=0) = \sum_{z_2=0}^{2} P(Z_1=0, Z_2=z_2) = 6/30$

5. Let X and Y be two discrete random variables such that

$$E(X|Y=0) = 2/3, \quad E(X|Y=1) = 1/2 \quad \text{and} \quad Y \sim Bin(n=1, p=1/4).$$

Find $E(Y)$ and $Var(Y)$. Compute $E(X)$.

Since $Y \sim Bin(n=1, p=1/4)$

$E(Y) = p = 1/4$

$Var(Y) = p(1-p) = 1/4 \times 3/4 = 3/16$

$E(X) = E(E(X|Y))$

tower property

$W = E(X|Y)$ is a r.v. such that

$P(W = 2/3) = P(Y=0)$

$P(W = 1/2) = P(Y=1)$

$P(W=w) = 0$ when $w \notin \{1/2, 2/3\}$

$E(X) = E(X|Y=0)P(Y=0) + E(X|Y=1)P(Y=1)$

$= 2/3 \times 0 + 1/2 \times 1 = 1/2$

6. Let X be a random variable following a normal distribution with mean 100 and variance 225, which represents a person's result on a psychometric test.

   (a) What percentage of people score between 85 and 115? On average, how many people does a psychiatrist have to apply the test to, in order to find one that scores between 85 and 115?

$$P(85 < X < 115) = P\left(\frac{85-100}{15} < \frac{X-100}{15} < \frac{115-100}{15}\right)$$

$$Z = \frac{X-100}{15} \sim N(0,1)$$

$$= P(-1 < Z < 1) = \Phi(1) - \Phi(-1)$$

$$= \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1$$

$$= 0.6827$$

About 68% of people score between 85 and 115

Y is the R.V. that represents the number of people that complete the test to find one that scores between 85 and 115

$Y \sim Geo(p)$, $p = 0.6827$ $E(Y) = \frac{1}{p} \simeq 1.46$

Therefore, one has to apply the test to, at least, two people.

   (b) In a scientific study, 25 people have completed the psychometric test, in an independent way. What is the probability that, on average, one person from this group scores more than 106.

$X_1, X_2, \ldots, X_{25}$ are independent and identically distributed to X

$\bar{X} = \frac{1}{25}\sum_{i=1}^{25} X_i$ represents the average score of one person in the group

$\bar{X} \sim N(100, 9)$

$$E(\bar{X}) = \frac{1}{25} E\left(\sum_{i=1}^{25} X_i\right)$$

$$= \frac{1}{25} \times 25 \times E(X) = 100$$

$$Var(\bar{X}) = \frac{1}{25^2} Var\left(\sum_{i=1}^{25} X_i\right)$$

$$= \frac{1}{25^2} \times 25 \times Var(X) = 9$$

$$P(\bar{X} > 106) =$$

$$= P\left(\frac{\bar{X}-100}{3} > \frac{106-100}{3}\right)$$

$$= P(Z > 2) = 1 - \Phi(2)$$

$$= 0.0228$$

$$Z = \frac{\bar{X}-100}{3} \sim N(0,1)$$