20205 2307 ①

Saathi

Date ___/___/___

## MA201: Probability and Statistics

### Tutorial - 2

1. A computer virus is trying to corrupt two files. The first file will be corrupted with probability 0.4. Independently of it, the second file will be corrupted with probability 0.3. Compute the pmf and cdf of X, the number of corrupted files

random variable X, the no. of corrupted can have following values: 0, 1 and 2.

$P(X=0) = P(\text{first file is not corrupted}) \times P(\text{second file is not corrupted})$

$P(X=0) = 0.6 \times 0.7 = 0.42$

$P(X=1) = P(\text{first file is not corrupted}) \times P(\text{second file is corrupted}) + P(\text{first file is corrupted}) \times P(\text{second file is not corrupted})$

$P(X=1) = 0.6 \times 0.3 + 0.4 \times 0.7$
$= 0.46$

$P(X=2) = P(\text{first file is corrupted}) \times P(\text{second file is corrupted})$
$= 0.4 \times 0.3$
$= 0.12$

202052307

| X | pmf (X) | cdf (X) |
|---|---|---|
| 0 | 0.42 | 0.42 |
| 1 | 0.46 | 0.88 |
| 2 | 0.12 | 1.00 |

Answer

2 a) Every day, the number of network blackouts has a distribution (probability mass function)

| x | 0 | 1 | 2 |
|---|---|---|---|
| P(x) | 0.7 | 0.2 | 0.1 |

A small internet trading company estimates that each network blackout results in a \$ 500 loss. Compute expectation and variance of this company's daily loss due to blackouts.

The expectation value of company's daily loss ($\mu$ or $E(X)$) can be calculated as:

$$\mu = E(X) = \sum_{x=0}^{2} x \cdot P(x) (\$500)$$

$$= 0(0.7)(\$500) + 1(0.2)(\$500) + 2(0.1)(\$500)$$

$$\mu = E(X) = 0 + \$100 + \$100$$

$$\therefore \boxed{\mu = E(X) = \$200}$$

Page No.

202052307

Saathi

Date __/__/__

the variance of company's daily loss, Var(X), can be calculated as:

$$Var(X) = \sum_{x=0}^{2} (\$500 \cdot x - \mu)^2 \cdot P(x)$$

$$= (-200)^2 \times 0.7 + (300)^2 \times 0.2 + (800)^2 \times 0.1$$

$$= (28000 + 18000 + 64000) \text{ squared dollars}$$

$$\boxed{Var(X) = 110000 \text{ squared dollars}}$$

The expected value of company's daily loss due to blackouts is \$200 and variance is 1,10,000 squared dollars.

2 b) There is an error in one of the five blocks of a program. To find the error, we test three randomly selected blocks. Let X be the number of errors in these three blocks. Compute E(X).

$$P(X=0) = \frac{(\text{No. of ways to select 3 blocks out of 4 correct blocks})}{\text{No. of ways to select 3 blocks out of 5 blocks}}$$

$$\therefore P(X=0) = \frac{{}^4C_3}{{}^5C_3} = 0.4$$

$$P(X=1) = \frac{(\text{ways to select 2 block out of 4 correct blocks}) \times (\text{ways to select block with error})}{\text{no. of ways to select 3 blocks out of 5 blocks}}$$

2020052307

Saathi

Date ___ / ___ / ___

$$P(X=1) = \frac{{}^4C_2 \cdot {}^1C_1}{{}^5C_3} = 0.6$$

$\therefore$ $P(X=0) = 0.4$

$P(X=1) = 0.6$

$\therefore$ $E(X) = 0 \times 0.4 + 1 \times 0.6$

$\boxed{E(X) = 0.6}$ Answer

2 (c) Tossing a fair die is an experiment that can result in any integer number from 1 to 6 with equal probabilities. Let X be the number of dots on the top face of a die. Compute E(X) and Var (X)

$X = \{1, 2, 3, 4, 5, 6\}$

$P(X = 1, 2, 3, 4, 5, 6) = \frac{1}{6}$

$\therefore$ $E(X) = \frac{1}{6} \times (1+2+3+4+5+6)$

$= \frac{21}{6}$

$\boxed{E(X) = 3.5}$

$Var(X) = E(X^2) - (E(X))^2$

$= \frac{1}{6}(1+4+9+16+25+36) - \frac{441}{36}$

202052307

(5)

Date __/__/__

$$Var(X) = \frac{91}{6} - \frac{441}{36}$$

$$Var(X) = \frac{105}{36} = \frac{35}{12}$$

$$\boxed{Var(X) = 2.9167}$$

3 (a) A software package consists of 12 programs, five of which must be upgraded. If 4 programs are randomly chosen for testing. What is the expected number of program out of the chosen 4, that must be upgraded?

Let us define ~~a~~ four random variables $X_1, X_2, X_3$ and $X_4$ for each of the chosen file.

$$X_i \ (i = 1,2,3,4) = \begin{cases} 1 & ; \text{if program } i \text{ needs an upgrade} \\ 0 & ; \text{'' '' } i \text{ doesn't need an upgrade} \end{cases}$$

Let us define success as program i needs an upgrade. Therefore, probability of success $p$:

$\therefore$ $p = \frac{5}{12}$ (as 5 must be upgraded)

$$\therefore \ E(X_i \ (i=1,2,3,4)) = 0 \times \frac{7}{12} + 1 \times \frac{5}{12}$$

$$E(X_i) = \frac{5}{12}$$

202052307

$\therefore$ expected value of programs out of chosen 4 that must be upgraded is:

$$E(X) = E(X_1 + X_2 + X_3 + X_4)$$
$$= E(X_1) + E(X_2) + E(X_3) + E(X_4)$$
$$= 4 \times \frac{5}{12}$$

$$\boxed{\therefore E(X) = \frac{5}{3} = 1.6667}$$ Answer

3(b) A computer program contains one error. In order to find the error, we split the program into 6 blocks and test 2 of them, selected at random. Let X be the number of errors in these 2 blocks. Compute E(X).

Let us define random variables $X_1$ and $X_2$ for the two selected blocks.

$$X_i \ (i=1,2) = \begin{cases} 1 & ; \text{ if selected block has error} \\ 0 & ; \text{ if selected block has no error} \end{cases}$$

Let us define success as block has error.
$\therefore$ probability of success $p$:

$$p = \frac{1}{6} \quad (\because \text{there is only one error})$$

2020082307

Date __/__/__

Now, $E(X_i (i=1,2)) = 0 \times \frac{5}{6} + 1 \times \frac{1}{6}$

$$E(X_i) = \frac{1}{6}$$

$\therefore E(X) = E(X_1 + X_2)$

$= E(X_1) + E(X_2)$

$= \frac{1}{6} + \frac{1}{6}$

$$\boxed{E(X) = \frac{1}{3} = 0.333}$$

Answer

4. The number of home runs scored by a certain team in one baseball game is a random variable with the distrubution.

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(x)$ | 0.4 | 0.4 | 0.2 |

The team plays 2 games. The number of home runs scored in one game is independent of the number of home runs in ~~each~~ the other game. Let Y be the total number of home runs. Find $E(Y)$ and $Var(Y)$.

Let A be the first game and $A_i$ be i home runs scored in first game. Similarly, let B be the second game and $B_i$ be i home runs scored in the second game.

202052307

Saathi

Date ___/___/___

The total number of home runs scored ~~by~~ by the team in both the games, can have values ~~between~~ of 0, 1, 2, 3, 4.

$\therefore \quad Y = \{0, 1, 2, 3, 4\}$

$P(Y=0) = A_0 B_0$ (since, both matches are independent)
$= 0.16$

$P(Y=1) = A_1 B_0 + A_0 B_1$
$= 0.32$

$P(Y=2) = A_2 B_0 + A_1 B_1 + A_0 B_2$
$= 0.32$

$P(Y=3) = A_2 B_1 + A_1 B_2$
$= 0.16$

$P(Y=4) = A_2 B_2$
$= 0.04$

$\therefore \quad E(Y) = \sum_{y=0}^{4} y\, P(y)$

$= 0 \times 0.16 + 1 \times 0.32 + 2 \times 0.32 + 3 \times 0.16 + 4 \times 0.0$

$= 1.60$ home runs

$Var(Y) = E(Y^2) - (E(Y))^2$

$= (0 \times 0.16 + 1 \times 0.32 + 4 \times 0.32 + 9 \times 0.16 + 16 \times 0.04)$
$- (1.6)^2$

202052307

Date ___/___/___

$\therefore$ $Var(Y) = 3.68 - 2.56$
$= 1.12$.

$E(Y) = 1.6$ home runs
$Var(Y) = 1.12$ squared home runs

Answer

5. Let X and Y be the number of hardware failures in two computer labs in a given month. The joint distribution of X and Y is given in the table below.

| P(x, y) | | x = 0 | x = 1 | x = 2 |
|---|---|---|---|---|
| y | 0 | 0.52 | 0.20 | 0.04 |
| | 1 | 0.14 | 0.02 | 0.01 |
| | 2 | 0.06 | 0.01 | 0 |

(a) Compute the probability of at least one hardware failure.

(b) From the given distribution, are X and Y independent? Why or why not?

(a) P(at least one hardware failure) = 1 − P(0 hardware failure)
$= 1 - P(X=0, Y=0)$
$= 1 - 0.52$

$\therefore$ P(at least one hardware failure) = 0.48

Answer

202052307

(b)

| $P(x,y)$ | | $x$ | | | pmf(Y) |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | |
| $y$ | 0 | 0.52 | 0.20 | 0.04 | 0.76 |
| | 1 | 0.14 | 0.02 | 0.01 | 0.17 |
| | 2 | 0.06 | 0.01 | 0 | 0.07 |
| pmf (X) | | 0.72 | 0.23 | 0.05 | |

The probability mass function for X and Y are written in the table.

For X and Y to be independent,

$$P(x,y) = P(x) \cdot P(y) \quad \text{should be true}$$

for each $x$ in $\{0,1,2\}$ and $y$ in $\{0,1,2\}$

~~∴ P(2,2) ≠ 0 and P(2)~~

$\because P(X=2, Y=2) = 0$

and $P(X=2) \cdot P(Y=2) = 0.0035$

$\therefore P(x,y) \neq P(x) \cdot P(y)$ for $x=2, y=2$

$\boxed{\text{Hence, X and Y are dependent.}}$

6. Every day, the number of traffic accidents has the probability mass function.

| $x$ | 0 | 1 | 2 | $>2$ |
|---|---|---|---|---|
| $P(x)$ | 0.6 | 0.2 | 0.2 | 0 |

202052307

Date ___/___/___

Saathi

independently of other days. What is the probability that there are more accidents on Friday than on Thursday?

Let X be the number of accidents on Thursday and Y be the " " " on Friday.
We can make a joint distribution for these two days as follows:

| $P(x,y)$ | | Accidents on Friday (Y) | | |
|---|---|---|---|---|
| | | 0 | 1 | 2 |
| Accidents | 0 | 0.36 | 0.12 | 0.12 |
| on | 1 | 0.12 | 0.04 | 0.04 |
| Thursday | 2 | 0.12 | 0.04 | 0.04 |

$P(x,y)$ denotes the probability of $x$ accidents on Thursday and $y$ accidents on Friday.
Since, accidents on different days are independent of each other. Therefore,

$$P(x,y) = P(x) \cdot P(y)$$

Now, we can fill our table using this formula.

We need to find the probability that there are more accidents on Friday than Thursday.

Hence, P(more accidents on Friday than thursday)

$$= P(0,1) + P(0,2) + P(1,2)$$
$$= 0.12 + 0.12 + 0.04$$
$$\boxed{= 0.28}$$ Answer

Saathi

Date / /

7\. An internet service provider charges its customers for the time of the internet use rounding it up to the nearest hour. The joint distribution of the used time (X, hours) and the charge per hour (Y, cents) is given in the table below.

| $P(x,y)$ | | $x$ = 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| | 1 | 0 | 0.06 | 0.06 | 0.10 |
| $y$ | 2 | 0.10 | 0.10 | 0.04 | 0.04 |
| | 3 | 0.40 | 0.10 | 0 | 0 |

Each customer is charged $Z = X \cdot Y$ cents, which is number of hours multiplied by the price of each hour. Find the distribution of Z.

The possible values of Z are:

$\{1, 2, 3, 4, 6, 8, 9, 12\}$

To find $P\{Z=z\}$, we will find $x$ and $y$ such that $x \cdot y = z$, and choose the $(x,y)$ cell(s) and add their probabilities to find pmf(Z).

$P(Z=1) = P(1,1) = 0$

$P(Z=2) = P(1,2) + P(2,1)$
$= 0.16$

2020052307

Saathi

Date __/__/__

| z | pmf (z) | cdf (z) |
|---|---|---|
| 1 | 0 | 0 |
| 2 | 0.16 | 0.16 |
| 3 | 0.46 | 0.62 |
| 4 | 0.20 | 0.82 |
| 6 | 0.14 | 0.96 |
| 8 | 0.04 | 1.00 |
| 9 | 0 | 1.00 |
| 12 | 0 | 1.00 |

Answer

$P(z=3) = P(1,3) + P(3,1)$
$= 0.46$

$P(z=4) = P(4,1) + P(2,2)$
$= 0.20$

$P(z=6) = P(2,3) + P(3,2)$
$= 0.14$

$P(z=8) = P(4,2)$
$= 0.04$

$P(z=9) = P(3,3)$
$= 0$

$P(z=12) = P(4,3)$
$= 0$

202052307

Date

Saathi

Q. Two random variables X and Y have the joint distribution, $P(0,0) = 0.2$, $P(0,2) = 0.3$, $P(1,1) = 0.1$, $P(2,0) = 0.3$, $P(2,2) = 0.1$ and $P(x,y) = 0$ for all other $(x,y)$.

(a) Find the pmf of $Z = X+Y$

(b) Find the pmf of $U = X-Y$

(c) Find the pmf of $V = XY$.

Let us tabularize the given data

| $P(x,y)$ | | y: 0 | 1 | 2 |
|---|---|---|---|---|
| | 0 | 0.20 | 0 | 0.30 |
| x | 1 | 0 | 0.10 | 0 |
| | 2 | 0.30 | 0 | 0.10 |

(a) Find pmf of $Z = X+Y$

Z can have the following values: $0, 1, 2, 3, 4$

$P(Z=0) = P(0,0) = 0.20$

$P(Z=1) = 0 + 0 = 0$

~~$P(Z=2) = 0.30 + 0.30 + 0.10$~~

$P(Z=2) = P(0,2) + P(1,1) + P(2,0)$

$= 0.30 + 0.10 + 0.30$

$= 0.70$

2020523007

Saathi

Date ___/___/___

$P(Z=3) = P(1,2) + P(2,1) = 0$

$P(Z=4) = P(2,2) = 0.10$

| $Z=X+Y$ | pmf(z) |
|---|---|
| 0 | 0.20 |
| 1 | 0 |
| 2 | 0.70 |
| 3 | 0 |
| 4 | 0.10 |

Answer

(b) find pmf of $U = X - Y$

U can have the following values:

0, −1, −2, 1, 2

Now,

$P(U=-2) = P(0,2) = 0.30$

$P(U=-1) = P(0,1) + P(1,2) = 0$

$P(U=0) = P(0,0) + P(1,1) + P(2,2) = 0.40$

$P(U=1) = P(1,0) + P(2,1) = 0$

$P(U=2) = P(2,0) = 0.30$

202082307

Date ___/___/___

| $U = X - Y$ | pmf(U) |
|---|---|
| -2 | 0.30 |
| -1 | 0 |
| 0 | 0.40 |
| 1 | 0 |
| 2 | 0.30 |

Answer

c) find pmf of $V = X \cdot Y$

V can have following values:
0, 1, 2, 4

~~P(0) =~~

$P(V=0) = P(0,0) + P(0,1) + P(0,2) + P(1,0) + P(2,0)$
$= 0.80$

$P(V=1) = P(1,1) = 0.10$

$P(V=2) = P(1,2) + P(2,1) = 0$

$P(V=4) = P(2,2) = 0.10$

| $V = X \cdot Y$ | pmf(V) |
|---|---|
| 0 | 0.80 |
| 1 | 0.10 |
| 2 | 0 |
| 4 | 0.10 |

Answer