GATE

DATA SCIENCE + CS & IT

Engineering Mathematics

SUPER 1500

Lec : 01 Statistics-1

By - Dr. Puneet Sharma Sir

Topics *to be covered*

Statistic -1

(Discrete Random Variable)

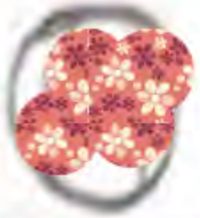
$X : \; : \; x_1 \quad x_2 \; — \; x_n$

$P(X) : \; p_1 \quad p_2 \; - - \; p_n$

① $E(X) = \sum p_i x_i$

② $Var(X) = E(X^2) - E^2(X)$

Four bad oranges are mixed accidently with 16 good oranges. Find the probability distribution of the number of bad oranges in a draw of two oranges. and it's Expected Value is ______



$X = \{\text{No. of Bad oranges in a draw of two oranges}\} = \{0, 1, 2\}$

X :	0	1	2
$P(X)$:	$\dfrac{{}^{16}C_2}{{}^{20}C_2}$	$\dfrac{{}^{4}C_1 \times {}^{16}C_1}{{}^{20}C_2}$	$\dfrac{{}^{4}C_2}{{}^{20}C_2}$

So $E(X) = \sum p_i x_i = p_1x_1 + p_2x_2 + p_3x_3$

$$= 0 + \left(\frac{\cancelto{4}{16} \times 4 \times 2}{\cancelto{5}{20} \times 19}\right) \times 1 + \left(\frac{\cancel{4} \times 3}{\cancelto{5}{20} \times 19}\right) \times 2$$

$$= \frac{32}{95} + \frac{6}{95} = \frac{38}{95} = 0.4$$

The following sequence of numbers is arranged in increasing order : 1, x, x, x, y, y, 9,16,18. Given that the mean and median are equal, and are also equal to twice the mode, the value of y is

good

(a) 5 (b) 6

(c) 7 (d) 8

Let Mean = ? = α

Mode = x

$Md = \left(\frac{N+1}{2}\right)^{th} = \left(\frac{9+1}{2}\right)^{th} = y$

ATQ, $\alpha = y$, $y = 2x$ — (1)

$\Rightarrow \boxed{\alpha = 2x}$

$$\frac{1+3x+2y+43}{9} = 2x$$

$44+3x+2y = 18x$

$2y = 15x - 44$ — (2)

Solving (1) & (2):

$x = 4$, $y = 8$

If X and Y are random variable such that $E[2X + Y] = 0$ and $E[X + 2Y] = 33$, then

$$E[X] + E[Y] = \_\_\_\_\_.$$

(a) 33 (b) 1

(c) 0 (d) 11

$2E(X) + E(Y) = 0$

$E(X) + 2E(Y) = 33$

$+$

$3(E(X) + E(Y)) = 33$

$E(X) + E(Y) = 11$

A random variate has the following distribution:

x :	0	1	2	3	4	5	6	7
$p(x)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

The value of k is ______.

[Ans:]

$$\Sigma p_i = 1$$

$$9k + 10k^2 = 1$$

$$10k^2 + 9k - 1 = 0$$

$$k = \frac{-9 \pm \sqrt{81 + 40}}{2 \times 10} = \frac{-9 \pm 11}{20} = \frac{2}{20} = \frac{1}{10} = 0.1$$

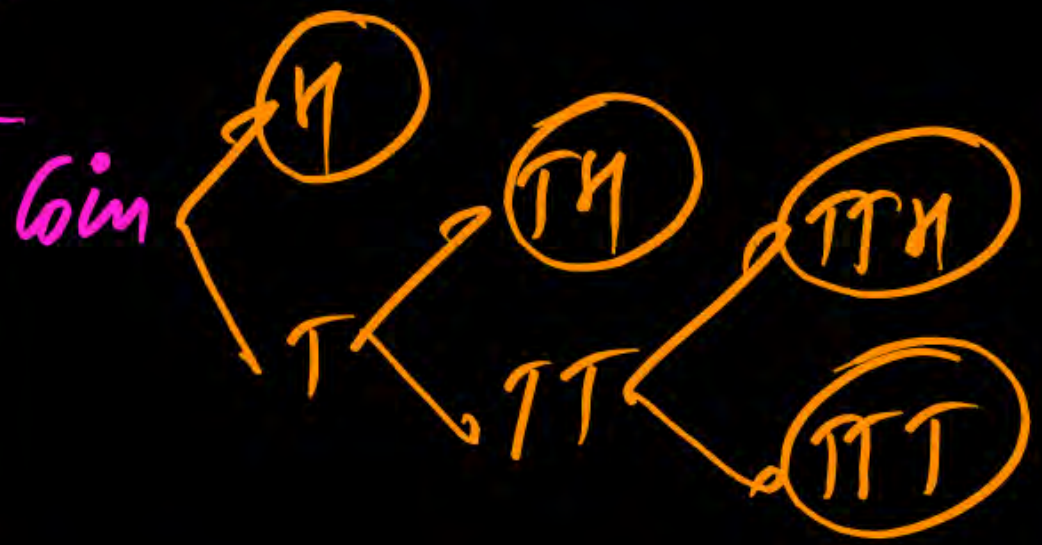
A person decides to toss a fair coin repeatedly until he gets a head. He will make at most 3 tosses. Let the random variable Y denote the number of heads. The value of var(Y), where var(.) denotes the variance, equals:

(a) $\frac{7}{8}$

(b) $\frac{49}{64}$

(c) $\frac{7}{64}$

(d) $\frac{105}{64}$



$Y = \{\text{No. of Heads}\} = \{0, 1\}$

Y:	0	1
$P(Y)$:	$1/8$	$7/8$

$p_1 = P(Y = 0H) = P(TTT) = 1/8$

$p_2 = P(Y = 1H) = P[H \text{ or } TH \text{ or } TTH] = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$

$$\boxed{\begin{array}{c|cc} y: & 0 & 1 \\ p(y) & \frac{1}{8} & \frac{7}{8} \end{array}} \rightarrow E(y) = p_1 y_1 + p_2 y_2 = 0 + \frac{7}{8} \times 1 = 7/8$$

$$E(y^2) = p_1 y_1^2 + p_2 y_2^2 = 0 + \frac{7}{8} \times (1)^2 = 7/8$$

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

$$= \frac{7}{8} - \frac{49}{64} = \frac{7}{8}\left(1 - \frac{7}{8}\right) = \frac{7}{64}$$

A fair coin is tossed repeatedly till both head and tail appear at least once. The average number of tosses required is______.

$X = \{\text{No. of tosses Required}\} = \{2, 3, 4, 5, 6 \ldots\ldots\}$

X:	2	3	4	5	------
$P(X)$:	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	------

$$E(x) = p_1x_1 + p_2x_2 + p_3x_3 + \cdots\cdots$$

$$= \frac{1}{2}(2) + \frac{1}{4}(3) + \frac{1}{8}(4) + \frac{1}{16}(5) + \cdots\cdots$$

$$= 2\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + 4\left(\frac{1}{2}\right)^3 + 5\left(\frac{1}{2}\right)^4 + \cdots\cdots$$

$$= \left[1 + 2\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + 4\left(\frac{1}{2}\right)^3 + \cdots\cdots\right] - 1$$

$$= (1-x)^{-2} - 1 \qquad \left[\because (1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \cdots\cdots\right]$$

$$= \left(1 - \frac{1}{2}\right)^{-2} - 1$$

Av No of tosses $= \frac{1}{2^{-2}} - 1 = 2^2 - 1 = 4 - 1 = 3$

A fair die with faces {1, 2, 3, 4, 5, 6} is thrown repeatedly till '3' is observed for the first time. Let X denote the number of times the die is thrown.

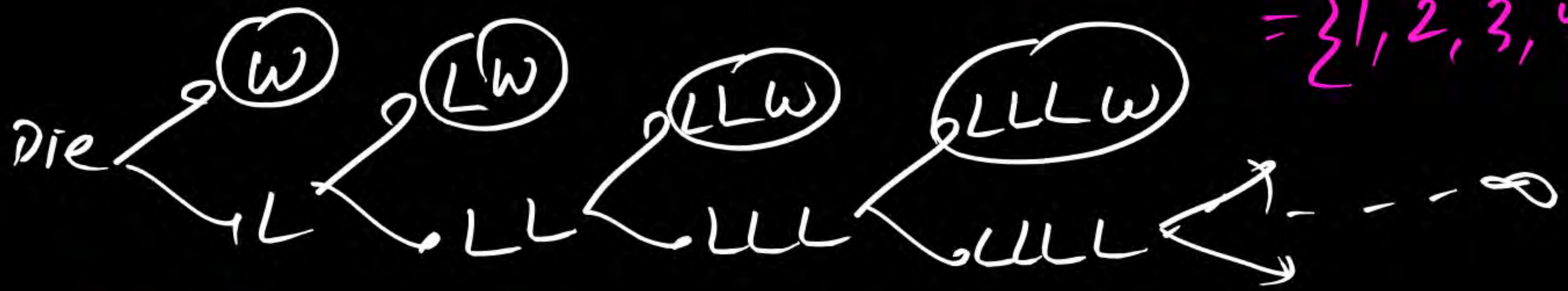
then $E(X) =$ ______

$P(3) = P(W) = \frac{1}{6}$

$P(\bar{3}) = P(L) = \frac{5}{6}$

$X = \{$No. of times die is thrown$\}$

$= \{1, 2, 3, 4, 5 \ldots\}$



X:	1	2	3	4	----
$P(X)$:	$\frac{1}{6}$	$\frac{5}{6}\cdot\frac{1}{6}$	$\left(\frac{5}{6}\right)^2\frac{1}{6}$	$\left(\frac{5}{6}\right)^3\frac{1}{6}$	----

$$E(X) = \sum p_i x_i = \frac{1}{6}(1) + \frac{5}{6}\frac{1}{6}(2) + \left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)\cdot(3) + \cdots$$

$$= \frac{1}{6}\left[1 + 2\left(\frac{5}{6}\right) + 3\left(\frac{5}{6}\right)^2 + 4\left(\frac{5}{6}\right)^3 + \cdots\right]$$

$$= \frac{1}{6}\left[1 - \left(\frac{5}{6}\right)\right]^{-2} = \frac{1}{6}\left(\frac{1}{6}\right)^{-2} = \frac{1}{6}\,\frac{1}{6^{-2}} = \frac{6^2}{6} = 6$$

BINOMIAL
(with Replacement) ① Bernoullie Trial → success = p, failure = q & $\boxed{q+p=1}$

② $\boxed{P(x=r) = {}^nC_r\, p^r q^{n-r}}$, Mean $= np$, Var $= npq$, SD $= \sqrt{npq}$

③ n = finite, n = Ind, p = Const for each R Exp, R Exp → Success, failure

X = {which is Required} → Success.

$n=5$, $r=3$, $p=\frac{3}{5}$, $q=\frac{2}{5}$

$P(x=3) = {}^5C_3\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 = {}^5C_3\left(\frac{3}{5}\times\frac{3}{5}\times\frac{3}{5}\right)\times\left(\frac{2}{5}\times\frac{2}{5}\right)$

In a binomial distribution, the mean is 9 and the standard deviation (σ) is $\sqrt{6}$. The value of n (total number of trials) and q (probability of failure of the event in each trial) respectively are:

(a) $27, \frac{1}{3}$

(b) $27, \frac{2}{3}$

(c) $36, \frac{3}{4}$

(d) $18, \frac{1}{2}$

$(n, q) = ?$

$np = 9 \quad , \quad npq = 6$

$q = \frac{6}{9} = \frac{2}{3}$

$n\left(\frac{1}{3}\right) = 9 \qquad p = \frac{1}{3}$

$n = 27$

A fair dice is tossed eight times. The probability that a third six is observed on the eight throw is ____

$p = P(\text{getting six}) = \frac{1}{6}$, $q = P(\text{not getting six}) = \frac{5}{6}$,

Req Prob = $P(\text{getting exactly 2 six in } 1^{st}\ 7 \text{ throws}) \times P(\text{six in } 8^{th} \text{ throw})$

$$= \left({}^7C_2\, p^2 q^5\right) \times \frac{1}{6}$$

$$= 0.039$$

A die has four blank faces and two faces marked 3. The chance of getting a total of 12 in 5 throws is

(a) ${}^5C_4\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)$ (b) ${}^5C_4\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4$

(c) ${}^5C_4\left(\frac{1}{6}\right)^5$ (d) none of these

Die: 3 occurs $= p = \frac{2}{6} = \frac{1}{3}$

3 Not occurs $= q = \frac{4}{6} = \frac{2}{3}$

Sum $= 12 = \{(33330), ?, ?, ?\}$

$= \{$Exactly Four times 3 must occur$\}$

$X = \{$No. of times 3 is occurring$\}$ success, $n = 5$

$P(X=4) = {}^5C_4\, p^4 q^1$

A man takes a step forward with probability 0.4 and backward with probability 0.6. The probability that at the end of 11 steps he is one step away from the starting point is

(a) $\left(\frac{6}{25}\right)^5$ (b) $462\left(\frac{6}{25}\right)^5$

(c) $538\left(\frac{1}{25}\right)^5$ (d) $\left(\frac{1}{25}\right)^5$ [JNU]

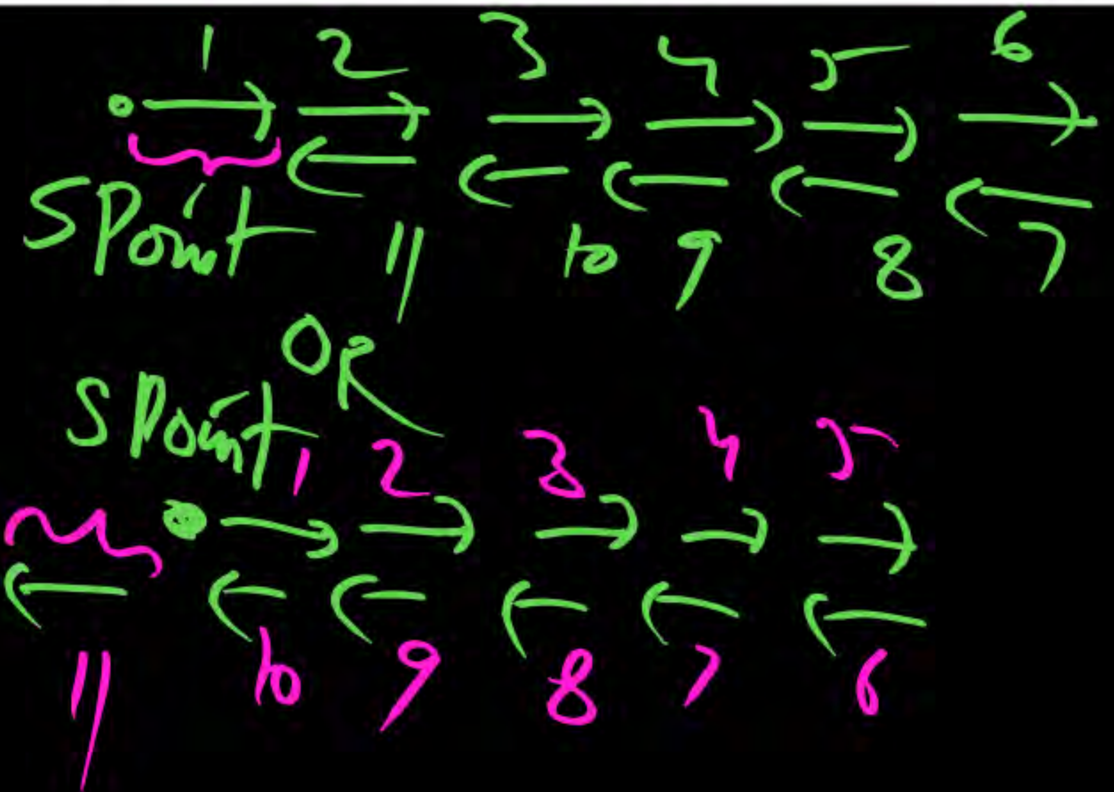
$x = \{\text{No. of F. steps}\} \rightarrow$ success

Man: F Step $\Rightarrow p = 0.4$, B Step $\Rightarrow q = 0.6$, $n = 11$ steps

Req Prob $= P(x = 5 \text{ F Steps or } 6 \text{ F Steps})$

$= {}^{11}C_5 (0.4)^5 (0.6)^6 + {}^{11}C_6 (0.4)^6 (0.6)^5$

$= (0.6)^5 [\ ?\ ] = (b)$



The probability that a screw manufactured by a company is defective is 0.1. The company sells screws in packets containing 5 screws and gives a guarantee of replacement if one or more screws in the packet are found to be defective. The probability that a packet would have to be replaced is ______.

$X = \{\text{No. of Defective Screws}\}$

Screw → Def $\Rightarrow p = 0.1$, $n = 5$

Screw → Non Def $\Rightarrow q = 0.9$

$P(\text{Packet would have to be Replaced})$

$= P(X \geq 1) = 1 - P(X=0)$

$= 1 - {}^5C_0 (0.1)^0 (0.9)^5$

$= 1 - (0.9)^5 = 0.41$

With & W/O Replacement →

There are 25 calculators in a box. Two of them are defective. Suppose 5 calculators are randomly picked for inspection (i.e., each has the same chance of being selected), what is the probability that only one of the defective calculators will be included in the inspection?

(a) $\frac{1}{2}$ (b) $\frac{1}{3}$

(c) $\frac{1}{4}$ (d) $\frac{1}{5}$

Correct App : (Hypergeometric)

25 Calci → 2 Def, 5 Calci, 23 Non Def

2 Def → 1 Def

23 Non Def → 4 N.D

$$\text{Req Prob} = \frac{\text{fav}}{\text{Total}} = \frac{{}^2C_1 \times {}^{23}C_4}{{}^{25}C_5}$$

$$= \frac{2 \times \frac{23 \times 22 \times 21 \times 20}{4 \times 3 \times 2 \times 1}}{25 \times 24 \times 23 \times 22 \times 21} \times 5 \times 4 \times 3 \times 2 \times 1$$

w/o Replacement

~~WRONG APP: $X = \{$No. of Def Calci$\}$~~ $p = \frac{2}{25}$, $q = \frac{23}{25}$

success

~~$n = 5, \ r = 1$~~

~~$P(X = 1 \text{ Def Calci}) = {}^5C_1 \left(\frac{2}{25}\right)^1 \left(\frac{23}{25}\right)^4 = ??$~~

∵ we are drawing Calci from finite Population.

A consignment of 15 record players contains 4 defectives. The record players are selected at random, one by one, and examined. The ones examined are not put back. What is the probability that 9th one examined is the last defective? ________ Sol:

→ w/o Replacement

15 R.P

4 Def   11 Non Def

$$P(\text{Def}) = \frac{4}{15}$$

$$\text{Req Prob} = P(\text{exactly 3 def in 1}^{st}\text{ 8 test}) \times P(\text{Def in 9}^{th}\text{ test})$$

$$= \left(\frac{\text{fav}}{\text{Total}}\right) \times \left(\frac{1}{7}\right)$$

$$= \left(\frac{{}^4C_3 \times {}^{11}C_5}{{}^{15}C_8}\right) \times \left(\frac{1}{7}\right)$$

Explanation

$15\ \text{R.P} \begin{cases} 4\ \text{Def} \\ 11\ \text{Non Def} \end{cases}$

Hypergeometric Dist.

$$\text{Req Prob} = P(\text{getting exactly 3 Def in } 1^{st}\ 8\ \text{test}) \times P(\text{Def in } 9^{th}\ \text{test})$$

$$= \left(\frac{{}^{4}C_{3} \times {}^{11}C_{5}}{{}^{15}C_{8}} \right) \times \frac{1}{7}$$

POISSON

$\lambda \to$ Average per unit Data, $\boxed{\text{Mean} = \text{Var} = \lambda}$

$$\boxed{P(x=r) = \frac{e^{-\lambda}\lambda^{r}}{r!}}$$

Cond^n: $n \to \infty$

$p \to 0$

$np \to$ const $(=\lambda)$

(*) n is not sure but we know λ then we should apply P. Dist.

(*) B. Dist $\longrightarrow$ P. Dist

(n, p) $\not\longleftarrow$ (λ)

If a random variable X satisfies the Poisson's distrubution with a $\underbrace{\text{mean value of 2}}_{\lambda =}$, then the probability that $X > 2$ is

(a) $2e^{-2}$ (b) $1-2e^{-2}$ (e) $1-5e^{-2}$

(c) $3e^{-2}$ (d) $1-3e^{-2}$

$\lambda = 2$, $P(x>2) = P(x \geq 3) = 1 - P(x \leq 2)$

$= 1 - [P(0)+P(1)+P(2)]$

$= 1 - e^{-\lambda}\left(\frac{\lambda^0}{0!}+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}\right) = 1 - e^{-2}(1+2+2)$

$= 1 - \frac{5}{e^2}$

It is estimated that the average number of events during a year is three. What is the probability of occurrence of not more than two events over a two-year duration? Assume that the number of events follow a poisson distribution.

(a) 0.052
(b) 0.062
(c) 0.072
(d) 0.082

X = { No. of Events in 2yrs }

Min Events = 0

Av " = 3

Max " = ∞

$\underbrace{P(0)+P(1)+P(2)}_{X\le 2}+P(3)+P(4)+\cdots=1$

$\lambda = 3$ Events/yr $= 6$ Events/2yrs.

$P(X\le 2) = ? = P(X=0 \text{ or } 1 \text{ or } 2) = e^{-\lambda}\left[\frac{\lambda^0}{0!}+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}\right]$

$= e^{-6}(1+6+18) = \frac{25}{e^6} = 0.062$

The second moment of a Poisson-distributed random variable is 2. The mean of the random variable is ....

w.k. that in P. Dist, Mean = Var = λ

$\leftarrow \ Var(x) = E(x^2) - E^2(x)$

$\lambda = 2 - (\lambda)^2$

$\boxed{\lambda^2 + \lambda - 2 = 0} \Rightarrow \lambda = -2, 1$

So $\lambda = 1$ $\because \lambda$ cannot -ve

If a random variable X has a Poisson distribution with mean 5, then the expectation $E\left[(X+2)^2\right]$ equals ______.

$Mean = Var = 5 = \lambda$

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

$5 = E(x^2) - (5)^2$

$E(x^2) = 30$

$E(x+2)^2 = E(x^2 + 4 + 4x)$

$= E(x^2) + E(4) + 4E(x)$

$= 30 + 4 + 4(5)$

$= 54$

The average amount earned by a employee is 2 rupees per day. What is the probability that 3 rupees will be earned tomorrow?

(a) 0.85

(b) 0.75

(c) 0.18

(d) 0.32

$\lambda = 2$ Rs / day

$X = \{$ Amount earned by employee tomorrow $\}$

So Av for tomorrow $(\lambda) = 2$ Rs / day

$$P(X=3) = \frac{e^{-\lambda}\lambda^3}{3!} = \frac{e^{-2} \cdot 2^3}{3!} = \frac{8}{6}\frac{1}{e^2} = \frac{4}{3e^2} = 0.18$$

A car manufacturing unit produces an average of 12 cars per hour. What is the probability that no car is manufactured in a particular 2 hr period?

(a) e^{-8} (b) e^{-24}

(c) e^{-4} (d) e^{-12}

$\lambda = 12/\text{hr}$

$= 24 \text{ cars}/2\text{hrs}.$

$$P(X=0) = \frac{e^{-\lambda}\cdot\lambda^0}{0!} = e^{-24}$$

$X = \{\text{No. of cars in 2 hrs}\} \rightarrow \text{success}$

$\lambda = 12 \text{ cars}/\text{hr} = 24 \text{ cars}/\text{two hrs}$

$$P(X=r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

$$P(X=0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-24}$$

Thank you

Keep Hustling!