Probability

G V V Sharma*

Contents

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

***Abstract*—This book provides solved examples on Probability from IES stats question papers.**

1 Poisson

1.1. Using Central Limit theorem, show that

$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\frac{1}{2} \tag{1.1.1}$$

Solution:

Definition 1. *Let a discrete rv X having poission distribution , then **PMF** is given by*

$$f(k;\lambda)=\Pr(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \tag{1.1.2}$$

Let X_1,X_2,X_3......X_n be the i.i.d rv with $X_i \sim$ Pois(1).

$$E(X_i)=\mu=\lambda=1 \tag{1.1.3}$$
$$var(X_i)=\sigma^2=\lambda=1 \tag{1.1.4}$$
$$\tag{1.1.5}$$

Let a random variable,

$$X=X_1+X_2+......+X_n \tag{1.1.6}$$
$$\implies X\sim Pois(n) \tag{1.1.7}$$

Theorem 1.1 (Classical central limit theorem)**.** *Let X_n be a sequence of independent, identically distributed (i.i.d.) random variables. Assume each X has finite mean, E(X) = μ, and finite variance, Var(X) = σ^2. Let Z_n be the normalized average of the first n random variables.*

$$Z_n=\frac{X-n\mu}{\sqrt{n\sigma^2}} \tag{1.1.8}$$

then Z_n converges in distribution to a standard normal distribution

For X

$$\lambda=n \tag{1.1.9}$$
$$\Pr(X=k)=\frac{n^k e^{-n}}{k!} \tag{1.1.10}$$

Corollary 1.2. *By theorem , Z_n converges to standard normal distribution as n goes to infinity i.e CDF of Z_n converges to CDF of standard normal distribution.*

$$\Pr(Z_n \le x) \xrightarrow{n\to\infty} = \Phi(x) \tag{1.1.11}$$

Proof.

$$\implies e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} = \sum_{k=0}^{n} e^{-n}\frac{n^k}{k!} \tag{1.1.12}$$

$$= \sum_{k=0}^{n}\Pr(X=k) \tag{1.1.13}$$

$$= \Pr(X \le n) \tag{1.1.14}$$

$$= \Pr\left(\frac{X-n\mu}{\sqrt{n\sigma^2}} \le \frac{n-n\mu}{\sqrt{n\sigma^2}}\right) \tag{1.1.15}$$

$$= \Pr(Z_n \le 0) \tag{1.1.16}$$

Using (1.1.11),

$$\Pr(Z_n \le 0) \xrightarrow{n\to\infty} \Phi(0) = \frac{1}{2} \tag{1.1.17}$$

□

2 2015

2.1. For random variables X and Y, show that: $Var[Y] = E[Var(Y|X)] + Var[E(Y|X)]$ **Solution:**

Let the abbreviations LE and LIE denote linearity of expectations and law of iterated expectations respectively.

$$Var[Y] = E[Y^2] - [E(Y)]^2 \text{ (definition)} \tag{2.1.1}$$

$$= E[E(Y^2|X)] - (E[E(Y|X)])^2 \text{ (LIE)} \tag{2.1.2}$$

$$= E[E(Y^2|X)] - (E[E(Y|X)])^2 - E([E(Y|X)]^2) + E([E(Y|X)]^2) \tag{2.1.3}$$

$$= E[E(Y^2|X)] - E([E(Y|X)]^2) + E([E(Y|X)]^2) - (E[E(Y|X)])^2 \text{ (LE \& LIE)} \tag{2.1.4}$$

$$= Var[E(Y|X)] + E[Var(Y|X)] \text{ (definition)} \tag{2.1.5}$$

Hence, proved.

2.2. Let X be a Random Variable with $E[X] = 3$, $E[X^2] = 13$. Use Chebyshev's Inequality to obtain $\Pr(-2 < X < 8)$

Solution: Let X be a random variable with finite expected value $E[X]$ and finite non-zero variance σ^2. Then for any real number k > 0,

$$\Pr(|X - E[X]| \ge k\sigma) \le \frac{1}{k^2} \tag{1}$$

computing the variance,

$$\sigma^2 = E[X^2] - E[X]^2$$
$$\implies \sigma^2 = 13 - 9 = 4 \tag{2}$$
$$\sigma = 2 \tag{3}$$

using (3),

$$\Pr(-2 < X < 8) = 1 - \Pr(|X-3| > 5) \tag{4}$$
$$\Pr(|X-3| > 5) = \Pr(|X - E[X]| > k\sigma) \tag{5}$$
$$k\sigma = 5$$
$$\implies 2k = 5$$
$$\therefore k = \frac{5}{2} \tag{6}$$

Using (1) , (5) and (6) in (4),

$$\Pr(-2 < X < 8) \ge 1 - \left(\frac{2}{5}\right)^2$$
$$\implies \Pr(-2 < X < 8) \ge \frac{21}{25} \tag{7}$$

2.3. Three points are chosen on the line of unit length.Find the probability that each the 3 line segments have length greater than $\frac{1}{4}$.

Solution:

Let $X, Y \in \{0, 1\}$ be the random variables which represent the position of two points on the line of unit length.

Conditions which should be satisfied to have three line segments with length greater than $\frac{1}{4}$ are given in the below table.

Then the required event which solves the prob-

Event	Condition
A	$\frac{1}{4} < X < \frac{3}{4}$
B	$\frac{1}{4} < Y < \frac{3}{4}$
C	$\frac{1}{4} < X - Y$
D	$\frac{1}{4} < Y - X$

TABLE 2.3.1: Events and their conditions

lem is ABC+ABD.

$$\Pr(ABC) = \Pr\left(\frac{1}{4} + Y < X, \frac{1}{4} < X, Y < \frac{3}{4}\right) \tag{2.3.1}$$

$$= \sum \Pr\left(Y = y | \frac{1}{4} < X, Y < \frac{3}{4}\right) \times \Pr\left(\frac{1}{4} + y < X, \frac{1}{4} < X < \frac{3}{4}\right) \tag{2.3.2}$$

$$= \int_{\frac{1}{4}}^{\frac{3}{4}} dy f_Y(y) \times \Pr\left(\frac{1}{4} + y < X, \frac{1}{4} < X < \frac{3}{4}\right) \tag{2.3.3}$$

$$= \int_{\frac{1}{4}}^{\frac{3}{4}} dy f_Y(y) \Pr\left(\frac{1}{4} + y < X < \frac{3}{4}\right) \tag{2.3.4}$$

As X is distributed uniformly between 0 and 1.

$$\Pr\left(\frac{1}{4} + y < X < \frac{3}{4}\right) = \begin{cases} \frac{1}{2} - y & y \in \left(0, \frac{1}{2}\right) \\ 0 & \text{otherwise} \end{cases} \tag{2.3.5}$$

Using (2.3.5),(2.3.4) can be written as

$$\Pr(ABC) = \int_{\frac{1}{4}}^{\frac{1}{2}} dy f_Y(y) \left(\frac{1}{2} - y\right) \tag{2.3.6}$$

As y is distributed uniformly between 0 and 1.

$$\Pr(ABC) = \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{1}{2} - y \, dy \tag{2.3.7}$$

$$= \frac{1}{32} \tag{2.3.8}$$

Similarly,we can find,

$$\Pr(ABD) = \frac{1}{32} \tag{2.3.9}$$

As C and D are mutually exclusive events.

$$\Pr(ABC + ABD) = \Pr(ABC) + \Pr(ABD) \tag{2.3.10}$$

$$= \frac{1}{16} \tag{2.3.11}$$

$\therefore$ probability that each of the three line segments have length greater than $\frac{1}{4}$ is $\frac{1}{16}$.

2.4. Two points are chosen on a line of unit length.Find the probability that each of the 3 line segments will have length greater than $\frac{1}{4}$?
Solution: let us choose points X,Y on a line ,since we are picking points randomly they have a uniform distribution in $(0, 1)$.
let us take two random variables X,Y. This points divide in into segments $X, Y - X, 1 - X$
From question

$$X > \frac{1}{4} \tag{2.4.1}$$

$$Y - X > \frac{1}{4} \tag{2.4.2}$$

$$1 - X < 1/4 \tag{2.4.3}$$

$\binom{2}{1} \times \Pr\left(X + \frac{1}{4} < Y < \frac{3}{4}, \frac{1}{4} < X < \frac{1}{2}\right)$ is the required answer since given random variables

X,Y are interchangable.

$$\Pr\left(X+\frac{1}{4}<Y<\frac{3}{4},\frac{1}{4}<X<\frac{1}{2}\right) \quad (2.4.4)$$

$$=\int_{X=\frac{1}{4}}^{X=\frac{1}{2}}\int_{Y=X+\frac{1}{4}}^{Y=\frac{3}{4}} dYdX \quad (2.4.5)$$

$$=\int_{X=\frac{1}{4}}^{X=\frac{1}{2}}\left(\frac{1}{2}-X\right)dX \quad (2.4.6)$$

$$=\left(\frac{X}{2}-X^2\right)\Big|_{\frac{1}{4}}^{\frac{1}{2}} \quad (2.4.7)$$

$$=\frac{1}{8}-\frac{1}{8}+\frac{1}{32} \quad (2.4.8)$$

$$=\frac{1}{32} \quad (2.4.9)$$

So,req answer is

$$2\times\Pr\left(X+\frac{1}{4}<Y<\frac{3}{4},\frac{1}{4}<X<\frac{1}{2}\right)=\frac{1}{16}$$

Ans is $\frac{1}{16}$

3 2016

3.1. Let the random variable X have the distribution $P(X=0)=P(X=3)=p$, $P(X=1)=1-3p$ for $0\le p\le\frac{1}{2}$. What is the maximum value of V(X)?

A) 3
B) 4
C) 5
D) 6
E) none

Solution: Given, for $0\le p\le\frac{1}{2}$,

$$P(X=0)=p \quad (3.1.1)$$

$$P(X=1)=1-3p \quad (3.1.2)$$

$$P(X=3)=p \quad (3.1.3)$$

Now consider $P(X=1)=1-3p$ for $p=\frac{1}{2}$. We get,

$$P(X=1)=1-3p \quad (3.1.4)$$

$$=1-(3)\left(\frac{1}{2}\right) \quad (3.1.5)$$

$$=1-\frac{3}{2} \quad (3.1.6)$$

$$=-\frac{1}{2}<0 \quad (3.1.7)$$

Probability cannot be negative. But in equation (0.0.7) probability is negative, which is not possible.
Therefore, the question is not a proper one.

Answer : Option E

3.2. X_1 and X_2 are independent Poisson variables such that $\Pr(X_1=2)=\Pr(X_1=1)$ and $\Pr(X_2=2)=\Pr(X_2=3)$. What is the variance of (X_1-2X_2) ?

a) 14
b) 4
c) 3
d) 2

Solution: For a Poisson variable X,

$$\Pr(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \quad (3.2.1)$$

Since $\Pr(X_1=2)=\Pr(X_1=1)$,

$$\frac{{\lambda_1}^2 e^{-\lambda_1}}{2!}=\frac{{\lambda_1}^1 e^{-\lambda_1}}{1!} \quad (3.2.2)$$

$$\lambda_1=2!/1!=2 \quad (3.2.3)$$

Similarly, as $\Pr(X_2=2)=\Pr(X_2=3)$,

$$\frac{{\lambda_2}^2 e^{-\lambda_2}}{2!}=\frac{{\lambda_2}^3 e^{-\lambda_2}}{3!} \quad (3.2.4)$$

$$\lambda_2=3!/2!=3 \quad (3.2.5)$$

Also we know for a Poisson variable X, the following holds true:

$$\mathrm{E}[X]=\lambda \quad (3.2.6)$$

$$\mathrm{Var}[X]=\lambda \quad (3.2.7)$$

$$\mathrm{Var}[X]=\mathrm{E}[X^2]-(\mathrm{E}[X])^2 \quad (3.2.8)$$

Now, for the variance of (X_1-2X_2)

$$\begin{aligned}\mathrm{Var}[X_1-2X_2]&=\mathrm{E}[(X_1-2X_2)^2]-(\mathrm{E}[X_1-2X_2])^2\\&=\mathrm{E}[X_1^2+4X_2^2-4X_1X_2]\\&-(\mathrm{E}[X_1]-2\mathrm{E}[X_2])^2\\&=\mathrm{E}[X_1^2]-(\mathrm{E}[X_1])^2+4\mathrm{E}[X_2^2]\\&-4(\mathrm{E}[X_2])^2)+4\mathrm{E}[X_1X_2]\\&+4\mathrm{E}[X_1]\mathrm{E}[X_2] \quad (3.2.9)\end{aligned}$$

Since the variables are independent:

$$\mathrm{E}[X_1X_2]=\mathrm{E}[X_1]\mathrm{E}[X_2] \quad (3.2.10)$$

Substituting equations (3.2.7) and (3.2.8), we get:

$$\begin{aligned}\mathrm{Var}[X_1 - 2X_2] &= \mathrm{Var}[X_1] + 4(\mathrm{Var}[X_2]) \\ &\quad - 4\mathrm{E}[X_1][X_2] + 4\mathrm{E}[X_1][X_2] \\ &= \lambda_1 + 4\lambda_2 = 2 + 4(3) = 14\end{aligned} \tag{3.2.11}$$

Hence option (a) 14 is correct.