PES UNIVERSITY, Bangalore
In-Semester Assessment
UE17CS203 - Introduction to Data Science

Assignment-1 Solution Set

1. Let *X* and *Y* be Bernoulli random variables. Let *Z* = *X* + *Y*.
a) Show that if *X* and *Y* cannot both be equal to 1, then *Z* is a Bernoulli random variable.
Soln: Let X and Y be Bernoulli random variables. Let Z = X+Y.
Case 1:
If X=1 and Y=0,
Z = X+Y
Z = 1+0
Z = 1
Case 2:
If X=0 and Y=1,
Z = X+Y
Z = 0+1
Z = 1
So in both the cases,
Z = 1
The random variable X is said to have the Bernoulli distribution when random event results in the success, then X=1. Otherwise X=0.
So from the definition we can say that Z also a Bernoulli random variable since value of Z=1.
Hence proved.

b) Show that if *X* and *Y* cannot both be equal to 1, then *pZ* = *pX* + *pY* .
Soln: Pz = Px+Py
From the addition rule of probability, we have
P(X=1 OR Y=1) = P(X=1) + P(Y=1) + P(X=1 AND Y=1) ………(1)
Since X and Y cannot both be equal to 1.Hence, P(X=1 AND Y=1) = 0

We can rewrite equation (1) as,
P(X=1 OR Y=1) = P(X=1) + P(Y=1) …….(2)
Renaming → Pz = P(X=1 OR Y=1) , Px= P(X=1), Py= P(Y=1)

Substitute above values in equation (2), we have
Pz = Px+Py
Hence proved.

c) Show that if *X* and *Y* *can* both be equal to 1, then *Z* is not a Bernoulli random variable..
Soln: X =1 and Y=1
Z=X+Y
Z=2
Since bernoulli random variable can be only 0 or 1, Z is not a bernoulli random variable.

2. Two dice are rolled. Let $X = 1$ if the dice come up doubles and let $X = 0$ otherwise. Let $Y = 1$ if the sum is 6, and let $Y = 0$ otherwise. Let $Z = 1$ if the dice come up both doubles and with a sum of 6 (that is, double 3), and let $Z = 0$ otherwise.
a) Let pX denote the success probability for X. Find pX.
b) Let pY denote the success probability for Y. Find pY.
c) Let pZ denote the success probability for Z. Find pZ.

Soln:

X	P(x)
0	30/36
1	6/36

Y	P(y)
0	31/36
1	5/36

Z	P(z)
0	35/36
1	1/36

d) Are X and Y independent?
Soln: X and Y are not independent.

e) Does $pZ = pX * pY$?
Soln: No, $P(Z = 1) \neq P(X = 1)P(Y = 1)$.

f) Does $Z = XY$?
Soln: No

3. Find the probability mass function of the random variable X if X □ Bin$(10, 0.4)$.
Find $P(X = 5)$.
Soln: $P(X = x) = C^n_x * (p^x) * (1 – p)^{(n-x)}$
$= 0.20$

4. A fair die is rolled eight times. Find the probability that no more than 2 sixes come up.
Soln: Each roll of the die is a Bernoulli trial with success probability 1/6. Let X denote the number of sixes in 8 rolls. Then $X \sim \text{Bin}(8, 1/6)$. We need to find $P(X \leq 2)$. Using the probability mass function,
$P(X \leq 2) = P(X = 0 \text{ or } X = 1 \text{ or } X = 2)$
$= P(X = 0) + P(X = 1) + P(X = 2)$
$= 0.2326 + 0.3721 + 0.2605$
$= 0.8652$

5. A large industrial firm allows a discount on any invoice that is paid within 30 days. Of all invoices, 10% receive the discount. In a company audit, 12 invoices are sampled at random. What is the probability that fewer than 4 of the 12 sampled invoices receive the discount?
Soln: Let X represent the number of invoices in the sample that receive discounts. Then $X \sim \text{Bin}(12, 0.1)$. The probability that fewer than four invoices receive discounts is $P(X \leq 3)$. Hence putting $n = 12$, $p = 0.1$, and $x = 3$. Therefore, $P(X \leq 3) = 0.974$

Note: Sometimes the best way to compute the probability of an event is to compute the probability that the event does not occur, and then subtract from 1.

6. What is the probability that more than 1 of the 12 sampled invoices receives a discount?
Soln: Let X represent the number of invoices in the sample that receive discounts. We wish to compute the probability $P(X > 1)$.Therefore we note that $P(X > 1) = 1 - P(X \leq 1)$.
Hence putting $n = 12$, $p = 0.1$, $x = 1$, we find that $P(X \leq 1) = 0.659$.
Therefore, $P(X > 1) = 1 - 0.659 = 0.341$.

7. Let *X* □ Bin*(n, p)*, and let *Y* = *n* − *X*. Show that *Y* □ Bin*(n,* 1 − *p)*
Soln:

$$P(x=k) = {}^nC_k * p^k * (1-p)^{n-k} \quad \text{......} \; (1)$$

$$P(y=k) = P((n-x)=k)$$

$$= p(x=(n-k))$$

$$= {}^nC_{n-k} * p^{n-k} * (1-p)^k$$

Rearranging & replacing

$$= {}^nC_k * (1-p)^k * p^{n-k} \quad \text{......} \; (2)$$

From (1) & (2), we conclude if $X \sim Bin(n,p)$,
$\Rightarrow Y \sim Bin(n, 1-p)$

8. Grandma bakes chocolate chip cookies in batches of 100. She puts 300 chips into the dough. When the cookies are done, she gives you one. What is the probability that your cookie contains no chocolate chips? Find the mean and variance
Soln: This is another instance of particles in a suspension. Let X represent the number of chips in your cookie. The mean number of chips is 3 per cookie, so X ~ Poisson(3).
It follows that $P(X = 0) =(e^{-3})*(3^0) /0! = 0.0498$.
Mean = Variance = 3

9. A test of weld strength involves loading welded joints until a fracture occurs. For a certain type of weld, 80% of the fractures occur in the weld itself, while the other 20% occur in the beam. A number of welds are tested. Let *X* be the number of tests up to and including the first test that results in a beam fracture. What is the distribution of *X*? find P(X=3)

Soln: Each test is a Bernoulli trial, with success defined as a beam fracture. The success probability is therefore p = 0.2. The number of trials up to and including the first success has a geometric distribution with parameter p = 0.2. Therefore X ~ Geom(0.2).
The event X = 3 occurs when the first two trials result in failure and the third trial results in success. It follows that
P(X = 3) = P(FFS)
= (0.8)(0.8)(0.2)
= 0.128

10. of customers ordering a certain type of personal computer, 20% order an upgraded graphics card, 30% order extra memory, 15% order both the upgraded graphics card and extra memory, and 35% order neither. Fifteen orders are selected at random. Let
*X*1*, X*2*, X*3*, X*4 denote the respective numbers of orders in the four given categories.
a. Find *P(X*1 = 3*, X*2 = 4*, X*3 = 2*,* and *X*4 = 6*)*.
Soln: $(15!/(3!4!2!6!)) * 0.2^3 * 0.3^4 * 0.15^2 * 0.35^6 = 0.0169$

b. Find *P(X*1 = 3*)*.
Soln: $(15!/(3!12!)) * 0.2^3 * 0.8^{12} = 0.250$