INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: Probability and Statistics

Chapter: Discrete Distributions

Category: Practice question – Extras

1. A discrete random variable XXX takes the values 0,1,2, 3 with probabilities: $P(X=0)=0.1$, $P(X=1)=0.3$, $P(X=2)=0.4$, $P(X=3)=0.2$

A. Verify that this defines a valid probability distribution.

B. Find the expected value E[X]

C. Find the variance Var(X)

Solutions:

A. For the validity of distribution, two conditions must be satisfied

All $0 < p < 1$, this is met here

$\sum p = 1$

$0.1 + 0.3 + 0.4 + 0.2 = 1$, this is also met here, so this is a valid probability distribution.

B. $E(X) = \sum XP(X = x)$

X	P(X=x)	XP(X=x)
0	0.1	0
1	0.3	0.3
2	0.4	0.8
3	0.2	0.6
Total	1	1.7

$E(X) = 1.7$

C. $V(X) = E(X^2) - [E(X)]^2$

$E(X^2) = \sum X^2 P(X = x)$

X	P(X=x)	X^2P(X=x)
0	0.1	0
1	0.3	0.3
2	0.4	1.6
3	0.2	1.8
Total	1	3.7

$V(X) = 3.7 - 1.7^2 = 0.81$

2. A company receives phone calls according to the following probability distribution for calls per hour (X):

P(X=0)=0.25, P(X=1)=0.35, P(X=2)=0.25, P(X=3)=0.15

Find the probability that the company receives at most 1 call in an hour.

Solution:

At most 1 in an hour means 1 or less than one

P(X <= 1)= P(x= 0) + P(X =1)

= 0.25 + 0.35

= 0.7

Discrete Uniform Distribution

1. A uniform distribution function has been defined as:

X={1,3,5,7,9}

What is the value of F(5) ?

Solution:

The probability of any point = 1/k, ...where 'k' is the outcomes

=1 / 5 = 0.2

F(5) is the cumulative probability of an outcome less than or equal to 5. Therefore,

F(5)=P(X≤5)

=P(X=1)+P(X=3)+P(X=5)

=0.2+0.2+0.2=0.6

2. A telephone number is selected at random from a directory. Suppose X denote the last digit of selected telephone number. Find the probability that the last digit of the selected number is

a. 6

b. less than 3

c. greater than or equal to 8

d. E(X) and V(X)

Solution

Let X denote the last digit of randomly selected telephone number. The possible values of X are $0, 1, 2, \cdots, 9$.

All the numbers $0, 1, 2, \cdots, 9$ are equally likely. Thus the random variable X follows a discrete uniform distribution $U(0, 9)$. The probability mass function of X is

P(X =x) = 1/10 , since there are 10 digits in 0 to 9

a. P(X = 6) = 1/10 = 0.1

b. The probability that the last digit of the selected telecphone number is less than 3

$$\boxed{\begin{aligned} P(X < 3) &= P(X \le 2) \\ &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{10} + \frac{1}{10} + \frac{1}{10} \\ &= 0.1 + 0.1 + 0.1 \\ &= 0.3 \end{aligned}}$$

c. The probability that the last digit of the selected telecphone number is greater than or equal to 8

$$\begin{aligned} P(X \ge 8) &= P(X = 8) + P(X = 9) \\ &= \frac{1}{10} + \frac{1}{10} \\ &= 0.1 + 0.1 \\ &= 0.2 \end{aligned}$$

d. E(X) = $\frac{k+1}{2}$

= 5.5

V(X) = $\frac{k^2-1}{12}$

= 8.25

3. Let when a fair die is rolled once. If before the die is rolled you are offered either (1/3.5) dollars or h(X) = 1/X dollars would you accept the guaranteed amount or would you gamble? [Note: It is not generally true that 1/E(X) = E(1/X).]

Discrete

PRACTICE QUESTION – Extras

Solution:
The guaranteed amount is 1/3.5 = 0.2857

The expected amount from gamble
E(gamble profit) = E(h(X))
= E(1/X)
Since this is discrete uniform we have
P(X = x) =1/6 ,   since there are 6 possible outcomes
= $\sum \frac{1}{X} P(X = x)$
= $\sum \frac{1}{X}$ (1/6)
= 1/6 *$(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6})$
= 49/120
= 0.40833

Since expected gamble profit is more than the guaranteed, we should accept the gamble.

Bernoulli and Binomial Distribution

1. Find the probability of getting five heads and seven tails in 12 flips of a balanced coin.
Solution:
X : no. Of heads on the flips
The probability of head on every flip of balanced coin = 0.5
X ~ Bin(n = 12, p = 0.5)
P(X =x) = $^{12}C_x$ $(0.5)^x$ $(0.5)^{(12-x)}$
P(X = 5) = $^{12}C_5$ $(0.5)^5$ $(0.5)^{(12-5)}$ = 0.1934

2. Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let X denote the number among 15 randomly selected copies that fail the test.
Then X has a binomial distribution with n = 15 and p = 0.2 .
a. The probability that at most 8 fail the test is
b. The probability that exactly 8 fail is
c. Finally, the probability that between 4 and 7, inclusive, fail is

Solution:

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

egP(X = 0) = $^{15}C_0\ (0.2)^0\ (0.8)^{(15-0)}$ = 0.0352

X	P(X=x)
0	0.0352
1	0.1319
2	0.2309
3	0.2501
4	0.1876
5	0.1032
6	0.0430
7	0.0138
8	0.0035
9	0.0007
10	0.0001
11	0.0000
12	0.0000
13	0.0000
14	0.0000
15	0.0000
total	1.0000

a. The probability that at most 8 fail the test is
   P(X <= 8) = P(X =0) + P(X=1)..+P(X =8)
   = 0.9992

b. The probability that exactly 8 fail is
   P(X= 8) = 0.0035

c. the probability that between 4 and 7, inclusive, fail is
   P(4 <= X <= 7) = P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)
   = 0.3476

Discrete

PRACTICE QUESTION – Extras

3. A particular telephone number is used to receive both voice calls and fax messages. Suppose that 25% of the incoming calls involve fax messages, and consider a sample of 25 incoming calls. What is the probability that

a. At most 4 of the calls involve a fax message?

b. Exactly 6 of the calls involve a fax message?

c. At least 2 of the calls involve a fax message?

d. More than 5 of the calls involve a fax message?

Solution:

This is binomial distribution since there are two possible outcomes , either calls have fax message or they don't, also each times it being a fax message has the same probability of 0.25. Each call's status is independent of the other.

X ~ Bin(n = 25, p = 0.25)

$P(X = x) = {}^{25}C_x\,(0.25)^x\,(0.75)^{(25-x)}$

eg$P(X = 0) = {}^{15}C_0\,(0.2)^0\,(0.8)^{(15-0)} = 0.0008$

X	P(X=x)
0	0.0008
1	0.0063
2	0.0251
3	0.0641
4	0.1175
5	0.1645
6	0.1828

a. At most 4 of the calls involve a fax message?

$P(X <= 4) = P(X = 0) + \ldots\ldots P(X= 4)$

$= 0.2138$

b. Exactly 6 of the calls involve a fax message?

$P(X = 6) = 0.1828$

c. At least 2 of the calls involve a fax message?

AT least 2 = $P(X >=2)$

$= P(X= 2) + \ldots.P(X=25)$

$=1 – P(X < 2)$

$=1 – [P(X =0) + P(X =1)]$

$=1 – 0.0071 = 0.9929$

d. More than 5 of the calls involve a fax message?
P(X > 5) = P(X = 6) + P(X = 7) + P(X=8)…P(X = 25)
= 1- P(X <= 5)
=1 – 0.3783
= 0.6217

4. A company that produces fine crystal knows from experience that 10% of its goblets have cosmetic flaws and must be classified as “seconds.” A random selection of 10 goblets is done.

a. Find the mean and SD

Solution:
X ~ Bin(n = 10, p = 0.1)
Mean = np = 1
Var = np(1-p) = 10*0.1*0.9 = 0.9
SD = $(Var)^{0.5}$ = 0.9487

Poisson distribution:

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

1. A certain kind of sheet metal has, on the average, five defects per 10 square feet, that is, it follows a poisson distribution with mean 5 per 10 square feet.

a. What is the probability that a 10-square-foot sheet of the metal will have at least 2 defects?

b. What is the probability that a 20-square-foot sheet of the metal will have at least 4 defects?

Solution:
a. X ~ Poi(λ = 5) per 10 sq feet
Since mean of poisson is λ
What is the probability that a 10-square-foot sheet of the metal will have at least 2 defects?
P(X >=2) = P(X = 2) + P(X= 3) + ……….
=1 – P(X < 2)
= 1- [P(X =0) + P(X = 1)]
= 1 – ($\frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!}$)

=1 – (0.0067 + 0.0337)
= 0.0404

b. What is the probability that a 20-square-foot sheet of the metal will have at least 4 defects

X ~ Poi(λ = 5) per 10 sq feet

X ~ Poi(λ = 5*2) per 20 (10*2) sq feet

λ = 10

P(X >= 4) = 1 – P(X < 4)

P(X =x) = $\frac{10^x e^{-10}}{x!}$

X	P(X=x)
0	0.0000
1	0.0005
2	0.0023
3	0.0076

P(X >= 4) = 1 – [P(X = 0) + P(X = 1) + P(X = 2) + P(X =3)]
=1 - 0.0103
= 0.9897

2. Customers arrive at a coffee shop according to a Poisson process with a mean of 3 arrivals per 10 minutes.

a. What is the probability that exactly 5 customers arrive in 10 minutes?
b. What is the probability that no customers arrive in 10 minutes?
c. Find the expected number and variance of arrivals in 15 minutes.

Solution:

a. X ~ Poi(λ = 3) per 10 minutes

P(X =x) = $\frac{3^x e^{-3}}{x!}$

P(X = 5) = $\frac{3^5 e^{-3}}{5!}$

= 0.1008

b. No customers

P(X= 0)= $\frac{3^0 e^{-3}}{0!}$

= 0.0498

c. $E(X) = \lambda$
For 1 min = 3/10
For 15 mins λ = 3/ 10 *15 = 4.5
E(X) = 4.5
V(X) = 4.5

3. A call center receives on average 12 calls per hour. Let X be the number of calls in one hour.

a. Find P(X=10)
b. Find P(X≤8)
c. What are the mean and variance of X for calls per 20 mins?

Solution:
Since mean calls per hour is given, we can assume poisson distribution
X : no. of calls per hour
X ~ Poi(λ = 12) per hour

a. P(X=10) = $\frac{12^{10}e^{-12}}{10!}$ = 0.1048

b. P(X <=8) = P(X= 0) + P(X= 1) ...+P(X = 8)

X	P(X=x)
0	0.0000
1	0.0001
2	0.0004
3	0.0018
4	0.0053
5	0.0127
6	0.0255
7	0.0437
8	0.0655

P(X <=8) = 0.845

c. Calls per minutes = 12/60
Calls per 20 mins = 12/60 * 20 = 4
E(X) = V(X) = 4 per 20 mins

Geometric Distribution: Type 1

X: no. Of trials on which the first success occurs / until the first success occurs

$P(X = x) = p(1-p)^{x-1}$, $x = 1,2,3, \dots$

1. We have 'X' the number of tosses until the first head of a biased coin with probability of head = 0.2.
   a. What is the probability of making 5 tosses?
   b. What is the probability of making at least 3 tossses ?

Solution:

a. $P(X = x) = 0.2(0.8)^{x-1}$

$P(X = 5) = 0.2(0.8)^{5-1}$

$= 0.08192$

b. At least 3 tosses

$P(X >= 3) = P(X= 3) + P(X=4) + \dots$

$=1 - P(X < 3)$

$=1 – [P(X = 1) + P(X = 2)]$

$= 1- 0.36$

$=0.64$

2. An instructor feels that 15% of students get below a C on their final exam. She decides to look at final exams (selected randomly and replaced in the pile after reading) until she finds one that shows a grade below a C. We want to know the probability that the instructor will have to examine at least ten exams until she finds one with a grade below a C.

Solution:

$P(X = x) = 0.15 (0.85)^{x -1}$

$P(X >= 10) = P(X=10) + \dots.$

$=1 – P(X < 10)$

$=1 – [P(X =1) + ..P(X = 9)]$

X	P(X=x)
1	0.1500
2	0.1275
3	0.1084
4	0.0921

5	0.0783
6	0.0666
7	0.0566
8	0.0481
9	0.0409

=1 – 0.7684

=0.2316

3. You need to find a store that carries a special printer ink. You know that of the stores that carry printer ink, 10% of them carry the special ink. You randomly call each store until one has the ink you need.

a. What are p and q?

b. Mean and variance

Solution:

a. 'p' probability of success that printer carries the special ink

p = 0.10

q = 1- p = 0.9

b. Mean = 1/ p = 10

Var(X) = $(1-p)/p^2$

= q/p^2

= $0.9/0.1^2$

= 90

4. The mean of a geometric distribution is given as 20. Let 'X' denote the number of trials until the first success.

a. Find the 'p'

b. Find the probability of X = 3

Solution:

a. Mean =1/p

p = 1 / mean = 1 / 20

= 0.05

b. P(X = 3) = $0.05*(0.95)^{3-1}$

= 0.0451

Discrete

PRACTICE QUESTION – Extras

Geometric Distribution: Type 2

Y: denotes the number of failures before the first success

$P(Y=y) = p(1-p)^y$, $y = 0, 1, 2, \ldots\ldots.$

1. The probability of a defective steel rod is 0.01. Steel rods are selected at random. 'Y' is the number of rods to be examined before finding the first defective rod. What is the probability of examining 4 rods before the first defective?

Solution:

$Y \sim Geo(p = 0.01)$

$P(Y = y) = 0.01(0.99)^y$

$P(Y = 4) = 0.01(0.99)^4$

$= 0.0096$

2. The lifetime risk of developing pancreatic cancer is about one in 78 (1.28%). Let Y = the number of people you ask before one says he or she has pancreatic cancer. Then Y is a discrete random variable with a geometric distribution: $Y \sim Geo(p = 1/78 = 0.0128)$

a. What is the probability of that you ask ten people before one says he or she has pancreatic cancer?

b. Find the (i) mean and (ii) standard deviation of Y.

Solution:

a. $P(Y=10) = 0.0128\ (1-0.0128)^{10}$

$= 0.0113$

b. Mean = $q/p = (1-0.0128) / 0.0128$ or $(1-1/78)/(1/78)$

$= 77$

SD = $\sqrt{\frac{q}{p^2}}$

$= 77.50$

Negative Binomial : type 1

X: is the number of trials on which the ‘r^{th} success occurs

$X \sim Neg(p,r)$, $X = r, r+1, \ldots$

$P(X = x) = {}^{x-1}C_{r-1}\, p^r\, (1-p)^{x-r}$

1. A marketing executive calls potential customers. The probability of a successful sale on any call is 0.2. The calls will be made until the 3^{rd} sale. Hence this $X \sim Neg(p= 0.2, r =3)$
What is the probability that the executive makes their 3rd sale on their 15th call?

Solution:

$P(X = 15) = {}^{14}C_2\, 0.2^3\, (1-0.2)^{12}$

= 0.05002

2. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil.

a. What is the probability that the third strike comes on the seventh well drilled?

b. What is the mean and variance of the number of wells that must be drilled if the oil company wants to set up three producing wells?

Solution:

a. $X \sim Neg(0.20, 3)$

$P(X =7) = {}^{6}C_2\, 0.2^3\, (1-0.2)^4$

= 0.049

b.

Solution

The mean number of wells is:

$$\mu = E(X) = \frac{r}{p} = \frac{3}{0.20} = 15$$

with a variance of:

$$\sigma^2 = Var(x) = \frac{r(1-p)}{p^2} = \frac{3(0.80)}{0.20^2} = 60$$

3. If the probability is 0.40 that a child exposed to a certain contagious disease will catch it, what is the probability that the 11th child exposed to the disease will be the fourth to catch it?

Solution :

$X \sim Neg(0.40, 4)$

$P(X = 11) = {}^{10}C_3\, 0.4^4\, (0.6)^7$

$= 0.086$

Negative Binomial : type 2

Y: is the number of trials before which the 'rth success occurs

$Y \sim Neg(p,r)$ $y = 0, 1, 2, \ldots\ldots.$

$P(Y = y) = {}^{y+r-1}C_{r-1}\, p^r\, (1-p)^y$

1. The probability is 0.75 that an applicant for a driver's license will pass the road test on any given try. To acquire the license the person will have to pass on 2 tests, what is the probability that an applicant will have to give 4 tries before he passes on the final one?

Solution:

Y: no. of failures before the 2nd success

$Y \sim Neg(0.75, 2)$

$P(Y = 4) = {}^5C_1\, 0.75^2\, (0.25)^4$

$= 0.011$

2. In Major League Baseball's Home Run Derby, each contestant is allowed to keep swinging the bat until they have made 3 "outs". (An "out" is anything that is not a home run.) If Barry Bonds has a 70% chance of hitting a home run on any given swing,

a. what is the probability that he hits at least 5 home runs before his turn is up?

b. Mean and Standard deviation

Solution:

Y: no. of failures(home runs) before the 3rd out

a. $Y \sim Neg(0.3, r = 3)$

$P(Y = y) = {}^{y+3-1}C_2\ 0.3^3\ (0.7)^y$
$P(Y >= 5) = 1 – P(X < 5)$

Y	P(Y=y)
0	0.0270
1	0.0567
2	0.0794
3	0.0926
4	0.0972

= 0.6471

b. Mean = rq/p
= 3*0.7/0.3
= 7

SD = $\sqrt{\frac{rq}{p^2}}$
= 4.8305

Hypergeometric distribution:
X~ H(n,N,M)
Where n : sample size
N : population size
M : number of success in the population
$P(X = x) = \frac{{}^{M}_{x}C * {}^{N-M}_{n-x}C}{{}^{N}_{n}C}$

1. As part of an air-pollution survey, an inspector decides to examine the exhaust of 6 of a company's 24 trucks. If 4 of the company's trucks emit excessive amounts of pollutants, what is the probability that none of them will be included in the inspector's sample?

Solution:
Substituting x = 0, n = 6, N = 24, and M = 4 into the formula for the hypergeometric distribution, we get

$$h(0; 6, 24, 4) = \frac{\binom{4}{0}\binom{20}{6}}{\binom{24}{6}}$$
$$= 0.2880$$

2. Among the 120 applicants for a job, only 80 are actually qualified. If 5 of the applicants are randomly selected for an in-depth interview, find the probability that only 2 of the 5 will be qualified for the job by using the formula for the hypergeometric distribution;

Solution:

Substituting $x = 2$, $n = 5$, $N = 120$, and $M = 80$ into the formula for the hypergeometric distribution, we get

$$h(2; 5, 120, 80) = \frac{\binom{80}{2}\binom{40}{3}}{\binom{120}{5}}$$
$$= 0.164$$

3. Five individuals from an animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into the population. After they have had an opportunity to mix, a random sample of 10 of these animals is selected. Let of tagged animals in the second sample. If there are actually 25 animals of this type in the region, what is the probability that

a. P(X = 2)

b. P(X<= 2)

c. Mean and SD

Solution:

The parameter values are $n = 10$, $M = 5$ (5 tagged animals in the population) and $N = 25$, so

$$h(x; 10, 5, 25) = \frac{\binom{5}{x}\binom{20}{10-x}}{\binom{25}{10}} \quad x = 0, 1, 2, 3, 4, 5$$

For part (a),

$$P(X = 2) = h(2; 10, 5, 25) = \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} = .385$$

For part (b),

$$P(X \le 2) = P(X = 0, 1, \text{ or } 2) = \sum_{x=0}^{2} h(x; 10, 5, 25)$$
$$= .057 + .257 + .385 = .699$$

c. Mean = nM/N
= 10*5/25
= 2

SD = $\text{Var}^{0.5}$

Var = $\frac{N-n}{N-1} * n * \frac{M}{N} * (1 - M/N)$

= 1

SD = 1