Random Variable And Probability Distribution

Introduction

Random Variable *(r.v.)*

Is defined as a real valued function defined on the sample space S. We denote it as X , Y , Z , T , … and denote the assumed values to be taken : x , y , z , t ,….. .

There are two different types of the random variables: Discrete r.v. and continuous r.v. as we mentioned before.

The Probability Distribution:

If the random variable X is discrete takes the assumed values :

X : $x_1 , x_{2,} \ x_3 , \ldots \ldots , x_n$, we define the probability distribution (or the probability mass) function of the random variable X as $f(x)$ such that :

$f(x_j) = P(X = x_j) \quad ; \ j = 1, 2, 3, \ldots \ldots, n$ or simply, we may write:

$$f(x) = P(X = x) \quad for \ any \ x$$

Which satisfies: $(1) \ f(x) = P(X = x) \geq 0 \quad for \ any \ real \ x \qquad (2) \sum_{all \ x} f(x) = 1$

Note: If the function f(x) does not satisfies any of the two conditions, it will be considered as a usual algebraic function.

Example (1):

Determine the value C so that each of the following functions can serve as a probability distribution function of the discrete random variable X:

(A) $f(x_j) = C(x^2 + 4) \quad ; x = 0,1, 2, 3 ;$

(B)

X	-2	0	3	4
f(x)	0.15	0.3	C	0.2

Solution:

(1) We note that C should be nonnegative; according to the first condition, and by using the second condition, we get

1 = f(0) + f(1) + f(2) + f(3) =C(4 + 5 + 8 + 13) , which gives C = 0.35

(B) Also; here we may say that C should be nonnegative value; according to the first condition and by using the second condition, we get

1 = 0.15 + 0.3 + C + 02 which gives C = 0.35

Cumulative Distribution Function (CDF):

If f(x) represents the probability distribution function of the random variable X , we define its Cumulative Distribution Function (CDF) such as:

$$\text{CDF} = F\ (x) = P(\ X \le x\ ) = \sum_{all\, X \le x} f(\ x\ )\ \ (discrete)\ or \int_{-\infty}^{x} f(t)\ dt \quad (continuous)$$

For instance; in example (1) part (B):

x	$x < -2$	$-2 \le x < 0$	$0 \le x < 3$	$3 \le x < 4$	$4 \le x$
$F(x)$	0	0.15	0.45	0.8	1

The Mathematical Expectation of a random variable:

If f(x) represents the probability distribution function of the random variable X, we define the mathematical Expectation (Or the expected value Or the mean value) of a random variable X as $\sum_{all\,x} x.f(x)$ and denote it as μ Or E(X). That is: μ = $E(X) = \sum_{all\,x} x.f(x)$ .and

satisfies the following properties:

(1)E (a) = a ; where a is any real constant,

(2)E (a X) = a E(X) for any real constant a ,

(3) $E(aX \pm b) = aE(X) \pm b, \quad where\ a\ and\ b\ are\ any\ arbitrary\ cons\tan ts$

(4) $E(g(x)) = \begin{cases} \sum_{all\ x} g(x)\cdot f(x) & (for\, discrete\ X) \\ \int_{-\infty}^{\infty} g(x)\cdot f(x)\ dx & (for\, continuous\ X) \end{cases}$

For instance; in example (1) part (B), we may write:

The mean value of the given random variable X will be:

μ= $E(X) = \sum_{all\,x} x.f(x) = (-2)(0.15) + (0)(0.3) + (3)(0.35) + (4)(0.2) = 1.55$

Also E (4 X+ 0.5) = 4 E(X) + 0.5 = 4 (1.55) + 0.5 = 6.7.

While, if g(x) = x^2 then

$E(X^2) = \sum x^2 f(x) = (-2)^2 (0.15) + (0)^2 (0.30) + (3^2)(0.35) + (4^2)(0.2) = 6.95$

While $E(0.8\ X^2) = (0.8)E(X^2) = (0.8)(6.95) = 5.56$.

<u>The Variance and Standard deviation of the random variable</u>

If f(x) represents the probability distribution function of the random variable X with finite mean μ , we define variance of a random variable X as:

$E(X-\mu)^2 = \sum_{all\,x} (X-\mu)^2.f(x)\ and\ we\ denote\ it\ as\ V(x)\ Or\ Var(x)\ Or\ \sigma^2$, So

$$V(x) = Var(x) = \sigma^2 = \begin{cases} \sum_{all\,x} (X-\mu)^2.f(x) & (for\ discrete\ X) \\ \int_{-\infty}^{\infty} (X-\mu)^2.f(x)\ dx & (for\ continuous\ X) \end{cases}$$

And to satisfy the following properties:

(1) $V(a) = \sigma_a^2 = Var(a) = 0$ (for a real constant a)

(2) $V(a\,x) = \sigma_{ax}^2 = Var(a\,x) = a^2 V(x) = a^2 \sigma_x^2$

(3) $V(a\,x \pm b) = \sigma_{ax\pm b}^2 = Var(a\,x \pm b) = a^2 V(x) = a^2 \sigma_x^2$ (for real constants a and b)

The standard deviation of the random variable X is the positive root square $\sigma = \sqrt{\sigma^2}$

Remark: We deduce from the variance definition that $\sigma^2 = V(x) = Var(x) = E(X^2) - \mu^2$:

$$E(X^2) = \sum_{all\ x} X^2 \cdot f(x)\ \ (for\ discrete)\ or\ \int_{-\infty}^{\infty} X^2 \cdot f(x)\ dx \quad (for\ continuous)$$

Again for instance; in example (1) part (B), we may write:

$$E(X^2) = \sum_{all\,x} x^2 . f(x) = (-2)2(0.15) + (0)2(0.3) + (3)2(0.35) + (4)2(0.2) = 6.95$$

$$\sigma^2 = E(X^2) - \mu^2 = 6.95 - (1.55)2 = 4.5475 \Rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{4.5475} = 2.13$$

For the **Continuous** random variable case we have:

The function f(x) is said to be a probability distribution (Or the probability density; pdf,) function of the random variable X if it satisfies:

(i) $f(x) \geq 0$ for any real x ; ($-\infty < x < \infty$)

(ii) $\int_{-\infty}^{\infty} f(x)\ dx = 1$ (The total area under the curve of $f(x)$ is exact value of one)

We may define the mathematical Expectation (Or the expected value or the mean value) of a random variable X as: $\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\ dx$

While the variance of the continuous random variable X is defined as:

$$\sigma^2 = V(x) = Var(x) = E(X - \mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = E(X^2) - \mu^2$$

EXERCISES

1. Let the random variable X having the probability distribution (p.m.f) as:

x	-3	6	9
f(x)	1/6	1/2	1/3

(A)Find the probability that: (i) The random variable X assumes a non-negative value.

(ii) The random variable X assumes a value less than 7.

(B)Find μ_x and σ_x^2 and then deduce each of μ_y, and σ_y^2; where $Y = 3X - 6$.

2. A large industrial firm purchases several word processors at the end of each year, The exact number depending of the frequency of repairs in the previous year. Suppose that the number of word processors, X that are purchased each year has The following probability distribution:

X	0	1	2	3
f(x)	1/10	3/10	2/5	1/5

(A)Find μ_x and σ_x^2

(B)If the cost of the desired model will remain fixed at 1200 dollars throughout this year and a discount of $50\ X^2$ dollars is credited toward any purchase, how much can this firm expect to spend on new word processors at the end of this year

3. Suppose that the number of cars X pass through a car wash between 4:00pm and 5:00 pm on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
f(x)	1/12	1/12	1/4	1/4	1/6	1/6

Let $g(x) = 2X - 1$ represent the amount of dollars, paid to the attendant by the manger. Find the attendant's expected earnings (and it variance) for this particular time period.

4. Let the random variable X represent the number of automobiles that are used for Official business purposes on a given workday. The probability distribution for

Company A is given as

X	1	2	3
f(x)	0.3	0.4	0.3

Company B is given as

X	0	1	2	3	4
f(x)	0.2	0.1	0.3	0.3	0.1

Show that the variance of the probability distribution for company B is greater than that for company A.

<u>Some Common Discrete Distributions</u>:

Let's to define what we call Bernoulli trials. A trial does mean a random experiment, if we repeat this random experiment (trial) for n times we can say that we have a set of n trials. This set of trials is said to be a set of Bernoulli trials if it satisfies the following (properties) conditions:

(1) All trials are independent ; i.e the outcome to be obtained from any trial is independent of that to be obtained from any other trial.

(2) In each trial there are <u>only</u> two possible outcomes to obtain one of them such as: patient or normal, male or female, defect or non-defect,…. we will denote them as (success , failure)

(3) In each trial, the probability of getting success is fixed (unchangeable) value , p .

If we define the random variable X such that:

X = No. of successes to be obtained within a set of n of Bernoulli trials

We may say that this random variable having what we call a Binomial probability distribution with the two parameters: n and *p* and write it as:

$X \sim Bin(n, p)$ And its pmf f(x) is defined as:

$$f(x) = P(X = x) = C_x^n p^x q^{n-x} \quad ; \quad x = 0,1,2,\ldots\ldots, n \quad and \quad p+q=1$$

Where : $C_x^n = \dfrac{n!}{x!\,(n-x)!}$; $n = positive\ integer$

Such that: $f(x) \geq 0$ & $\sum\limits_{all\ x} f(x) = \sum\limits_{all\ x} C_x^n p^x q^{n-x} = (p+q)^n = 1$

<u>Theorem</u>

If $X \sim Bin(n, p)$ then,

(A) $\mu = E(X) = n\,p$

(B) $\sigma^2 = V(x) = Var(x) = E(X-\mu)^2 = E(X^2) - \mu^2 = n\,p\,q$

<u>Example (2):</u>

Suppose that it is known that 30 % of a certain population are immune to some disease. If a random sample of size 10 persons to selected this population. Let X represents number of those who are immune to some disease within that selected sample.

(A)Write the probability distribution of X

(B) What the probability that will be

(i) Exactly four immune persons,

(ii) At least two immune persons

(C) What is expected number immune persons in this sample and its variance?

Solution:

(A) We note that $X \sim Bin(\, n = 10,\ p = 0.3\,)$, so we may write the pmf as

$$f(x) = P(X = x) = C_x^{10} (0.3)^x\ (0.7)^{10-x} \quad ; \quad x = 0,1,2, \ldots\ldots, n = 10$$

(B) Here

(i) P(exactly four immune persons) =

$$f(4) = P(X = 4) = C_4^{10} (0.3)^4\ (0.7)^6 = 0.2001$$

(ii) P(at least two immune persons) =

$$P(X \geq 2) = \sum_{x=2}^{10} C_x^{10} (0.3)^x\ (0.7)^{10-x} = f(2) + f(3) + \ldots f(10)$$

$$= 1 - [\ f(0) + f(1)\ ] = 1 - [\ 0.0282 + 0.1211] = 0.8507$$

EXERCISE

Q1. Suppose that 30% of the buildings in a certain city violate the building code. A building engineer randomly inspects a sample of 3 new buildings in the city.

(a) Find the probability distribution function of the random variable X representing the

(b) Number of buildings that violate the building code in the sample.

(c) Find the probability that:

(i) None of the buildings in the sample violating the building code.

(ii) One building in the sample violating the building code.

(iii) At least one building in the sample violating the building code.

(d) Find the expected number of buildings in the sample that violate the building code (E(X)).

(e) Find σ^2=Var(X).

Q2. A missile detection system has a probability of 0.90 of detecting a missile attack. If 4 detection systems are installed in the same area and operate independently, then

(i) The probability that at least two systems detect an attack is

(A) 0.9963 (B) 0.9477 (C) 0.0037 (D) 0.0523 (E) 0.5477

(ii) The average (mean) number of systems detect an attack is

(A) 3.6 (B) 2.0 (C) 0.36 (D) 2.5 (E) 4.0

Q3. Suppose that the probability that a person dies when he or she contracts a certain disease is 0.4. A sample of 10 persons who contracted this disease is randomly chosen.

(1) What is the expected number of persons who will die in this sample?

(2) What is the variance of the number of persons who will die in this sample?

(3) What is the probability that exactly 4 persons will die among this sample?

(4) What is the probability that less than 3 persons will die among this sample?

(5) What is the probability that more than 8 persons will die among this sample?

Q4. Suppose that the percentage of females in a certain population is 50%. A sample of 3 people is selected randomly from this population.

(a) The probability that no females are selected is

(A) 0.000 (B) 0.500 (C) 0.375 (D) 0.125

(b) The probability that at most two females are selected is

(A) 0.000 (B) 0.500 (C) 0.875 (D) 0.125

(c) The expected number of females in the sample is

(A) 3.0 (B) 1.5 (C) 0.0 (D) 0.50

(d) The variance of the number of females in the sample is

(A) 3.75 (B) 2.75 (C) 1.75 (D) 0.75

Q5. 20% of the trainees in a certain program fail to complete the program. If 5 trainees of this program are selected randomly,

(i) Find the probability distribution function of the random variable X, where:

X = number of the trainees who fail to complete the program.

(ii) Find the probability that all trainees fail to complete the program.

(iii) Find the probability that at least one trainee will fail to complete the program.

(iv)How many trainees are expected to fail completing the program?

(v) Find the variance of the number of trainees who fail completing the program.

Q6. In a certain industrial factory, there are 7 workers working independently. The probability of accruing accidents for any worker on a given day is 0.2, and accidents are independent from worker to worker.

(a) The probability that at most two workers will have accidents during the day is

(A) 0.7865 (B) 0.4233 (C) 0.5767 (D) 0.6647

(b) The probability that at least three workers will have accidents during the day is:

(A) 0.7865 (B) 0.2135 (C) 0.5767 (D) 0.1039

(c) The expected number workers who will have accidents during the day is

(A) 1.4 (B) 0.2135 (C) 2.57 (D) 0. 59

Q7. From a box containing 4 black balls and 2 green balls, 3 balls are drawn independently in succession, each ball being replaced in the box before the next draw is made. The probability of drawing 2 green balls and 1 black ball is:

(A) 6/27 (B) 2/27 (C) 12/27 (D) 4/27

Q8. The probability that a lab specimen is contaminated is 0.10. Three independent samples are checked.

1) The probability that none is contaminated is:

(A) 0.0475 (B) 0.001 (C) 0.729 (D) 0. 3

2) The probability that exactly one sample is contaminated is:

(A) 0.243 (B) 0.081 (C) 0.757 (D) 0. 3

Q9. If X~ Binomial (n,p), E(X)=1, and Var(X)=0.75, find P(X=1).

Q10. Suppose that X~ Binomial (3, 0.2). Find the cumulative distribution function (CDF) of X.

Q11. A traffic control engineer reports that 75% of the cars passing through a checkpoint are from Riyadh city. If at this checkpoint, five cars are selected at random.

(1) The probability that none of them is from Riyadh city equals to:

(A) 0.00098 (B) 0.9990 (C) 0.2373 (D) 0.7627

(2) The probability that four of them are from Riyadh city equals to:

(A) 0.3955 (B) 0.6045 (C) 0 (D) 0.1249

(3) The probability that at least four of them are from Riyadh city equals to:

(A) 0.3627 (B) 0.6328 (C) 0.3955 (D) 0.2763

(4) The expected number of cars that are from Riyadh city equals to:

(A) 1 (B) 3.75 (C) 3 (D) 0

Poisson Probability Distribution

Let we have an infinite number of Bernoulli trials whereas the random event of success is a rare to occur (so that $n \to \infty$, *and* $p \to 0$), let's to define the random variable X such that:

X = number of successes to be occurred within interval of time or space (or some volume of matter) for simplicity let it to be within one unite; i.e t=1 unite. So X = { x: x=0 , 1 , 2 , ... }, the random variable X is said to have a Poisson probability distribution with the parameter λ where $\lambda = n\ p$ = the average number of successes to occur during a unit of time,

and we denote it as $X \sim Poisson\ (\lambda)$ and its p.m.f to be defined as "

$$f(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!} \quad ; \quad x = 0,1,2,\ldots\ldots \qquad \text{Which satisfies:}$$

$$f(x) \geq 0 \quad \& \quad \sum_{all\ x} f(x) = \sum_{all\ x} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1$$

Theorem

If $X \sim Poisson\ (\lambda)$ then: (A) $\mu = E(X) = \lambda$ and (B) $\sigma^2 = \lambda$

Example (3):

A hospital administrator, who has been studying daily emergency admissions over a period of several years has concluded that they are distributed according to the Poisson law. Hospital records reveal that emergency admissions have averaged three per a day. Find the probability that in a given day, there will be:

(i) exactly four emergency admissions (ii) at least two emergency admissions

Solution:

Let X = No. of emergency admissions to be admitted per a day to that hospital. So, we may write $X \sim Poisson\ (\ \lambda = 3\ )$, Or $f\,(x)\ = P(\,X = x) = \ e^{-3}\ \dfrac{3^{\,x}}{x!};\ X = 1,\ 2,\ 3,\ \ldots,\ \infty$, so

(i) $P(\ X = 4\ ) = e^{-3}\ \dfrac{3^{\,4}}{4!}\ =\ 0.168$

(ii) $P(\,X\ \geq 2\ )\ = \sum_{x=2}^{\infty}\ f\,(x) = 1 -\ \{\,f\,(0) +\ f\,(1)\ \} = 1 - e^{-3}\,\{\ \dfrac{3^0}{0!} + \dfrac{3^1}{1\,!}\ \} =\ 0.801$

<u>Example (4)</u>:

Suppose that the number of snake bites cases at KKUH in a year has a Poisson probability distribution with average 6 bite cases.

(A)What is the probability that in a given year:

(i) There will be 7 snake bite cases,

(ii) There will be at least two snake bite cases.

(B)What is the probability that will be 10 snake bite cases in the next two years?

(C) What is the probability that will be no snake bite cases in a month?

Solution

(A)Let X = No. of snake bites cases at KKUH in a year; so we may write $X\ \sim\ Poisson\ (\ \lambda = 6\ )$

And $f\,(x)\ = P(\,X = x) = \ e^{-6}\ \dfrac{6^{\,x}}{x!}\quad ;\quad x = 1,\ 2,\ 3,\ \ ..\ ,\ \infty$, so

(i) $f\,(7)\ = P(\,X = 7) = \ e^{-6}\ \dfrac{6^{\,7}}{7!}\ \ =\ 0.1377$

(ii) $P(X \geq 2) = \sum_{X=2}^{\infty} f(x) = 1 - \{f(0) + f(1)\} = 1 - e^{-6}\{\frac{6^0}{0!} + \frac{6^1}{1!}\} = 0.983$

(B)Let U = No. of snake bites cases at KKUH in two years; so we may

Write: $U \sim Poisson\ (\lambda^* = 2\lambda = 2 \times 6 = 12)$

$$f(u) = P(U = u) = e^{-12}\frac{12^u}{u!}\ ;\ \ U = 1, 2, 3, \ldots\ldots, \infty \Rightarrow f(10) = P(U = 10) = e^{-12}\frac{12^{10}}{10!} = 0.1048$$

(C) Let Y = No. of snake bites cases at KKUH in a month; so we may

Write: $Y \sim Poisson\ (\lambda^{**} = (1/12)\lambda = (1/12) \times 6 = 2)$

$$f(y) = P(Y = u) = e^{-2}\frac{2^y}{y!}\ ;\ \ Y = 1, 2, 3, \ldots\ldots, \infty \Rightarrow f(0) = P(Y = 0) = e^{-2}\frac{2^0}{0!} = 0.1353$$

Approximating ***Bin(n, p)*** by $Poisson(\lambda = n\,p)$

In case $X \sim Bin\ (n, p)$ and ***n*** is so large and ***p*** is so small then $Bin\ (n, p) \approx Poisson(\lambda = n\,p)$.

For instance $Bin\ (5000, 0.001) \approx Poisson(\lambda = 5)$

EXERCISE

Q1. On average, a certain intersection results in 3 traffic accidents per day. Assuming Poisson Distribution,

(i) What is the probability that at this intersection:

(1) No accidents will occur in a given day?

(2) More than 3 accidents will occur in a given day?

(3) Exactly 5 accidents will occur in a period of two days?

(ii) What is the average number of traffic accidents in a period of 4 days?

Q2. At a checkout counter, customers arrive at an average of 1.5 per minute. Assuming Poisson distribution, then

(1) The probability of no arrival in two minutes is

(A) 0.0 (B) 0.2231 (C) 0.4463 (D) 0.0498 (E) 0.2498

(2) The variance of the number of arrivals in two minutes is

(A) 1.5 (B) 2.25 (C) 3.0 (D) 9.0 (E) 4.5

Q3. Suppose that the number of telephone calls received per day has a Poisson distribution with mean of 4 calls per day.

(a) The probability that 2 calls will be received in a given day is

(A) 0.546525 (B) 0.646525 (C) 0.146525 (D) 0.746525

(b) The expected number of telephone calls received in a given week is

(A) 4 (B) 7 (C) 28 (D) 14

(c) The probability that at least 2 calls will be received in a period of 12 hours is

(A) 0.59399 (B) 0.19399 (C) 0.09399 (D) 0.29399

Q4. The average number of car accidents at a specific traffic signal is 2 per a week. Assuming Poisson distribution, then the probability that:

(i)there will be no accident in a given week.

(A) 0.865 (B) 0.092 (C) 0.135 (D) 0.908

(ii)there will be at least two accidents in a period of two weeks.

(A) 0.865 (B) 0.092 (C) 0.135 (D) 0.908

Q5. The average number of airplane accidents at an airport is two per a year. Assuming Poisson distribution, So

(i)the probability that there will be one accident in a given year.

(A) 0.271 (B) 0.092 (C) 0.135 (D) 0.908

(ii)the average number of airplane accidents at this airport in a period of two years.

(A) 3.5 (B) 6 (C) 2 (D) 4

(iii)the probability that there will be at least two accidents in a period of 18 months.

(A) 0.865 (B) 0.331 (C) 0.135 (D) 0.669

Q6. Suppose that X~ Binomial (1000,0.002). By using Poisson approximation, P(X=3) is approximately equal to (choose the nearest number to your answer):

(A) 0.62511 (B) 0.72511 (C) 0.82511 (D) 0.92511 (E) 0.18045

Q7. The probability that a person dies when he or she contracts a certain disease is 0.005. A sample of 1000 persons who contracted this disease is randomly chosen.

(1) What is the expected number of persons who will die in this sample?

(2) What is the probability that exactly 4 persons will die among this sample?

Q8. The number of faults in a fiber optic cable follows a Poisson distribution with an average of 0.6 per 100 feet. (1) The probability of 2 faults per 100 feet of such cable is:

(A) 0.0988 (B) 0.9012 (C) 0.3210 (D) 0.5

(2) The probability of less than 2 faults per 100 feet of such cable is:

(A) 0.2351 (B) 0.9769 (C) 0.8781 (D) 0.8601

(3) The probability of 4 faults per 200 feet of such cable is:

(A) 0.02602 (B) 0.1976 (C) 0.8024 (D) 0.9739

Hypergeomtric Distribution

Suppose that we have a finite population of N distinct subjects which could be classified into two groups: K successes and the remaining (N-K) failures. Let us to draw a random sample of size n different subjects out of this population <u>without replacement</u>. If we define the random variable X such that X = number of successes to be obtained within this selected sample. Then,

The random variable X is said to have a hypergeomtric distribution with the parameters: N, n, K and to be denoted as X ~ ***h (N, n, K)*** with the probability distribution of the form

$$f(x) = f(x; N, n, K) = \begin{cases} \dfrac{C_x^K \; C_{n-x}^{N-K}}{C_n^N} & x = 0, 1, 2, \ldots\ldots, \min(n, k) \\ 0 & O..W \end{cases}$$

Which satisfies: $f(x) \geq 0 \quad \& \sum_{all\ x} f(x) = \sum_{all\ x} \dfrac{C_x^K \; C_{n-x}^{N-K}}{C_n^N} = 1$

<u>Theorem</u>: If X ~ ***h (N, n, K)*** then $\mu = E(X) = \sum_{all\,x} x\, f(x) = n\dfrac{K}{N} \quad \& \quad \sigma^2 = n\dfrac{K}{N}(1-\dfrac{K}{N})(\dfrac{N-n}{N-1})$

<u>Example (5)</u>:

Lots of 40 components are called acceptable if a sample of five components without replacements from the lot is to be selected and none of them is found defective. What is the probability that the lot to be rejected if there are three defectives in the entire contains?

<u>Solution</u>:

If we define the random variable X as No. of defective units within the selected random sample,

so we may say that X ~ ***h (N=40, n=5, K=3)*** that is $f(x) = f(x; N, n, K) = \dfrac{C_x^3 \; C_{5-x}^{37}}{C_5^{40}} \; ; \; x = 0, 1, 2, 3$

And therefore the required probability will be $P(X \geq 1) = 1 - P(X = 0) = 1 - \dfrac{C_0^3 \; C_5^{37}}{C_5^{40}} = 0.3376$.

Note: we may calculate the expected number of defective to be obtained and their variance as

$$\mu = 5\times(\frac{3}{40}) = 0.375 \text{ and } \sigma^2 = 5\times(\frac{3}{40})\times(1-3/40)\times(\frac{40-5}{40-1}) = 0.3113 .$$

<u>Approximating</u> *h (N, n, K)* by ***Bin(n, p)***

In case X ~ ***h (N, n, K)*** and ***N*** and ***K*** are so large then ***h (N, n, K)*** $\approx Bin\ (n, p = K/N)$.

For instance ***h (N=1000, n=10, K=400)*** $\approx Bin\ (10, p = 0.4)$

<u>EXERCISE</u>

Q1. A shipment of 7 television sets contains 2 defective sets. A hotel makes a random purchase of 3 of the sets.

(i) Find the probability distribution function of the random variable X representing the number of defective sets purchased by the hotel.

(ii) Find the probability that the hotel purchased no defective television sets.

(iii) What is the expected number of defective television sets purchased by the hotel?

(iv) Find the variance of X.

Q2. Suppose that a family has 5 children, 3 of them are girls and the rest are boys. A sample of 2 children is selected randomly and without replacement.

(1) The probability that no girls are selected is

(A) 0.0 (B) 0.3 (C) 0.6 (D) 0.1

(2) The probability that at most one girls are selected is

(A) 0.7 (B) 0.3 (C) 0.6 (D) 0.1

(3) The expected number of girls in the sample is

(A) 2.2 (B) 1.2 (C) 0.2 (D) 3.2

(4) The variance of the number of girls in the sample is

(A) 36.0 (B) 3.6 (C) 0.36 (D) 0.63

Q3. A committee of size 4 is selected from 2 chemical engineers and 8 industrial engineers.

(1)Write a formula for the probability distribution function of the random variable X representing the number of chemical engineers in the committee.

(2)Find the probability that there will be no chemical engineers in the committee.

(3)Find the probability that there will be at least one chemical engineer in the committee.

(4)What is the expected number of chemical engineers in the committee?

(5)What is the variance of the number of chemical engineers in the committee?

Q4. A box contains 2 red balls and 4 black balls. Suppose that a sample of 3 balls were selected randomly and without replacement. Find,

(1)The probability that there will be 2 red balls in the sample.

(2)The probability that there will be 3 red balls in the sample.

(3)The expected number of the red balls in the sample.

Q5. A lot of 8 missiles, 3 are selected at random and fired. The lot contains 2 defective missiles that will not fire. Let X be a random variable giving the number of defective missiles selected.

(1)Find the probability distribution function of X.

(2)What is the probability that at most one missile will not fire?

(3)Find $\mu = E(X)$ and $\sigma^2 = Var(X)$.

Q6. A particular industrial product is shipped in lots of 20 items. Testing to determine whether an item is defective is costly; hence, the manufacturer samples production rather than using 100% inspection plan. A sampling plan constructed to minimize the number of defectives Shipped to consumers calls for sampling 5 items from each lot and rejecting the lot if more

Than one defective is observed. (If the lot is rejected, each item in the lot is then tested.) If a lot contains 4 defectives, what is the probability that it will be accepted.

Q7. Suppose that X~h(x; 100, 2, 60); i.e., X has a hyper geometric distribution with parameters N=100, n=2, and K=60. Calculate the probabilities P(X=0), P(X=1), and P(X=2) as follows:

(a) Exact probabilities using hyper geometric distribution.

(b) Approximated probabilities using binomial distribution.

Q8. A particular industrial product is shipped in lots of 1000 items. Testing to determine whether an item is defective is costly; hence, the manufacturer samples production rather than using 100% inspection plan. A sampling plan constructed to minimize the number of defectives shipped to consumer's calls for sampling 5 items from each lot and rejecting the lot if more than one defective is observed. (If the lot is rejected, each item in the lot is then tested.) If a lot contains 100 defectives, calculate the probability that the lot will be accepted using:

(a) Hyper geometric distribution (exact probability.)

(b) Binomial distribution (approximated probability.)

Q9. A shipment of 20 digital voice recorders contains 5 that are defective. If 10 of them are randomly chosen (without replacement) for inspection, then:

(1) The probability that 2 will be defective is:

(A) 0.2140 (B) 0.9314 (C) 0.6517 (D) 0.3483

(2) The probability that at most 1 will be defective is:

(A) 0.9998 (B) 0.2614 (C) 0.8483 (D) 0.1517

(3) The expected number of defective recorders in the sample is:

(A) 1 (B) 2 (C) 3.5 (D) 2.5

(4) The variance of the number of defective recorders in the sample is:

(A) 0.9868 (B) 2.5 (C) 0.1875 (D) 1.875

Q10. A box contains 4 red balls and 6 green balls. The experiment is to select 3 balls at random. Find the probability that all balls are red for the following cases:

(1) If selection is without replacement

(A) 0.216 (B) 0.1667 (C) 0.6671 (D) 0.0333

(2) If selection is with replacement

(A) 0.4600 (B) 0.2000 (C) 0.4000 (D) 0.0640