Lecture-14

Binomial Distribution

Bernoulli trails:

A Series of trails that satisfies the following assumptions is known as Bernoulli trail.

1. There are only two possible outcomes for each trail (success and failure)
2. The probability of success is same for each trail
3. The outcomes from different trails are independent

Example-1:

Tossing a Coin 100 times is a Bernoulli trail.

There are only 2 outcomes, Head and Tail. Here getting Head can be considered as success and Tail as failure. Probability of getting Head or success remains same though out the process and the events are independent.

Binomial distribution

Consider a Bernoulli trail which results in success with probability p and a failure with probability $q = 1 - p$.

Let X: the number of success in n trails.

Then the probability distribution of the binomial random variable X is

$$P(X = x) = f(x) = b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, 2, ...n$$

Example-2:

If Probability of hitting the target is $\frac{3}{4}$ and three shots are fired, then

(i)Find the probability of hitting the target 2 times.

(ii)Formulate the binomial Probability distribution function

Ans:

Total no. of trails n=3.

Let X = no. of times hitting the target (no. of success), $x = 0, 1, 2, 3$

P (success) = P (hitting the target) $=\frac{3}{4}$

P (failure) = $\frac{1}{4}$

Therefore

(i) $P(X = 2) = f(2) = b\left(2; 3, \frac{3}{4}\right) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^{3-2}$

(ii) $P(X = x) = f(x) = b\left(x; 3, \frac{3}{4}\right) = \begin{pmatrix} 3 \\ x \end{pmatrix} \left(\frac{3}{4}\right)^x \left(\frac{1}{4}\right)^{3-x}, x = 0, 1, 2, .3$

Note: The mean and variance of binomial distribution $\boldsymbol{f(x) = b\,(x; n, p)}$ are $\boldsymbol{\mu = np}$ and $\boldsymbol{\sigma^2 = npq}$

(Q.11) The probability that a patient recovers from a delicate heart operation is 0.9. What is the probability that exactly 5 of the next 7 patients having this operation survive?

Ans:

Let X = no. of patients recovered from the heart operation i.e. $x = 0, 1, 2, ...7$

Here, $n = 7$, $p = 0.9$, $q = 0.1$

hence,

$$P(X = 5) = f(5) = b\,(5; 7, 0.9) = \begin{pmatrix} 7 \\ 5 \end{pmatrix} (0.9)^5 \,(0.1)^{7-5}$$

$$= 0.1240$$

Binomial distribution Table:

The cumulative distribution for the binomial distribution is pre-calculated and given in the form of a table.

Examples-3:(Use of binomial distribution Table)

We know

$$\begin{aligned} P(X \leq 4) &= F(4) \\ &= P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4) \end{aligned}$$

Suppose, $n=5$, $p=0.6$

Hence,

$$\begin{aligned} P(X \leq 4) &= F(4) \\ &= \binom{5}{0}(0.6)^0(0.4)^{5-0}+\binom{5}{1}(0.6)^1(0.4)^{5-1}+.......+\binom{5}{4}(0.6)^4(0.4)^{5-4} \\ &= \sum_{x=0}^{4} b(x;5,0.6) = 0.9222 \end{aligned}$$

$P(X \leq 4) = B(4;5,0.6) = 0.9222$ (from binomial distribution table)

In general $P(X \leq x) = B(x;n,p)$ and $b(x;n,p) = B(x;n,p) - B(x-1;n,p)$

(Q.15) It is known that 60% of mice inoculated with a serum are protected from a certain disease. If 5 mice are inoculated, find the probability that (a) none contracts the disease; (b) fewer than 2 contract the disease; (c) more than 3 contract the disease.

Ans:

Let X = no. of mice from the disease after inoculated, $x = 0,1,2,3,4,5$

Here, $n=5$, $p=0.4$, $q=0.6$

(i) $P(X=0) = f(0) = \binom{5}{0}(0.4)^0(0.6)^5 = 0.0778$

(ii) $P(X < 2) = P(X \leq 1)$

$$= P(X = 0) + P(X = 1)$$
$$= \begin{pmatrix} 5 \\ 0 \end{pmatrix} (0.4)^0 (0.6)^5 + \begin{pmatrix} 5 \\ 1 \end{pmatrix} (0.4)^1 (0.6)^{5-1} = 0.3370$$

(iii) $P(X > 3) = P(X = 4) + P(X = 5)$

$$= \begin{pmatrix} 5 \\ 4 \end{pmatrix} (0.4)^4 (0.6)^{5-4} + \begin{pmatrix} 5 \\ 5 \end{pmatrix} (0.4)^5 (0.6)^{5-5} = 0.087$$

(Q.16) Suppose that airplane engines operate independently and fail with probability equal to 0.4. Assuming that a plane makes a safe fight if at least one-half of its engines run, determine whether a 4-engine plane or a 2 engine plane has the higher probability for a successful fight.

Ans:

Case-1

The plane has 4 engines; $n = 4$

The plane will make a safe flight if 2 or more engines are working.

$$P(X \geq 2) = 1 - P(X \leq 1)$$
$$= 1 - [P(X = 0) + P(X = 1)]$$
$$= 1 - \left[\begin{pmatrix} 4 \\ 0 \end{pmatrix} (0.6)^0 (0.4)^{4-0} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} (0.6)^1 (0.4)^{4-1} \right] = 0.8208$$

Case-2

The plane has 2 engines; $n = 2$

The plane will make a safe flight if 1 or more engines are working.

$$P(X \geq 1) = 1 - P(X = 0)$$
$$= 1 - \left[\begin{pmatrix} 2 \\ 0 \end{pmatrix} (0.6)^0 (0.4)^{2-0} \right] = 0.84$$

Conclusion: comparing the above two cases, a 2 engine flight has higher probability for a successful flight.

Lecture-15

Multinomial Distribution

Multinomial distribution:

Consider a trial which results k outcomes, E_1, E_2, ...E_k with probabilitites p_1, p_2, ..., p_k respectively such that

$$\sum_{i=1}^{k} p_i = 1$$

Let

X_1 = no. of times E_1 occurs in n independent trials

X_2 = no. of times E_2 occurs in n independent trials

....

X_k = no. of times E_k occurs in n independent trials

Now,

$$P(X_1 = x_1, X_2 = x_2, ...X_k = x_k) = f(x_1, x_2, ...x_k; p_1, p_2, ...p_k)$$
$$= \begin{pmatrix} n \\ x_1, x_2, ...x_k \end{pmatrix} {p_1}^{x_1} {p_2}^{x_2} .....{p_k}^{x_k}$$

With $\sum_{i=1}^{n} x_i = n, \quad \sum_{i=1}^{n} p_i = 1$

(Q.19) As a student drives to school, he encounters a trafic signal. This trafic signal stays green for 35 seconds, yellow for 5 seconds, and red for 60 seconds. Assume that the student goes to school each weekday between 8:00 and 8:30 a.m. Let X_1 be the number of times he encounters a green light, X_2 be the number of times he encounters a yellow light, and X_3 be the number of times he encounters a red light. Find the joint distribution of X_1, X_2, and X_3

Ans:

Let

X_1 = no. of times he encounters a green light

X_2 = no. of times he encounters a yellow light

X_3 = no. of times he encounters a red light

Given $p_1 = 0.35, p_2 = 0.05, p_3 = 0.60$

Therefore,

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, X_3 = x_3) &= f(x_1, x_2, .x_3; n, 0.35, 0.05, 0.60) \\ &= \begin{pmatrix} n \\ x_1, x_2, ...x_k \end{pmatrix} (0.35)^{x_1} (0.05)^{x_2} (0.60)^{x_3} \end{aligned}$$

Where $x_1 + x_2 + x_3 = n$

(Q.22) According to a genetics theory, a certain cross of guinea pigs will result in red, black, and white ofspring in the ratio 8:4:4. Find the probability that among 8 ofspring, 5 will be red, 2 black, and 1 white.

Ans:

Let

X_1 = no. of red guinea pigs

X_2 = no. of black guinea pigs

X_3 = no. of white guinea pigs

It is given that the ratio of red, black, and white guinea pigs is 8:4:4

Hence,

P (guinea pig is red) = 8/16 =0.5

P (guinea pig is black) = 4/16 =0.25

P (guinea pig is white) = 4/16 =0.25

$$\begin{aligned} P(X_1 = 5, X_2 = 2, X_3 = 1) &= f(5, 2, .1; 8, 0.5, 0.25, 0.25) \\ &= \begin{pmatrix} 8 \\ 5, 2, 1 \end{pmatrix} (0.5)^5 (0.25)^2 (0.25)^1 \\ &= 21/256 \end{aligned}$$

Lecture-16

Hypergeometric Distribution

General discussion:

Suppose total number of items in a a bag $= N$

Total number of defective items (out of N) $= k$

Total number of items selected $= n$

lets discuss the probability that x out of n $(x \leq n)$ items selected is defective.

Now, total number of ways n items can be selected out of N items $= \begin{pmatrix} N \\ n \end{pmatrix}$.

Our requirement is x out of n are defective i.e. remaining $(n-x)$ are non-defective.

Number of ways x defective items can be selected from k defective items $= \begin{pmatrix} k \\ x \end{pmatrix}$.

Number of ways $n-x$ defective items can be selected from $N-k$ defective items $= \begin{pmatrix} N-k \\ n-x \end{pmatrix}$.

Probability of selecting x defectives

$$= \frac{all\ favorable\ cases}{all\ possible\ cases}$$

$$= \frac{\begin{pmatrix} k \\ x \end{pmatrix} \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Definition:

Let $X =$ The number of successes in a random sample size n selected from N items of which k are labeled success and $(N-k)$ labeled failure.Then probability distribution of the above hypergeometric random variable is

$$f(x) = h(x; N, n, k) = \frac{\begin{pmatrix} k \\ x \end{pmatrix} \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Such that $\max\{0, n-(N-k)\} \le x \le \min(n,k)$

(Q.30) A random committee of size 3 is selected from 4 doctors and 2 nurses. Write a formula for the probability distribution of the random variable X representing the number of doctors on the committee. Find $P(2 \le X \le 3)$.

Ans:

There are 4 doctors and 2 nurses

3 persons will be selected out of 4+2=6 persons

Let X: number of doctors in the committee which consists of 3 persons

So $x = 1, 2, 3$ ($x \neq 0$ why?)

$$P(X = x) = f(x) = \frac{\begin{pmatrix} 4 \\ x \end{pmatrix} \begin{pmatrix} 2 \\ 3-x \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}}$$

Now,

$P(2 \le X \le 3) = P(X = 2) + P(X = 3)$

$= \frac{\begin{pmatrix} 4 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 3-2 \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}} + \frac{\begin{pmatrix} 4 \\ 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3-3 \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}} = \frac{4}{5}$

(Q.32) From a lot of 10 missiles, 4 are selected at random and fired. If the lot contains 3 defective missiles that will not fire, what is the probability that (a) all 4 will fire? (b) at most 2 will not fire?

Ans:

Total number of missiles = 10

Total number of defective missiles = 3

Hence, total number of non-defective missiles = 7

4 missiles will be fired

Let X: number of non-defective missiles fired

(a) $P(X=0) = \dfrac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{1}{6}$

(b) At most 2 will not fire means 2 or more will fire

$$P(X \geq 2) = P(X=2) + P(X=3) + P(X=4)$$
$$= \frac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}} + \frac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}} + \frac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{29}{30}$$

Multivariate Hyper-geometric Distribution:

If N items can be partitioned into k cells $A_1, A_2, ...A_k$ with $a_1, a_2, ...a_k$elements, respectively then probability distribution of the random variables $X_1, X_2, ...X_k$representing the number of elements selected from $A_1, A_2, ...A_k$in a random sample of size n is

$$f(x_1, x_2, ...x_k; a_1, a_2, ...a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}.......\binom{a_k}{x_k}}{\binom{N}{n}}$$

With $\sum_{i=1}^{n} x_i = n,\ \sum_{i=1}^{n} a_i = N$

(Q.43) A foreign student club lists as its members 2 Canadians, 3 Japanese, 5 Italians, and 2 Germans. If a committee of 4 is selected at random, ?nd the probability that (a) all nationalities are represented; (b) all nationalities except Italian are represented.

Ans:

Total number of members = 2+3+5+2 12

Total number of members selected = 4

(a) All nationalities represented means one from each country.

One person can be selected from 2 Canadians in $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ways

One person can be selected from 3 Japanies in $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ ways

One person can be selected from 5 Italians in $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$ ways

One person can be selected from 2 Germans in $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ways

Four persons can be selected from 12 persons in $\begin{pmatrix} 12 \\ 4 \end{pmatrix}$ ways

$$P(all\ nationalities\ are\ represented)$$
$$= \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} = \tfrac{4}{33}$$

(b) All nationalities except Italians are represented, then 3 cases arise;

Case-I: 2 Canadians + 1 Japanies+ 0 Italian + 1 German

Case-2: 1 Canadians + 2 Japanies+ 0 Italian + 1 German

Case-3: 1 Canadians + 1 Japanies+ 0 Italian + 2 German

$$P(all\ nationalities\ are\ represented)$$
$$= \frac{\begin{pmatrix} 2 \\ 2 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} + \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} + \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} = \tfrac{8}{165}$$

(Q.44) An urn contains 3 green balls, 2 blue balls, and 4 red balls. In a random sample of 5 balls, find the probability that both blue balls and at least 1 red ball are selected.

Ans:

Total number of balls = 3+2+4 = 9

Total number of balls selected = 5

Number of blue balls = 2, green balls = 3, red balls = 4

$$P(2\,blue\ balls\ and\ atleast\ 1\ red\ ball)$$
$$= \frac{\binom{2}{2}\binom{4}{1}\binom{3}{2}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{2}\binom{3}{1}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{3}\binom{3}{0}}{\binom{9}{5}} = \frac{17}{63}$$

Negative Binomial Distribution:

Let repeated independent trials results in a success with probability p and a failure with probability $q = 1 - p$.

Where X: number of trials in which the kth success occurs,

Then,

$$P(X = x) = f(x) = b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}$$

Where $x = k, k + 1, k + 2....$

Geometric Distribution:

A particular case of negative binomial distribution for $k = 1$ is known as geometric distribution.

Here, X= the number of trials on which the first success occurs.

$$P(X = x) = f(x) = g(x; p) = pq^{x-1}; \quad x = 1, 2, 3........$$

Where $q = 1 - p$

(Q.49) The probability that a person living in a certain city owns a dog is estimated to be 0.3. Find the probability that the tenth person randomly interviewed in that city is the fifth one to own a dog.

Ans:

Here, $p = 0.3, q = 1 - p = 0.7$

$X=$ The number of persons interviewed in which kth person own a dog.

Given $x = 10, k = 5$

$$
\begin{aligned}
b^*(10; 5, 0.3) &= \begin{pmatrix} 10 - 1 \\ 5 - 1 \end{pmatrix} p^5 q^{10-5} \\
&= \begin{pmatrix} 9 \\ 4 \end{pmatrix} (0.3)^5 \, (0.7)^{10-5} = 0.0515
\end{aligned}
$$

(Q.50) Find the probability that a person flipping a coin gets (a) the third head on the seventh flip; (b) the first head on the fourth flip.

Ans:

Here, $p = 0.5, q = 1 - p = 0.5$

$X=$ The number of trials in which kth head occurs.

(a)Third head in 7^{th} flip means $x = 7, k = 3$

Hence,

$$
b^*(7; 3, 0.5) = \begin{pmatrix} 7 - 1 \\ 3 - 1 \end{pmatrix} (0.5)^3 (0.5)^4 = 0.1172
$$

(b)First head in the fourth flip means$x = 4, k = 1$

Using negative binomial or geometric distribution

$$
g(x; p) = g(4, 0.5) = 0.5(1 - 0.5)^{4-1} = (0.5)^4
$$

Lecture-17

Poisson Distribution

Poisson Distribution:

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring a given time interval t is

$$P(x; \lambda t) = \frac{e^{-\lambda t}\,(\lambda t)^x}{x!}, \qquad x = 0, 1, 2, ...$$

For $t = 1$

$$P(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, ...$$

The Poission random variable X, has the mean $\mu = \lambda$ and variance $\sigma^2 = \lambda$.

Note: For the purpose of minimizing calculation, refer Poisson distribution table in the problems.

(Q.58) A certain area of the eastern United States is, on average, hit by 6 hurricanes a year. Find the probability that in a given year that area will be hit by (a) fewer than 4 hurricanes; (b) anywhere from 6 to 8 hurricanes.

Ans:

The average number of hurricane hits in a year is 6 i.e. $\lambda = 6$

Let X : The number of hurricane hits in a year

a.

P(fewer than 4 hurricanes) means

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$
$$= \tfrac{e^{-6}6^0}{0!} + \tfrac{e^{-6}6^1}{1!} + \tfrac{e^{-6}6^2}{2!} + \tfrac{e^{-6}6^3}{3!} = 0.1512$$

b.

P(anywhere from 6 to 8 hurricanes) means

$$P(6 \le X \le 8) = P(X = 6) + P(X = 7) + P(X = 8)$$
$$= \frac{e^{-6}6^6}{6!} + \frac{e^{-6}6^7}{7!} + \frac{e^{-6}6^8}{8!} = 0.4015$$
$$Or$$
$$= F(8) - F(5) = 0.4015 \ (using \ table)$$

(Q.60) The average number of field mice per acre in a 5-acre wheat field is estimated to be 12. Find the probability that fewer than 7 field mice are found (a) on a given acre; (b) on 2 of the next 3 acres inspected.

Ans:

The average number of field mice per acre is 12 i.e. $\lambda = 12$

(a) Let X : The number of mice per acre

Hence,

$$P(X < 7) = P(X \le 6) = 0.0458 \ \ (using \ table)$$

(b) Let Y : The number of acres of land inspected.

Here Y follows binomial distribution.

Given $n = 3, y = 2$

Hence, $P(Y = 2) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} p^2 q^{3-2}, \ \ with \ p = 0.0458$ (from part a), $q = 1 - p$

(Q.69) The probability that a person will die when he or she contracts a virus infection is 0.001. Of the next 4000 people infected, what is the mean number who will die?

Ans:

Given $p = 0.001, \ n = 4000$

Therefore, $\mu = np = 4000 \times 0.001 = 4$

Lecture-18

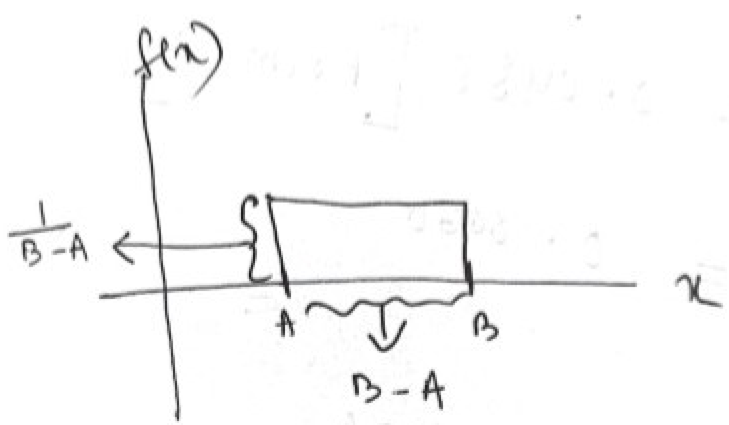
Uniform / Rectangular Distribution

Defination:

The probability density function of the continuous uniform random variable X on the interval $[A, B]$ is

$$f(x; A, B) = f(x) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B \\ 0, & otherwise \end{cases}$$

Geometrical interpretation: The density function $f(x)$ forms a rectangle with base



$B - A$ and constant height $\frac{1}{B-A}$ as shown in the figure given above.

Example: The density function of the uniform distribution in the interval $[2, 5]$is

$$f(x) = \begin{cases} \frac{1}{5-2} = \frac{1}{3}, & 2 \leq x \leq 5 \\ 0, & otherwise \end{cases}$$

Theorem-1: Prove that the mean and variance of the uniform distribution are

$$\mu = \frac{A+B}{2}, \ \ \sigma^2 = \frac{(B-A)^2}{12}$$

Proof: Here,

mean $\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{A}^{B} x \frac{1}{B-A} dx = \frac{A+B}{2}$

Variance $\sigma^2(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \int_{-\infty}^{\infty} \left(x - \frac{A+B}{2}\right)^2 \frac{1}{B-A} dx = \frac{(B-A)^2}{12}$

(Q.2) Suppose X follows a contineous uniform distribution from 1 to 5. Determine the conditional probability $P(X > 2.5|X \leq 4)$

Ans:

The Probability density function,

$$f(x) = \begin{cases} \frac{1}{5-1} = \frac{1}{4}, & 1 \leq x \leq 5 \\ 0, & otherwise \end{cases}$$

Therefore,

$$P(X > 2.5|X \leq 4) = \frac{P(X>2.5 \ \wedge \ X\leq4)}{P(X\leq4)}$$
$$= \frac{P(2.5<X\leq4)}{P(X\leq4)} = \frac{\int_{2.5}^{4} f(x)dx}{\int_{-\infty}^{4} f(x)dx} = \frac{\int_{2.5}^{4} \frac{1}{4}dx}{\int_{1}^{4} \frac{1}{4}dx} = \frac{1}{2}$$

(Q4) A bus arrives every 10 minutes at a bus stop.It is assumed that the waiting time for a particular individual is a random variable with a contineous uniform distribution .

(a) What is the probability that the individual waits more than 7 minutes ?

(b) What is the probability that the individual waits between 2 and 7 minutes?

Ans:

Let$X$: The waiting time for a particular individuals

Here, $f(x) = \begin{cases} \frac{1}{10-0} = \frac{1}{10}, & 0 \leq x \leq 10 \\ 0, & otherwise \end{cases}$

(a) $P(X > 7) = \int_{7}^{\infty} f(x)dx = \int_{7}^{10} \frac{1}{10}dx = \frac{3}{10}$

(b) $P(2 < X < 7) = \int_{2}^{7} f(x)dx = \int_{2}^{7} \frac{1}{10}dx = \frac{5}{10}$

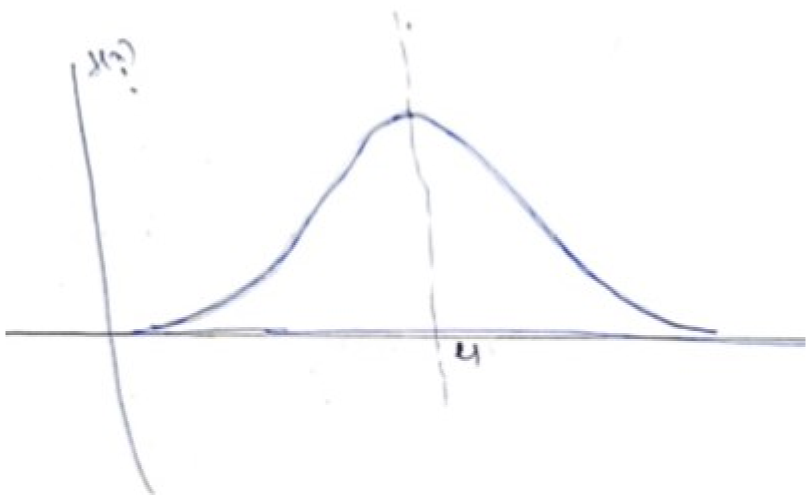
Lecture-19

Normal Distribution

The probability density function of the normal random variable X, with mean μ and variance σ^2 is

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \quad -\infty < x < x$$

Observations:



1. **T**he mode, which is the point on the horizontal axis, where the curve is a maximum, occurs at $x = \mu$.

2. The curve is symmetric about the vertical axis through the mean.

3. The normal curve approaches the horizontal axis asymptotically as we procced in the either direction away from the mean.

4. The total area under the curve and above the horizontal axis is 1 (why?).

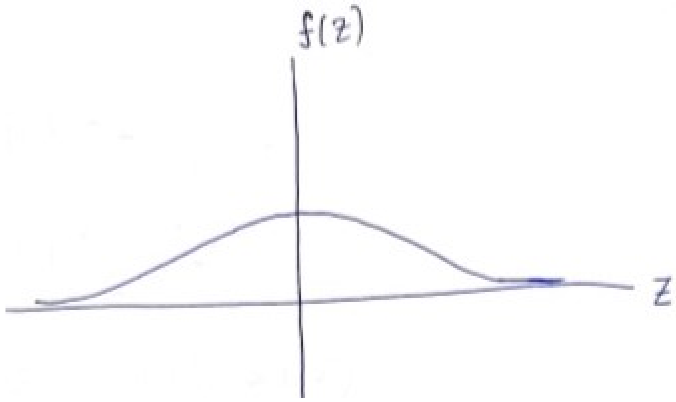
Special case: $(\mu = 0, \sigma^2 = 1)$

Standard Normal Distribution:

Let X be the normal random variable, with mean $\mu = 0$ and variance $\sigma^2 = 1$; let us specially denote it as Z; then the density function

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}; \quad -\infty < z < \infty$$

Properties:



1. **G**raph is symmetric at $z = 0$, about $\mu = 0$.

2. The area under the curve fro $-\infty \ to \ -z$ is same as $z \ to \ \infty$ i.e. $P(Z \leq -z) = P(Z \geq z)$

Cumulative Distribution Function

The cumulative distribution function

$$F(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

Two important results:

If Z is the standard normal random variable then,

1. $P(a < Z \leq b) = F(b) - F(a)$

2. $F(z) + F(-z) = 1 \ / \ F(c) + F(-c) = 1$

Proof: $P(a < Z \leq b)$ represents area under the curve from $a \ to \ b$; i.e. same as, area up to b from $-\infty$ minus the area up to a from $-\infty$, so,

$$P(a < Z \leq b) = P(Z \leq b) - P(Z \leq a) = F(b) - F(a)$$

Total area under the curve is 1 so

$$P(Z \leq c) + P(Z > c) = 1$$

As area fro $c\ to\ \infty$ is same as $-\infty\ to\ -c$ ie,

$$P(Z > c) = P(Z < -c)$$

,

$$P(Z \leq c) + P(Z \leq -c) = 1$$

$$F(c) + F(-c) = 1$$

Example: Given a standard normal distribution,

1. Find the area to the right of $z = 1.84$ i.e. $P(Z > 1.84)$
2. Find the area between $z = -1.97$ and $z = 0.86$ i.e. $P(-1.97 \leq Z \leq 0.86)$

Ans:

1. $P(Z > 1.84) = 1 - P(Z \leq 1.84) = 1 - 0.9671 = 0.0329$
2. $P(-1.97 \leq Z \leq 0.86) = P(Z \leq 0.86) - P(Z \leq -1.97) = 0.8051 - 0.0244 = 0.7807$

Example: Given a standard normal distribution, find the value of k such that

1. $P(Z \leq k) = 0.9671$
2. $P(Z > k) = 0.3015$
3. $P(k < Z < -0.18) = 0.4197$

Ans:

1. $P(Z \leq k) = 0.9671 \Rightarrow k = 1.84$

2.
$$P(Z > k) = 0.3015$$
$$\Rightarrow 1 - P(Z \leq k) = 0.3015$$
$$\Rightarrow P(Z \leq k) = 1 - 0.3015 = 0.6985$$
$$\Rightarrow k = 0.52$$

3. $P(k < Z < -0.18) = 0.4197$

$$\Rightarrow F(-0.18) - F(k) = 0.4197$$
$$\Rightarrow 0.4286 - F(k) = 0.4197$$
$$\Rightarrow F(k) = 0.4286 - 0.4197 = 0.0089$$
$$\Rightarrow k = -2.37$$

(Q7) Given a standard normal distribution, find the value of k such that

(a) $P(Z > k) = 0.2946$ (b) $P(Z \leq k) = 0.0427$ (c) $P(-0.93 < Z < k) = 0.7235$

Ans:

(a) $P(Z > k) = 0.2946$

$$\Rightarrow 1 - P(Z \leq k) = 0.2946$$
$$\Rightarrow P(Z \leq k) = 1 - 0.2946 = 0.7054$$
$$\Rightarrow k = 0.54$$

(b) $P(Z \leq k) = 0.0427 \Rightarrow k = -1.72$

(c) $P(-0.93 < Z < k) = 0.7235$

$$\Rightarrow F(k) - F(-0.93) = 0.7235$$
$$\Rightarrow F(k) = 0.4197 + F(-0.93)$$
$$\Rightarrow F(k) = 0.4197 + 0.1762 = 0.8997$$
$$\Rightarrow k = 1.28$$

Lecture-20

Normal Distribution(continuing)

Working with arbitrary mean and arbitrary variance

If X is any random variable with mean μ and variance σ^2 then,

$Z = \frac{X-\mu}{\sigma}$ with have mean $\mu = 0$ and variance $\sigma^2 = 1$

Proof: Given X has meanμ and variance σ^2

$$Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} + \left(-\frac{\mu}{\sigma}\right)$$

Mean(Z) =$E(Z) = \frac{1}{\sigma}E(X) + -\left(\frac{\mu}{\sigma}\right) = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$

Variance(Z)=$\sigma^2(Z) = \frac{1}{\sigma^2}Variance(X) = \frac{1}{\sigma^2}.\sigma^2 = 1$

Example: If X has the mean$\mu = 2$ and variance $\sigma^2 = 9$ then,

$Z = \frac{X-\mu}{\sigma} = \frac{X-2}{3}$ has the mean $\mu = 0$ variance $\sigma^2 = 1$.

(Q8) Given a normal distribution with $\mu = 30$ and $\sigma = 6$, find (a) the normal curve area to the right of $x = 17$; (b) the normal curve area to the left of $x = 22$; (c) the normal curve area between $x = 32$ and $x = 41$; (d) the value of x that has 80% of the normal curve area to the left; (e) the two values of x that contain the middle 75% of the normal curve area.

Ans: Given $\mu = 30$ and $\sigma = 6$,

(a) $P(X > 17) = 1 - P(X \leq 17)$

$$= 1 - P\left(Z \leq \tfrac{17-30}{6}\right)$$
$$= 1 - P(Z \leq -2.17)$$
$$= 1 - 0.0150 = 0.9850$$

(b) $P(X < 22) = P\left(Z < \frac{22-30}{6}\right) = P(Z < -1.33) = 0..0918$

(c) $P(32 < X < 41) = P\left(X < 41\right) - P\left(X < 32\right)$

$$= P\left(Z < \tfrac{41-30}{6}\right) - P\left(Z < \tfrac{32-30}{6}\right)$$
$$= P(Z < 1.83) - P(Z < 0.33)$$
$$= 0.9664 - 0.6293 = 0.3371$$

(d) Here we have to find the value of x(or c)such that $P(X < x) = 80\%$

$$P(X \leq x) = 80\% \Rightarrow P(X \leq x) = 0.8$$
$$\Rightarrow P\left(Z \leq \tfrac{x-\mu}{\sigma}\right) = 0.8$$
$$\Rightarrow \tfrac{x-\mu}{\sigma} = 0.84$$
$$\Rightarrow \tfrac{x-30}{6} = 0.84$$
$$\Rightarrow x = 6 \times 0.84 + 30 = 35.04$$

(e) Here we have to find two values c_1 and c_2 such that $P(c_1 < X < c_2) = 75\%$

$$P(c_1 < X < c_2) = 75\% \Rightarrow P(-z_2 < Z < z_2) = 0.75$$
$$\Rightarrow P(Z < z_2) - P(Z < -z_2) = 0.75$$
$$\Rightarrow F(z_2) - (1 - F(z_2)) = 0.75$$
$$\Rightarrow 2F(z_2) = 1.75$$
$$\Rightarrow F(z_2) = 0.875$$

$$\Rightarrow z_2 = 1.15$$
$$\Rightarrow \tfrac{c_2-\mu}{\sigma} = 1.15$$
$$\Rightarrow \tfrac{c_2-30}{6} = 1.15$$
$$\Rightarrow c_2 = 6 \times 1.15 + 30 = 36.9$$

$$-z_2 = -1.15$$
$$\Rightarrow \tfrac{c_1-\mu}{\sigma} = -z_2 = -1.15$$
$$\Rightarrow \tfrac{c_1-30}{6} = -1.15$$
$$\Rightarrow c_1 = -6 \times 1.15 + 30 = 23.1$$

(Q10) According to Chebyshev's theorem, the probability that any random variable assumes a value within 3 standard deviations of the mean is at least $\frac{8}{9}$. If it is known that the probability distribution of a random variable X is normal with mean μ and variance σ^2, what is the exact value of $P(\mu - 3\sigma < X < \mu + 3\sigma)$?

Ans:

$$\begin{aligned}
&P(\mu - 3\sigma < X < \mu + 3\sigma) \\
&= P\left(\tfrac{\mu-3\sigma-\mu}{\sigma} < \tfrac{X-\mu}{\sigma} < \tfrac{\mu+3\sigma-\mu}{\sigma}\right) \\
&= P(-3 < Z < 3) \\
&= F(3) - F(-3) = 0.9987 - (0.0013) = 0.9974
\end{aligned}$$

(Q15) A lawyer commutes daily from his suburban home to his midtown office. The average time for a one-way trip is 24 minutes, with a standard deviation of 3.8 minutes. Assume the distribution of trip times to be normally distributed. (a) What is the probability that a trip will take at least $\frac{1}{2}$ hour? (b) If the office opens at 9:00 A.M. and the lawyer leaves his house at 8:45 A.M. daily, what percentage of the time is he late for work?

Ans: Let$X$: time taken for a one way trip

Given mean $\mu = 24$ and $\sigma = 3.8$ then

(a) $P(X \geq 30) = 1 - P(X < 30) = 1 - P\left(Z < \frac{30-24}{3.8}\right) = 1 - P(Z < 1.58) = 0.0571$

(b) As the person starts at 8.45 A.M. and office starts at 9 A.M., if he takes more than 15 minutes for the trip,

$$P(X > 15) = 1 - P(X \leq 15) = 0.9911$$

(c) $P(X > 25) = 1 - P(X \leq 25) = 0.3974$

(d)

$$\begin{aligned}
&P(X \geq c) = 0.15 \\
&\Rightarrow 1 - P(X < c) = 0.15 \\
&\Rightarrow 1 - P\left(Z < \tfrac{c-24}{3.8}\right) = 0.15 \\
&\Rightarrow P\left(Z < \tfrac{c-24}{3.8}\right) = 0.85 \\
&\Rightarrow \tfrac{c-24}{3.8} = 1.14 \\
&\Rightarrow c = 3.8 \times 1.14 + 24 = 27.95
\end{aligned}$$

(e) Let Y: be the no. of trips that takes at least half an hour

Y follows binomil distribution

Here, $n = 3, p = 0.0571\ (From\ (a))$

$$P(Y = 2) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} p^2 q^{3-2} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} (0.0521)^2 (1 - 0.0521)^{3-2} = 0.0092$$

(Q22) If a set of observations is normally distributed, what percent of these differ from the mean by (a) more than 1.3σ? (b) less than 0.52σ?

(a) $P(X < \mu - 1.3\sigma) + P(X > \mu + 1.3\sigma)$

$$\begin{aligned} &= P(Z < -1.3) + P(Z > 1.3) \\ &= P(Z < -1.3) + 1 - P(Z \leq 1.3) \\ &= 0.1936 = 19.36\% \end{aligned}$$

(b) $P(\mu - 0.52\sigma < X < \mu + 0.52\sigma)$

$$\begin{aligned} &= P(-0.52 < Z < 0.52) \\ &= 0.3970 = 39.7\% \end{aligned}$$

Lecture-21

Normal Approximation to Binomial Distribution

we want to find the probability of getting (a) exactly 80 heads (b) at most 80 heads in a tossing a coin 200 times.

Using binomial distribution

$$P(X = 80) = f(80) = \left(\begin{array}{c} 200 \\ 80 \end{array} \right) \left(\frac{1}{2}\right)^{80} \left(\frac{1}{2}\right)^{200-80}$$

= math error will pop up in calculator

Again if try to calulate $P(X \leq 80)$ using binomial distribution, calculation will be very difficult or you will get math error in your calculator . That happens because n is very large, This promots us to use Noramal distribution in stead of binomial distribution when n is large.

It is advisable to use normal approximation to binomial distribution when np and npq are greater than 15.

Theorem

if X is a binomial random variable with mean $\mu = np$ and variance $\sigma^2 = npq$ then the limiting form of the distribution of
$Z = \frac{X-np}{\sqrt{npq}}$
As $n \longrightarrow \infty$, the standardized normal distribution is denoted as $n(Z; 0, 1)$
working rule
suppose X follows the binomial distribution with mean $\mu = np$ and variance $\sigma^2 = npq$ then

$P(X \leq c) = PZ \leq (\frac{c+0.5-np}{\sqrt{npq}})$

Example

$P(X \leq 50) = P(Z \leq \frac{50.5-np}{\sqrt{npq}})$

$P(X < 50) = P(Z \leq \frac{49.5-np}{\sqrt{npq}})$

$P(50 < X \leq 60) = P(\frac{50.5-np}{\sqrt{npq}} \leq Z \leq \frac{60.5-np}{\sqrt{npq}})$

$P(X > 50) = 1 - P(X \leq 50) = 1 - (Z \leq \frac{50.5-np}{\sqrt{npq}})$

(Q24) A coin is toseed 400 times. Use the normal curve apnroximation to find the probability of obtaining (a) between 185 and 210 heads incluscve; (b) exactly 205 heads; (c) fewer than 176 or more than 227 heads.

Ans:

Here,$n = 400, p = q = \frac{1}{2}$

Hence, mean $\mu = np = 400 * \frac{1}{2} = 200$ and $\sigma = \sqrt{npq} = \sqrt{400 * \frac{1}{2} * \frac{1}{2}} = 10$

Let X : no of heads turned up

Therefore,

(a) $P(185 \leq X \leq 210)$

$= P(\frac{185-0.5-np}{\sqrt{npq}} \leq Z \leq \frac{210+0.5-np}{\sqrt{npq}})$

$= P(\frac{184.5-200}{10} \leq Z \leq \frac{210.5-200}{10})$

$= P(-1.55 \leq Z \leq 1.05)$

$= F(1.05) - F(-1.55) = 0.8531 - 0.0606 = 0.7925$

(b) $P(X = 205)$

$= P(204 < X < 206)$

$= P(\frac{204+0.5-np}{\sqrt{npq}} \leq Z \leq \frac{206-0.5-np}{\sqrt{npq}})$

$= P(0.45 \leq Z \leq 0.55) = 0.0352$

(c) $P(X < 176) + P(X > 227)$
$= P(Z < \frac{176-0.5-200}{10}) + P(Z > \frac{227+0.5-200}{10})$

$= P(Z < -2.45) + P(Z > 2.75)$

$= P(Z < -2.45) + [1 - P(Z \leq 2.75)]$

$= 0.0071 + 1 - 0.9970 = 0.0101$

(Q26) A process yields 10% defective items. If 100 items are randomly seeected from the procees, what is the probability that the numeer of defectives
(a) exceeds 13? (b) is less than 8?

Ans: Given,$n = 100, p = 0.1, q = 1 - 0.1 = 0.09$

Let X : no. of defective

Hance , meen $\mu = np = 100 * 0.1 = 10$ and $\sigma = \sqrt{npq} = \sqrt{100 * 0.1 * 0.9} = 3$

(a)

$P(X > 13) = 1 - P(X \leq 13)$

$= 1 - P(Z \leq \frac{13+0.5-np}{\sqrt{npq}})$

$= 1 - P(Z \leq \frac{13+0.5-np}{3}) = 1 - P(Z \leq 1.17) = 1 - 0.8790 = 0.1210$

(b)

$P(X < 8) = P(Z \leq \frac{8-0.5-np}{\sqrt{npq}}) = P(Z \leq \frac{7.5-10}{3}) = P(Z \leq -0.83) = 0.2033$

Lecture-22

Gamma Distribution

Gamma Function:

We know that $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx \; ; \alpha > 0$

Properties Of Gamma function:

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$

   Proof:

$$\begin{aligned}\Gamma(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}dx \\ &= -x^{\alpha-1}e^{-x}\Big|_0^\infty + \int_0^\infty (\alpha-1)e^{-x}x^{\alpha-2}dx \\ &= (\alpha-1)\int_0^\infty e^{-x}x^{\alpha-2}dx = (\alpha-1)\Gamma(\alpha-1)\end{aligned}$$

2. $\Gamma 1 = 1$

3. If $\alpha = n$ $(possitive\ integer)$,

$$\Gamma(n) = (n-1)!$$

4. $\Gamma(1/2) = \sqrt{\pi}$

Gamma Distribution

The continuous random variable X has a gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if its density function is given by

$$f(x) = f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}; x > 0 \\ 0, \qquad\qquad\qquad elsewhere \end{cases}$$

Particular case: Exponential distribution $(\alpha = 1)$

The density function of the exponential distribution is given by

$$f(x) = \begin{cases} \frac{1}{\beta}e^{-1/\beta} & ; x > 0 \\ 0, & elsewhere \end{cases}$$

Note:

1. The mean and variance of the gamma distribution are, $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$

2. The mean and variance of the exponential distribution are, $\mu = \beta$ and $\sigma^2 = \beta^2$

Example: If a random variable has the gamma distribution with $\alpha = 2$ and $\beta = 3$, then find the mean and standard deviation of this distribution.

Ans: Given $\alpha = 2$ and $\beta = 3$

Hence, $\mu = \alpha\beta = 6$ and $\sigma^2 = \alpha\beta^2 = 18 \Rightarrow \sigma = 3\sqrt{2}$

(Q41) If a random variable X has the gamma distribution with $\alpha = 2$ and $\beta = 1$, find $P(1.8 < X < 2.4)$.

Ans: Given X follows gamma distribution with $\alpha = 2$ and $\beta = 1$.

Hence, $f(x) = \begin{cases} xe^{-x} & ; x > 0 \\ 0, & elsewhere \end{cases}$

Now, $P(1.8 < X < 2.4) = \int_{1.8}^{2.4} f(x)dx = \int_{1.8}^{2.4} xe^{-x}dx = 0.1545$

(Q46) The life, in years, of a certain type of electrical switch has an exponential distribution with an average life $\beta = 2$. If 100 of these switches are installed in different systems, what is the probability that at most 30 fail during the first year?

Ans: Let X: life of the electric switches in years

Given X follows exponential distribution with $\beta = 2$

Hence, $f(x) = \begin{cases} \frac{1}{2}e^{-x/2} & ; x > 0 \\ 0, & elsewhere \end{cases}$

Now, $P(X < 1) = \int_0^1 f(x)dx = \frac{1}{2}\int_0^1 e^{-x/2}dx = 0..3935$

Let Y: the number of electric switches fails during the first year

Given Y follows binomial distribution where $p = 0.3935, n = 100$ and we have to calculate $P(Y \leq 30)$ but to simplify calculation we will use normal approximation to binomial distribution.

Hence,

$$\mu = np = 100 \times 0.3935 = 39.35$$
$$\sigma = \sqrt{npq} = \sqrt{100 \times 0.3935 \times 0.6065} = 4.885$$

Now,

$$P(Y \leq 30) = P\left(Z \leq \tfrac{30.5-np}{\sqrt{npq}}\right)$$
$$= P\left(Z \leq \tfrac{30.5-39.35}{4.885}\right)$$
$$= P\left(Z \leq -1.81\right) = 0.035$$

(Q54) The lifetime, in weeks, of a certain type of transistor is known to follow a gamma distribution with mean 10 weeks and standard deviation $\sqrt{50}$ weeks. (a) What is the probability that a transistor of this type will last at most 50 weeks? (b) What is the probability that a transistor of this type will not survive the first 10 weeks?

Ans:

Let X: life time of the transistor in a weeks

Given X follows gamma distribution with

$\mu = \alpha\beta = 10$, and $\sigma = \sqrt{\alpha\beta^2} = \sqrt{50}$

Hence, $f(x) = \begin{cases} \frac{1}{25} x e^{-x/5} & ; x > 0 \\ 0, & elsewhere \end{cases}$

(a) $P(X \leq 50) = \int_0^{50} f(x)dx = \frac{1}{25}\int_0^{50} x e^{-x/5} dx = 0.9995$

(b) $P(X < 10) = \int_0^{10} f(x)dx = \frac{1}{25}\int_0^{10} x e^{-x/5} dx = 0.5940$