17/1/12

Distributions

Distribution is a mathematical functional relationship b/w values of x and is of corresponding probabilities.

1. **Binomial distribution**: A random v. X is said to follow binomial distribution if its probability law is given as

$$p(x) = \begin{cases} {}^nC_x\, p^x q^{n-x}, & \text{where } x = 0, 1, \ldots n \\ 0 & \text{otherwise} \end{cases}$$

It gives the probability for x no of success out of n trials.

Conditions: no. of trials (n) must be finite.

- Each trial must result in 2 mutually disjoint ways (success & failure) i.e; $p+q=1$ [p = probability of success in one trial].
- p is constant in all the trials.
- All the trials are independent.

The binomial distribution can also be represented as

$X \sim B(n,p)$ & $(q+p)^n$ where n, p are parameters (constants)

Moment Generating function:

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx} \cdot p(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum \binom{n}{x} (p \cdot e^t)^x q^{n-x} = (q + p \cdot e^t)^n \quad [\because \text{by expansion}]$$

$$E(x) = \frac{d}{dt} M_x(t) \Big/ t=0 = n(q+pe^t)^{n-1} pe^t \Big/ t=0 \text{ (given)}$$

$$= np \quad (\because p+q=1).$$

$$E(x^2) = \frac{d^2}{dt^2} M_x(t) \Big/ t=0 = \frac{d}{dt}\left(n(q+pe^t)^{n-1} pe^t\right)$$

$$np\frac{d}{dt}\left((q+pe^t)^{n-1}e^t\right) = np\left[e^t\left((n-1)(q+pe^t)^{n-2}pe^t\right) + (q+pe^t)^{n-1}e^t\right]$$

$$= np\left[(n-1)(q+p)^{n-2}p + (q+p)^{n-1}\right].$$

$$= np\left[(p+q)^{n-2}\left\{(n-1)p + (p+q)\right\}\right]$$

$$= np\left[(p+q)^{n-2}\left\{np - \cancel{p} + \cancel{p} + q\right\}\right]$$

$$= np\left\{np + q\right\} = E(x^2).$$

$$np(np+q) - (np)^2 \Rightarrow V(x) = E(x^2) - (E(x))^2$$

$$\cancel{n^2p^2} + npq - \cancel{n^2p^2} = npq$$

$\therefore$ Mean of binomial distribution $= np$. $\Big\{$ $\boxed{\text{Mean} > \text{Variance}}$ **
Variance of binomial distribution $= npq$

→ For binomial distribution, mean is 3 and variance is 2.
Find the distribution & calculate $P(X>1)$

$$np = 3,\quad npq = 2 \Rightarrow q = \frac{2}{3} \Rightarrow p = 1/3.$$

$$n\left(\frac{2}{9}\right) = 2 \Rightarrow n = 9 \Rightarrow P(x) = \binom{9}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{9-x}$$

$$P(X>1) = 1 - P(X \le 1) = 1 - \left[P(0) + P(1)\right].$$

$$P(0) = (1)\left(\frac{2}{3}\right)^9, \quad P(1) = 9\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^8$$

$=$ $=$

Q The mean and variance of binomial distributions are 4 & 2
Find the distribution & $P(1 < X \le 3)$

2) The probability that a man hitting a target is 1/3, if he fires 6 times find the probability that he fires (i) atmost 5 times (ii) exactly once.

1) $np = 4, \quad npq = 2 \implies q = 1/2 \rightarrow p = 1/2. \quad n\left(\frac{1}{4}\right) = 2$

$n = 8.$

$P(x) = \binom{8}{x}(1/2)^x(1/2)^{8-x}.$

$P(1 < x \leq 3) = P(2) + P(3) = \binom{8}{2}(1/2)^2(1/2)^6 + \binom{8}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^5$

$7/64 + (7/32) = \frac{21}{64}$

2) $p = 1/3, \quad n = 6, \quad q = 2/3.$

$P(1) = \binom{6}{1}(1/3)^1(2/3)^5. =$

$P(x \leq 5) = 1 - P(6) =$

18/1/12.

Unit II

Poisson's distribution:

The probability law of poisson distribution is

$$P(x) = \begin{cases} e^{-\lambda}\lambda^x/x! & \text{where } x = 0, 1, \ldots \infty \\ 0 & \text{otherwise.} \end{cases}$$

The poisson distribution can be represented as $X \sim P(\lambda)$

It gives the probability for x no of success out of n trials where (i) no of trials (n) is indefinitely large.

(ii) Probability of success p is small 0

(iii) np = finite = λ (say)

The mean and ~~variate~~ variances of the poisson's distribution are equal ~~and~~ i.e λ

1) If x is a poisson variate such that $P(X=1) = \frac{3}{2} P(X=3)$

Find (i) $P(X<1)$, $P(X>3)$, $P(2<X<5)$.

$P(X=1) = \frac{3}{2} P(X=3) \Rightarrow \frac{e^{-\lambda}\lambda^1}{1!} = \frac{\cancel{3}}{2} \frac{e^{-\lambda}\lambda^3}{\cancel{3} \times 2}$

$\Rightarrow \lambda^2 = 4 \Rightarrow \lambda = 2$

$P(X<1) = P(X=0) = \frac{e^{-2}(2)^0}{0!} = \underline{0.135}$

$P(X>3) = 1 - \left[P(X=0) + P(X=1) + P(X=2)\right] - P(X=3)$

$= 1 - \left[0.135 + 0.271 + 0.271 + 0.180\right] = \underline{\underline{0.143}}$

$P(2<X<5) = P(X=3) + P(X=4)$

$= 0.180 + 0.090 = \underline{0.270}$

2) For a poisson variate, if the mean is 2, find std. deviation, $P(X \le 2)$.

Standard deviation $= \sqrt{\text{mean}} = \sqrt{\lambda} = \sqrt{2} = 1.414$

$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$= 0.135 + 0.271 + 0.271 = \underline{0.677}$

3) It is observed that 2% suffered due to bad reaction of an injection. The injection is given to 200 people. Find out the probability atleast 2 persons will suffer due to bad reaction.

$p = 2\%$ $\quad n = 200$ $\quad \lambda = np = 200\left(\frac{2}{100}\right) = 4$

$P(X \ge 2) = 1 - P(X<2) = 1 - [P(X=0) + P(X=1)]$

$= 1 - [0.018 + 0.073] = \underline{0.909}$

Success = suffering due to bad reaction.

4) Average no. of accidents on any day on a national highway is 3. Determine the probability that no. of accidents is atleast 2, atmost 3.

$P(X \ge 2)$ $\quad P(X \le 3)$

$\lambda = 3$

(i) $1 - [P(X=0) + P(X=1)] = 0.801$

(ii) $P(X=0) + P(X=1) + P(X=2) + P(X=3) = \underline{0.647}$

5) A manufacturer of pins knows that 2% of products is defective. If he sells pins in the boxes of hundred pins and guarantee that not more than 4 pins will be defective. What is the probability that the manufacturer to meet the guarantee and fails to meet the guarantee.

$\lambda = 2$ $\left[\because p = 2\%, n = 100 \rightarrow \lambda = np = 100\left(\frac{2}{100}\right)\right]$

Meets the guarantee level: $P(X \le 4) = 0.947$

Does not meet the guarantee level: $1 - P(X \le 4) = 0.053$

Fitting of Poisson's distribution:

Fit a poisson distribution to following data

X	0	1	2
f	50	40	10

X	f	$f_i(x_i)$
0	50	0
1	40	40
2	10	20
	$\Sigma f = 100$	$\Sigma f(x) = 60$

$$\bar{x} = \frac{\Sigma f(x)}{\Sigma f} = \frac{60}{100} = 0.6 = \lambda$$

exp. frequency.

$$f = N \times P(x) \qquad P(x) = \frac{e^{-0.6}(0.6)^x}{x!}, \quad x = 0, 1, 2$$

is the following poisson's distribution.

x	P(x)	f
0	0.549	54.9
1	0.329	32.9
2	0.099	9.9

- 27/1/15 $M_x(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} P(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \cdot \lambda)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$$\text{mean} = E(x) = \frac{d}{dt} m_x(t) / t = 0$$

$$= e^{\lambda(e^t - 1)} \lambda e^t / t=0 = \lambda.$$

$$V(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \Sigma x^2 P(x) = \frac{d^2}{dt^2} m_x(t) / t = 0.$$

$$= \frac{d}{dt}\left[\lambda e^t e^{\lambda(e^t - 1)}\right] / t = 0$$

$$= \lambda\left[e^t \cdot \lambda e^t \cdot e^{\lambda(e^t - 1)} + e^t \cdot e^{\lambda(e^t - 1)}\right] // t = 0$$

$$= \lambda[\lambda + 1]$$

$V(x) = E(x^2) - (E(x))^2 = \lambda^2 + \lambda - \lambda^2 = \underline{\lambda}$

<u>RR for probabilities:</u> Consider $\dfrac{P(x+1)}{P(x)} = \dfrac{e^{-\lambda}\lambda^{x+1}}{(x+1)!} \Big/ \dfrac{e^{-\lambda}\lambda^{x}}{x!}$

$$\frac{P(x+1)}{P(x)} = \lambda/_{x+1} \Rightarrow P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-1.5}(1.5)^0}{0!}$$

$$= \underline{e^{-1.5}} = \cancel{0.2231}$$

$$P(1) = e^{-1.5}(1.5)/1!$$

$$P(x+1) = \frac{x}{x+1} P(x).$$

<u>Normal Distribution:</u> The pdf of the normal distribution is given as $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$-\infty < x < \infty$

$-\infty < \mu < \infty$

$\sigma > 0$

$$X \sim N(\mu, \sigma^2)$$

Let $z = \dfrac{x-\mu}{\sigma}$, then z is called a standard normal variate and it follows standard Normal distribution as

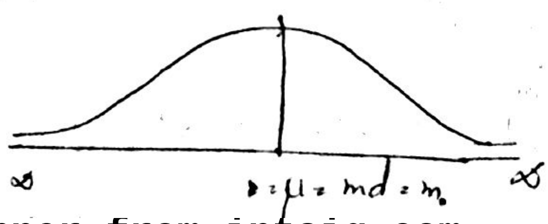
$$f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

The mean and variences of the standard Normal distribution are '0' and '1' respectively.

<u>Properties of Normal distribution.</u>

(i) The curve of N.D is bell shaped

(ii) The curve is symmetrical because mean, median and modes are equal.



iii) It is neither peaked nor flattened

Q.D : M.D : S.D = 10 : 12 : 15

iv) 'x' axis is an asymptote to the curve.

(v) The area of the curve will be above the x axis.

Total area = 1.

vii) Area Property

(i) $P(\mu-\sigma < x < \mu+\sigma)$

(ii) $P(\mu-2\sigma < x < \mu+2\sigma)$

iii) $P(\mu-3\sigma < x < \mu+3\sigma)$.

Moment generating function:

$$M_x(t) = E(e^{tx})$$

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

$x = \sigma z + \mu$

$dx = \sigma dz$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad , \quad z = \frac{x-\mu}{\sigma}$$

$$M_x(t) = \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma z) - \frac{1}{2}z^2} \cancel{\sigma}\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma z) - \frac{1}{2}z^2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2\sigma z t - 2t\mu)} dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{\mu t} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2\sigma z t)} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} e^{\sigma^2 t^2/2} \cdot \int_{-\infty}^{\infty} e^{-1/2(z-\sigma t)^2} \, dz$$

$$= \frac{e^{\mu t + \sigma^2 t^2/2}}{\sqrt{2\pi}} \cdot 2\int_0^{\infty} e^{-p^2/2} \, dp \qquad \left\{ \begin{array}{l} z - \sigma t = p \\ \Rightarrow p^2 = (z-\sigma t)^2 \end{array} \right.$$

$$\left\{ \begin{array}{l} p^2/2 = y \\ \frac{\cancel{2}p}{\cancel{2}} dp = dy \\ dp = \frac{1}{\sqrt{2y}} dy \end{array} \right.$$

$$= \frac{e^{\mu t + \sigma^2 t^2/2}}{\sqrt{\cancel{2}\pi}} \times \cancel{2} \int_0^{\infty} e^{-y} \frac{1}{\sqrt{\cancel{2}y}} \, dy$$

$$= \frac{e^{\mu t + \sigma^2 t^2/2}}{\sqrt{\cancel{\pi}}} (\cancel{\pi}^{1/2}) = \underline{\underline{e^{\mu t + \sigma^2 t^2/2}}}$$

$$\frac{d}{dt} M_x(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)\left(\mu + \frac{\sigma^2}{2}(2t)\right) \Big/ t = 0$$

$$E(x^2) = \frac{d^2}{dt^2} M_x(t) / t = 0$$

$$= \frac{d}{dt}\left[e^{\mu t + \sigma^2 t^2/2} \cdot (\mu + \sigma^2 t) \right] / t = 0$$

$$= e^{\mu t + \sigma 2 t^2/2} \cdot (\mu + 2\sigma^2 t/2)(\mu + \sigma^2) t$$

$$= e^{\mu t + \sigma^2 t^2/2} (\sigma^2) / t = 0$$

$$= \mu^2 + \sigma^2$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= \mu^2 + \sigma^2 - \mu^2$$

$$= \underline{\sigma^2}$$

14/2/13

Uniform Distribution

A random variable $X \in (a, b)$ is said to follow ~~exponential~~ uniform distribution if its probability law is given as

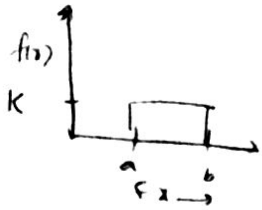
$$f(x) \begin{cases} = k \; ; & a < x < b \\ = 0 & \text{otherwise} \end{cases}$$

To find the constant k, $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

$$\int_a^b k\,dx = 1 \Rightarrow k(x)_a^b = k(b-a) = 1$$

$$\Rightarrow k = 1/(b-a)$$

The graph of uniform distribution is



As the probability curve is rectangle, it is known as rectangular distribution.

Moment Generating Function: $M_x(t) = E(e^{tx})$

$$X \sim U(a,b) \qquad = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_a^b e^{tx} \cdot \frac{1}{(b-a)}\,dx = \frac{1}{(b-a)} \frac{(e^{bt} - e^{at})}{(t)}$$

Note: The MGF of uniform distribution cannot generate $E(X)$, $E(X^2)$ etc. Then $E(X) = \int_a^b x f(x)\,dx$.

$$= \int_a^b x \left(\frac{1}{b-a}\right) = \left[\frac{x^2}{2}\right]_a^b \left(\frac{1}{b-a}\right)$$

$$= \left(\frac{b^2 - a^2}{2}\right)\left(\frac{1}{b-a}\right) = \frac{b+a}{2}$$

$$E(X^2) = \int_a^b x^2 f(x)\,dx = \frac{b^2 + ab + a^2}{3}$$

$$V(X) = E(x^2) - (E(X))^2.$$

Exponential distribution: Probability density function

is $\quad f(x) = \theta e^{-\theta x} \quad , \theta > 0, x > 0.$

Moment Generating function: $E(e^{tx})$.

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = \int_{0}^{\infty} e^{tx} \theta e^{-\theta x} dx.$$

$$= \theta \int_{0}^{\infty} e^{(tx - \theta x)} dx = \theta \int_{0}^{\infty} e^{x(t-\theta)} dx.$$

$$= \theta \int_{0}^{\infty} e^{t} = \frac{\theta}{t-\theta} \left[e^{x(t-\theta)} \right]_{0}^{\infty}$$

$$= \frac{\theta}{t-\theta} \left[e^{-x(\theta - t)} \right]_{0}^{\infty} = \frac{\theta}{t-\theta} \left[e^{-\infty} - e^{-0} \right] = \frac{-\theta}{t-\theta}$$

$$= \frac{1}{1 - t/\theta} = (1 - t/\theta)^{-1}$$

$$E(x) = \frac{d}{dt} M_x(t) / t = 0. = 1 + \left(\frac{t}{\theta}\right) + \left(\frac{t}{\theta}\right)$$

28/1/13

Sampling Distribution:

Population (popn) : The whole group of individuals under study.

popn size: (N) : It is no. of units in the population

parameter: The statistical constant of the population. General notation is θ

popn mean (μ), popn variance (σ^2)

Sample: Finite subgroup of population.

Sample size: (n) : No. of observations in the sample.

Statistics: The statistical constant calculated for sample. observations represented by t.

sample mean ($\bar{x}$), sample variance (s^2)

When the population is divided into k possible samples & for all the samples if we calculate a statistic ($t_1, t_2 \dots t_k$) the distribution followed by these statistics is known as sampling distribution. ~~Examples~~

Ex: Sampling distribution of means.

S.No.	1	2	3	.	k.
Sample mean	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$		$\bar{x}_k$

then the distribution followed by these k sample means is called as sampling distribution of means for which

mean $= \mu_{\bar{x}} = \bar{\bar{x}} = \frac{1}{k} \Sigma \bar{x}_i$.

Variance $= \sigma^2_{\bar{x}} = \frac{1}{k} \Sigma (\bar{x}_i - \bar{\bar{x}})^2$

we can verify that $\mu_{\bar{x}} = \mu$ (where μ is popn mean).

$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$ (where σ^2 = popn variance) &

$\therefore$ $SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}$. But in finite popn. we use.

finite popn crctn factor (FPC) as $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

Sampling distribution of proportions:

Consider

S.No.	1	2	3	......	k
Sample proportion	p_1	p_2	p_3		p_k.

The distribution followed by these sample proportions is called as sampling distribution of proportions for which mean $\mu_p = \bar{p}$; Variance $= \sigma^2_p = \frac{1}{k}\sum(p_i - \bar{p})^2$

We can verify that $\mu_p = P$ (popn proportion)

$\sigma^2_p = \frac{PQ}{n}$ where $Q = 1-P$.

Thus $SE(p) = \sqrt{\frac{PQ}{n}}$.

1. If $N=5$, $n=2$, find FPC $= \frac{\cancel{\sigma^2}}{\cancel{n}} \frac{N-n}{N-1}$

$= \frac{5-2}{5-1} = \frac{3}{4}$.

2. What happens to SE when (i) the size of the sample is increased from 100 to 200 (ii) decreased from 400 to 100

(i) $SE(\bar{x}) = \sigma/\sqrt{n}$

$SE(\bar{x})_{①} = \sigma/\sqrt{100}$ $\quad SE(\bar{x})_{②} = \sigma/\sqrt{200}$

$\frac{SE(\bar{x})_{②}}{SE(\bar{x})_{①}} = \frac{\sigma/\sqrt{200}}{\sigma/\sqrt{100}} = \sqrt{\frac{1}{2}} = (0.707)$.

It decreased by 0.707.

(ii) $\frac{SE(\bar{x})_{②}}{SE(\bar{x})_{①}} = \frac{\sigma/\sqrt{100}}{\sigma/\sqrt{400}} = \sqrt{4} = 2$.

It increased by 2 times or doubled.