MAT2001
Statistics for Engineers

Module 4
Probability Distributions

Special Probability Distributions

Discrete Probability Distributions

1. *Binomial Distribution*
2. *Poisson Distribution*

Continuous Probability Distributions

1. *Normal Distribution*
2. *Exponential Distribution*
3. *Gamma Distribution*
4. *Weibull Distribution*

Binomial Distribution

Definition: Let A be some event associated with a random experiment E, such that $P(\text{A}) = p$ and $P(\overline{\text{A}}) = 1 - p = q$. Assuming that p remains the same for all repetitions, if we consider n independent repetitions (or trials) of E and if the random variable (RV) X denotes the number of times the event A has occurred, then X is called a *binomial random variable* with parameters n and p or we say that X follows a *binomial distribution* with parameters n and p, or symbolically $B(n, p)$. Obviously the possible values that X can take, are $0, 1, 2, \ldots, n$.

Probability Mass Function of the Binomial Distribution

$$P(X = r) = nC_r\, p^r q^{n-r};\ \ r = 0, 1, 2, \ldots, n \ \text{ where } \ p + q = 1$$

(i) Binomial distribution is a legitimate probability distribution since

$$\sum_{r=0}^{n} P(X = r) = \sum_{r=0}^{n} nC_r\, q^{n-r} p^r$$
$$= (q + p)^n = 1$$

Binomial Distribution $B(n, p)$

A DRV 'X' is said to follows B·D if the pmf of X is defined as

where

$$\boxed{p_r = P(X = x_r) = P(X = r) = {}^nC_r \, p^r \cdot q^{n-r}}$$

X represents the trails of the experiment.

$X = 0, 1, 2, 3, \ldots, n$

$r \in \{0, 1, 2, \ldots, n\}$

$${}^nC_r = \frac{n!}{(n-r)!\, r!}$$

p = probability of success

q = probability of failure

$\ni \quad p + q = 1 \quad (\text{i.e. } q = 1 - p)$

For pmf

(i). +ve ness

$p_r \geq 0, \; \forall \; r = 0, 1, 2, \ldots, n$

(ii).
$$\sum_{r=0}^{n} p_r = \sum_{r=0}^{n} {}^nC_r \cdot p^r \cdot q^{n-r}$$

$$= (p+q)^n = 1$$

$$\boxed{(x+y)^n = \sum_{r=0}^{n} {}^nC_r \, x^r \cdot y^{n-r}}$$

Mean and Variance of the Binomial Distribution

$$E(X) = \sum_r x_r p_r$$

$$= \sum_{r=0}^{n} r \cdot nC_r \, p^r q^{n-r}$$

$$= \sum_{r=0}^{n} r \cdot \frac{n!}{r!(n-r)!} p^r q^{n-r} \qquad (1)$$

$$= np \cdot \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!\{(n-1)-(r-1)\}!} p^{r-1} q^{(n-1)-(r-1)}$$

$$= np \sum_{r=1}^{n} (n-1) C_{r-1} \cdot p^{r-1} \cdot q^{(n-1)-(r-1)}$$

$$= np\,(q+p)^{n-1}$$

$$= np \qquad (2)$$

$$E(X^2) = \sum_r x_r^2 p_r = \sum_0^n r^2 p_r$$

$$= \sum_{r=0}^{n} \{r(r-1) + r\} \frac{n!}{r!(n-r)!} p^r q^{n-r}$$

$$= n(n-1)p^2 \sum_{r=2}^{n} (n-2)C_{r-2} p^{r-2} q^{n-r} + np,$$

[by (1) and (2)]

$$= n(n-1)p^2 (q+p)^{n-2} + np$$

$$= n(n-1)p^2 + np$$

$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$

$$= n(n-1)p^2 + np - n^2p^2$$

$$= np(1-p)$$

$$= npq$$

Poisson Distribution

Definition: If X is a discrete RV that can assume the values 0, 1, 2, …, such that its probability mass function is given by

$$P(X = r) = \frac{e^{-\lambda} \lambda^r}{r!}; \quad r = 0, 1, 2, \ldots; \ \lambda > 0$$

then X is said to follow a *Poisson distribution* with parameter λ or symbolically X is said to follow $P(\lambda)$.

(**Note:** Poisson distribution is a legitimate probability distribution, since

$$\sum_{r=0}^{\infty} P(x = r) = \sum_{r=0}^{\infty} \frac{e^{-\lambda} \lambda^r}{r!}$$

$$= e^{-\lambda} e^{\lambda} = 1)$$

Mean and Variance of the Poisson Distribution

$$E(X) = \sum_{r} x_r p_r$$

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

$$= \lambda e^{-\lambda} \sum_{r=1}^{\infty} \frac{\lambda^{r-1}}{(r-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \lambda \qquad (2)$$

$$E(X^2) = \sum_{r} x_r^2 pr$$

$$= \sum_{r=0}^{\infty} \{r(r-1) + r\} e^{-\lambda} \frac{\lambda^r}{r!}$$

$$= \lambda^2 e^{-\lambda} \sum_{r=2}^{\infty} \frac{\lambda^{r-2}}{(r-2)!} + \lambda \qquad \text{[by (1) and (2)]}$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

Poisson Distribution as Limiting Form of Binomial Distribution

Poisson distribution is a limiting case of binomial distribution under the following conditions:

(i) n, the number of trials is indefinitely large, i.e., $n \to \infty$.

(ii) p, the constant probability of success in each trial is very small, i.e., $p \to 0$.

(iii) $np (= \lambda)$ is finite or $p = \dfrac{\lambda}{n}$ and $q = 1 - \dfrac{\lambda}{n}$, where λ is a positive real number.

$$\boxed{\lim_{n \to \infty} B(n, p) = P(\lambda)}$$

Note:

if X is the Poisson RV

$$E(X) = \lim_{\substack{n \to \infty \\ np = \lambda}} (np) = \lambda$$

and $$\operatorname{Var}(X) = \lim_{\substack{p \to 0 \\ np = \lambda}} (npq) = \lim_{p \to 0} [\lambda(1 - p)] = \lambda.$$

Example:

N n

Out of 800 families with 4 children each, how many families would be expected to have (i) 2 boys and 2 girls, (ii) at least 1 boy, (iii) at most 2 girls and (iv) children of both sexes. Assume equal probabilities for boys and girls.

Soln.:

Let X represents the no. of ~~boys~~ (girls)

$X = 0, 1, 2, 3 \& 4$ n.

$\Rightarrow X \to$ DRV

$n = 4$ & $p =$ probability of success

$p = 1/2 \Rightarrow q = 1/2$

$B(n, p)$

(i) P(2 Bs & 2 Gs)

$$P(X = 2) = 4C_2 \left(1/2\right)^2 \cdot \left(1/2\right)^{4-2}$$

Solution:

Considering each child as a trial, $n = 4$. Assuming that birth of a boy is a success, $p = \frac{1}{2}$ and $q = \frac{1}{2}$. Let X denote the number of successes (boys).

(i) $P(\text{2 boys and 2 girls}) = P(X = 2)$

$$= 4C_2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{4-2}$$

$$= 6 \times \left(\frac{1}{2}\right)^4 = \frac{3}{8}$$

$\therefore$ No. of families having 2 boys and 2 girls

$= N \cdot (P(X = 2)$ (where N is the total no. of families considered)

$= 800 \times \frac{3}{8}$

$= 300.$

(ii) $P(\text{at least 1 boy}) = P(X \geq 1)$

$= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$= 1 - P(X = 0)$

$$= 1 - 4C_0 . \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^4$$

$$= 1 - \frac{1}{16} = \frac{15}{16}$$

$\therefore$ No. of families having at least 1 boy

$$= 800 \times \frac{15}{16} = 750.$$

(iii) $P(\text{at most 2 girls}) = P(\text{exactly 0 girl, 1 girl or 2 girls})$

$= P(X = 4, X = 3 \text{ or } X = 2)$

$= 1 - \{P(X = 0) + P(X = 1)\}$

$$= 1 - \left\{ 4C_0 \cdot \left(\frac{1}{2}\right)^4 + 4C_1 \cdot \left(\frac{1}{2}\right)^4 \right\}$$

$$= \frac{11}{16}$$

$\therefore$ No. of families having at most 2 girls

$$= 800 \times \frac{11}{16} = 550.$$

(iv) P (children of both sexes)

$= 1 - P(\text{children of the same sex})$

$= 1 - \{P(\text{all are boys}) + P(\text{all are girls})\}$

$= 1 - \{P(X = 4) + P(X = 0)\}$

$$= 1 - \left\{ 4C_4 \cdot \left(\frac{1}{2}\right)^4 + 4C_0 \cdot \left(\frac{1}{2}\right)^4 \right\}$$

$$= 1 - \frac{1}{8} = \frac{7}{8}$$

$\therefore$ No. of families having children of both sexes

$$= 800 \times \frac{7}{8} = 700.$$

Example:

An irregular 6-faced die is such that the probability that it gives 3 even numbers in 5 throws is twice the probability that it gives 2 even numbers in 5 throws. How many sets of exactly 5 trials can be expected to give no even number out of 2500 sets?

n

N

Soln/.

Let X represents the no/. of even faces of a die.

$X = 0, 1, 2, 3, 4, 5 \longrightarrow B(n, p)$

$n = 5$, $r \in \{0, 1, \ldots, 5\}$

$p = ?$

Given,

$P(X=3) = 2 \times P(X=2)$ ✓

$^5C_3\, p^3 \cdot (1-p)^2 = 2 \times {^5C_2} \cdot p^2 \cdot (1-p)^3$

$\Rightarrow p = ?$ & $q = ?$

$P(X=0) = {^5C_0} \cdot p^0 \cdot q^5$

$= ?$

$N \times P(X=0) = 2500 \times \underline{\ ?\ }$

Solution:

Let the probability of getting an even number with the unfair die be p.

Let X denote the number of even numbers obtained in 5 trials (throws).

Given: $P(X = 3) = 2 \times P(X = 2)$

i.e., $5C_3\, p^3 q^2 = 2 \times 5C_2\, p^2 q^3$

i.e., $p = 2q = 2(1 - p)$

$\therefore$ $3p = 2$ or $p = \dfrac{2}{3}$ and $q = \dfrac{1}{3}$

Now P(getting no even number)

$$= P(X = 0)$$

$$= 5C_0 \cdot p^0 \cdot q^5 = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$$

$\therefore$ Number of sets having no success (even number) out of N sets $= N \times P(X = 0)$

$\therefore$ Required number of sets $= 2500 \times \dfrac{1}{243}$

$= 10$, nearly

Example:

Two dice are thrown 120 times. Find the average number of times in which the number on the first dice exceeds the number on the second dice.

Mean

Soln/. $B(n,p)$

$n = 120$

$p = ?$

$S = \{(1,1), (1,2), \ldots, (1,6),$
$(2,1), (2,2), \ldots, (2,6),$
$\vdots$
$(6,1), (6,2), \ldots, (6,6)\}$

$|S| = 36$.

Expected Cases $= \{(2,1), (3,1), (3,2), \ldots, (6,5)\}$

$|E| = 15$

$p = \frac{|E|}{|S|} = \frac{15}{36}$

Average $= E(X) = n \times p = 120 \times \frac{15}{36}$

$= \underline{\underline{50}}$

Solution:

The number on the first dice exceeds that on the second die, in the following combinations:

(2, 1); (3, 1), (3, 2); (4, 1), (4, 2), (4, 3); (5, 1), (5, 2), (5, 3); (5, 4); (6, 1), (6, 2), (6, 3), (6, 4), (6, 5),

where the numbers in the parentheses represent the numbers in the first and second dice respectively.

$\therefore \quad P(\text{success}) = P(\text{no. in the first die exceeds the no. in the second die})$

$$= \frac{15}{36} = \frac{5}{12}$$

This probability remains the same in all the throws that are independent.

If X is the no. of successes, then X follows a binomial distribution with parameters $n\,(= 120)$ and $p\left(= \frac{5}{12}\right)$.

$$\therefore \quad E(X) = np = 120 \times \frac{5}{12} = 50$$

Example:

Fit a binomial distribution for the following data:

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80

frequency

Solution:

Fitting a binomial distribution means assuming that the given distribution is approximately binomial and hence finding the probability mass function and then finding the theoretical frequencies.

To find the binomial frequency distribution $N(q + p)^n$, which fits the given data, we require N, n and p. We assume N = total frequency = 80 and n = no. of trials = 6 from the given data.

To find p, we compute the mean of the given frequency distribution and equate it to np (mean of the binomial distribution).

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80
fx:	0	18	56	36	28	30	24	192

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4$$

i.e., $np = 2.4$ or $6p = 2.4$

$\therefore$ $p = 0.4$ and $q = 0.6$

If the given distribution is nearly binomial, the theoretical frequencies are given by the successive terms in the expansion of $80(0.6 + 0.4)^6$. Thus we get,

x:	0	1	2	3	4	5	6
Theoretical f:	3.73	14.93	24.88	22.12	11.06	2.95	0.33

Converting these values into whole numbers consistent with the condition that the total frequency is 80, the corresponding binomial frequency distribution is as follows:

x:	0	1	2	3	4	5	6	Total
f:	4	15	25	22	11	3	0	80

Example:

The number of monthly breakdowns of a computer is a RV having a Poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month

(a) without a breakdown,

(b) with only one breakdown and

(c) with atleast one breakdown.

Soln/.

Let $X = 0, 1, 2, 3, \ldots$

$X \sim PD(\lambda)$. with the mean $= \lambda = 1.8$

$\lambda = 1.8$.

$$p_r = P(X = r) = \frac{e^{-\lambda} \cdot \lambda^r}{r!}, \quad r = 0, 1, 2 \ldots$$

$$P(X = r) = \frac{e^{-1.8} \cdot (1.8)^r}{r!}.$$

Solution:

Let X denote the number of breakdowns of the computer in a month. X follows a Poisson distribution with mean (parameter) $\lambda = 1.8$.

$$\therefore P\{X = r\} = \frac{e^{-\lambda}.\lambda^r}{r!} = \frac{e^{-1.8}.(1.8)^r}{r!}$$

(a) $P(X = 0) = e^{-1.8} = 0.1653$

(b) $P(X = 1) = e^{-1.8}\,(1.8) = 0.2975$

(c) $P(X \geq 1) = 1 - P(X = 0) = 0.8347$

Example:

Fit a Poisson distribution for the following distribution: n

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400

$\rightarrow$ Frequency Distribution

N

x_i :

f_i :

$f_i x_i$:

$$\lambda = \text{Mean} = \frac{\sum x_i f_i}{\sum f_i}$$

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

$N \times P(x=0) =$

$N \times P(x=1) =$

$\vdots \qquad \vdots$

$N \times P(x=5) =$

Solution:

Fitting a Poisson distribution for a given distribution means assuming that the given distribution is approximately Poisson and hence finding the probability mass function and then finding the theoretical frequencies.

To find the probability mass function

$$P\{X = r\} = \frac{e^{-\lambda} \cdot \lambda^r}{r!} \quad r = 0, 1, 2, \ldots, \infty$$

of the approximate Poisson distribution, we require λ, which is the mean of the Poisson distribution.

We find the mean of the given distribution and assume it as λ.

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400
fx:	0	156	138	81	20	5	400

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{400}{400} = 1 = \lambda$$

The theoretical frequencies are given by

$$\frac{N e^{-\lambda} \cdot \lambda^r}{r!} \text{ where } N = 400\text{, obtained from the given distribution.}$$

$$= \frac{400\, e^{-1}}{r!}, \quad r = 0, 1, 2, \ldots, \infty$$

Thus, we get

	x:	0	1	2	3	4	5
Theoretical	f:	147.15	147.15	73.58	24.53	6.13	1.23

The theoretical frequencies for $x = 6, 7, 8, \ldots$ are very small and hence neglected.

Converting the theoretical frequencies into whole numbers consistent with the condition that the total frequency = 400, we get the following Poisson frequency distribution which fits the given distribution:

	x:	0	1	2	3	4	5
Theoretical	f:	147	147	74	25	6	1

Exercise:

It is known that the probability of an item produced by a certain machine will be defective is 0.05. If the produced items are sent to the market in packets of 20, find the number of packets containing at least, exactly and at most 2 defective items in a consignment of 1000 packets using (i) binomial distribution and (ii) Poisson approximation to binomial distribution.

Exercise:

If a fair coin is tossed at random 5 independent times, find the conditional probability of 5 heads relative to the hypothesis that there are at least 4 heads.

Exercise:

A car hire firm has 2 cars which it hires out day by day. The number of demands for a car on each day follows a Poisson distribution with mean 1.5. Calculate the proportion of days on which (i) neither car is used and (ii) some demand is not fulfilled.

Exponential Distribution ED(λ)

Definitions: A continuous RV X is said to follow an *exponential distribution* or *negative exponential distribution* with parameter $\lambda > 0$, if its probability density function is given by

$$\text{pdf} = f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We note that $\int_0^\infty f(x)\,dx = \int_0^\infty \lambda e^{-\lambda x}\,dx = 1$ and hence $f(x)$ is a legitimate density function.

For pdf:

① $f(x) \geq 0$ (+ve non)

② $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_0^\infty \lambda e^{-\lambda x}\,dx = 1$

Mean and Variance of the Exponential Distribution

$$E(X) = \text{Mean of the exponential distribution}$$
$$= \mu_1' = \frac{1}{\lambda}$$

$$E(X^2) = \mu_2' = \frac{2}{\lambda^2}$$

$$\therefore \ \mathrm{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Gamma Function

$$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx, \quad 1 \le n \le 2.$$

Some properties of the gamma function:

$$\Gamma(n+1) = n\Gamma(n), n > 0,$$

and when n = integer > 0, we have $\Gamma(n) = (n - 1)!$

Beta Function

The gamma function is related to the beta function, $B(m,n)$, as follows:

$$B(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}\, dx$$

$$B(m,n) = B(n,m) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

Table for Gamma Function

n	$\Gamma(n)$	n	$\Gamma(n)$	n	$\Gamma(n)$	n	$\Gamma(n)$
1.00	1.00000	1.25	0.90640	1.50	0.88623	1.75	0.91906
1.01	0.99433	1.26	0.90440	1.51	0.88659	1.76	0.92137
1.02	0.98884	1.27	0.90250	1.52	0.88704	1.77	0.92376
1.03	0.98355	1.28	0.90072	1.53	0.88757	1.78	0.92623
1.04	0.97844	1.29	0.89904	1.54	0.88818	1.79	0.92877
1.05	0.97350	1.30	0.89747	1.55	0.88887	1.80	0.93138
1.06	0.96874	1.31	0.89600	1.56	0.88964	1.81	0.93408
1.07	0.96415	1.32	0.89464	1.57	0.89049	1.82	0.93685
1.08	0.95973	1.33	0.89338	1.58	0.89142	1.83	0.93969
1.09	0.95546	1.34	0.89222	1.59	0.89243	1.84	0.94261
1.10	0.95135	1.35	0.89115	1.60	0.89352	1.85	0.94561
1.11	0.94739	1.36	0.89018	1.61	0.89468	1.86	0.94869
1.12	0.94359	1.37	0.88931	1.62	0.89592	1.87	0.95184
1.13	0.93993	1.38	0.88854	1.63	0.89724	1.88	0.95507
1.14	0.93642	1.39	0.88785	1.64	0.89864	1.89	0.95838
1.15	0.93304	1.40	0.88726	1.65	0.90012	1.90	0.96177
1.16	0.92980	1.41	0.88676	1.66	0.90167	1.91	0.96523
1.17	0.92670	1.42	0.88636	1.67	0.90330	1.92	0.96878
1.18	0.92373	1.43	0.88604	1.68	0.90500	1.93	0.97240
1.19	0.92088	1.44	0.88580	1.69	0.90678	1.94	0.97610
1.20	0.91817	1.45	0.88565	1.70	0.90864	1.95	0.97988
1.21	0.91558	1.46	0.88560	1.71	0.91057	1.96	0.98374
1.22	0.91311	1.47	0.88563	1.72	0.91258	1.97	0.98768
1.23	0.91075	1.48	0.88575	1.73	0.91466	1.98	0.99171
1.24	0.90852	1.49	0.88595	1.74	0.91683	1.99	0.99581
						2.00	1.00000

General Gamma or Erlang Distribution

Definition: A continuous RV X is said to follow an *Erlang distribution* or *General Gamma distribution* with parameters $\underline{\lambda > 0}$ and $\overline{k} > 0$, if its probability density function is given by

$Er(k, \lambda)$

$$\text{pdf} = f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, & \text{for } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

We note that $\displaystyle\int_0^\infty f(x)\,dx = \frac{\lambda^k}{\Gamma(k)} \int_0^\infty x^{k-1} e^{-\lambda x}\,dx$

$$= \frac{1}{\Gamma(k)} \int_0^\infty t^{k-1}\, e^{-t}\,dt, \text{ [on putting } \lambda x = t]$$

$$= 1$$

Hence $f(x)$ is a legitimate density function.

Mean and Variance of the General Gamma or Erlang Distribution

$$\text{Mean} = E(X) = \frac{1}{\lambda} \cdot \frac{\Gamma(k+1)}{\Gamma(k)} = \frac{k}{\lambda}$$

$$\begin{aligned}\text{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{1}{\lambda^2} \cdot \frac{\Gamma(k+2)}{\Gamma(k)} - \left(\frac{k}{\lambda}\right)^2 \\ &= \frac{1}{\lambda^2} \{k(k+1) - k^2\} = \frac{k}{\lambda^2}\end{aligned}$$

Gamma Distribution

GD(k)

Note

1. When $\lambda = 1$, the Erlang distribution is called Gamma distribution or simple Gamma distribution with parameter k whose density function is $f(x) = \frac{1}{\Gamma(k)} x^{k-1} e^{-x}; x \geq 0; k > 0.$
2. When $k = 1$, the Erlang distribution reduces to the exponential distribution with parameter $\lambda > 0$.
3. Sometimes, the Erlang distribution itself is called Gamma distribution.

Mean and Variance of the Gamma Distribution

$\text{Mean} = E(X) = k$

$\text{Var}(X) = k$

Weibull Distribution WD(α, β)

Definition: A continuous RV X is said to follow a *Weibull distribution* with parameters $\alpha, \beta > 0$, if the RV $Y = \alpha X^{\beta}$ follows the exponential distribution with density function $f_Y(y) = e^{-y}$, $y > 0$.

Density Function of the Weibull Distribution

Since $Y = \alpha \cdot X^{\beta}$, we have $y = \alpha \cdot x^{\beta}$.

By the transformation rule, derived in chapter 3, we have $f_X(x) = f_Y(y) \left| \dfrac{dy}{dx} \right|$,

where $f_X(x)$ and $f_Y(y)$ are the density functions of X and Y respectively.

$$\therefore \quad f_X(x) = e^{-y}\, \alpha \beta x^{\beta - 1}$$

pdf: $$= \alpha \beta x^{\beta - 1} e^{-\alpha x^{\beta}}; \; x > 0 \quad [\because y > 0]$$

Note *When $\beta = 1$, Weibull distribution reduces to the exponential distribution with parameter α.*

Mean and Variance of the Weibull Distribution

$$\text{Mean} = E(X) = \mu_1' = \boxed{\alpha^{-\frac{1}{\beta}}\,\Gamma\left(\frac{1}{\beta}+1\right)}$$

$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$

$$= \alpha^{-2/\beta}\left[\Gamma\left(\frac{2}{\beta}+1\right) - \left\{\Gamma\left(\frac{1}{\beta}+1\right)\right\}^2\right]$$

Normal (Gaussian) Distribution $N(\mu, \sigma)$

Definition: A continuous RV X is said to follow a *normal distribution* or *Gaussian distribution* with parameters μ and σ, if its probability density function is given by

Pdf =
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2}; \quad -\infty < x < \infty$$
$$-\infty < \mu < \infty \quad \sigma > 0 \qquad (1)$$

Symbolically 'X follows $N(\mu, \sigma)$'. Sometimes it is also given as $N(\mu, \sigma^2)$.

Note:

$f(x)$ is a legitimate density function, as

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2}\,\sigma\sqrt{2}\,dt, \quad \left(\text{on putting } t = \frac{x-\mu}{\sigma\sqrt{2}}\right)$$
$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2}\,dt$$
$$= \frac{1}{\sqrt{\pi}}\, 2\int_{0}^{\infty} e^{-t^2}\,dt = \frac{1}{\sqrt{\pi}} \cdot \Gamma\left(\frac{1}{2}\right) = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1$$

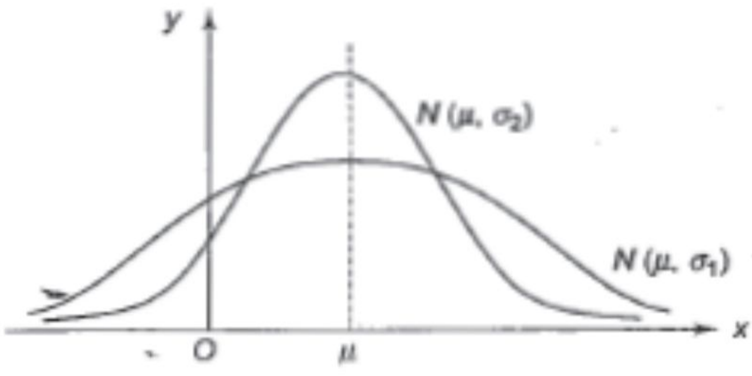
Standard Normal Distribution

The normal distribution $N(0, 1)$ is called the standardised or simply the standard normal distribution, whose density function is given by

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

This is obtained by putting $\mu = 0$ and $\sigma = 1$ and by changing x and f respectively into z and ϕ. If X has distribution $N(\mu, \sigma)$ and if $\boxed{Z = \frac{X-\mu}{\sigma}}$, then we can prove that Z has distribution $N(0, 1)$.

Normal Probability Curve



AREAS UNDER NORMAL CURVE

Normal probability curve is given by

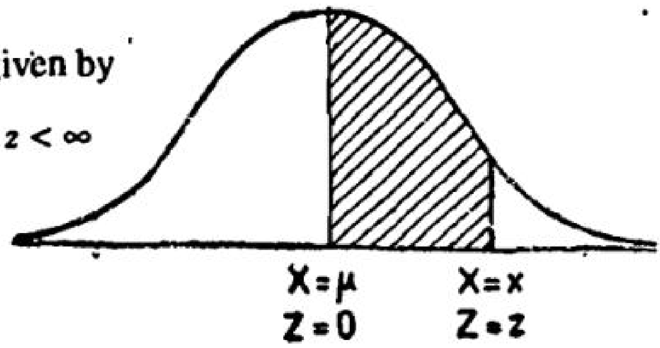
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad -\infty < x < \infty$$

and standard normal probability curve is given by

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right), \quad -\infty < z < \infty$$

where $Z = \dfrac{X - E(X)}{\sigma_X} \sim N(0, 1)$

Areas under Nomal Curve



The following table gives the shaded area in the diagram *viz*.. $P(0<Z<z)$ for different values of z.

TABLE OF AREAS

↓Z→	0	1	2	3	4	5	6	7	8	9
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0759
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3655	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4678	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4959	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.1960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Mean and Variance of the Normal Distribution

If X follows $N(\mu, \sigma)$, then $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$ and $S \cdot D = \sigma$

Mean and Variance of the Standard Normal Distribution

$\text{Mean}(Z) = \mu = 0$

$\text{Var}(Z) = \sigma^2 = \underline{1}$

Example:

The mileage which car owners get with a certain kind of radial tire is a RV having an exponential distribution with mean 40,000 km. Find the probabilities that one of these tires will last (i) at least 20,000 km and (ii) at most 30,000 km.

Soln/.

$X \sim ED(\lambda)$ with Mean $= 1/\lambda = 40{,}000$

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & ; x \geq 0 \\ 0 & ; \text{otherwise} \end{cases} \qquad \Rightarrow \lambda = \frac{1}{40{,}000}$$

(i). $P(X \geq 20{,}000) = \int_{20{,}000}^{\infty} \frac{1}{40{,}000} \cdot e^{-\frac{1}{40{,}000}(x)} \, dx$

$=$

(ii) $P(X \leq 30{,}000) = \int_{-\infty}^{30{,}000} f(x)\, dx$

$$= \int_{-\infty}^{0} (0)\, dx + \int_{0}^{30{,}000} f(x)\, dx$$

Solution:

Let X denote the mileage obtained with the tire

$$f(x) = \frac{1}{40,000} e^{-x/40,000} \quad x > 0$$

(i) $$P(X \geq 20,000) = \int_{20,000}^{\infty} \frac{1}{40,000} e^{-x/40,000} dx$$

$$= \left[-e^{-x/40,000}\right]_{20,000}^{\infty}$$

$$= e^{-0.5} = 0.6065$$

(ii) $$P(X \leq 30,000) = \int_{0}^{30,000} \frac{1}{40,000} e^{-x/40,000} dx$$

$$= \left[-e^{-x/40,000}\right]_{0}^{30,000}$$

$$= 1 - e^{-0.75} = 0.5270$$

Exercise:

The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = 1/2$.

(a) What is the probability that the repair time exceeds 2 h?

(b) What is the conditional probability that a repair takes at least 10 h given that its duration exceeds 9 h?

HINT:

$X \sim ED(\lambda)$, with $\lambda = 1/2$

(a) $P(X > 2)$

(b) $P(X \geq 10 / X > 9) = \dfrac{P(X \geq 10, X > 9)}{P(X > 9)}$

Example:

In a certain city, the daily consumption of electric power in millions of kilowatt-hours can be treated as a RV having an Erlang distribution with parameters $\lambda = \frac{1}{2}$ and $k = 3$. If the power plant of this city has a daily capacity of 12 millions kilowatt-hours, what is the probability that this power supply will be inadequate on any given day.

Soln.

$X \sim Er(\lambda, k)$

$P(X > 12)$

Solution:

Let X represent the daily consumption of electric power (in millions of kilo-watt-hours). Then the density function of X is given as

$$f(x) = \frac{\left(\frac{1}{2}\right)^3}{\Gamma(3)} x^2 \; e^{-x/2}, x > 0$$

P(the power supply is inadequate)

$$= P(X > 12) = \int_{12}^{\infty} f(x)\, dx \quad [\because \text{The daily capacity is only 12}]$$

$$= \int_{12}^{\infty} \frac{1}{\Gamma(3)} \cdot \frac{1}{2^3} x^2 e^{-x/2} dx$$

$$= \frac{1}{16}\left[x^2\left(\frac{e^{-x/2}}{-\frac{1}{2}}\right) - 2x\left(\frac{e^{-x/2}}{\frac{1}{4}}\right) + 2\left(\frac{e^{-x/2}}{-\frac{1}{8}}\right)\right]_{12}^{\infty}$$

$$= \frac{1}{16} e^{-6} (288 + 96 + 16)$$

$$= 25\, e^{-6} = 0.0625$$

Example:

If the life X (in years) of a certain type of car has a Weibull distribution with the parameter $\beta = 2$, find the value of the parameter α, given that probability that the life of the car exceeds 5 years is $e^{-0.25}$. For these values of α and β, find the mean and variance of X.

Soln.

$X \sim WD(\alpha, \beta)$,

$\beta = 2$

$\alpha = ?$

$P(X > 5) = e^{-0.25}$

Solution:

The density function of X is given by

$$f(x) = 2\alpha x\, e^{-\alpha x^2},\ x > 0 \quad [\because \beta = 2]$$

Now $P(X > 5) = \int_5^\infty 2\alpha x\, e^{-\alpha x^2}\, dx$

$$= \left(-e^{-\alpha x^2}\right)_5^\infty$$

$$= e^{-25\alpha}$$

Given that $P(X > 5) = e^{-0.25}$

$\therefore \quad e^{-25\alpha} = e^{-0.25}$

$\therefore \quad \alpha = \frac{1}{100}$

For the Weibull distribution with parameters α and β, $E(X) = \alpha^{-1/\beta}\, \Gamma\left(\frac{1}{\beta} + 1\right)$

$\therefore$ Required mean $= \left(\frac{1}{100}\right)^{-\frac{1}{2}} \cdot \Gamma\left(\frac{3}{2}\right)$

$$= 10 \times \frac{1}{2}\, \Gamma\left(\frac{1}{2}\right)$$

$$= 5\sqrt{\pi}.$$

$$\text{Var}(X) = \alpha^{-\frac{2}{\beta}}\left[\Gamma\left(\frac{2}{\beta} + 1\right) - \left\{\Gamma\left(\frac{1}{\beta} + 1\right)\right\}^2\right]$$

$$= \left(\frac{1}{100}\right)^{-1}\left[\Gamma(2) - \left\{\Gamma\left(\frac{3}{2}\right)\right\}^2\right]$$

$$= 100\left[1 - \left(\frac{1}{2}\sqrt{\pi}\right)^2\right]$$

$$= 100\left(1 - \frac{\pi}{4}\right)$$

Exercise:

Each of the 6 tubes of a radio set has a life length (in years) which may be considered as a RV that follows a Weibull distribution with parameters $\alpha = 25$ and $\beta = 2$. If these tubes function independently of one another, what is the probability that no tube will have to be replaced during the first 2 months of service?

Example:

The marks obtained by a number of students in a certain subject are approximately normally distributed with mean 65 and standard deviation 5. If 3 students are selected at random from this group, what is the probability that at least 1 of them would have scored above 75?

Solution:

If X represents the marks obtained by the students, X follows the distribution $N(65, 5)$.

P(a student scores above 75)

$$= P(X > 75) = P\left(\frac{75-65}{5} < \frac{X-65}{5} < \infty\right)$$

$= P(2 < Z < \infty)$, (where Z is the standard normal variate)

$= 0.5 - P(0 < Z < 2)$

$= 0.5 - 0.4772$, (from the table of areas)

$= 0.0228$

Let $p = P$(a student scores above 75) $= 0.0228$ then $q = 0.9772$ and $n = 3$.

Since p is the same for all the students, the number Y, of (successes) students scoring above 75, follows a binomial distribution.

P(at least 1 student scores above 75)

$= P$(at least 1 success)

$= P(Y \geq 1) = 1 - P(Y = 0)$

$= 1 - nC_0 \times P^0 q^n$

$= 1 - 3C_0\,(0.9772)^3$

$= 1 - 0.9333$

$= 0.0667$

Example:

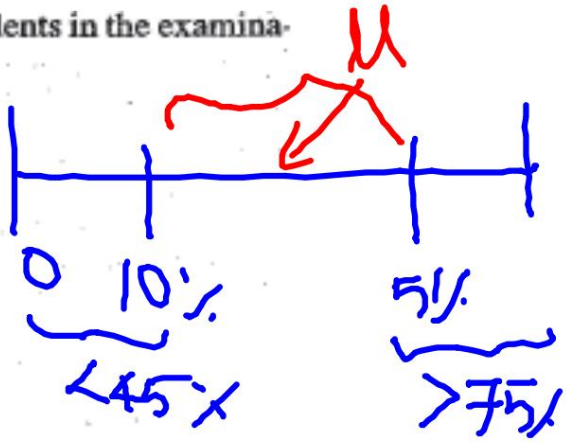
In an engineering examination, a student is considered to have failed, secured second class, first class and distinction, according as he scores less than 45%, between 45% and 60%, between 60% and 75% and above 75% respectively. In a particular year 10% of the students failed in the examination and 5% of the students got distinction. Find the percentages of students who have got first class and second class. (Assume normal distribution of marks).

Solution:

Let X represent the percentage of marks scored by the students in the examination.

Let X follow the distribution $N(\mu, \sigma)$.

Given: $P(X < 45) = 0.10$ and $P(X > 75) = 0.05$



i.e., $$P\left(-\infty < \frac{X-u}{\sigma} < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

$$P\left(\frac{75-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \infty\right) = 0.05$$

-ve

i.e., $$P\left(-\infty < Z < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

+ve

$$P\left(\frac{75-\mu}{\sigma} < Z < \infty\right) = 0.05$$

$45 < \mu < 75$

$\therefore$ $$P\left(0 < Z < \frac{\mu-45}{\sigma}\right) = 0.40 \text{ and}$$

$$P\left(0 < Z < \frac{75-\mu}{\sigma}\right) = 0.45$$

From the table of areas, we get

$$\frac{\mu-45}{\sigma} = 1.28 \quad \text{and} \quad \frac{75-\mu}{\sigma} = 1.64$$

i.e., $\mu - 1.28\,\sigma = 45$ (1)

and $\mu + 1.64\,\sigma = 75$ (2)

Solving equations (1) and (2), we get

$\mu = 58.15$ and $\sigma = 10.28$

Solution (Continued):

Now P (a student gets first class)

$$= P(60 < X < 75)$$

$$= P\left\{\frac{60 - 58.15}{10.28} < Z < \frac{75 - 58.15}{10.28}\right\}$$

$$= P\{0.18 < Z < 1.64\}$$

$$= P\{0 < Z < 1.64\} - P\{0 < Z < 0.18\}$$

$$= 0.4495 - 0.0714 = 0.3781$$

$\therefore$ Percentage of students getting first class = 38 (approximately)

Now percentage of students getting second class

= 100 – (sum of the percentages of students who have failed, got first class and got distinction)

= 100 – (10 + 38 + 5), approximately.

= 47 (approximately)

Exercise:

If the actual amount of instant coffee which a filling machine puts into '6-ounce' jars is a RV having a normal distribution with SD = 0.05 ounce and if only 3% of the jars are to contain less than 6 ounces of coffee, what must be the mean fill of these jars?

Exercise:

The local corporation authorities in a certain city install 10,000 electric lamps in the streets of the city with the assumption that the life of lamps is normally distributed. If these lamps have an average life of 1,000 burning hours with a standard deviation of 200 hours, then how many lamps might be expected to fail in the first 800 burning hours and also how many lamps might be expected to fail between 800 and 1,200 burning hours.

Exercise:

The marks obtained by a number of students in a certain subject are assumed to be approximately normally distributed with mean 55 and a SD of 5. If 5 students are taken at random from this set, then what is the probability that 3 of them would have scored marks above 60?