Subject Name: Probability and Statistics.

Unit-III

Continuous probability distributions

Topics

1. **Uniform distribution.**
2. **Exponential distribution.**
3. **Normal distribution.**
4. **Mean, variance, moment generating function and evaluation of statistical parameters for these distributions.**

Topic-I

Uniform distribution

Definition: A random variable X is said to have a continuous uniform distribution over an interval (a, b) if its probability density function is constant = k (say), over the entire range of X, i.e.,

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

Since total probability is always unity, we have

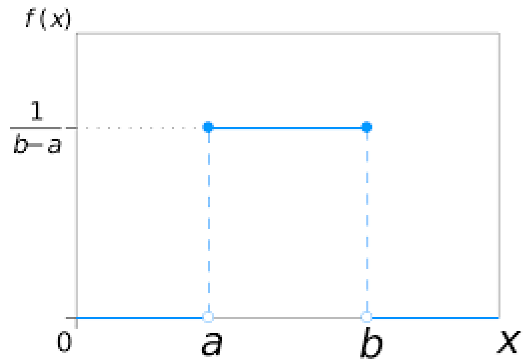
$$\int_a^b f(x)dx = 1 \Longrightarrow \int_a^b k.dx = 1$$

$$k = \frac{1}{b-a}$$

Therefore we can define uniform distribution as

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & otherwise \end{cases}$$

rfsayeed@mjcollege.ac.in



Question: Subway trains on a certain line run every half hour between mid-night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes?

Solution: Let the r.v. X denote the waiting time (in minutes) for the next train. Under the assumption that a man arrives at the station at random, X is distributed uniformly on (0, 30), with p.d.f.

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & otherwise \end{cases}$$

$$f(x) = \begin{cases} \dfrac{1}{30-0} = \dfrac{1}{30} & 0 < x < 30 \\ 0 & otherwise \end{cases}$$

The probability that he has to wait at least 20 minutes is

$$p(X > 20) = \int_{20}^{30} \frac{1}{30} dx = \frac{1}{30}[x]_{20}^{30}$$

$$p(X > 20) = \frac{1}{30}(30 - 20) = \frac{10}{30} = \frac{1}{3}$$

Moment Generating Function:

Moment Generating Function of a uniform probability distribution is given by

$$M_x(t) = \int_a^b e^{tx} f(x) dx$$

rfsayeed@mjcollege.ac.in

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

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

$$M_x(t) = \frac{1}{b-a} \left[\frac{e^{tx}}{t}\right]_a^b$$

$$\boldsymbol{M_x(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}}$$

Mean and Variance of Uniform distribution:

We know that the first moment about origin is mean

$$Mean = \mu_1' = \int_a^b x f(x) dx$$

$$Mean = \int_a^b x \frac{1}{b-a} dx$$

$$Mean = \frac{1}{b-a} \int_a^b x dx$$

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

$$\boldsymbol{Mean} = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \boldsymbol{\frac{b+a}{2}}$$

Similarly the second moment about origin is given by

$$\mu_2' = \int_a^b x^2 f(x) dx$$

$$\mu_2' = \int_a^b x^2 \frac{1}{b-a} dx$$

rfsayeed@mjcollege.ac.in

$$\mu_2' = \frac{1}{b-a}\int_a^b x^2\,dx$$

$$\mu_2' = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$\mu_2' = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)(b^2+a^2+ab)}{2(b-a)}$$

$$\mu_2' = \frac{b^2+a^2+ab}{\mathbf{3}}$$

Now we evaluate variance of uniform distribution as

$$V(X) = \mu_2' - (\mu_1')^2$$

$$V(X) = \mu_2 = \frac{b^2+a^2+ab}{\mathbf{3}} - \left(\frac{b+a}{2}\right)^2$$

$$V(X) = \mu_2 = \frac{b^2+a^2+ab}{\mathbf{3}} - \frac{b^2+a^2+2ab}{4}$$

$$V(X) = \frac{4b^2+4a^2+4ab-(3b^2+3a^2+6ab)}{12}$$

$$V(X) = \frac{b^2+a^2-2ab}{12} = \frac{(b-a)^2}{12}$$

Question: If X is uniformly distributed with mean 1 and variance 4/3, then find P(X<0).

Solution: Let X be a random variable uniformly distributed and given that

$$Mean = 1 = \frac{b+a}{2}$$

$$b + a = 2 \text{ ---------------(1)}$$

$$V(X) = \frac{4}{3} = \frac{(b-a)^2}{12}$$

$$(b-a)^2 = \frac{48}{3} = 16$$

$$(b-a)^2 = 16$$

$$b - a = \pm 4\text{-------------}(2)$$

On solving equation (1) and (2), we get

$$a = -1 \text{ and } b = 3$$

Then p.d.f of uniform distribution is given by

$$f(x) = \int_a^b \frac{1}{b-a} dx$$

$$f(x) = \frac{1}{b-a} = \frac{1}{3-(-1)} = \frac{1}{4}$$

$$P(X<0) = \int_{-1}^{0} f(x)dx = \int_{-1}^{0} \frac{1}{4} dx = \frac{1}{4} * 1 = \frac{1}{4}$$

Topic-II

Exponential Distribution

Definition: A continuous random variable X assuming non-negative values is said to have an exponential distribution with parameter θ>0, if its p.d.f. is given by

$$f(x) = \begin{cases} \theta * e^{-\theta x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Moment Generating Function:

Moment Generating Function of an exponential probability distribution is given by

$$M_x(t) = \int_0^\infty e^{tx} f(x) dx$$

$$M_x(t) = \int_0^\infty e^{tx} (e^{-\theta x})\, dx$$

rfsayeed@mjcollege.ac.in

$$M_x(t) = \theta \int_0^{\infty} e^{tx} \mathrm{e}^{-\theta \mathrm{x}}\, dx$$

$$M_x(t) = \theta \int_0^{\infty} e^{-(\theta - t)x}\, dx$$

$$M_x(t) = \theta \left[\frac{e^{-(\theta - t)x}}{\theta - t} \right]_0^{\infty}$$

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

$$\boldsymbol{M_x(t)} = \theta \left[\frac{0}{\theta - t} - \frac{1}{\theta - t} \right]$$

$$\boldsymbol{M_x(t)} = \left[\frac{\theta}{\theta - t} \right]$$

$$\boldsymbol{M_x(t)} = \left[\frac{1}{1 - \frac{t}{\theta}} \right]$$

$$\boldsymbol{M_x(t)} = \left(1 - \frac{t}{\theta}\right)^{\boldsymbol{-1}}$$

$$\boldsymbol{M_x(t)} = \boldsymbol{1} + \frac{t}{\theta} + \left(\frac{t}{\theta}\right)^2 + \left(\frac{t}{\theta}\right)^3 \cdots\cdots\cdots\cdots \infty$$

Mean and Variance of exponential distribution:

We know that the first moment about origin is

$$Mean = \mu_1' = \left(\frac{dM_x(t)}{dt} \right)_{t=0}$$

$$Mean = \mu_1' = \left(\frac{d}{dt} \left(1 - \frac{t}{\theta}\right)^{-1} \right)_{t=0}$$

$$Mean = \mu_1' = \left(-\left(1-\frac{t}{\theta}\right)^{-2}\left(-\frac{1}{\theta}\right)\right)_{t=0}$$

$$Mean = \mu_1' = \left(-\left(-\frac{1}{\theta}\right)\right) = \frac{1}{\theta}$$

$$\boldsymbol{Mean = \mu_1' = \frac{1}{\theta}}$$

Similarly

$$\mu_2' = \left(\frac{d^2}{dt^2}M_x(t)\right)_{t=0}$$

$$\mu_2' = \left(\frac{d^2}{dt^2}\left(1-\frac{t}{\theta}\right)^{-1}\right)_{t=0}$$

$$\mu_2' = \left(-1*-2*\left(1-\frac{t}{\theta}\right)^{-3}\left(-\frac{1}{\theta}\right)\left(-\frac{1}{\theta}\right)\right)_{t=0}$$

$$\mu_2' = -1*-2*\left(-\frac{1}{\theta}\right)\left(-\frac{1}{\theta}\right) = \frac{2}{\theta^2}$$

$$\boldsymbol{\mu_2' = \frac{2}{\theta^2}}$$

Now we evaluate variance of exponential distribution as

$$V(X) = \mu_2' - (\mu_1')^2$$

$$V(X) = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$$

$$\textbf{Variance} = \boldsymbol{V(X) = \frac{1}{\theta^2}}$$

Topic-III

Normal distribution

(The normal model has nevertheless become the most important probability model in statistical analysis)

Definition: A random variable X is said to have a normal distribution with parameters μ (called "mean") and σ^2 (called "variance") if its density function is give by the probability law:

$$\boldsymbol{f(x;\mu,\sigma) = N(\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}}$$

$(-\infty < x < +\infty)$

As well as $(-\infty < \mu < +\infty)$ and $\sigma > 0$

Note: Total probability of normal distribution

Since total probability is always unity, we have

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

Standard Normal Variable:

If $X = \boldsymbol{N(\mu,\sigma)}$ then $Z = \boldsymbol{\frac{x-\mu}{\sigma}}$ is a standard normal variable with mean (Z) = 0 and Variance (Z) = 1.

If $X = N(\mu,\sigma)$ then $Z = N(0,1)$

Definition: A random variable Z is said to have a standard normal distribution (mean is 0, variance is 1) if its density function is give by the probability law:

$$\boldsymbol{\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}(z)^2\right)}} \qquad (-\infty < z < +\infty)$$

Note: Total probability of standard normal distribution

Since total probability is always unity, we have

$$\int_{-\infty}^{+\infty} \varphi(z)dz = 1 \Longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\left(-\frac{1}{2}(z)^2\right)} dz = 1$$

rfsayeed@mjcollege.ac.in

Chief characteristics of the normal distribution and normal probability curve.

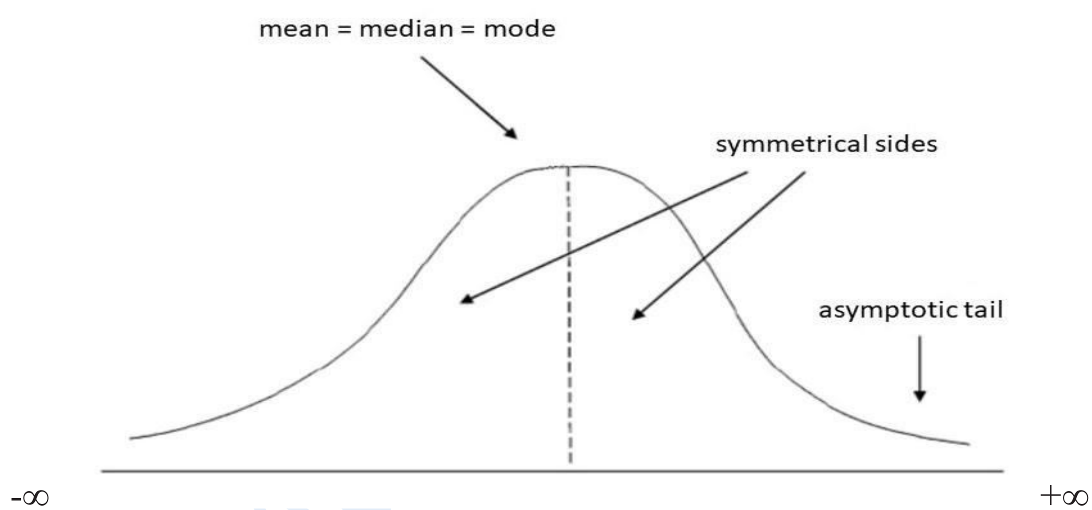
The normal probability curve with mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ is given by the equation

$$\boldsymbol{f(x;\mu,\sigma) = N(\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}}$$

$$(-\infty < x < +\infty)$$

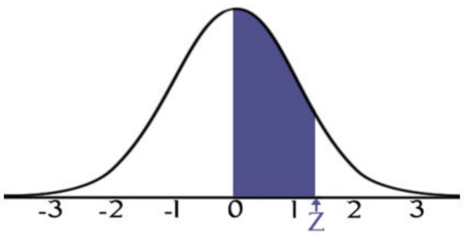
and has the following properties:

1. The curve is bell shaped and symmetrical about the line $x = \mu$



$$x = \mu \; or \; z = 0$$

2. Mean, median and mode of the distribution coincide.
3. As x increases numerically, f(x) decreases rapidly, the maximum probability occurring at the point $x = \mu$
4. Since *f(x)* being the probability, can never be negative, no portion of the curve lies below the x-axis.
5. Skewness= $\beta_1 = 0$(Not skewed)
6. Kurtosis= $\beta_2 = 3$
7. X-axis is an asymptote to the curve.



STANDARD NORMAL TABLE (Z)

Entries in the table give the area under the curve between the mean and z standard deviations above the mean. For example, for $z = 1.25$ the area under the curve between the mean (0) and z is 0.3944.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

Problem: X is a normal variate with mean 12 and standard deviation 4, find probabilities of

(i) X>20, (ii) X<20, (iii) 0<X<12

(iv) Find x', when P(X>x')>0.24

Solution: Given that X is a random variable follow normal distribution where $\mu = 12$ and $\sigma = 4$

(i) P(X>20),
Then we find for X=20, the value of standard normal variable
$Z = \frac{x-\mu}{\sigma} = \frac{20-12}{4} = \mathbf{2}$

$$P(X > 20) = P(Z > 2) = \int_2^\infty \varphi(z)dz$$

$$= \int_0^\infty \varphi(z)dz - \int_0^2 \varphi(z)dz =$$

$$0.5 - 0.4772 = 0.0288$$

(ii) P(X<20),
Then we find for X=20, the value of standard normal variable
$Z = \frac{x-\mu}{\sigma} = \frac{20-12}{4} = \mathbf{2}$

$$P(X < 20) = P(Z < 2) = \int_{-\infty}^2 \varphi(z)dz$$

$$= \int_{-\infty}^0 \varphi(z)dz + \int_0^2 \varphi(z)dz =$$

$$0.5 + 0.4772 = 0.9772$$

(iii) P(0<X<12),
Then we find for X=0, the value of standard normal variable
$Z = \frac{0-12}{4} = -\mathbf{3}$
Then we find for X=12, the value of standard normal variable
$Z = \frac{12-12}{4} = \mathbf{0}$

rfsayeed@mjcollege.ac.in

$$P(0 < X < 12) = P(-3 < Z < 0) = \int_{-3}^{0} \varphi(z)dz$$

$$= \int_{0}^{3} \varphi(z)dz = 0.4987$$

Example2:

The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year?

Given data is $\mu = 4300$ and $\sigma = 750$

At $x = 2500, z = \frac{2500-4300}{750} = -2.40$

Similarly At $x = 4200, z = \frac{4200-4300}{750} = -0.133$

Now $P(2500 < X < 4200) = P(-2.40 < Z < -0.133)$

$$P(-2.40 < Z < -0.133) = \int_{-2.40}^{-0.133} \varphi(z)dz$$

$$P(-2.40 < Z < -0.133) = \int_{0.133}^{2.40} \varphi(z)dz$$

$$P(-2.40 < Z < -0.133) = \int_{0}^{2.40} \varphi(z)dz - \int_{0}^{0.133} \varphi(z)dz$$

$$P(2500 < X < 4200) = P(-2.40 < Z < -0.133)$$

$$P(2500 < X < 4200) = 0.4488 - 0.0082 = 0.4401$$

Example3:

The time taken for a data file to travel from the source to the company is Gaussian distributed. If 6.8% of the files take over 200 ms, and 3.0% take under 140 ms to complete the journey, then find out the mean and standard deviation of the distribution.

Solution: Given that time taken of transfer of data file is normally distributed i.e. $X(\mu, \sigma)$.

Given that at $x = 200, z = \frac{x-\mu}{\sigma} = \frac{200-\mu}{\sigma}$

$$P(X > 200) = P\left(Z > \frac{200-\mu}{\sigma}\right) = 0.5 - P(0 < Z < \frac{200-\mu}{\sigma}) = 0.068$$

$$0.5 - \varphi\left(\frac{200-\mu}{\sigma}\right) = 0.068$$

$$0.5 - 0.068 = \varphi\left(\frac{200-\mu}{\sigma}\right)$$

$$\varphi\left(\frac{200-\mu}{\sigma}\right) = 0.43$$

Using standard normal table $\frac{200-\mu}{\sigma} = 1.49$---------------(1)

Similarly at $x = 140, z = \frac{x-\mu}{\sigma} = \frac{140-\mu}{\sigma}$

$$P(X < 140) = P\left(Z < \frac{140-\mu}{\sigma}\right) = P\left(Z > \frac{140-\mu}{\sigma}\right) = 0.03$$

$$0.5 - P\left(0 < Z < \frac{140-\mu}{\sigma}\right) = 0.03$$

$$P\left(0 < Z < \frac{140-\mu}{\sigma}\right) = 0.5 - 0.03 = 0.47$$

$$\varphi\left(\frac{140-\mu}{\sigma}\right) = 0.47$$

$\frac{140-\mu}{\sigma} = 1.88$-----------------(2)

Solving equation (1) and (2) we get $\mu = 173.5 \quad and \quad \sigma = 17.8$