[Academic Script]

Continuous Distributions

Subject:	Business Economics
Course:	B. A. (Hons.), 1st Semester, Undergraduate
Paper No. & Title:	Paper – 102 Statistics for Business Economics
Unit No. & Title:	Unit – 5 Probability and Distribution
Lecture No. & Title:	Lecture – 3 Continuous Distributions

Academic Script

1. Introduction

Continuous Distributions

In the discrete case, the random variable takes a limited number of values, but there are many situations where the variable of interest may take infinitely many values. i.e. all the values within a certain interval. i.e. A random variable X which can assume any value in R or an interval (a, b), of R with a < b is called a continuous random variable and corresponding distribution is called continuous distribution.

e.g.: (1) Weight of a person

(2) Maximum temperature of the year

(3) Breaking strength of yarn.

In case of such continuous random variable, we do not talk of probability at a particular point(which is always zero) but we always talk of probability in an interval.

If f(x)dx is the probability that the random variable X takes the values in small interval of magnitude dx, eg. .$\left(x-\frac{dx}{2}, x+\frac{dx}{2}\right)$ then f(x) is called the probability density function(pdf) of the random variable X.

Here $\int_{-\infty}^{\infty} f(x)dx$ must be equal to one.

e.g. Let f(x)=$\frac{1}{3}$; 2 < X < 5.

we can see here $\int_{2}^{5} f(x)dx$ = $\int_{2}^{5}\frac{1}{3}dx = \frac{1}{3}\left[x\right]_{2}^{5}$ = 1.

Thus f(x) is a pdf of a random variable X.

In earlier talk we have studied some discrete distribution. Now we will discuss some continues distributions. For continuous distribution a random variable assumes values in an interval like (a, b), a < b; (0, ∞) or (-∞, ∞) etc.

2. Continuous Uniform Distribution

Consider the random variable *X* representing the Volvo Luxury bus time of travelling from Ahmedabad to Rajkot. Suppose the bus time can be any value in the interval from 200 minutes to 230 minutes. Because the random variable *X* can assume any value in that interval, *X is* a continuous rather than a discrete random variable. Let us assume that the probability of a bus time within any 1-minute interval is the same as the probability of a bus time within any other 1-minute interval contained in the larger interval from 200 to 220 minutes. With every 1-minute interval being equally likely, the random variable *X* is said to have a uniform probability distribution.

The probability density function, which defines the uniform distribution for the bus-time random variable, is

$$f(x) = \frac{1}{20}, \qquad 200 \le x \le 220$$

The probability density function of a uniform distribution over the interval (a, b) is given by

$$f(x) = \frac{1}{b-a}, \qquad a \le x \le b$$

The mean and variance of the distribution are given as follows:

$$\text{Mean} = \frac{(b+a)}{2}$$

and

$$Variance = \frac{(b-a)^2}{12}$$

For the above example average bus time becomes 210 minutes and its variance 400/12 = 33.3333 minutes.

For this distribution for any constants c and k , $a \le c \le k \le b$ the probability

i. $P(X \le c) = \frac{c-a}{b-a}$

ii. $P(X \ge c) = \frac{b-c}{b-a}$

iii. $P(c \le X \le k) = \frac{k-c}{b-a}$

Example 1. The probability density function of the time Required to complete an assembly operation is f(x) = 1/15 for 30 < x < 45 seconds.

(a) Determine the proportion of assemblies that requires more than 35 seconds to complete.

(b) Determine the mean and variance of time of assembly.

Solution:

Suppose X denotes the time required to complete an assembly operation then,

(a) The proportion of assemblies that requires more than 35 seconds to complete = P(X>35)

Using the above results we say that

P(X>35) = (45-35)/(45-30) = 2/3

(b) Mean = (45+30)/2 = 37.5 seconds and

Variance = $(45-30)^2$/ 12 = 225/12 seconds

3. Exponential Distribution

The exponential distribution has been found useful for the time duration of telephone calls, breaking strength of yarn, life of electrical components. It is widely used as inter arrival time distribution in queuing theory.

The probability density function of the exponential distribution is given as

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

θ is known as number of events happened per time unit.

The mean and variance of the distribution are given by

Mean = $1/\theta$ and variance = $1/\theta^2$.

For this distribution

i. $P(X \le c) = 1 - e^{-c\theta}$

ii. $P(X \geq c) = e^{-c\theta}$

iii. For a < b, $P(a \leq X \leq b) = e^{-a\theta} - e^{-b\theta}$

iv. Variance = (mean)^2

Example 1. A highway petrol pump can serve on an average 10 cars per hour. What is the probability that for a particular car, the time taken will be (i) less than 4 minutes? (ii) more than 2 minutes?

Here we assume exponential distribution for the time taken by the petro pump for a car.

Here mean time taken is 1/θ = 1/10 = 0.1 hours = 6 minutes.

Let us assume X = time taken in minute to serve a car

We are interested in finding

(i) P(X <4) , where $= 1 - e^{-4/6} = 1 - e^{-2/3}$ = 0.4866

(ii) P(X > 2) $= e^{-2/6}$ = 0.7165

Example 2 .

The daily consumption of milk (in litre) per family in a given city is approximately exponentially distributed with mean 3.0 litres. What is the probability that the selected family has the requirement of milk on a particular day is (i) between 3 to 5 litres, (ii) more than five litres?

Let X denotes the daily consumption of milk (in litre) per family in a given city. Then the pdf of X is given as

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

Here mean consumption = 1/θ = 3 litters. i. e. θ = 1/3

(i) We required probability that the selected family has the requirement of milk on a particular day is between 3 to 5 litres. That is

$$P(3 \le X \le 5) = F(5) - F(3) = (1 - e^{-5\theta}) - (1 - e^{-3\theta})$$
$$= e^{-1} - e^{-5/3}$$
$$= 0.179004$$

(ii) probability that the selected family has the requirement of milk on a particular day is more than five litres. That is

$$P(X > 5) = e^{-5\theta} = e^{-5/3}$$
$$= 0.188876$$

4. Normal distribution

The normal distribution is the most versatile of all the continuous probability distributions. It is found useful in the distribution of errors in Astronomy, in the theory of accidental errors of measurements involved in the calculation of orbits of heavenly bodies.

It is also useful in statistical theory of inference and in approximating other probability distributions. The dimensions of a product like height, weight etc are found to be normally distributed continuous random variables.

The probability density function of the normal distribution is given by

$$f(x;\, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2};$$
$$-\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0.$$

where $\pi \approx 3.14159$ and e = 2.71828 are constants and μ *and* σ are the parameters of the distribution. Here μ is mean and σ is the standard deviation (σ^2 is the variance) of the normal distribution.

When $\mu = 0$ *and* $\sigma = 1$ the distribution is known as standard normal distribution and its probability density function reduces to

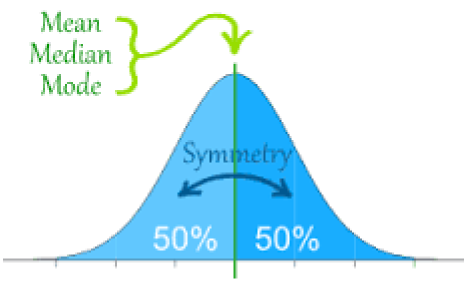
$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2}, -\infty < z < \infty$ and $z = \frac{x-\mu}{\sigma}$ is called standard normal variate.

Properties of the normal distribution:

(i) The distribution is symmetric about mean μ.

(ii) The curve of the probability density function is bell shape and the right and left tails of the curve extend indefinitely without touching the horizontal line.

(iii) Mean = median = mode.



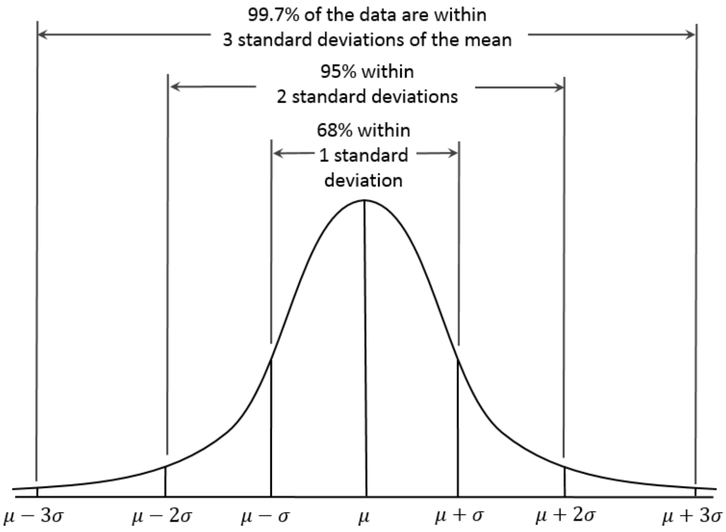
(iii) The values of the areas of the normal curve on both sides of the vertical line X = μ are equal and each is equal to 0.5

(iv) $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.6826$

(v) $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9545$

(vi) $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9973$

Area under the standard normal curve:



Example 1. In a manufacturing organization the distribution of wages was perfectly normal and the number of workers employed in the organization was 3000. The mean wages of the workers were found as Rs. 5500 per month with standard deviation of Rs. 500. On the basis of this information estimate the following.

(i) The number of workers getting salary between Rs. 6400 to Rs. 6800

(ii) The number of workers getting salary above Rs. 6800

(iii) The number of workers getting salary below Rs. 5800.

Let X is the salary per month of a worker in the given organization. Here $X \sim N(\mu, \sigma^2), \mu = 5500 \; and \; \sigma = 500$.

(i) To find out the number of workers getting salary between Rs. 6400 to Rs. 6800, we first find out the probability of this event. That is

$$P(6400 < X < 6800).$$

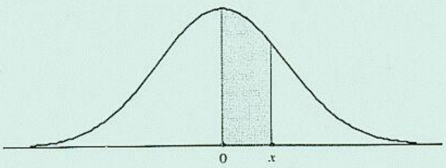
$$= P(\frac{6400-5500}{500} < \frac{X-\mu}{\sigma} < \frac{6800-5500}{500})$$

$= P(1.8 < Z < 2.6)$

Now, from the table of standard normal distribution we find that $P(0 < Z < 1.8) = 0.4641$ and

$P (0 < Z < 2.6) = 0.4953$

Appendix 4 – Values of the Standard Normal distribution



The table gives the probability that a standard Normal variable lies between 0 and x (which is equivalent t shaded area on the figure).

x	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	0.0199	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	0.0754
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	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.2	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.3	0.2881	0.2910	0.2939	0.2967	0.2996	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	0.3531	0.3554	0.3577	0.3599	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	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
0.4987	0.4990	0.4993	0.4995	0.4997	0.4998	0.4998	0.4999	0.4999	0.5000

Hence the required probability can be obtained as

$$P(1.8 < Z < 2.6) = 0.4953 - 0.4641 = 0.0312 .$$

Thus, the number of workers getting the salary between Rs 6400 to Rs. 6800 = Total number of workers × probability

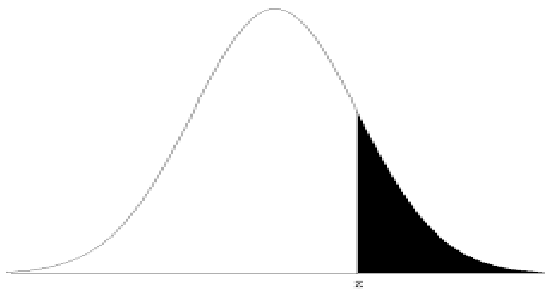
$= 3000 \times 0.0312$

$= 93.6 \approx 94$ workers.

(ii) To find out the number of workers getting salary above Rs. 6800, we first find out the probability of this event. That is

$$P(X > 6800) = P\left(\frac{X-\mu}{\sigma} > \frac{6800-5500}{500}\right)$$

$$= P(Z > 2.6)$$



From the table of standard normal distribution we find the probability

$$= 0.5 - 0.4953$$

$$= 0.0047$$

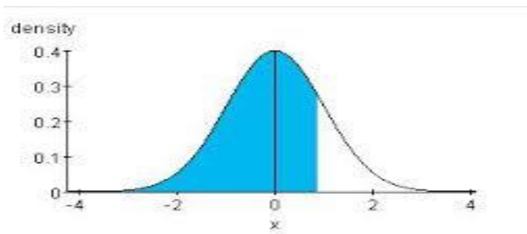
Hence the number of workers getting the salary more than Rs. 6800 $= 0.0047 \times 3000$

$$= 14.1 \approx 14 \text{ workers.}$$

(iii) To find out the number of workers getting salary below Rs. 5800. We first find out the probability of this event. That is

$$P(X < 5800) = P\left(\frac{X-\mu}{\sigma} < \frac{5800-5500}{500}\right)$$

$$= P(Z<0.60)$$



From the table of standard normal distribution we find the probability

$$= 0.5 + 0.2257 = 0.7257$$

Hence the number of workers getting the salary less than Rs. 5800

$= 0.7257 \times 3000 = 2177.1 \approx 2177$ workers.

Example 2. A minimum height is to be prescribed for eligibility to government services such that 75.80% of the young men will have a fair chance of coming up to that standard. The height of young men is normally distributed with mean 62 inches and standard deviation 3.5 inches. Determine the minimum specification.

Let X be the height of young man. Here $X \sim N(\mu, \sigma^2)$, $\mu = 62$ and $\sigma = 3.5$

We wish to find minimum specification (height) K such that 75.80% of the young men will have a fair chance of coming up to that standard. That is

$$P(X > K) = 0.7580$$

$$=> \quad P\left(\frac{X-\mu}{\sigma} > \frac{K-62}{3.5}\right) = 0.7580$$

$$=> \quad P(Z > Z1) = 0.7580, \quad \text{where} \quad Z1 = \frac{K-62}{3.5}.$$

From the table of standard normal distribution we find that

Z1 = -0.70.

Thus, $\frac{K-62}{3.5} = -0.70$ => K = 62 – 0.70(3.5) = 59.55.

That is minimum height should be 59.55 inches for 75.80% of the young men to be eligible for the government services.

Example 3. According to the sleep foundation, the average night's sleep is 6.8 hours (Fortune, March 20, 2006).
Assume the standard deviation is 0.6 hours and that the probability distribution of sleeping hours is normal,

(a) What is the probability that a randomly selected person sleeps more than 8 hours?

(b) What is the probability that a randomly selected person sleeps 6 or less hours?

(c) Doctor suggests getting between 7 and 8 hours sleep each night, what percentage of the population gets this much sleep?

Let X = sleeping hours of a person at night. Here $X \sim N(\mu, \sigma^2)$, $\mu = 6.8$ and $\sigma = 0.6$.

(a) Probability that a randomly selected person sleeps more than 8 hours

= P(X > 8)

= P($\frac{X-\mu}{\sigma} > \frac{8-6.8}{0.6}$)

= P(Z > $\frac{1.2}{0.6}$)

= P(Z > 2)

From the table of standard normal distribution we find the probability

$= 0.0228$

(b)Probability that a randomly selected person sleeps 6 or less hours

$= P(X < 6)$

$= P(\frac{X-\mu}{\sigma} < \frac{6-6.8}{0.6})$

$= P(Z < -1.3333)$

From the table of standard normal distribution we find the probability

$= 0.0918$

(c) Probability of getting between 7 and 8 hours sleep each night

$= P(7 < X < 8)$

$= P(\frac{7-6.8}{0.6} < \frac{X-\mu}{\sigma} < \frac{6-6.8}{0.6})$

$= P(0.3333 < Z < 2)$

From the table of standard normal distribution we find the probability

$= 0.3479$

Remarks:

(1)Normal approximation to a binomial probability distribution.

When the number of trials becomes large, evaluating the binomial probability by manually becomes difficult. When np ≥ 5 or n(1-p) ≥ 5, the normal distribution becomes a good approximation over the binomial distribution. We use μ = np and σ = $\sqrt{np(1-p)}$.

(2) Normal approximation to a Poisson distribution. When λ, mean of Poisson distribution becomes very large, the normal distribution becomes a good approximation over the Poisson distribution. Here we use μ = λ and σ = $\sqrt{\lambda}$.

5. Summary

We have studied the continuous distributions like, continuous uniform, exponential and normal distributions. We have also seen the real life situations where such distributions can be employed. In uniform distribution probability remains uniform over the given interval. Exponential distribution is a skewed distribution where as normal distribution is a symmetric distribution. It possess many important properties.