MAT2001-Moudule-4-Distributions

Dr. Nalliah M

Assistant Professor

Department of Mathematics

School of Advanced Sciences

Vellore Institute of Technology

Vellore,Tamil Nadu,India.

nalliah.moviri@vit.ac.in

September 6, 2020

Introduction

No matter whether a discrete probability distribution is represented graphically by a histogram, in tabular form, or by means of a formula, the behavior of a random variable is described. Often, the observations generated by different statistical experiments have the same general type of behavior. Consequently, discrete random variables associated with these experiments can be described by essentially the same probability distribution and therefore can be represented by a single formula. In fact, one needs only a handful of important probability distributions to describe many of the discrete random variables encountered in practice.

Intoduction

Such a handful of distributions describe several real-life random phenomena. For instance, in a study involving testing the effectiveness of a new drug, the number of cured patients among all the patients who use the drug approximately follows a binomial distribution.

In an industrial example, when a sample of items selected from a batch of production is tested, the number of defective items in the sample usually can be modeled as a hypergeometric random variable.

In a statistical quality control problem, the experimenter will signal a shift of the process mean when observational data exceed certain limits. The number of samples required to produce a false alarm follows a geometric distribution, which is a special case of the negative binomial distribution.

Intoduction

On the other hand, the number of white cells from a fixed amount of an individuals blood sample is usually random and may be described by a Poisson distribution.

In this module, we study the following distributions

- Binomial distribution
- Poission distribution
- Normal distribution
- Gamma distribution
- Exponential distribution
- Weibull distribution

Binomial distribution

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled success or failure. One obvious application deals with the testing of items as they come off an assembly line, where each trial may indicate a defective or a nondefective item. We may choose to define either outcome as a success. The process is referred to as a Bernoulli process. Each trial is called a Bernoulli trial.

Example: Observe that, for example, if one is drawing cards from a deck, the probabilities for repeated trials change if the cards are not replaced.

That is,the probability of selecting a heart on the first draw is $1/4$, but on the second draw it is a conditional probability having a value of $13/51$ or $12/51$, depending on whether a heart appeared on the first draw: this, then, would no longer be considered a set of Bernoulli trials.

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

- The experiment consists of repeated trials.
- Each trial results in an outcome that may be classified as a success or a failure.
- The probability of success, denoted by p, remains constant from trial to trial.
- The repeated trials are independent.

Consider the set of Bernoulli trials where three items are selected at random from a manufacturing process, inspected, and classified as defective or nondefective. A defective item is designated a success. The number of successes is a random variable X assuming integral values from 0 through 3. The eight possible outcomes and the corresponding values of X are

Outcome	NNN	NDN	NND	DNN	NDD	DND	DDN	DDD
x	0	1	1	1	2	2	2	3

Since the items are selected independently and we assume that the process produces 25% defectives, we have

$P(NDN) = P(N)P(D)P(N) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{9}{64}$.

Similar calculations yield the probabilities for the other possible outcomes.

The probability distribution of X is therefore

x	0	1	2	3
P(X=x)	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

Binomial Distribution

The number X of successes in n Bernoulli trials is called a binomial random variable. The probability distribution of this discrete random variable is called the binomial distribution, Let us now generalize the above illustration to yield a formula for binomial distribution. That is, we wish to find a formula that gives the probability of x successes in n trials for a binomial experiment.

First, consider the probability of x successes and $n - x$ failures in a specified order.

Since the trials are independent, we can multiply all the probabilities corresponding to the different outcomes.

Each success occurs with probability p and each failure with probability $q = 1 - p$.

Therefore, the probability for the specified order is $p^x q^{n-x}$.

We must now determine the total number of sample points in the experiment that have x successes and $n - x$ failures.

This number is equal to the number of partitions of n outcomes into two groups with x in one group and $n - x$ in the other and is written $\binom{n}{x}$. Because these partitions are mutually exclusive, we add the probabilities of all the different partitions to obtain the general formula, or simply multiply $p^x q^{n-x}$ by $\binom{n}{x}$.

Binomial Distribution

A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \ldots, n.$$

Where Does the Name Binomial Come From?

The binomial distribution derives its name from the fact that the $n + 1$ terms in the binomial expansion of $(q + p)^n$ correspond to the various values of $P(X = x)$ for $x = 0, 1, 2, \ldots, n$. That is,

$$\begin{aligned}(q+p)^n =& \binom{n}{0} q^n + \binom{n}{1} pq^{n-1} + \binom{n}{2} p^2 q^{n-2} +, \ldots, + \binom{n}{n} p^n. \\ &= P(X=0) + P(X=1) + P(X=2) + \ldots, + P(X=n).\end{aligned}$$

Since $p + q = 1$, we get

$$\sum_{x=0}^{n} P(X = x) = 1,$$

a condition that must hold for any probability distribution.

Frequently, we are interested in problems where it is necessary to find $P(X < r)$ or $P(a \leq X \leq b)$. Binomial sums

$$P(X < r) = \sum_{x=0}^{r} P(X = x).$$

Theorem: The mean and variance of the binomial distribution $b(x; n, p)$ are $\mu = np$ and $\sigma^2 = npq$.

A binomial experiment consist of n idependent trials. -If we consider N sets of n independent trials, then the number of times we get x success is $N \bullet \binom{n}{x} p^x q^{n-x}$, which gives the expected number of successes. This is called the expected frequency of n sucesses in N experiments, each conisisting of n trials.

Problem: For a binomial distribution with parameters n=5, p=0.3. Find the probabilities of getting

- at least 3 successes
- at most 3 Sucesses
- exactly 3 failures.

Solution: Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \ldots, n.$$

Here $n = 5, p = 0.3$.

- $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) = 0.1631,$
- $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.9692$
- $P(X = 2 \text{ sucesses}) = P(X = 3 \text{ failures}) = 0.3087.$

Problem: If the chance of running a bus service according to schedule is 0.8, Calculate the probability on a day schedule with 10 services:

- exactly one is late
- at least one is late.

Solution: Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \ldots, n.$$

Probability of a bus running according to schedule=0.8=q

therefore, the probability that a bus is late is 0.2=p

Here $n = 10, p = 0.2$.

- $P(X = 1) = \binom{10}{1}(0.2)^1(0.8)^{10-1} = 0.2684$
- $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$
  $1 - \binom{10}{0}(0.2)^0(0.8)^{10-0} = 0.8926$

Problem: The 10% of the screws produced by an automatic machine are defective, find the probability that os 20 screws selected at random, there are

- exactly two defectives
- at the most three defectives
- at least two defectives and
- between one and three defectives(inclusive).

Problem: 4 coins are tossed and number of heads noted. The experiment is repeated 200 times and the following distribution is obtained.

X:Number of heads	0	1	2	3	4
f:frequencies	62	85	40	11	2

Find the binomial distribution and expected frequency distribution.

Solution:Let X be a binomial random variable with parameters n and p. Then

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, 3, \ldots, n.$$

Here$n = 4$ and $N = 200$. Since the mean of the binomial distribution is np.

To find mean:

$$\text{Mean} = \frac{\sum_{i=0}^{4} f_i x_i}{\sum f_i} = \frac{206}{200}$$

$$np = 1.03$$

$$4p = 1.03$$

$$p = 0.257$$

Since $q = 1 - p = 1 - 0.257 = 0.743$

X	f	fx	Expected frequency$= N \bullet P(X = x)$	
0	62	0	$200 \bullet \binom{4}{0}(0.257)^0(0.743)^{4-0} = 60.95$	60
1	85	85	$200 \bullet \binom{4}{1}(0.257)^1(0.743)^{4-1} = 84.33$	84
2	40	80	$200 \bullet \binom{4}{2}(0.257)^2(0.743)^{4-2} = 43.75$	44
3	11	33	$200 \bullet \binom{4}{3}(0.257)^3(0.743)^{4-3} = 10.8$	11
4	2	8	$200 \bullet \binom{4}{4}(0.257)^4(0.743)^{4-4} = 0.872$	1
	N=200	206		200

Poisson Distribution and the Poisson Process

Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called Poisson experiments.

The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

Example: A Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, the number of days school is closed due to snow during the winter, or the number of games postponed due to rain during a baseball season.

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

In such instances, X might represent the number of field mice per acre, the number of bacteria in a given culture, or the number of typing errors per page. A Poisson experiment is derived from the Poisson process and possesses the following properties.

- The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region. In this sense we say that the Poisson process has no memory.

- The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.
- The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Poisson Distribution: The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval.

The formula for the Poisson probability mass function is
$P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, 3, \ldots,$ where λ is the average number of outcomes. Mean $= \lambda =$ variance; so that standard deviation $= \sqrt{\lambda}$.

Here λ is known as parameter of the distribution so that $\lambda > 0$. Since number of trials is very large and the probability of success p is very small, it is clear that the event is a rare event. Therefore, Poisson distribution relates to rare events.

Examples

1. The number of blinds born in a town in a particular year.
2. Number of mistakes committed in a typed page.
3. The number of students scoring very high marks in all subjects
4. The number of plane accidents in a particular week.
5. The number of defective screws in a box of 100, manufactured by a reputed company.
6. Number of suicides reported in a particular day

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

- The number of trials n is indefinitely large i.e., $n \to \infty$.
- The probability of success p for each trial is very small; i.e., $p \to 0$.
- $np = \lambda$ (say) is finite, $\lambda > 0$.

Problem: Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year?

Solution: Clearly, the event of this problem is rare event. Therefore, the randaom variable X follows a Poisson distribution.

Here $p = P(\text{fire accident during a year}) = \frac{1}{1000} = 0.001$ and n=total number of houses=2000.

Therefore, $\lambda = np = 2$. Hence $P(X = 5) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2}2^5}{5!}$.

Problem: If 2% of electric bulbs manufactured by a certain company are defective. Find the probability that in a sample of 200 bulbs

- less than 2 bulbs defective.
- more than 3 bulbs are defective.

Problem: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Problem: In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

- What is the probability that in any given period of 400 days there will be an accident on one day?
- What is the probability that there are at most 3 days with an accident?

Problem: 100 car radios are inspected as they come off the production line and number of defects per set is recorded below

Number of defects	0	1	2	3	4
Number of sets	79	18	2	1	0

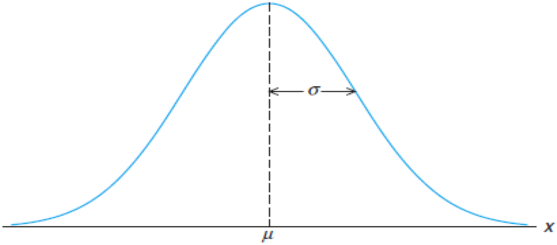
Fit a Poisson distribution and find expected frequencies.

Normal Distribution

We have discussed the discrete distributions, the Binomial and Poisson distribution. The normal probability distribution or simply normal distribution-continuous distribution was first discovered-DeMoivre (English Mathematician) in 1733 as limiting case of binomial distribution. Later it was applied in natural and social science-Laplace (French Mathematician) in 1777. The normal distribution is also known as Gaussian distribution in honour of Karl Friedrich Gauss(1809).

A continuous random variable X having the bell-shaped distribution of Figure is called a normal random variable. The mathematical equation for the probability distribution of the normal variable depends on the two parameters μ and σ, its mean and standard deviation, respectively.

Hence, we denote the values of the density of X by $f(x)$.



Definition: A continuous random variable X is said to follow normal distribution with mean μ and standard deviation σ, if its probability density function

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

Properties of the normal curve

- The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
- The curve is symmetric about a vertical axis through the mean μ.
- The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < X < \mu + \sigma$ and is concave upward otherwise.
- The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.
- The total area under the curve and above the horizontal axis is equal to 1.

Theorem: The mean and variance of $f(x)$ are μ and σ^2, respectively.

Hence, the standard deviation is σ.

Properties of the normal curve

- The normal curve is bell shaped and is symmetric at $x = \mu$.
- Mean, median, and mode of the distribution are coincide i.e., Mean $=$ Median $=$ Mode $= \mu$.
- It has only one mode at $x = \mu$ (i.e., unimodal)
- Since the curve is symmetrical, Skewness $= \beta_1 = 0$ and Kurtosis $= \beta_2 = 3$.
- The points of inflection are at $x = \mu \pm \sigma$.
- The maximum ordinate occurs at $x = \mu$ and its value is $= \frac{1}{\sigma\sqrt{2\pi}}$.
- The x axis is an asymptote to the curve (i.e. the curve continues to approach but never touches the x axis)
- The first and third quartiles are equidistant from median.

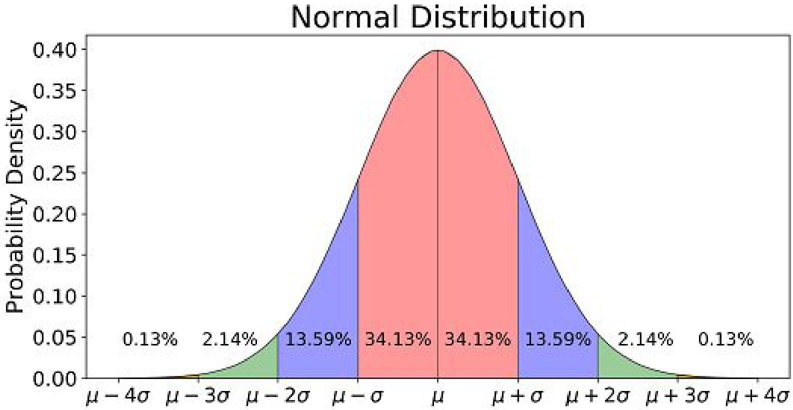
Properties of the normal curve

- The mean deviation about mean is 0.8σ
- Quartile deviation $= 0.6745\sigma$
- If X and Y are independent normal variates with mean μ_1 and μ_2, and variance σ_1^2 and σ_2^2 respectively then their sum $(X+Y)$ is also a normal variate with mean $(\mu_1+\mu_2)$ and variance $(\sigma_1^2+\sigma_2^2)$
- Area Property

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

$$P(\mu-2\sigma < X < \mu+2\sigma) = 0.9544$$

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


Normal Distribution
Probability Density
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.13%
2.14%
13.59%
34.13%
34.13%
13.59%
2.14%
0.13%
$\mu-4\sigma$
$\mu-3\sigma$
$\mu-2\sigma$
$\mu-\sigma$
μ
$\mu+\sigma$
$\mu+2\sigma$
$\mu+3\sigma$
$\mu+4\sigma$

The mean μ and standard deviation σ are called the parameters of Normal distribution. The normal distribution is expressed by $X \ \ N(\mu, \sigma^2)$.

Condition of Normal Distribution:

- Normal distribution is a limiting form of the binomial distribution under the following conditions.
  - n, the number of trials is indefinitely large ie., $n \to \infty$ and
  - Neither p nor q is very small.
- Normal distribution can also be obtained as a limiting form of Poisson distribution with parameter $\ \to \infty$
- Constants of normal distribution are mean $= \mu$, variation $= \sigma^2$, Standard deviation $= \sigma$.

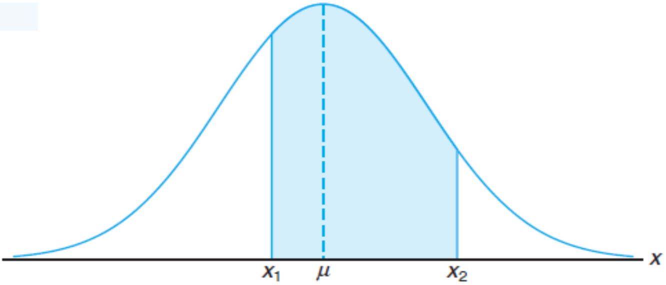
Areas under the Normal Curve

The curve of any continuous probability distribution or density function is constructed so that the area under the curve bounded by the two ordinates $x = x_1$ and $x = x_2$ equals the probability that the random variable X assumes a value between $x = x_1$ and $x = x_2$. Thus, for the normal curve in Figure,

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi\sigma}} \int_{x_1}^{x_2} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx$$

is represented by the area of the shaded region.

Areas under the Normal Curve



There are many types of statistical software that can be used in calculating areas under the normal curve. The difficulty encountered in solving integrals of normal density functions necessitates the tabulation of normal curve areas for quick reference.

However, it would be a hopeless task to attempt to set up separate tables for every conceivable value of μ and σ. Fortunately, we are able to transform all the observations of any normal random variable X into a new set of observations of a normal random variable Z with mean 0 and variance 1. This can be done by means of the transformation

$$Z = \frac{X - \mu}{\sigma}.$$

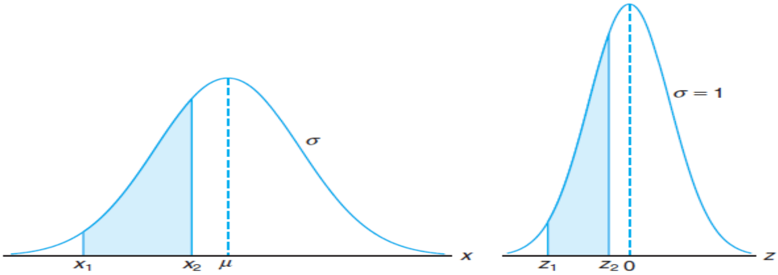
Whenever X assumes a value x, the corresponding value of Z is given by $z = \frac{x\mu}{\sigma}$. Therefore, if X falls between the values $x = x_1$ and $x = x_2$, the random variable Z will fall between the corresponding values $z_1 = \frac{x_1 - \mu}{\sigma}$ and $z_2 == \frac{x_2 - \mu}{\sigma}$. Consequently, we may write

$$
\begin{aligned}
P(x_1 < X < x_2) &= \frac{1}{\sqrt{2\pi\sigma}} \int_{x_1}^{x_2} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{\frac{-1}{2}z^2} dz \\
&= P(z_1 < Z < z_2),
\end{aligned}
$$

where Z is seen to be a normal random variable with mean 0 and variance 1.

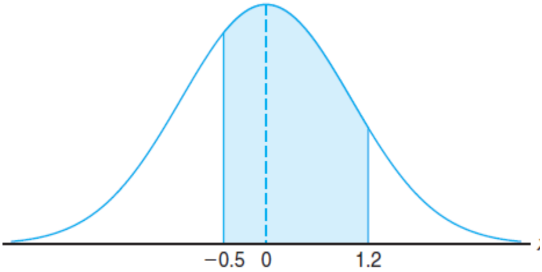
Definition: The distribution of a normal random variable with mean 0 and variance 1 is called a standard normal distribution.

The original and transformed distributions are illustrated in the following Figure. Since all the values of X falling between x_1 and x_2 have corresponding z values between z_1 and z_2, the area under the X-curve between the ordinates $x = x_1$ and $x = x_2$ in the following Figure equals the area under the Z-curve between the transformed ordinates $z = z_1$ and $z = z_2$.



We have now reduced the required number of tables of normal-curve areas to one, that of the standard normal distribution. the following Table indicates the area under the standard normal curve corresponding to $P(0 < Z < z)$ for values of z ranging from 0 to 3.

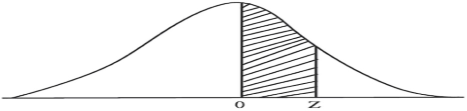
Problem: Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$, find the probability that X assumes a value between 45 and 62. Solution: The z values corresponding to $x_1 = 45$ and $x_2 = 62$ are $z_1 = \frac{45-50}{10} = -0.5$ and $z_2 = \frac{62-50}{10} = 1.2$. Therefore,
$P(45 < X < 62) = P(0.5 < Z < 1.2)$.
$P(0.5 < Z < 1.2)$ is shown by the area of the shaded region in the following Figure.



Using the normal Table, we have

$$
\begin{aligned}
P(45 < X < 62) =& P(0.5 < Z < 1.2) \\
&= P(0 < Z < 1.2) + P(-0.5 < Z < 0) \\
&= 0.3849 + P(0 < Z < 0.5) \quad \text{by symetry} \\
&= 0.3849 + 0.1915 = 0.5764
\end{aligned}
$$

AREA UNDER STANDARD NORMAL CURVE



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.4593	.1331	.1368	.14706	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.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	.3665	.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	.4956	.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	.4878	.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	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.5959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	04977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4989	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Problem: A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Solution: To find $P(X < 2.3)$. Hence, we find that $z = \frac{2.3-3}{0.5} = -1.4$, and then, usingnormal Table, we have

$$
\begin{aligned}
P(X < 2.3) = P(Z < -1.4) &= P(-\infty < Z < 0) - P(-1.4 < Z < 0) \\
&= P(Z > 0) - P(0 < Z < 1.4) \\
&= 0.5 - 0.4192 = 0.0808.
\end{aligned}
$$

Problem: A certain machine makes electrical resistors having a mean resistance of 40 ohms and a standard deviation of 2 ohms. Assuming that the resistance follows a normal distribution and can be measured to any degree of accuracy, what percentage of resistors will have a resistance exceeding 43 ohms?

Problem:Students of a class were given an aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percentage of students scored.

- More than 60 marks
- Less than 56 marks
- Between 45 and 65 marks

Problem: The mean inside diameter of a sample of 200 washers produced by a machine is 0.502 cm and S.D is 0.005 cm. The prupose for which these washers are intended allows a maximum tolerance in diameter of 0.496 cm to 0.508 cm. Otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine,assuming the diameter's are normally distributed.

Problem: In a normal distribution 31% of the items are under 45 and 8% are over 64. Find mean and the standard deviation.

Problem: A manufacturer produces air mail envelopes where weight is normal with mean 1.950 gm and S.D is 0.025 gm. The envelope are sold in lots of 1000. How many enevelopes in a lot may be heavies than 2 grams?

Gamma Distribution

The continuous random variable X has a gamma distribution, with parameters α and β, if its density function is given by

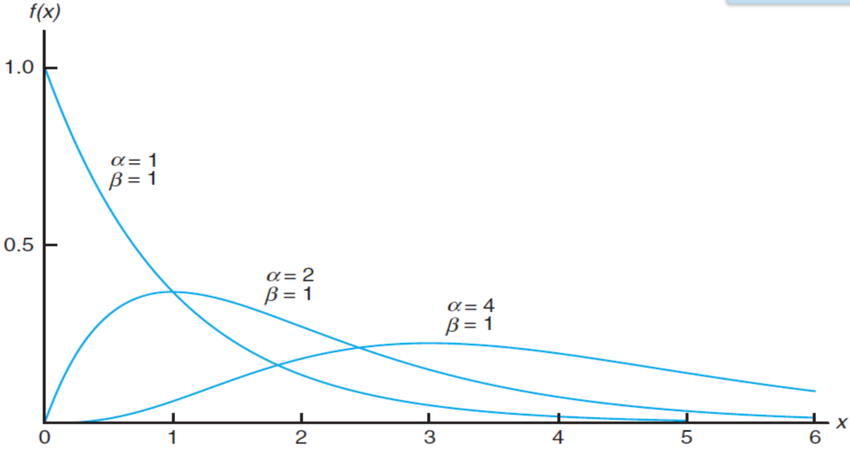
$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}}, & x > 0 \\ 0, & elsewhere, \end{cases}$$

where $\alpha > 0$ and $\beta > 0$.

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

Graphs of several gamma distributions are shown in Figure for certain specified values of the parameters α and β.

The special gamma distribution for which $\alpha = 1$ is called an exponential distribution.



Exponential distribution

The continuous random variable X has an exponential distribution, with parameter β, if its density function is given by

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

where $\beta > 0$. The mean and variance of the exponential distribution are $\mu = \beta$ and $\sigma^2 = \beta^2$. Put $\lambda = \frac{1}{\beta}$in the equation (1). Then

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

Problem: Suppose that a system contains a certain type of component whose time, in years, to failure is given by T. The random variable T is modeled nicely by the exponential distribution with mean time to failure $\beta = 5$. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

Solution: The probability that a given component is still functioning after 8 years is given by $P(T > 8) = \frac{1}{5}\int_8^{\infty} e^{\frac{-t}{5}} dt = 0.2$

Let X represent the number of components functioning after 8 years.

Then using the binomial distribution, we have $n = 5, P(T > 8) = p = 0.2$

Therefore, $P(X \geq 2) = 1 - P(X < 2) = 1 - (P(x = 0) + P(x = 1)) =$

$1 - 0.7373 = 0.2627.$

Problem: Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse by the time 2 calls have come in to the switchboard?

Solution: The Poisson process applies, with time until 2 Poisson events following a gamma distribution with $\beta = 1/5$ and $\alpha = 2$. Denote by X the time in minutes that transpires before 2 calls come. The required probability is given by $P(X \leq 1) = 0.96$.

Weibull distribution

The continuous random variable X has a Weibull distribution, with parameters α and β, if its density function is given by

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

where $\alpha > 0$ and $\beta > 0$.

The mean $\mu = \alpha^{\frac{-1}{\beta}} \Gamma(1 + \frac{1}{\beta})$, and variance $\alpha^2 = \alpha^{\frac{-2}{\beta}} \left\{ \Gamma(1 + \frac{2}{\beta}) - \left[\Gamma(1 + \frac{1}{\beta}) \right]^2 \right\}$.

Problem: The life time of a component measured in hours follows Weibull distribution with parameter $\alpha = 0.2, \beta = 0.5$. Find the mean life time of the component.

Answer: mean=50 hours.

Problem: 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, give 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.

Thank you