Probability Distributions

Prob. and Stats.,
CSVTU Bhilai

December 30, 2022

Table of Content

Bernoulli Trials and Binomial Distribution

Trials of a random experiment are called Bernoulli trials, if they satisfy the following conditions :

- There should be a finite number of trials.
- The trials should be independent.
- Each trial has exactly two outcomes : success or failure.
- The probability of success remains the same in each trial.

For example, throwing a die 50 times is a case of 50 Bernoulli trials, in which each trial results in success (say an even number) or failure (an odd number) and the probability of success (p) is same for all 50 throws. Obviously, the successive throws of the die are independent experiments. If the die is fair and have six numbers 1 to 6 written on six faces, then $p = \frac{1}{2}$ and $q = 1 - \frac{1}{2} = \frac{1}{2}$

Example

Six balls are drawn successively from an urn containing 7 red and 9 black balls. Tell whether or not the trials of drawing balls are Bernoulli trials when after each draw the ball drawn is

- replaced.
- not replaced in the urn.

The number of trials is finite. When the drawing is done with replacement, the probability of success (say, red ball) is $p = \frac{7}{16}$ which is same for all six trials (draws). Hence, the drawing of balls with replacements are Bernoulli trials.

When the drawing is done without replacement, the probability of success (i.e., red ball) in first trial is $\frac{7}{16}$, in 2nd trial is $\frac{6}{15}$ if the first ball drawn is red or $\frac{7}{15}$ if the first ball drawn is black and so on. Clearly, the probability of success is not same for all trials, hence the trials are not Bernoulli trials.

Binomial Distribution

X	0	1	2	...	x	...	n
P(X)	${}^nC_0\, q^n$	${}^nC_1\, q^{n-1}p^1$	${}^nC_2\, q^{n-2}p^2$		${}^nC_x\, q^{n-x}p^x$		${}^nC_n\, p^n$

The above probability distribution is known as *binomial distribution* with parameters n and p, because for given values of n and p, we can find the complete probability distribution.

The probability of x successes $P(X = x)$ is also denoted by $P(x)$ and is given by

$$P(x) = {}^nC_x\, q^{n-x}p^x, \quad x = 0, 1, \ldots, n. \ (q = 1 - p)$$

This $P(x)$ is called the *probability function* of the binomial distribution.

A binomial distribution with n-Bernoulli trials and probability of success in each trial as p, is denoted by $B(n, p)$.

Examples

If a fair coin is tossed 10 times, find the probability of

- exactly six heads
- at least six heads
- at most six heads

The repeated tosses of a coin are Bernoulli trials. Let X denote the number of heads in an experiment of 10 trials.
Clearly, X has the binomial distribution with $n = 10$ and $p = \frac{1}{2}$
Therefore

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

Here $n = 10, p = \frac{1}{2}, q = \frac{1}{2}$

Continue

- $P(X = 6) = (105/512)$
- $P(X >= 6) = (193/512)$
- $P(X <= 6) = (53/64)$

Questions

- Ten eggs are drawn successively with replacement from a lot containing 10% defective eggs. Find the probability that there is at least one defective egg.
- A bag consists of 10 balls each marked with one of the digits 0 to 9. If four balls are drawn successively with replacement from the bag, what is the probability that none is marked with the digit 0?
- Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. What is the probability that
  1. all the five cards are spades?
  2. only 3 cards are spades?
  3. none is a spade?

Poisson Distribution

Let X be a discrete random variable that can take on the values $0, 1, 2, \ldots$ such that the probability function of X is given by

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!},\ x = 0, 1, 2, \ldots$$

where λ is a given positive constant. This distribution is called the Poisson distribution, and a random variable having this distribution is said to be Poisson distributed.

When to use

- The probability of success is very low
- The number of trials is very large.

Here λ represents the mean of the random variable X, and it is given by

$$\lambda = n * p$$

where n represents the number of trials and p represents the probability of success.
The variance of Poisson Distribution is λ.

Example 1

Find the Probability that at most 5 defective fuses will be found in a box of 200 fuses, if experience shows that 2% of such fuses are defective.

Solution

Here we have, $p = 2\% = 0.02, n = 200,$
So mean $= \lambda = np = 200 * 0.02 = 4$
We are asked to find 'P(getting at most 5 defective fuses)'
Since it satisfies the conditions for Poisson Distribution, we solve it by poisson probability distribution.
$P(X \leq 5) = (P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$
$P(X \leq 5) = \frac{4^0 e^{-4}}{0!} + \frac{4^1 e^{-4}}{1!} + \frac{4^2 e^{-4}}{2!} + \frac{4^3 e^{-4}}{3!} + \frac{4^4 e^{-4}}{4!} + \frac{4^5 e^{-4}}{5!}$
$P(X \leq 5) = e^{-4}(1 + 4 + 8 + \frac{64}{6} + \frac{256}{24} + \frac{1024}{120})$
$P(X \leq 5) = 0.7845$

Example 2

If the probability that an individual suffer a bad reaction from a certain injection is 0.001, determine the probability that out of 2000 individuals-

- Exactly 3
- more that 2
- none
- more than 1

individuals will suffer a bad reaction.

Solution

Here $n = 2000$, $p = 0.001$, And it satisfies the requirement of Poisson Probability Distribution.
So mean $= \lambda = n * p = 2000 * 0.001 = 2$
Now, We know that Poisson Distribution is given by $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$, so

- Probability that exactly 3 individual will suffer a bad reaction is given by $P(X = 3) = \frac{2^3 e^{-2}}{3!} = \frac{8e^{-2}}{6} = 0.180$
- Probability that more than 2 individuals will suffer bad reaction is given by $1 - [P(X = 0) + P(X = 1) + P(X = 2)]$. Hence
  $P(X \geq 2) = 1 - [\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!}]$
  $P(X \geq 2) = 1 - e^{-2}[1 + 2 + 2]$
  $P(X \geq 2) = 0.323$

Continue

- Probability that no individual will suffer a bad reaction is $P(X=0)$, $P(X=0)=\frac{e^{-2}2^0}{0!}=\frac{1}{e^2}=0.135$
- Probability that more than 1 individual will suffer a bad reaction is given by $P(X\geq 1)$,
  $P(X\geq 1)=1-[P(X=0)+P(X=1)]$
  $P(X\geq 1)=1-[\frac{2^0e^{-2}}{0!}+\frac{2^1e^{-2}}{1!}]$
  $P(X\geq 1)=1-e^{-2}(1+2)$
  $P(X\geq 1)=0.594$

Try Yourself

For health reasons, homes need to be inspected for radon gas which decays and produces alpha particles. One device counts the number of alpha particles that hit its detector. To a good approximation, in one area, the count for the next week follows a Poisson distribution with mean 1.3. Determine

- the probability of exactly one particle next week. (0.3543)
- the probability of one or more particles next week. (0.727)
- the probability of at least two but no more than four particles next week. (0.362)
- The variance of the Poisson distribution. (1.3)

Hint: Refer to page no. 119, Probability and Statistics for Engineers, Miller and Freunds.

Geometric Distribution

Suppose that in a sequence of trials we are interested in the number of the trial on which **the first success occurs**. The three assumptions for Bernoulli trials are satisfied but the extra assumption underlying the binomial distribution is not. In other words, n is not fixed.
Clearly, if the first success is to come on the xth trial, it has to be preceded by (x-1) failures, and if the probability of a successes is p, the probability of (x-1) failures in (x-1) trials is (1-p)$^{x-1}$.

Then, if we multiply this expression by the probability p of a success on the xth trial, we find that the probability of getting the first success on the xth trial is given by

$$p(X = x) = \begin{cases} q^{(x-1)}p & x = 1, 2, 3... \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

such that $p + q = 1,\ p \geq 0$

- mean $= \frac{1}{p}$
- standard deviation $= \frac{\sqrt{q}}{p}$

Example

Question1

If the probability is 0.05 that a certain kind of measuring device will show excessive drift, what is the probability that the sixth measuring device tested will be the first to show excessive drift?

Sol:
Substituting x = 6 and p = 0.05 into the formula for the geometric distribution, we get

$$p(X = 6) = (1 - p)^5 * p = (1 - 0.05)^5 * 0.05 = 0.039$$

Example

Question 2

Find the probability that in successive tosses of a fair die, a 3 will come up for the first time on the fifth toss.

Sol:

The probability of not getting a 3 on the first toss is $\frac{5}{6}$. Similarly, the probability of not getting a 3 on the second toss is $\frac{5}{6}$, etc. Then the probability of not getting a 3 on the first 4 tosses is $(\frac{5}{6})^4$.

Therefore, since the probability of getting a 3 on the fifth toss is $\frac{1}{6}$, the required probability is

$$p(X = 5) = (1 - p)^4 * p = \left(\frac{5}{6}\right)^4 * \frac{1}{6}$$

Try it Yourself

Suppose that a trainee solder shoots a target in an independent fashion. If the probability that the target is hit in any shot is 0.7, what is the

- probability that first success comes in 3rd shot. (0.063)
- average number of shots needed to hit the target. ($\frac{1}{p} = 1.4286$)

Normal Distribution

One of the most important examples of a continuous probability distribution is the normal distribution, sometimes called the Gaussian distribution. The density function for this distribution is given by

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

where μ represents the mean and σ represents the standard deviation of the random variable X.

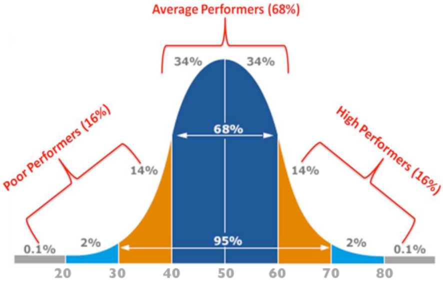
Example

The Distribution of Height and Weight of adults follow the normal distribution.

Standard Normal Distribution

In the Normal Distribution, if $\mu = 0$ and $\sigma = 1$, then it is called the standard normal distribution and the probability density function is given by

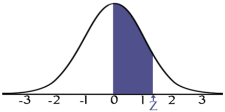
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$



.

The area under the curve gives the probability. Observe that the shape of the graph is bell shaped.

Standard Normal Distribution Table



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

Working Rule

- Convert x into a standard normal variate by the formula $z = \frac{x-\mu}{\sigma}$.
- Find the limit of z corresponding to the limit of x, when $x = a$ then $z = \frac{a-\mu}{\sigma}$. Similarly when $x = b$ then $z = \frac{b-\mu}{\sigma}$.
- $P(a < x < b) = P(\frac{a-\mu}{\sigma} < x < \frac{b-\mu}{\sigma})$. Use the standard normal table.

Example

Question

The mean height of 500 students is 151 cm and the standard deviation is 15 cm. Assuming that the heights are normally distributed, find how many students have height between 120 and 155 cm.

Solution: Here we have, mean $\mu = 151$, standard deviation $\sigma = 15$.
Now $z = \frac{x-\mu}{\sigma}$
So when $x = 120$, we have $z = \frac{120-151}{15} = -2.067$
Similarly when $x = 155$, we have $z = \frac{155-151}{15} = 0.267$.

Continue

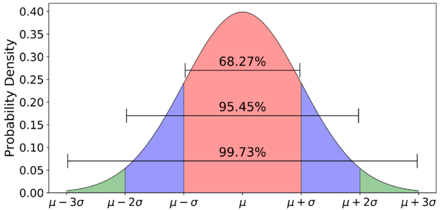
So, $P(120 < x < 155) = P(-2.067 < z < 0.267) = P(-2.067 < z < 0) + P(0 < z < 0.267)$
$P(120 < x < 155) = 0.4808 + 0.1064$
$P(120 < x < 155) = 0.5872$
Hence the number of students whose height lies between 120 cm to 155 cm $= N * P(120 < x < 155) = 500 * 0.5872 = 294$

Try it Yourself

If a random variable has a normal distribution, what are the probabilities that it will take on a value within

- 1 standard deviation of the mean (z score for z=1 is 0.3413); (0.6827)
- 2 standard deviations of the mean(z score for z=2 is 0.0.4772); (0.9545)
- 3 standard deviations of the mean(z score for z=3 is 0.4987); (0.9973)

Hint: Refer to Question no. 5.22 on page no. 149, Probability and Statistics for Engineers, Miller and Freunds.



Exponential Probability Distribution

A random variable X is said to have an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{2}$$

This distribution is used where we are concerned with the exponential growth or decay of organisms or matters.

Probability Distribution Function

$F(x) = P(X \geq x) = \int_{-\infty}^{x} f(x)dx = \int_{0}^{x} \lambda e^{-\lambda x} dx.$
After integrating and putting the limits, we obtain

$$F(x) = 1 - e^{-\lambda x}$$

For example,

$$P(X \geq 3) = 1 - e^{-3\lambda}$$

So we have,

$$F(x) = P(X \geq x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \\ 1 & x = \infty \end{cases} \tag{3}$$

Mean and Variance

We already know that

$$\begin{aligned} \mu &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{0}^{\infty} x\lambda e^{-\lambda x}dx \\ &= \lambda\left[\frac{xe^{-\lambda x}}{-\lambda} - \frac{e^{-\lambda x}}{\lambda^2}\right]_0^{\infty} \\ &= \frac{1}{\lambda} \end{aligned}$$

Similarly calculating variance gives us $Var[X] = \frac{1}{\lambda^2}$ and Standard Deviation is $\sigma = \frac{1}{\lambda}$

Example 1

Question 1

A random variable has an exponential distribution with probability density function

$$f(x) = \begin{cases} 2e^{-2x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

. Compute the probability that X is not less than 3. Also find the mean and standard deviation.

Solution: Here we can observe $\lambda = 2$,
Now

$$P(X \geq 3) = \int_3^\infty f(x)dx = \int_3^\infty 2e^{-2x}dx$$
$$= 2\left[\frac{e^{-2x}}{-2}\right]_3^\infty$$
$$= -(0 - e^{-6}) = e^{-6}$$

Continue

Mean $= \frac{1}{\lambda} = \frac{1}{2}$
Standard Deviation $= \sigma = \frac{1}{\lambda} = \frac{1}{2}$

Example 2

Question 2

Suppose the life of a mobile battery is exponentially distributed with parameter $\lambda = 0.001$ day. What is the probability that a battery will last more than 1200 days.

Solution: The probability distribution of exponential distribution is given by

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

Here $\lambda = 0.001$ days

Continue

Now

$$
\begin{aligned}
P(X > 1200) &= \int_{1200}^{\infty} f(x)dx \\
&= \int_{1200}^{\infty} \lambda e^{-\lambda x} dx \\
&= \lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_{1200}^{\infty} \\
&= e^{-0.001*1200} = e^{-1.2} = 0.301
\end{aligned}
$$

Uniform Distribution

A continuous random variable X is said to follow a continuous uniform or rectangular distribution over an interval (a, b) if its probability density function is given by:

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

Here k is a constant which is equal to $\frac{1}{b-a}$

Value of k

We know

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$\int_{a}^{b} kdx = 1$$

$$[kx]_a^b = k(b-a) = 1$$

This gives $k = \frac{1}{b-a}$
The probability density function is given as :

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

Mean of Uniform Distribution

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f(x) dx \\
&= \int_a^b \frac{x}{b-a} dx \\
&= \frac{1}{2(b-a)} [x^2]_a^b \\
&= \frac{1}{2(b-a)} (b^2 - a^2) \\
&= \frac{b+a}{2}
\end{aligned}
$$

Hence the mean of uniform distribution is $\mu = \frac{b+a}{2}$.

Variance of Uniform Distribution

$$
\begin{aligned}
\sigma^2 &= E[X^2] - (E[X])^2 \\
&= \int_{-\infty}^{\infty} x^2 f(x) dx - (\frac{b+a}{2})^2 \\
&= \frac{1}{3(b-a)}[x^3]_a^b - (\frac{b+a}{2})^2 \\
&= \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} \\
&= \frac{(b-a)^2}{12}
\end{aligned}
$$

Hence the variance of uniform distribution is given by $\sigma^2 = \frac{(b-a)^2}{12}$.

Example 1

Example

If X is uniformly distributed with mean 1 and variance $\frac{4}{3}$, find $P(X < 0)$.

Given: mean $= \frac{a+b}{2} = 1 \implies a + b = 2$ and
variance $= \frac{(b-a)^2}{12} = \frac{4}{3} \implies (a-b)^2 = 16$ which gives $a - b = \pm 4$
On solving both the equalities, we get $a = -1, b = 3$. Hence the Probability density function is given by:

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

which gives

$$f(x) = \begin{cases} \frac{1}{4} & \text{if } -1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Continue

$$
\begin{aligned}
P(X < 0) &= \int_{-\infty}^{0} f(x)dx \\
&= \int_{-1}^{0} \frac{1}{4} dx \\
&= \frac{1}{4}[x]_{-1}^{0} \\
&= \frac{1}{4}
\end{aligned}
$$

Example 2

Example

The number of PC's sold daily at Alfa Computers is uniformly distributed with a minimum of 2000 PC's and a maximum of 5000 PC's.

1. Find the probability that the daily sales will fall between 2500 and 3000 PC's.
2. Find the probability that the Alfa computers will sell at least 4000 PC's.
3. Find the probability that the Alfa computers will sell exactly 4000 PC's.

Let X denotes the number of PC's sold daily at alfa computers, then X follows uniform distribution over the interval (2000,5000). Hence the probability distribution function is given by:

continue

$$f(x) = \begin{cases} \frac{1}{3000} & \text{if } 2000 < x < 5000 \\ 0 & \text{otherwise} \end{cases}$$

1.

$$\begin{aligned} P(2500 < X < 3000) &= \int_{2500}^{3000} \frac{1}{3000} dx \\ &= \frac{1}{3000}(3000 - 2500) \\ &= \frac{1}{6} \end{aligned}$$

continue

- $$\begin{aligned} P(X \geq 4000) &= \int_{4000}^{\infty} f(x)dx \\ &= \int_{4000}^{5000} f(x)dx + \int_{5000}^{\infty} f(x)dx \\ &= \int_{4000}^{5000} \frac{1}{3000} dx + 0 \\ &= \frac{1}{3000}(5000 - 4000) \\ &= \frac{1}{3} \end{aligned}$$

- $$\begin{aligned} P(X = 4000) &= \int_{4000}^{4000} f(x)dx \\ &= 0 \end{aligned}$$

Gamma Function

Definition

For a positive integer n, the $\Gamma(n)$ is defined as

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx$$

- For $n = 1$,

$$\begin{aligned}\Gamma(1) &= \int_0^\infty x^{1-1} e^{-x} dx \\ &= \int_0^\infty e^{-x} dx \\ &= [-e^{-x}]_0^\infty = 1\end{aligned}$$

Properties of Gamma Function

- $\Gamma(n+1) = n\Gamma(n)$

We have

$$\begin{aligned}
\Gamma(n+1) &= \int_0^\infty x^n e^{-x} dx \\
&= x^n \int_0^\infty e^{-x} dx + \int_0^\infty n x^{n-1} e^{-x} dx \\
&= \left| -x^n e^{-x} \right|_0^\infty + n\Gamma(n) \\
&= n\Gamma(n)
\end{aligned}$$

Factorial Notation

We have $\Gamma(n+1) = n\Gamma(n)$, and $\Gamma(1) = 1$. Now
$\Gamma(2) = \Gamma(1+1) = 1\Gamma(1) = 1 * 1 = 1$. Similarly
$\Gamma(3) = \Gamma(2+1) = 2\Gamma(2) = 2 * 1 = 2$. And
$\Gamma(4) = \Gamma(3+1) = 3\Gamma(3) = 3 * 2 = 6$.
From the above we can observe that $\Gamma(n) = (n-1)!$

Gamma Distribution

We have $\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx$. This gives us

$$\int_0^\infty \frac{x^{n-1}e^{-x}dx}{\Gamma(n)} = 1 \tag{4}$$

So the function on the left hand side can be used as a probability distribution function for $x > 0$ and for n being a positive integer. This is the idea behind the Gamma Distribution.

There are two types of Gamma Distribution based on the number of parameters used.

- 1st Kind (Standard Gamma Distribution)
- 2nd Kind (Two Parameter Gamma Distribution)

1st kind

A Continuous random variable X is said to follow Gamma Distribution of 1st kind if its probability density function is given by

$$f(x) = \begin{cases} \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

2nd kind

A Continuous random variable X is said to follow Gamma Distribution of 2nd kind if its probability density function is given by

$$f(x) = \begin{cases} \frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where α is shape parameter and β is scale parameter.

Observations

We have

$$f(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- If $\beta = 1$ then 2nd kind reduces to first kind.
- If $\alpha = 1$ and $\beta = \lambda$, then 2nd kind reduces to exponential distribution $\lambda e^{-\lambda x}$.

Mean and Variance of Gamma Distribution

- For First kind:

We have

$$
\begin{aligned}
E[X] &= \int_0^\infty x f(x) dx \\
&= \int_0^\infty x * \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} dx \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-x} dx \\
&= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \\
&= \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \\
&= \alpha
\end{aligned}
$$

Variance of first kind Gamma Distribution

We know $\text{Var}[X] = E[X^2] - (E[X])^2$. So we have

$$\begin{aligned}
Var[X] &= \int_0^\infty x^2 f(x)dx - (\alpha)^2 \\
&= \int_0^\infty x^2 * \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} dx - (\alpha)^2 \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1}e^{-x} dx - (\alpha)^2 \\
&= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} - (\alpha)^2 \\
&= \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} - (\alpha)^2 \\
&= \alpha
\end{aligned}$$

Observe that mean and variance of Gamma Distribution of first kind are equal $(= \alpha)$.

Mean and Variance of Gamma Distribution of 2nd kind

- For Second Kind:

We have

$$
\begin{aligned}
E[X] &= \int_0^\infty x f(x) dx \\
&= \int_0^\infty x * \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} dx \\
&= \int_0^\infty \frac{\beta^\alpha x^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx = \int_0^\infty \frac{(\beta x)^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx \\
&= \frac{1}{\beta \Gamma(\alpha)} \int_0^\infty t^\alpha e^{-t} dt, \ (\because t = \beta x) \\
&= \frac{\Gamma(\alpha+1)}{\beta \Gamma(\alpha)} = \frac{\alpha \Gamma(\alpha)}{\beta \Gamma(\alpha)} \\
&= \frac{\alpha}{\beta}
\end{aligned}
$$

Variance of Gamma Distribution of 2nd kind

We know $\text{Var}[X] = E[X^2] - (E[X])^2$. So we have

$$\begin{aligned}
Var[X] &= \int_0^\infty x^2 f(x)dx - \left(\frac{\alpha}{\beta}\right)^2 \\
&= \int_0^\infty x^2 * \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} dx - \left(\frac{\alpha}{\beta}\right)^2 \\
&= \int_0^\infty \frac{\beta^{\alpha+1} x^{\alpha+1} e^{-\beta x}}{\beta\Gamma(\alpha)} dx - \left(\frac{\alpha}{\beta}\right)^2 \\
&= \frac{1}{\beta\Gamma(\alpha)} \int_0^\infty (\beta x)^{\alpha+1} e^{-\beta x} dx - \left(\frac{\alpha}{\beta}\right)^2 \\
&= \frac{\Gamma(\alpha+2)}{\beta^2\Gamma(\alpha)} - \left(\frac{\alpha}{\beta}\right)^2 \\
&= \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\beta^2\Gamma(\alpha)} - \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2}
\end{aligned}$$

References I

- Miller and Freund's Probability and Statistics for Engineers, Richard A. Johnson, Ninth Edition, Pearson.
- Schaum's Outline Probability and Statistics, 3rd Edition