WILEY

Applied Statistics and Probability for Engineers

Sixth Edition

Douglas C. Montgomery   George C. Runger

Chapter 4

Continuous Random Variables and Probability Distributions

4 Continuous Random Variables and Probability Distributions

CHAPTER OUTLINE

Learning Objectives for Chapter 4

After careful study of this chapter, you should be able to do the following:

1. Determine probabilities from probability density functions.
2. Determine probabilities from cumulative distribution functions, and cumulative distribution functions from probability density functions, and the reverse.
3. Calculate means and variances for continuous random variables.
4. Understand the assumptions for continuous probability distributions.
5. Select an appropriate continuous probability distribution to calculate probabilities for specific applications.
6. Calculate probabilities, means and variances for continuous probability distributions.
7. Standardize normal random variables.
8. Use the table for the cumulative distribution function of a standard normal distribution to calculate probabilities.
9. Approximate probabilities for Binomial and Poisson distributions.

Continuous Random Variables

- A continuous random variable is one which takes values in an uncountable set.
- They are used to measure physical characteristics such as height, weight, time, volume, position, etc...

Examples

1. Let Y be the height of a person (a real number).
2. Let X be the volume of juice in a can.
3. Let Y be the waiting time until the next person arrives at the server.

Probability Density Function

For a continuous random variable X, a probability density function is a function such that

(1) $f(x) \geq 0$ means that the function is always non-negative.

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

(3) $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx =$ area under $f(x)dx$ from a to b

(4) $f(x) = 0$ means there is no area exactly at x.

Example 4-1: Electric Current

Let the continuous random variable *X* denote the current measured in a thin copper wire in milliamperes(mA). Assume that the range of *X* is $4.9 \le x \le 5.1$ and $f(x) = 5$. What is the probability that a current is less than 5mA?

Answer:

$$P(X < 5) = \int_{4.9}^{5} f(x)dx = \int_{4.9}^{5} 5\,dx = 0.5$$

$$P(4.95 < X < 5.1) = \int_{4.95}^{5.1} f(x)dx = 0.75$$

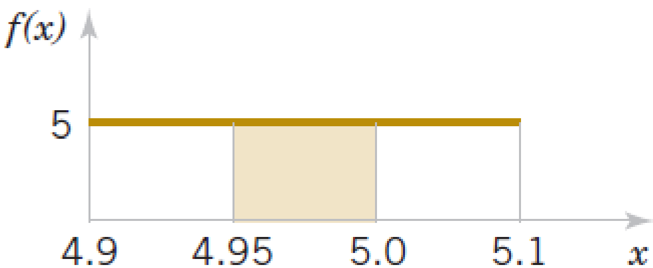


Figure 4-4 P(X < 5) illustrated.

Cumulative Distribution Functions

The cumulative distribution function of a continuous random variable X is,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \quad \text{for} \quad -\infty < x < \infty$$

The cumulative distribution function is defined for all real numbers.

Example 4-3: Electric Current

For the copper wire current measurement in Exercise 4-1, the cumulative distribution function consists of three expressions.

	0	$x < 4.9$
$F(x) =$	$5x - 24.5$	$4.9 \le x \le 5.1$
	1	$5.1 \le x$

The plot of *F(x) is shown in Figure 4-6.*

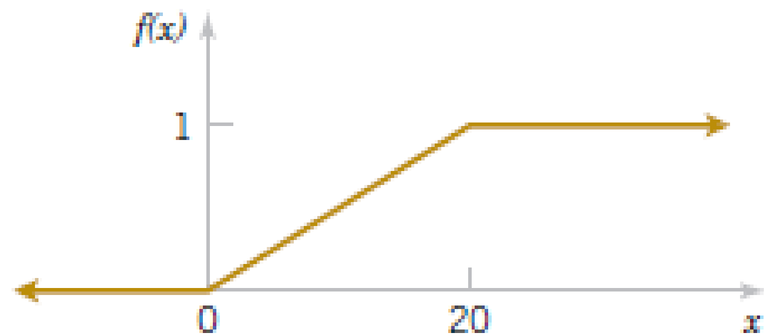


Figure 4-6 Cumulative distribution function

Probability Density Function from the Cumulative Distribution Function

- The probability density function (PDF) is the derivative of the cumulative distribution function (CDF).
- The cumulative distribution function (CDF) is the integral of the probability density function (PDF).

Given $F(x)$, $f(x) = \dfrac{dF(x)}{dx}$ as long as the derivative exists.

Exercise 4-5: Reaction Time

- The time until a chemical reaction is complete (in milliseconds, ms) is approximated by this cumulative distribution function:

$$F(x)=\begin{cases}0 & \text{for } x<0\\ 1-e^{-0.01x} & \text{for } 0\le \mathrm{x}\end{cases}$$

- What is the Probability density function?

$$f(x)=\frac{dF(x)}{dx}=\frac{d}{dx}\begin{cases}0\\ 1-e^{-0.01x}\end{cases}=\begin{cases}0 & \text{for } x<0\\ 0.01e^{-0.01x} & \text{for } 0\le \mathrm{x}\end{cases}$$

- What proportion of reactions is complete within 200 ms?

$$P(X<200)=F(200)=1-e^{-2}=0.8647$$

WILEY

Mean & Variance

Suppose X is a continuous random variable with probability density function $f(x)$. The mean or expected value of X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)\,dx \qquad (4\text{-}4)$$

The variance of X, denoted as $V(X)$ or σ^2, is

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

The standard deviation of X is $\sigma = \sqrt{\sigma^2}$.

WILEY

Example 4-6: Electric Current

For the copper wire current measurement, the PDF is $f(x) = 0.05$ for $0 \le x \le 20$. Find the mean and variance.

$$E(X) = \int_0^{20} x \cdot f(x)\,dx = \left.\frac{0.05x^2}{2}\right|_0^{20} = 10$$

$$V(X) = \int_0^{20} (x-10)^2 f(x)\,dx = \left.\frac{0.05(x-10)^3}{3}\right|_0^{20} = 33.33$$

Mean of a Function of a Continuous Random Variable

If X is a continuous random variable with a probability density function $f(x)$,

$$E\left[h(x)\right] = \int_{-\infty}^{\infty} h(x) f(x)\, dx$$

<u>Example 4-7:</u>

Let X be the current measured in mA. The PDF is $f(x) = 0.05$ for $0 \le x \le 20$. What is the expected value of power when the resistance is 100 ohms? Use the result that power in watts $P = 10^{-6}RI^2$, where I is the current in milliamperes and R is the resistance in ohms. Now, $h(X) = 10^{-6}100X^2$.

$$E\left[h(x)\right] = 10^{-4}\int_0^{20} x^2 dx = 0.0001 \frac{x^3}{3}\Bigg|_0^{20} = 0.2667 \text{ watts}$$

Continuous Uniform Distribution

- This is the simplest continuous distribution and analogous to its discrete counterpart.
- A continuous random variable X with probability density function

$$f(x) = 1 / (b-a) \text{ for } a \le x \le b$$

Figure 4-8 Continuous uniform Probability Density Function

Mean & Variance

- Mean & variance are:

$$\mu = E(X) = \frac{(a+b)}{2}$$

and

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

Example 4-9: Uniform Current

The random variable *X* has a continuous uniform distribution on [4.9, 5.1]. The probability density function of *X is f*(x) = 5, 4.9 ≤ x ≤ 5.1. What is the probability that a measurement of current is between 4.95 & 5.0 mA?

$$P(4.95 < x < 5.0) = \int_{4.95}^{5.0} f(x)\,dx = 5(0.05) = 0.25$$

The mean and variance formulas can be applied with *a = 4.9 and b = 5.1. Therefore,*

$$\mu = E(X) = 5 \text{ mA and } V(X) = \frac{(0.2)^2}{12} = 0.0033 \text{ mA}^2$$

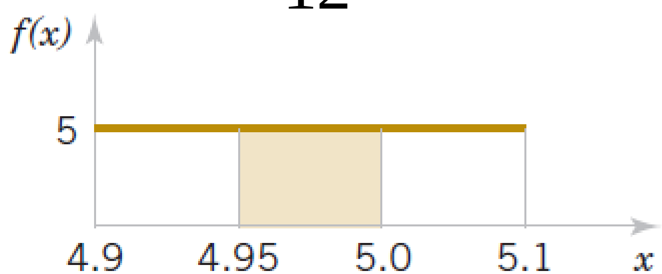


Figure 4-9

WILEY

Cumulative distribution function of Uniform distribution

$$F(x) = \int_a^x \frac{1}{(b-a)} du = \frac{x-a}{b-a}$$

The Cumulative distribution function is

$$F(x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & b \le x \end{cases}$$

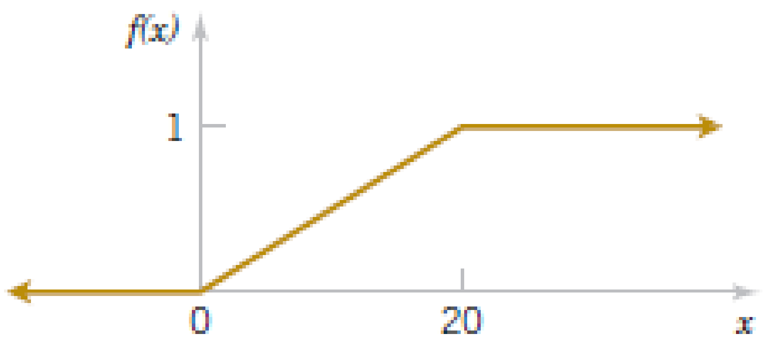


Figure 4-6 Cumulative distribution function

Normal Distribution

A random variable X with probability density function

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

is a normal random variable with parameters μ,
where $-\infty < \mu < \infty$, and $\sigma > 0$. Also,

$$E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2$$

and the notation $N(\mu, \sigma^2)$ is used to denote the distribution.

Empirical Rule

For any normal random variable,

$P(\mu - \sigma < X < \mu + \sigma) = 0.6827$

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

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

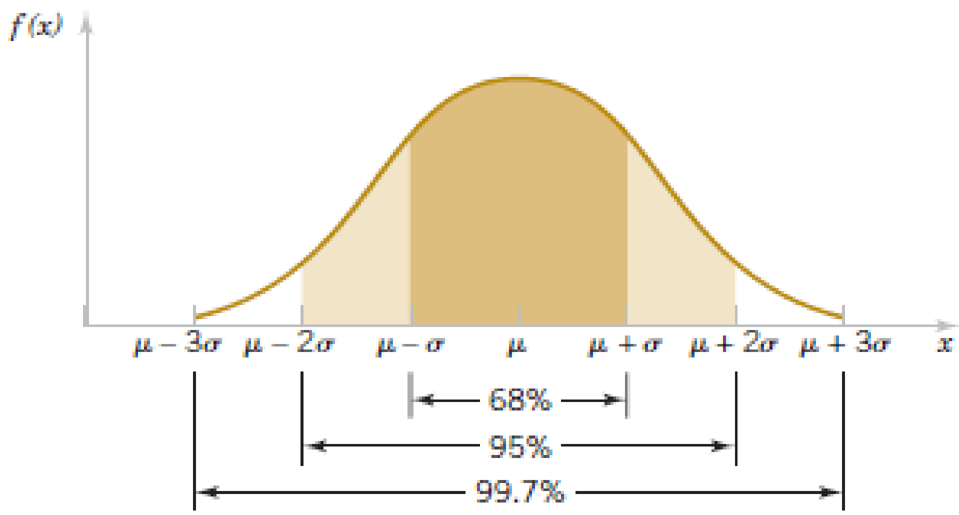


Figure 4-12 Probabilities associated with a normal distribution

WILEY

Standard Normal Random Variable

A normal random variable with

$$\mu = 0 \text{ and } \sigma^2 = 1$$

is called a standard normal random variable and is denoted as Z. The cumulative distribution function of a standard normal random variable is denoted as:

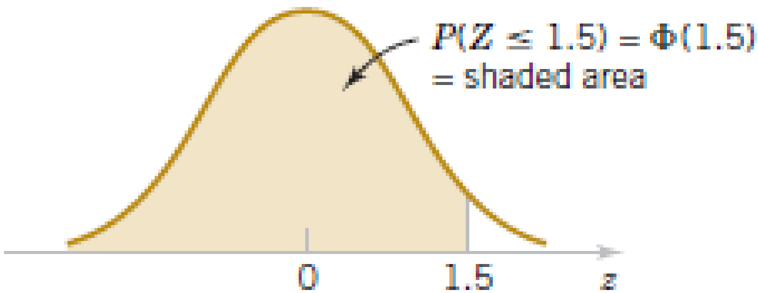
$$\Phi(z) = P(Z \leq z)$$

Values are found in Appendix Table III and by using Excel and Minitab.

Example 4-11: Standard Normal Distribution

Assume *Z* is a standard normal random variable.

Find $P(Z \leq 1.50)$. Answer: 0.93319



z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Figure 4-13 Standard normal Probability density function

Find $P(Z \leq 1.53)$. Answer: 0.93699

Find $P(Z \leq 0.02)$. Answer: 0.50398

<u>NOTE :</u> The column headings refer to the hundredths digit of the value of z in $P(Z \leq z)$.
For example, $P(Z \leq 1.53)$ is found by reading down the z column to the row 1.5 and then selecting the probability from the column labeled 0.03 to be 0.93699.

Standardizing a Normal Random Variable

Suppose X is a normal random variable with mean μ and variance σ^2, the random variable

$$Z = \frac{(X - \mu)}{\sigma}$$

is a normal random variable with $E(Z) = 0$ and $V(Z) = 1$.

The probability is obtained by using Appendix Table III with $z = \frac{(x - \mu)}{\sigma}$.

WILEY

Example 4-14: Normally Distributed Current-1

Suppose that the current measurements in a strip of wire are assumed to follow a normal distribution with μ = 10 and σ = 2 mA, what is the probability that the current measurement is between 9 and 11 mA?

Answer:

$$P(9 < X < 11) = P\left(\frac{9-10}{2} < \frac{x-10}{2} < \frac{11-10}{2}\right)$$

$$= P(-0.5 < z < 0.5)$$

$$= P(z < 0.5) - P(z < -0.5)$$

$$= 0.69146 - 0.30854 = 0.38292$$

Using Excel	
0.38292	= NORMDIST(11,10,2,TRUE) - NORMDIST(9,10,2,TRUE)

Example 4-14: Normally Distributed Current-2

Determine the value for which the probability that a current measurement is below 0.98.

Answer:

$$P(X < x) = P\left(\frac{X - 10}{2} < \frac{x - 10}{2}\right)$$

$$= P\left(Z < \frac{x - 10}{2}\right) = 0.98$$

$$z = 2.05 \text{ is the closest value.}$$

$$z = 2(2.05) + 10 = 14.1 \text{ mA.}$$

Using Excel	
14.107	=NORMINV(0.98,10,2)

WILEY

Normal Approximations

- The binomial and Poisson distributions become more bell-shaped and symmetric as their mean value increase.
- For manual calculations, the normal approximation is practical – exact probabilities of the binomial and Poisson, with large means, require technology (Minitab, Excel).
- The normal distribution is a good approximation for:
  - Binomial if $np > 5$ and $n(1-p) > 5$.
  - Poisson if $\lambda > 5$.

Normal Approximation to the Binomial Distribution

If X is a binomial random variable with parameters n and p,

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

is approximately a standard normal random variable. To approximate a binomial probability with a normal distribution, a continuity correction is applied as follows:

$$P(X \leq x) = P(X \leq x + 0.5) \approx P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

and

$$P(x \leq X) = P(x - 0.5 \leq X) \approx P\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}} \leq Z\right)$$

The approximation is good for $np > 5$ and $n(1-p) > 5$.

Refer to Figure 4-19 to see the rationale for adding and subtracting the 0.5 continuity correction.

Example 4-18: Applying the Approximation

In a digital communication channel, assume that the number of bits received in error can be modeled by a binomial random variable. The probability that a bit is received in error is 10^{-5}. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur?

$$P\left(X \le 150\right) = P\left(X \le 150.5\right)$$

$$= P\left(\frac{X - 160}{\sqrt{160\left(1 - 10^{-5}\right)}} \le \frac{150.5 - 160}{\sqrt{160\left(1 - 10^{-5}\right)}}\right)$$

$$= P\left(Z \le \frac{-9.5}{12.6491}\right) = P\left(Z \le -0.75104\right) = 0.2263$$

Using Excel	
0.2263	= NORMDIST(150.5, 160, SQRT(160*(1-0.00001)), TRUE)
-0.7%	= (0.2263-0.228)/0.228 = percent error in the approximation

Normal Approximation to Hypergeometric

Recall that the hypergeometric distribution is similar to the binomial such that $p = K / N$ and when sample sizes are small relative to population size. Thus the normal can be used to approximate the hypergeometric distribution.

hypergeometric	≈	binomial	≈	normal
distribution		distribution		distribution
	$n / N < 0.1$		$np < 5$	
			$n(1\text{-}p) < 5$	

Figure 4-21 Conditions for approximating hypergeometric and binomial probabilities

Normal Approximation to the Poisson

If X is a Poisson random variable with $E(X) = \lambda$ and $V(X) = \lambda$,

$$Z = \frac{X - \lambda}{\sqrt{\lambda}}$$

is approximately a standard normal random variable. The same continuity correction used for the binomial distribution can also be applied. The approximation is good for $\lambda \geq 5$

Example 4-20: Normal Approximation to Poisson

Assume that the number of asbestos particles in a square meter of dust on a surface follows a Poisson distribution with a mean of 1000. If a square meter of dust is analyzed, what is the probability that 950 or fewer particles are found?

$$P(X \leq 950) = \sum_{x=0}^{950} \frac{e^{-1000} 1000^x}{x!} \quad \text{... too hard manually!}$$

The probability can be approximated as

$$P(X \leq 950) = P(X \leq 950.5)$$

$$\approx P\left(Z \leq \frac{950.5 - 1000}{\sqrt{1000}}\right)$$

$$= P(Z \leq -1.57) = 0.058$$

Using Excel	
0.0578	= POISSON(950,1000,TRUE)
0.0588	= NORMDIST(950.5, 1000, SQRT(1000), TRUE)
1.6%	= (0.0588 - 0.0578) / 0.0578 = percent error

Exponential Distribution Definition

The random variable *X* that equals the distance between successive events of a Poisson process with mean number of events $\lambda > 0$ per unit interval is an exponential random variable with parameter λ. The probability density function of *X* is:

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \leq x < \infty$$

Exponential distribution - Mean & Variance

If the random variable X has an exponential distribution with parameter λ,

$$\mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2} \qquad \text{(4-15)}$$

Note:

- Poisson distribution : Mean and variance are same.
- Exponential distribution : Mean and standard deviation are same.

Example 4-21: Computer Usage-1

In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour. What is the probability that there are no log-ons in the next 6 minutes (0.1 hour)?

Let *X* denote the time in hours from the start of the interval until the first log-on.

$$P(X > 0.1) = \int_{0.1}^{\infty} 25e^{-25x}dx = e^{-25(0.1)} = 0.082$$

The cumulative distribution function also can be used to obtain the same result as follows

$$P(X > 0.1) = 1 - F(0.1) = 0.082$$

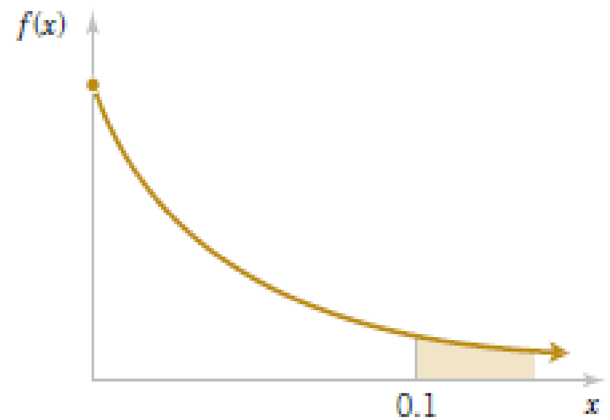


Figure 4-23 Desired probability

Using Excel	
0.0821	= 1 - EXPONDIST(0.1,25,TRUE)

Example 4-21: Computer Usage-2

Continuing, what is the probability that the time until the next log-on is between 2 and 3 minutes (0.033 & 0.05 hours)?

$$P(0.033 < X < 0.05) = \int_{0.033}^{0.05} 25e^{-25x}dx$$

$$= -e^{-25x}\Big|_{0.033}^{0.05} = 0.152$$

An alternative solution is

$$P(0.033 < X < 0.05) = F(0.05) - F(0.033) = 0.152$$

Using Excel	
0.148	= EXPONDIST(3/60, 25, TRUE) - EXPONDIST(2/60, 25, TRUE)
	(difference due to round-off error)

Example 4-21: Computer Usage-3

- Continuing, what is the interval of time such that the probability that no log-on occurs during the interval is 0.90?

$$P(X > x) = e^{-25x} = 0.90, \quad -25x = \ln(0.90)$$

$$x = \frac{-0.10536}{-25} = 0.00421 \text{ hour} = 0.25 \text{ minute}$$

- What is the mean and standard deviation of the time until the next log-in?

$$\mu = \frac{1}{\lambda} = \frac{1}{25} = 0.04 \text{ hour } = 2.4 \text{ minutes}$$

$$\sigma = \frac{1}{\lambda} = \frac{1}{25} = 0.04 \text{ hour } = 2.4 \text{ minutes}$$

Lack of Memory Property

An interesting property of an exponential random variable concerns conditional probabilities.

For an exponential random variable *X*,

$$P(X<t_1+t_2|X>t_1)= P(X < t_2)$$

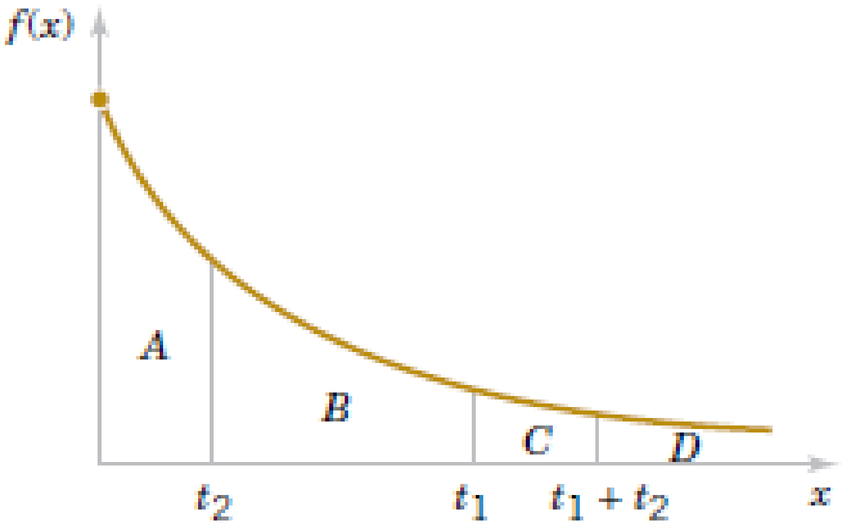


Figure 4-24 Lack of memory property of an exponential distribution.

WILEY

Example 4-22: Lack of Memory Property

Let *X* denote the time between detections of a particle with a Geiger counter. Assume *X* has an exponential distribution with E(*X*) = 1.4 minutes. What is the probability that a particle is detected in the next 30 seconds?

Using Excel	
0.300	= EXPONDIST(0.5, 1/1.4, TRUE)

$$P(X < 0.5) = F(0.5) = 1 - e^{-0.5/1.4} = 0.30$$

No particle has been detected in the last 3 minutes. Will the probability increase since it is “due”?

$$P(X < 3.5 | X > 3) = \frac{P(3 < X < 3.5)}{P(X > 3)} = \frac{F(3.5) - F(3)}{1 - F(3)} = \frac{0.035}{0.117} = 0.30$$

No, the probability that a particle will be detected depends only on the interval of time, not its detection history.

Erlang & Gamma Distributions

- The Erlang distribution is a generalization of the exponential distribution.
- The exponential distribution models the interval to the 1st event, while the Erlang distribution models the interval to the r^{th} event, i.e., a sum of exponentials.
- If r is not required to be an integer, then the distribution is called gamma.
- The exponential, as well as its Erlang and gamma generalizations, is based on the Poisson process.

Example 4-23: Processor Failure

The failures of CPUs of large computer systems are often modeled as a Poisson process. Assume that units that fail are repaired immediately and the mean number of failures per hour is 0.0001. Let *X* denote the time until 4 failures occur. What is the probability that *X* exceed 40,000 hours?

Let the random variable *N* denote the number of failures in 40,000 hours. The time until 4 failures occur exceeds 40,000 hours if and only if the number of failures in 40,000 hours is ≤ 3.

$$P(X > 40{,}000) = P(N \leq 3)$$

The assumption that the failures follow a Poisson process implies that N has a Poisson distribution with

$$E(N) = 40{,}000(0.0001) = 4 \text{ failures per 40,000 hours}$$

$$P(N \leq 3) = \sum_{k=0}^{3} \frac{e^{-4} 4^k}{k!} = 0.433$$

Using Excel	
0.433	= POISSON(3, 4, TRUE)

Erlang Distribution

Generalizing from the prior exercise:

$$P\left(X > x\right) = \sum_{k=0}^{r-1} \frac{e^{\lambda x}\left(\lambda x\right)^{k}}{k!} = 1 - F\left(x\right)$$

Now differentiating $F\left(x\right)$:

$$f\left(x\right) = \frac{\lambda^{r} x^{r-1} e^{-\lambda x}}{\left(r-1\right)!} \quad \text{for } x > 0 \text{ and } r = 1, 2, \ldots$$

Gamma Function

The gamma function is the generalization of the factorial function for r > 0, not just non-negative integers.

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx, \quad \text{for } r > 0$$

Properties of the gamma function

$$\Gamma(r) = (r-1)\Gamma(r-1) \quad \text{recursive property}$$

$$\Gamma(r) = (r-1)! \qquad \text{factorial function}$$

$$\Gamma(1) = 0! = 1$$

$$\Gamma(1/2) = \pi^{1/2} = 1.77 \qquad \text{useful if manual}$$

Gamma Distribution

The random variable *X* with probability density function:

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}, \text{ for } x > 0$$

is a gamma random variable with parameters $\lambda > 0$ and $r > 0$. If r is an integer, then *X* has an Erlang distribution.

Mean & Variance of the Gamma

If X is a gamma random variable with parameters λ and r,

$$\mu = E(X) = r / \lambda$$

and

$$\sigma^2 = V(X) = r / \lambda^2$$

Example 4-24: Gamma Application-1

The time to prepare a micro-array slide for high-output genomics is a Poisson process with a mean of 2 hours per slide. What is the probability that 10 slides require more than 25 hours?

Let *X* denote the time to prepare 10 slides. Because of the assumption of a Poisson process, *X* has a gamma distribution with $\lambda = \frac{1}{2}$, $r = 10$, and the requested probability is $P(X > 25)$.

Using the Poisson distribution, let the random variable *N* denote the number of slides made in 10 hours. The time until 10 slides are made exceeds 25 hours if and only if the number of slides made in 25 hours is ≤ 9.

$$P(X > 25) = P(N \le 9)$$

$$E(N) = 25(1/2) = 12.5 \text{ slides in 25 hours}$$

$$P(N \le 9) = \sum_{k=0}^{9} \frac{e^{-12.5}(12.5)^k}{k!} = 0.2014$$

Using Excel	
0.2014	= POISSON(9, 12.5, TRUE)

Example 4-24: Gamma Application-2

Using the gamma distribution, the same result is obtained.

$$P(X > 25) = 1 - P(X \le 25)$$
$$= 1 - \int_0^{25} \frac{0.5^{10} x^9 e^{-0.5x}}{\Gamma(10)} dx$$
$$= 1 - 0.7986$$
$$= 0.2014$$

Using Excel	
0.2014	= 1 - GAMMADIST(25,10,2,TRUE)

What is the mean and standard deviation of the time to prepare 10 slides?

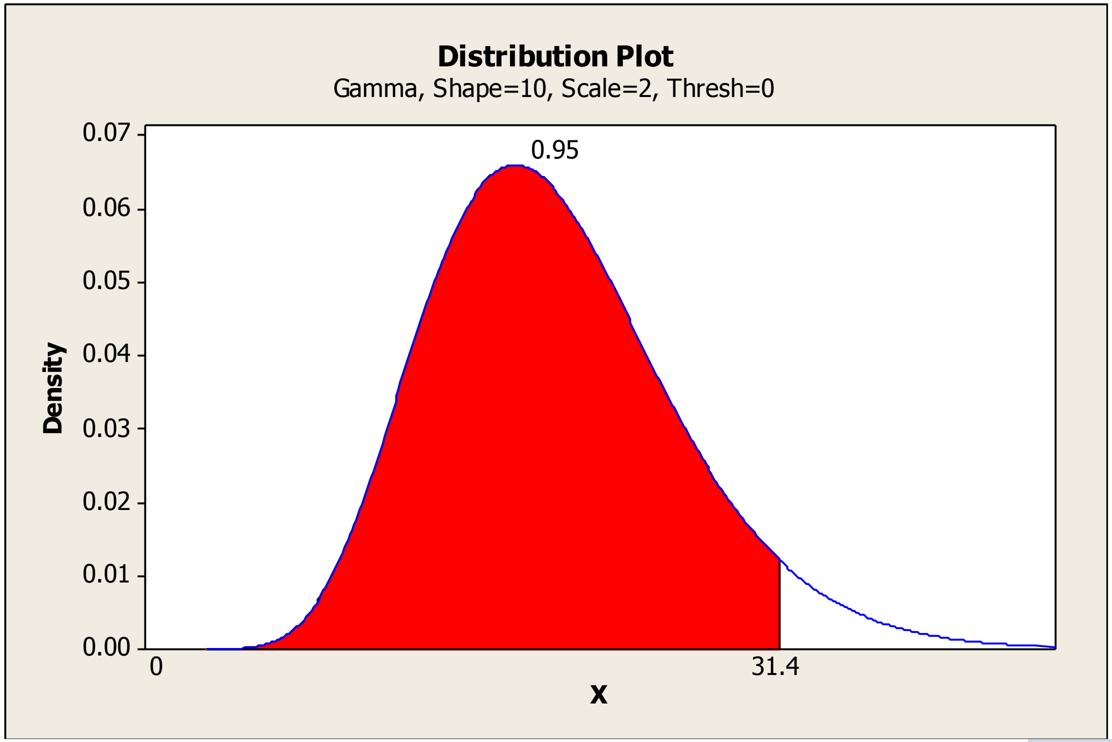
$$E(X) = \frac{r}{\lambda} = \frac{10}{0.5} = 20 \text{ hours}$$

$$\mathrm{V}(X) = \frac{r}{\lambda^2} = \frac{10}{0.25} = 40 \text{ hours}$$

$$SD(X) = \sqrt{V(X)} = \sqrt{40} = 6.32 \text{ hours}$$

Example 4-24: Gamma Application-3

The slides will be completed by what length of time with 95% probability? That is: $P(X \leq x) = 0.95$



Minitab: Graph > Probability Distribution Plot > View Probability

Using Excel	
31.41	=GAMMAINV(0.95, 10, 2)

Weibull Distribution

The random variable X with probability density function

$$f(x) = \frac{\beta}{\delta}\left(\frac{x}{\delta}\right)^{\beta-1} e^{-(x/\delta)^{\beta}} \quad , \text{ for } x > 0$$

is a Weibull random variable with

scale parameter $\delta > 0$ and shape parameter $\beta > 0$.

The cumulative distribution function is:

$$F(x) = 1 - e^{-(x/\delta)^{\beta}}$$

The mean and variance is given by

$$\mu = E(X) = \delta \cdot \Gamma\left(1 + \frac{1}{\beta}\right) \quad \text{and}$$

$$\sigma^2 = V(X) = \delta^2\left[\Gamma\left(1+\frac{2}{\beta}\right)\right] - \delta^2\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2$$

Example 4-25: Bearing Wear

- The time to failure (in hours) of a bearing in a mechanical shaft is modeled as a Weibull random variable with $\beta = ½$ and $\delta = 5{,}000$ hours.
- What is the mean time until failure?

$$E(X) = 5000 \cdot \Gamma(1 + 1/2) = 5000 \cdot \Gamma(1.5)$$
$$= 5000 \cdot 0.5\sqrt{\pi} = 4{,}431.1 \text{ hours}$$

Using Excel	
4,431.1	= 5000 * EXP(GAMMALN(1.5))

- What is the probability that a bearing will last at least 6,000 hours?

$$P(X > 6{,}000) = 1 - F(6{,}000) = e^{-\left(\frac{6000}{5000}\right)^{0.5}}$$
$$= e^{-1.0954} = 0.334$$

Using Excel	
0.334	= 1 - WEIBULL(6000, 1/2, 5000, TRUE)

Only 33.4% of all bearings last at least 6000 hours.

Lognormal Distribution

Let W denote a normal random variable with mean θ and variance ω^2, then $X = \exp(W)$ is a lognormal random variable with probability density function

$$f(x) = \frac{1}{x\omega\sqrt{2\pi}} e^{-\left[\frac{(\ln(x)-\theta)^2}{2\omega^2}\right]} \qquad 0 < x < \infty$$

The mean and variance of X are

$$E(X) = e^{\theta+\omega^2/2} \qquad \text{and}$$

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

Example 4-26: Semiconductor Laser-1

The lifetime of a semiconductor laser has a lognormal distribution with θ = 10 and ω = 1.5 hours.
What is the probability that the lifetime exceeds 10,000 hours?

$$\begin{aligned}
P(X > 10{,}000) &= 1 - P\left[\exp(W) \le 10{,}000\right] \\
&= 1 - P\left[W \le \ln(10{,}000)\right] \\
&= 1 - \Phi\left(\frac{\ln(10{,}000) - 10}{1.5}\right) \\
&= 1 - \Phi(-0.5264) \\
&= 1 - 0.30 \\
&= 0.701
\end{aligned}$$

1 - NORMDIST(LN(10000), 10, 1.5, TRUE) =	0.701

Example 4-26: Semiconductor Laser-2

- What lifetime is exceeded by 99% of lasers?

$$P(X > x) = P[\exp(W) > x] = P[W > \ln(x)]$$

$$= 1 - \Phi\left(\frac{\ln(x) - 10}{1.5}\right) = 0.99$$

From Appendix Table III, $1 - \Phi(z) = 0.99$ when $z = -2.33$

Hence $\frac{\ln(x) - 10}{1.5} = -2.33$ and $x = \exp(6.505) = 668.48$ hours

-2.3263	= NORMSINV(0.99)
6.5105	= -2.3263 * 1.5 + 10 = ln(x)
672.15	= EXP(6.5105)
(difference due to round-off error)	

- What is the mean and variance of the lifetime?

$$E(X) = e^{\theta + \varpi^2/2} = e^{10 + 1.5^2/2}$$

$$= \exp(11.125) = 67{,}846.29$$

$$V(X) = e^{2\theta + \omega^2}\left(e^{\omega^2} - 1\right) = e^{2 \cdot 10 + 1.5^2}\left(e^{1.5^2} - 1\right)$$

$$= \exp(22.25) \cdot [\exp(2.25) - 1] = 39{,}070{,}059{,}886.6$$

$$SD(X) = 197{,}661.5$$

Beta Distribution

The random variable *X* with probability density function

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\ x^{\alpha-1}(1-x)^{\beta-1}\ ,\ \text{for } x \text{ in } [0,\ 1]$$

is a beta random variable with parameters $\alpha > 0$ and $\beta > 0$.

Example 4-27: Beta Computation-1

Consider the completion time of a large commercial real estate development. The proportion of the maximum allowed time to complete a task is a beta random variable with α = 2.5 and β = 1. What is the probability that the proportion of the maximum time exceeds 0.7?

Let *X* denote the proportion.

$$P(X > 0.7) = \int_{0.7}^{1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx$$

$$= \frac{\Gamma(3.5)}{\Gamma(2.5)\cdot\Gamma(1)} \int_{0.7}^{1} x^{1.5}dx$$

$$= \frac{2.5(1.5)(0.5)\sqrt{\pi}}{1.5(0.5)\sqrt{\pi}} \bullet \left.\frac{x^{2.5}}{2.5}\right|_{0.7}^{1}$$

$$= 1-(0.7)^{2.5} = 0.59$$

Using Excel	
0.590	= 1 - BETADIST(0.7,2.5,1,0,1)

Mean & Variance of the Beta Distribution

If *X* has a beta distribution with parameters α and β,

$$\mu = E(X) = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = V(X) = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

Example 4-28: In the above example, α = 2.5 and β = 1. What are the mean and variance of this distribution?

$$\mu = \frac{2.5}{2.5+1} = \frac{2.5}{3.5} = 0.71$$

$$\sigma^2 = \frac{2.5(1)}{(2.5+1)^2 (2.5+1+1)} = \frac{2.5}{3.5^2 (4.5)} = 0.045$$

Mode of the Beta Distribution

If α >1 and β > 1, then the beta distribution is mound-shaped and has an interior peak, called the mode of the distribution. Otherwise, the mode occurs at an endpoint.

General formula:

$$\text{Mode} = \frac{\alpha - 1}{\alpha + \beta - 2} \quad , \quad \text{for } [0, 1]$$

For the above Example 4.28 the mode is

$$\text{Mode} = \frac{\alpha - 1}{\alpha + \beta - 2} = \frac{2.5 - 1}{2.5 + 1 - 2}$$

$$= \frac{1.5}{1.5}$$

$$= 1$$

case	alpha	beta	mode	
Example 4-28	2.25	1	1.00	= (2.5-1) / (2.5+1.0-2)

Important Terms & Concepts of Chapter 4

Beta distribution
Chi-squared distribution
Continuity correction
Continuous uniform distribution
Cumulative probability distribution for a continuous random variable
Erlang distribution
Exponential distribution
Gamma distribution
Lack of memory property of a continuous random variable
Lognormal distribution
Mean for a continuous random variable
Mean of a function of a continuous random variable
Normal approximation to binomial & Poisson probabilities
Normal distribution
Probability density function
Probability distribution of a continuous random variable
Standard deviation of a continuous random variable
Standardizing
Standard normal distribution
Variance of a continuous random variable
Weibull distribution