IE6200: Engineernig Probability and Statistics

LAB 05: Probability Distributions

Prof. Mohammad Dehghani

Contents

1 Probability Distributions

In this lab, a few of the most common discrete and continuous probability distributions will be discussed. It will be followed examples solved using functions in R.

2 Discrete Probability Distributions

The *probability mass function (PMF)* is a function that describes the probability distribution of a discrete random variable X. It creates a list of each possible value of X together with the probability that X takes that value in one trial of the experiment. Properties of PMF are:

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$
3. $\mathrm{P}(X = x) = f(x)$

2.1 Binomial Distribution

The *binomial distribution* is a of discrete probability distribution, in which the probabilities of interest are those of receiving a certain number of successes in n independent trials each having only two possible outcomes with the same probability of success, p.

A descrete random variable X having a binomial distribution, with n independent trials, probability of success p, and probability of failure q = *1* -p , is given by:

$$\boxed{X \sim b(x; n, p)} \tag{1}$$

Its probability mass function is given by:

$$\boxed{P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \, where \, x = 1, 2, 3...n} \tag{2}$$

The mean and variance of the binomial distribution are:

- $\mu = np$
- $\sigma^2 = npq.$

Properties of binomial Distribution

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

Applications

1. To determine whether the product manufactured is defective or non-defective.
2. To determine whether an individual will have a certain genotype or phenotype.
3. To determine whether a person will gets the job after their interviews.

2.2 Multinomial Distribution

A *multinomial experiment* is almost identical to a binomial distribution with one main difference: a binomial experiment can have two outcomes, while a multinomial experiment can have multiple outcomes. In *multinomial distribution*, the probabilities of interest are those of receiving a certain number of successes in n independent trials each having more than two possible outcomes with the same probability of success, p.

The descrete random variables x_1,x_2,. . . . , x_k having a multinomial distribution, with k outcomes E_1,E_2, . . . , E_k and probabilities $p = p_1$, p_2, . . . , p_k for a given trial, is given by:

$$\boxed{X \sim Mult(n, p)} \tag{3}$$

and its probability mass function representing the number of occurrences for E_1,E_2, . . . , E_k in n independent trials, is given by:

$$\boxed{P_r = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}} \tag{4}$$

with

$$\sum_{i=1}^{k} x_i = n$$

and

$$\sum_{i=1}^{k} p_i = 1$$

Applications

1. To determine who will win the next election, out of five candidates, each with a fixed probability of receiving a particular vote.
2. To determine the result of a chess match, having a certain probability of player A winning, players B winning, or the game ending in a draw.

2.3 Geometric Distribution

The *geometric distribution* is a special case of the *negative binomial distribution*. It deals with the number of trials required for a single success.

The random variable X for a geometric distribution is given by:

$$\boxed{X \sim G(p)} \tag{5}$$

Its probability mass function is given by:

$$\boxed{P(X = x) = pq^{x-1}} \tag{6}$$

and its cumulative distribution function is given by:

$$\boxed{F(x) = 1 - (1 - p)^x} \tag{7}$$

Applications

1. To determine the first correct answer by an algorithm after running for 'n' number of times.
2. To analyzing the probability a batter earns a hit before he receives three strikes in a baseball match.
3. In cost-benefit analyses, such as a company deciding whether to fund research trials that, if successful, will earn the company some estimated profit, the goal is to reach a success before the cost outweighs the potential gain.

2.4 Hypergeometric Distribution:

The *hypergeometric distribution* is similar to binomial distribution with one main difference: independence among trials is required in a binomial experiment, while a hypergeometric experiment does not require independence among trials.

A discrete random variable X having a hypergeometric distribution, with the number of successes in a random sample of size n selected from N items of which k are labeled success is given by:

$$\boxed{X \sim h(x; N, n, k)} \tag{8}$$

And its probability mass function is given by:

$$\boxed{f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \text{for } max(0, n-(N-k)) \leq x \leq min(n, k)} \tag{9}$$

The mean and variance of the hypergeometric distribution are:

- $\mu = \frac{(nk)}{N}$
- $\sigma^2 = \frac{N-n}{N-1} n \frac{k}{N}(1 - \frac{k}{N})$

Applications

1. To determine the probability of finishing with a flush of spades if the two private cards are bith spades while playing Texas Hold'em card game.
2. To determine the probability of a woman for a jury from a pool of 'n' members out of which 'k' members are men.

2.5 Poisson Distribution

The *poisson distribution* expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.

A discrete random variable X having a poisson distribution, with average number of outcomes per unit time, distance, area, or volume denoted by λ in a given time interval, is given by:

$$\boxed{X \sim P(x;\lambda)} \tag{10}$$

Its probability mass function is given by:

$$\boxed{P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!} \; where\, x = 1, 2, 3, ...n} \tag{11}$$

and its cumulative distribution function is given by:

$$\boxed{F(x;\lambda) = \sum_{i=0}^{x} \frac{\lambda^x e^{-\lambda}}{x!}} \tag{12}$$

Both the mean and the variance of the poisson distribution p(x; λ) are λ.

Properties of Poission Distribution

1. 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.
2. 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.
3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Applications

1. To determine the number of calls coming per minute into a hotel's reservation center.
2. To determine the number of births expected during the night in a hospital.
3. To determine the number of automobiles arriving at a traffic light within the hour.

3 Continuous Probability Distributions

The *probability density function (PDF)* is a function which maps the values of the variable in a certain interval to the probability of their occurrence. The PDF characterizes the distribution of a continuous random variable. Properties of PDF are:

1. $f(x) \geq 0, \; for \; all \; x \in R$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$
3. $\mathrm{P}[a \leq X \leq b] = \int_a^b f(x)\, dx$

3.1 Uniform Distribution

The *uniform distribution* is a continuous probability distribution, for which all of the values that a random variable can take on occur with equal probability.

A continuous random variable X having a uniform distribution, on the interval [A, B], is given by:

$$\boxed{X \sim U(A, B)} \tag{13}$$

and its probability density function is given by:

$$\boxed{f(x) = \begin{cases} \frac{1}{B-A} & \text{for } A \leq x \leq B, \\ 0 & \text{for } x < A \text{ or } x > B \end{cases}} \tag{14}$$

The mean and variance of uniform distribution is given by:

- $\mu = (A + B)/2$
- $\sigma^2 = (B - A)^2/12$

Applications

1. Consider a "spinner": an object like an unmagnetized compass needle that can pivots freely around an axis, and is stable pointing in any direction. You give it a spin and see where it comes to rest, measuring the resulting angle (divided by 2π) as a number from 0 to 1.
2. In analog-to-digital conversion a quantization error occurs. This error is either due to rounding or truncation. When the original signal is much larger than one least significant bit (LSB), the quantization error is not significantly correlated with the signal, and has an approximately uniform distribution. The RMS error therefore follows from the variance of this distribution.
3. Consider throwing a dart at a dart board. Assuming that all directions are equally likely, the angle of deflection from the x-axis drawn through the bullseye should be uniformly distributed between 0 and 360° or 0 and 2π. Rescaling would produce a uniform [0, 1].

3.2 Triangular Distribution

The *triangular distribution* is shaped like a triangle with three parameters like minimum value a, peak value b, and maximum value c, is given by:

The continuous random variable of X having a trinagular distribution is given by

$$\boxed{X \sim triangular(a, b, c)} \tag{15}$$

And its probability density function is given by:

$$\boxed{f(x) = \begin{cases} 0 & \text{for } x < a, \\ \frac{2(x-a)}{(b-a)(c-a)} & \text{for } a \leq x \leq c, \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } c \leq x \leq b, \\ 0 & \text{for } x > b \end{cases}} \tag{16}$$

The mean and variance of the triangular distribution are:

- $\mu = \frac{(a+b+c)}{3}$
- $\sigma^2 = \frac{a^2+b^2+c^2-ab-ac-bc}{18}$

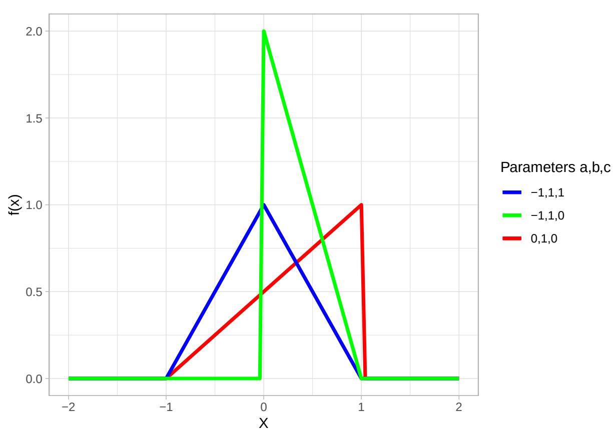


Figure 1: PDF of Triangular Distribution

3.3 Normal Distribution

The *normal distribution* is a continuous probability distribution, which has the same general shape, but differing in their location (that is, the mean or average) and scale parameters (that is, the standard deviation). The graph of its probability density function is a symmetric and bell-shaped curve.

A continuous random variable X having a normal distribution, with mean μ and variance σ^2, is given by:

$$\boxed{X \sim \mathcal{N}(\mu, \sigma)} \tag{17}$$

and the probability density of the normal distribution is:

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \tag{18}$$

A normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ is called the *standard normal distribution* and is shown in Figure 1.

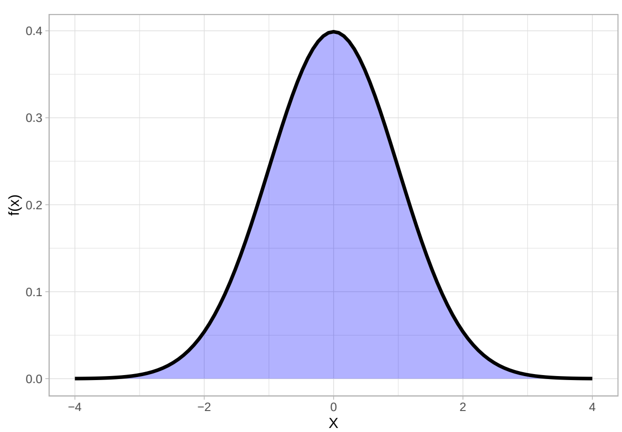


Figure 2: PDF of Standard Normal Distribution

Properties of Normal Distribution

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

Applications

1. Suppose a student ride a bus to school every morning. Barring a serious delay due to an accident or breakdown, there's a good chance that the time his bus arrives to pick him up in the morning will be well-modeled by a normal distribution. Because the arrival time will depend on things like times the bus spends at traffic lights and times it spends waiting for other students before it gets to him, and these individual variations are probably roughly independent and roughly similar in duration. If he write down the arrival time of the bus every day for a couple of months, he can accurately estimate the mean and standard deviation of the arrival time. He can figure out exactly how late he can wake up in the morning and still have a 99% chance of not missing the bus.
2. Suppose a student is out of school and trying to figure out a budget. It's probably the case that his monthly expenditures on groceries will be approximately normally distributed. His total monthly expenditures might not be normally distributed if he include things like the airplane ticket he buy in June to go home for a holiday or money he rack up on Christmas presents on Black Friday. Notice that his monthly grocery expense is the result of a large number of purchases that are similar in size and whose variations are due to independent factors like variations in produce prices and variations on what he pick out to eat.
3. Suppose you are a parent and your young kids just got a new video game system. If left to their own devices, the time they spend playing it each day will be approximately normal since the various times of each "session" they play will vary based on the time to get bored, to get hungry, to get mad at a sibling, etc. Each "session" will be approximately independent and approximately the same size, so their daily total will be approximately normal.

3.4 Log Normal Distribution

The *lognormal distribution* is a continuous probability distribution, with parameters $\mu \in R$ and $\sigma > 0$ if $ln(X)$ has the normal distribution with mean μ and standard deviation σ, is given by:

$$\boxed{X \sim logn(\mu, \sigma)} \tag{19}$$

and its probability density function is given by:

$$\boxed{P(X = x) = \frac{1}{x\sigma\sqrt{2\pi}}\; e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}} \tag{20}$$

The mean and variance of the lognormal distribution are given by:

- $\mu = e^{\mu + \sigma^2/2}$
- $\sigma^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$

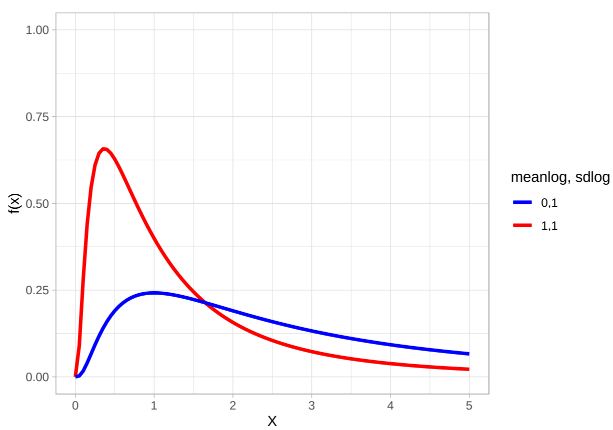


Figure 3: PDF of Log Normal Distribution

Applications

1. To determine the impact of an epidemics in a particular area.
2. To predict human behaviour by analysing the length of comments in internet discussion forums.
3. To model times to repair a maintainable system in reliability analysis.

3.5 Gamma Distribution

The *gamma distribution* is characterized with two parameters α and β, is given by:

$$\boxed{X \sim \Gamma(\alpha, \beta)} \tag{21}$$

And its probability density function is given by:

$$\boxed{f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0, \\ 0 & \text{for } elsewhere \end{cases}} \tag{22}$$

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

The mean and variance of the gamma distribution are:

- $\mu = \alpha\beta$
- $\sigma^2 = \alpha\beta^2$.

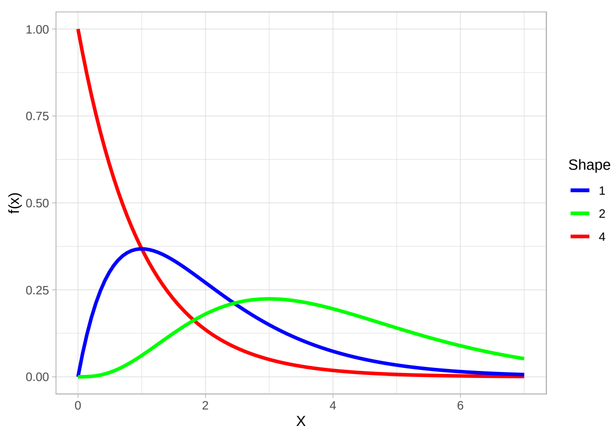


Figure 4: PDF of Gamma Distribution

3.6 Exponential Distribution

The *exponential distribution* is characterized by a single parameter called rate and is denoted by λ. Time between arrivals at service facilities and time to failure of component parts and electrical systems often are nicely modeled by the exponential distribution.

A continuous random variable X having a exponential distribution, with with parameters $\alpha = 1$ and $\beta > 0$, is given by:

$$\boxed{X \sim Exp(x; \beta)} \tag{23}$$

Its probability mass function is given by:

$$\boxed{f(x;\beta) = \begin{cases} \beta e^{-\beta x} & x \geq 0, \\ 0 & x < 0. \end{cases}} \tag{24}$$

And its cumulative distribution function is given by:

$$\boxed{F(x;\beta) = \begin{cases} 1 - e^{-\beta x} & x \geq 0, \\ 0 & x < 0. \end{cases}} \tag{25}$$

Its relationship with poisson distribution can be describe with the relation of rate $\lambda = 1/\ \beta$

The mean and variance of the exponential distribution are:

- $\mu = \beta$
- $\sigma^2 = \beta^2$.

Exponential distribution has memory less property, which tells that the distribution of remaining time does not depend on how long the component has been operating.

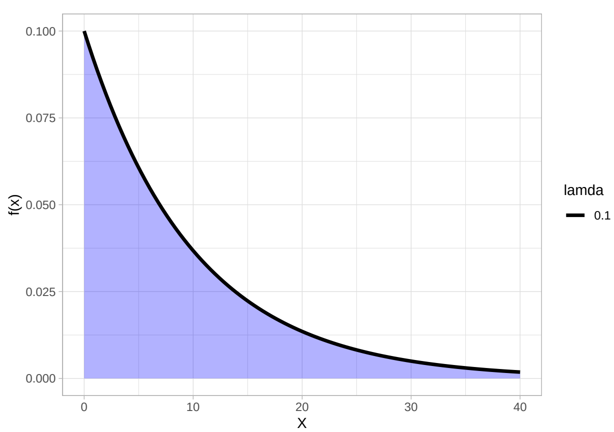


Figure 5: PDF of Expontential Distribution

Applications

1. To determine the probability of a battery dying over a period of time.
2. To determine the risk of a client getting in an accident for a auto insurers.
3. To determine the waiting time for a vehicle to develop a traffic simulation.

4 Probability Distributions in R

R comes with built-in implementations of many probability distributions. A few of the most commonly used distributions will be covered in this lab. Each probability distribution in R is associated with four functions which are summarized in the following table:

Table 1: Probability Distribution in R

Function	What it does
d function	Evaluates Probability Density (PMF/PDF)
p function	Evaluates Cumulative Distribution (CDF)
q function	Evaluates quantile (inverse CDF)
r function	Generate random variables

For example, for the normal distribution, these functions are dnorm(), pnorm(), qnorm(), and rnorm.()

Input arguements for these functions follow the same template but vary depending on the parameter of the distribution being evaluated. For example, parameters for normal distribution are *mean* and *standard deviation*, while for poisson distribution it is *lambda*.

- x or q represent the vector of values of R.V. to be evaluated
- p represents the vector of probabilites
- n represents the number of random values to return in case of the r function

To get a full list of the distributions available in R, use the following command ?Distributions. To get functions about a particular distribution, such as Normal distribution, use ?Normal

4.1 Discrete Distributions

4.1.1 Binomial Distribution

Example 1: A manufacturer of metal pistons finds that on the average, 12% of his pistons are rejected because they are either oversize or undersize. Given a batch of 10 pistons is selected, therefore

- $X \equiv$ R.V. of the number of rejected pistons
- $X \sim b(x; n = 10, p = 0.12)$

(a) What is the probability that the batch will contain 2 rejected pistons [P(X = 2)]

```r
# x is the value of the R.V. for which the probability is to be calculated
dbinom(x=2,size=10,prob=0.12)
```

```
[1] 0.2330432
```

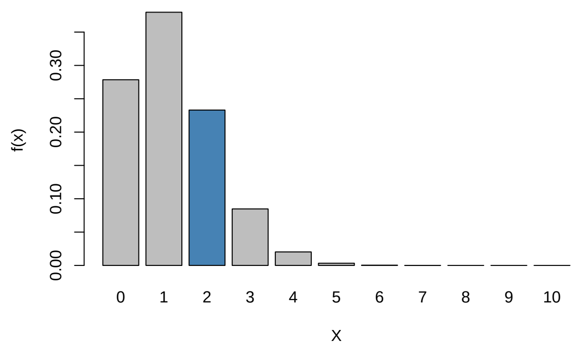


Figure 6: Example 1.a

(b) What is the probability that the batch will contain at least 2 rejects [P(X $\geq$ 2) = 1 - P(X $\leq$ 1) = 1 - $F(1)$]

```
1-pbinom(q=1,size=10,prob=0.12)
```

```
[1] 0.341725
```

Another way to calculate the above probability is to use the lower.tail arguement in the p function. Refer the help documentation and try it out.

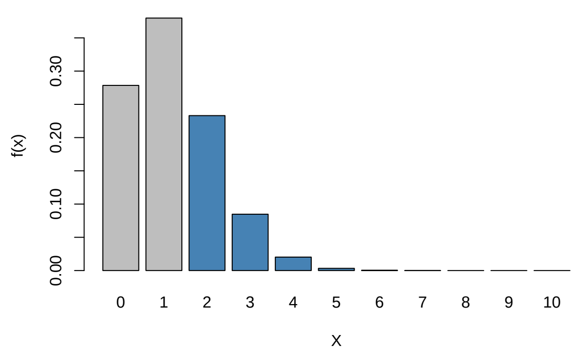


Figure 7: Example 1.b

4.1.2 Multinomial Distribution

Example 2: Three card players play a series of matches. The probability that player A wins any game is 20%, the probability that player B wins is 30%, and the probability player C wins is 50%.

- X: [X_1, X_2, X_3]
  - $X_1 \equiv$ R.V. of the number of successes (wins) for player A
  - $X_2 \equiv$ R.V. of the number of successes (wins) for player B
  - $X_3 \equiv$ R.V. of the number of successes (wins) for player C
- $X \sim Mult(x; n = 7, p = [0.2, 0.3, 0.5])$

(a) If they play 7 games, what is the probability that player A wins 4 games, player B wins 1 games, and player C wins 2? [P(X_1 = 4, X_2 = 1, X_3 = 2)]

```
dmultinom(x=c(4,1,2),,prob=c(0.2, 0.3, 0.5))
```

```
[1] 0.0126
```

(b) If they play 7 games, what is the probability that player B wins 3 game and C wins 4 games? [P(X_1 = 0, X_2 = 3, X_3 = 4)]

```
dmultinom(x=c(0,3,4),,prob=c(0.2, 0.3, 0.5))
```

```
[1] 0.0590625
```

Multinomial Distritution does not have a p function that evaluates CDF because of the complexity involved in calculating such a metric.

4.1.3 Geometric Distribution

Example 3: In an amusement fair, a competitor is entitled for a prize if he/she throws and lands a ring on a peg from a certain distance. It is observed that only 30% of the competitors are able to do this.

- $X \equiv$ R.V. of the number of throws before landing a ring on the peg for the *first* time
- $X \sim G(x; p = 0.3)$

(a) If someone is given 5 chances, what is the probability of his/her winning the prize when he/she has already missed 4 chances? [P(X = 5)]

```
dgeom(x=4, prob=0.3)
```

```
[1] 0.07203
```

In R, for geometric distribution, arguement for x is the number of failures before a success occurs, as such, to calculate probability for success in 5th trial, input for x is 4.

4.1.4 Hypergeometric Distribution

Example 4: A consignment of 100 microprocessors has arrived. 10 out of the 100 in the consignment are actually defective. To check the consignment the buyer randomly checks 15 microprocessors.

- $X \equiv$ R.V. of the number of defective microprocessors in the cosignment
- $X \sim h(x; N = 100, n = 15, k = 10)$

For Hypergeometric Distribution, there is a slight difference in notations usually found in textbooks and in R

Table 2: Hypergeometric Distriburtion Notations

Parameter	Textbook Notation	R Notation	Example Value
Population Size	N		100
Sample Size	n	k	15
Number of successes	k	m	10
Number of failures		n	90

(a) Find the probability that the buyer finds 4 defective processors in the check he conducts. [P(X = 4)]

```
dhyper(x=4, m=10, n=90, k=15)
```

```
[1] 0.0344874
```

(b) Find the probability that the buyer find two or defective processors in the check he conducts. [P(X $\geq$ 2) = 1 - P(X $\leq$ 1) = 1 - $F(1)$]

```
1-phyper(q=1, m=10, n=90, k=15)
```

```
[1] 0.4624509
```

4.1.5 Poisson Distribution

Example 5: The number of calls coming per minute into a hotel's reservation center is Poisson random variable with mean 3.

- $X \equiv$ R.V. of the number of calls received in the hotel in a minute
- $X \sim P(x; \lambda = 3)$

(a) Find the probability that 5 calls come in a given 1-minute period [P(X=5)]

```r
dpois(x=5,lambda=3)
```

```
[1] 0.1008188
```

(b) Find the probability that at most 5 calls come in a given 1-minute period [P(X $\leq$ 5)]

```r
ppois(q=5,lambda=3)
```

```
[1] 0.9160821
```

4.2 Continuous Distributions

4.2.1 Uniform Distribution:

Example 6: Suppose the delay in arrival of a train is uniformly distributed upto 60 minutes.

- $X \equiv$ R.V. of delay of train in minutes
- $X \sim U(x; A = 0, B = 60)$

Here A and B correspond to the minimum and maximum values of the uniform distribution.

(a) What is the probability that train will be delayed by 40 to 55 minutes? [P(40 $\leq$ X $\leq$ 55)]

```
punif(q=55,min=0,max=60) - punif(q=40,min=0,max=60)
```

```
[1] 0.25
```

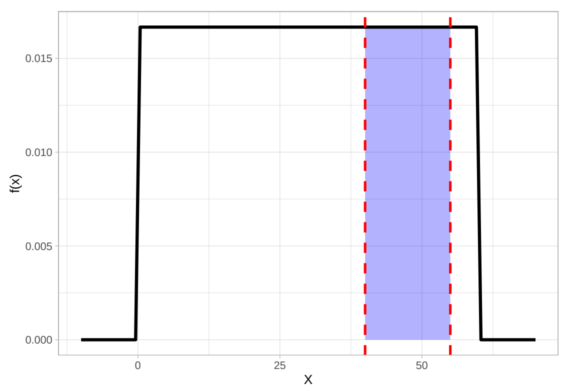


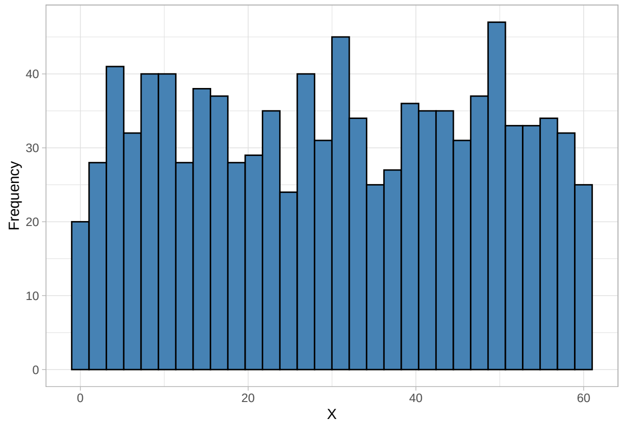
Figure 8: Example 6.a

(b) Generate 1,000 random numbers using the given uniform distribution and plot them into Histogram

```
y_unif <- runif(n=1000,min=0,max=60)
head(y_unif)
```

```
[1]  3.818183 51.091743 35.309816 56.818855 51.742231  6.749811
```

```
ggplot(data.frame(y_unif), aes(y_unif)) +
  geom_histogram(fill="steelblue", color = 'black') +
  theme_light() +
  ylab('Frequency') +
  xlab('X')
```



ϟ Try creating another histogram after generating 10,000 and 50,000 random numbers

4.2.2 Standard Normal Distribution

- $X \sim \mathcal{N}(\mu = 0, \sigma = 1)$

Example 7.a: Find P(Z<-1.68)

```
pnorm(q=-1.68, mean=0, sd=1)
```

```
[1] 0.04647866
```

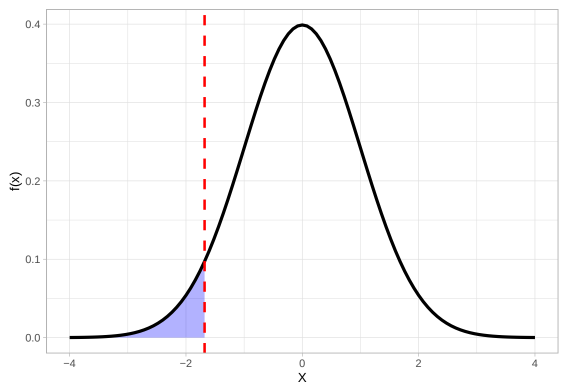


Figure 9: Example 7.a

(b) Find P(Z>-1.68)

```
1-pnorm(q=-1.68)
```

```
[1] 0.9535213
```

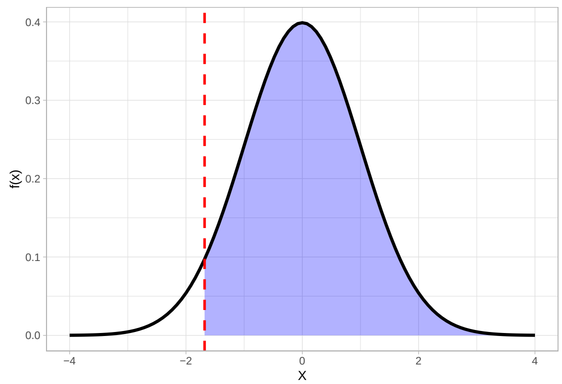


Figure 10: Example 7.b

(c) Find P(Z<0)

```r
pnorm(q=0)
```

```
[1] 0.5
```

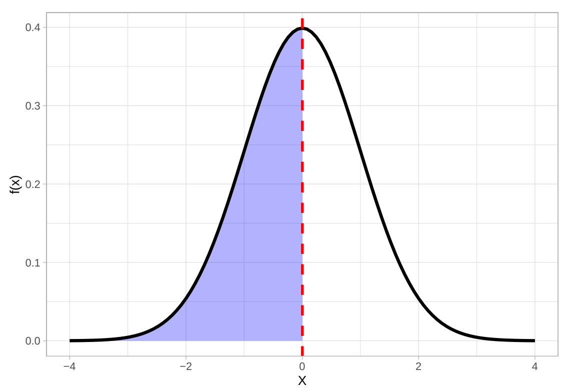


Figure 11: Example 7.c

(d) Find $Z_{0.025}$

```r
qnorm(p=0.025, mean=0,sd=1)
```

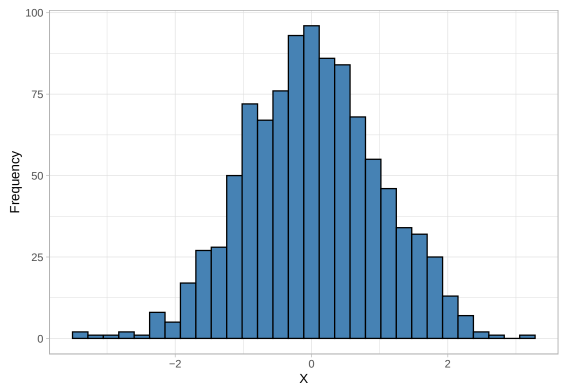
```
[1] -1.959964
```

(e) Generate 1,000 random numbers using Z distribution and plot them into Histogram

```r
x_norm <- rnorm(n=1000,mean=0,sd=1)
head(x_norm)
```

```
[1] -0.9184966 -0.9020393  0.3639336  0.2037506 -0.8308564  0.4195510
```

```r
ggplot(data.frame(x_norm), aes(x_norm)) +
  geom_histogram(fill="steelblue", color = 'black') +
  theme_light() +
  xlab('X') +
  ylab('Frequency')
```



Try creating another histogram after generating 10,000 and 50,000 random numbers

4.2.3 Normal Distribution

Example 8: It was found that the mean length of 100 parts produced by a lathe was 20.05 mm with a standard deviation of 0.02 mm.

- $X \equiv$ R.V. of the length of lathe machined parts in mm
- $X \sim \mathcal{N}(x; \mu = 20.05, \sigma = 0.02)$

(a) Find the probability that a part selected at random would have a length between 20.03 mm and 20.08 mm [P(20.03 $\leq$ X $\leq$ 20.08)]

```r
pnorm(q=20.08,mean=20.05,sd=0.02) - pnorm(q=20.03,mean=20.05,sd=0.02)
```

```
[1] 0.7745375
```

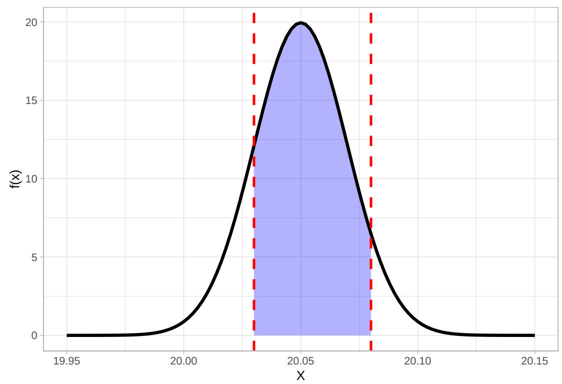


Figure 12: Example 8.a

(b) Find the probability that a part selected at random would have a length at most 20.06 mm [P(X $\leq$ 20.06)]

```
pnorm(q=20.06,mean=20.05,sd=0.02)
```

```
[1] 0.6914625
```

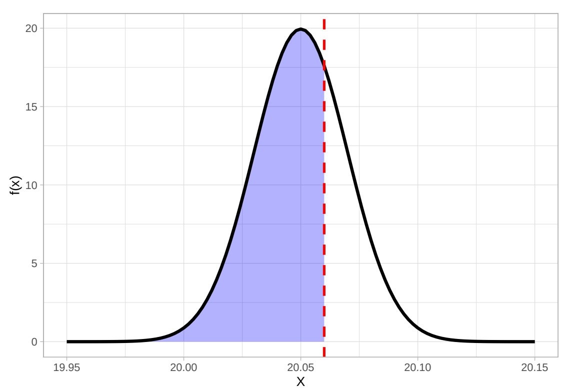


Figure 13: Example 8.b

(c) Find the probability that a part selected at random would have a length at least 20.07 mm [P(X $\geq$ 20.07)]

```
1-pnorm(q=20.07,mean=20.05,sd=0.02)
```

```
[1] 0.1586553
```

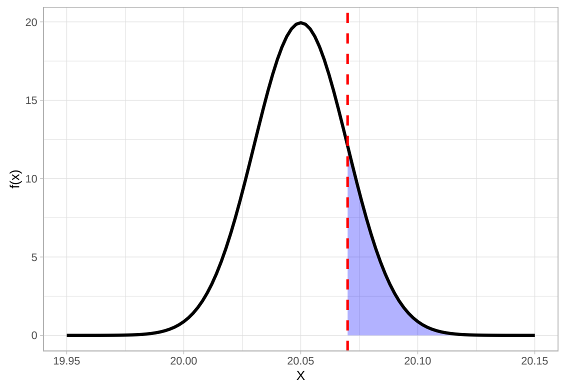


Figure 14: Example 8.c

4.2.4 Log Normal Distribution

Example 9: Suppose that the random variable X has a lognormal distribution with parameters $\mu = 1$ and $\sigma = 2$

- $X \sim logn(x; \mu = 1, \sigma = 2)$

(a) Find P(X $\leq$ 3)

```
plnorm(q=3,meanlog=1,sdlog=2)
```

```
[1] 0.5196623
```

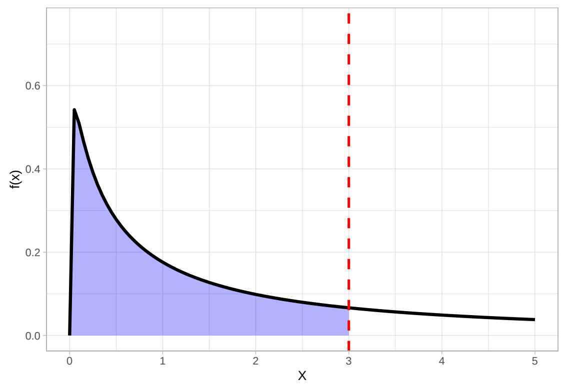


Figure 15: Example 9

4.2.5 Exponential Distribution

Example 10: The number of days ahead travelers purchase their airline tickets can be modeled by an exponential distribution with the average amount of time equal to 15.6 days.

- $X \equiv$ R.V. of the airline ticket purchased prior to travel in days
- $X \sim Exp(x; \beta = 0.064)$

(a) Find the probability that a traveler will purchase a ticket between three to seven days in advance [P(3 $\leq$ X $\leq$ 7)]

```
pexp(q=7, rate=1/15.6) - pexp(q=3, rate=0.064)
```

```
[1] 0.1868607
```

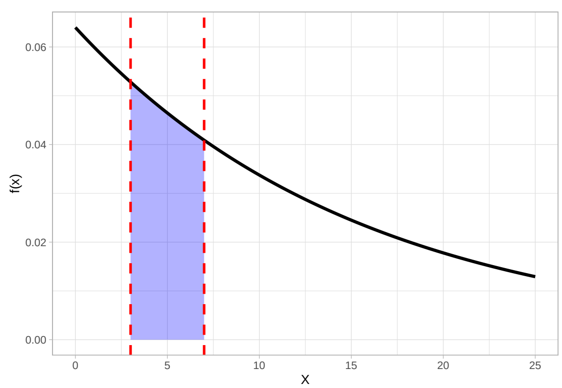


Figure 16: Example 10.a

(b) Find the probability that a traveler will purchase a ticket fewer than ten days in advance [P(X $\leq$ 10)]

```
pexp(q=10, rate=0.064)
```

```
[1] 0.4727076
```

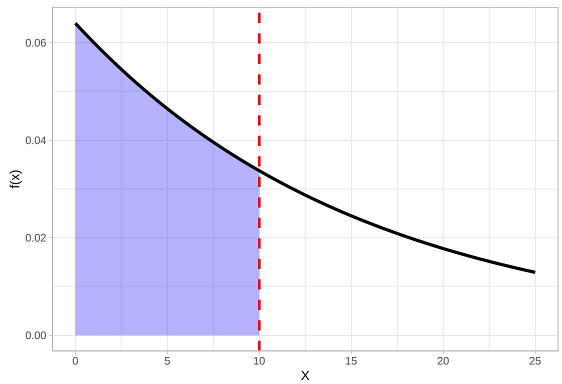


Figure 17: Example 10.b

(c) How many days do half of all travelers wait before purchasing the tickets?

```
qexp(p=0.5, rate=0.064)
```

```
[1] 10.83042
```

4.2.6 Triangular Distribution

Example 11: A burger franchise planning a new outlet in Auckland. They use a triangular distribution to model the future weekly sales with a minimum value of a=$1000, and maximum value of b=$6000 and a peak value of c=$3000.

- $X \equiv$ R.V. of the weekly sale of a new outlet
- $X \sim triangular(x; a = 1000, b, = 6000, c = 3000)$

where a = Minimum Value; b = Maximum Value, C = Mode of the distribution

(a) Determine the probability the new outlet will have weekly sales of less than $2000.

```
ptriangle(q = 2000, a = 1000, b = 6000, c = 3000)
```

```
[1] 0.1
```

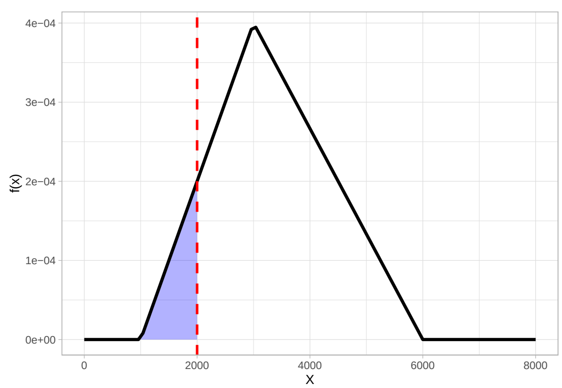


Figure 18: Example 11.a

(b) Determine the probability the new outlet will have weekly sales of more than $3500.

```
1 - ptriangle(q = 3500, a = 1000, b = 6000, c = 3000)
```

```
[1] 0.4166667
```

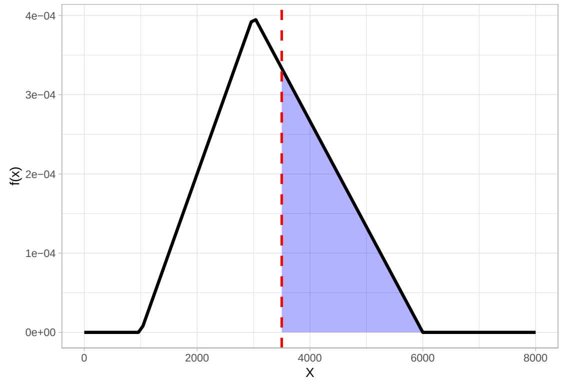


Figure 19: Example 11.b

(c) Determine the probability the new outlet will have weekly sales between \$2500 ,and \$3500.

```r
ptriangle(q = 3500, a = 1000, b = 6000, c = 3000) -
  ptriangle(q = 2500, a = 1000, b = 6000, c = 3000)
```

```
[1] 0.3583333
```

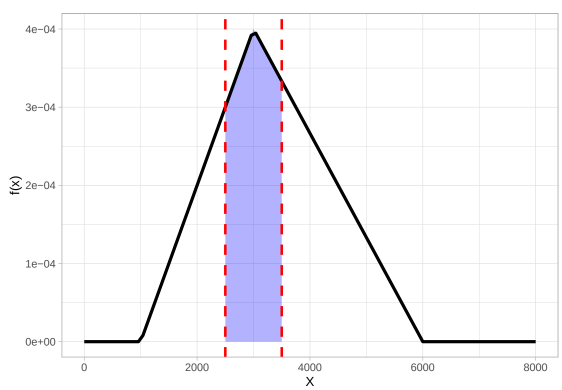


Figure 20: Example 11.c