STA 256: Statistics and Probability I

Al Nosedal.
University of Toronto.

Fall 2017

My momma always said: "Life was like a box of chocolates. You never know what you're gonna get."

Forrest Gump.

Exercise 4.1

Let X be a random variable with $p(x)$ given in the table below.

x	1	2	3	4
p(x)	0.4	0.3	0.2	0.1

a) Find an expression for the function $F(x) = P(X \leq x)$.
b) Sketch the function given in part a).

Solution

If $x < 1$, then
$F(x) = P(X \leq x) = 0$ because X does not assume values that are less than 1.

If $1 \leq x < 2$, then
$F(x) = P(X \leq x) = P(X = 1) = 0.4$.

Solution

If $2 \le x < 3$, then

$$F(x) = P(X \le x) = P(X = 1) + P(X = 2)$$
$$= 0.4 + 0.3 = 0.7.$$

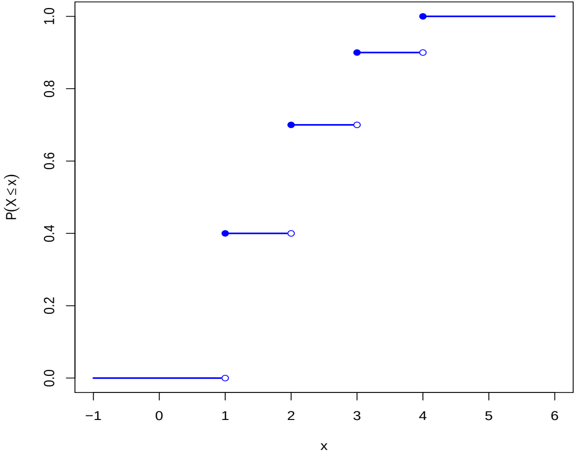
If $3 \le x < 4$, then

$$F(x) = P(X \le x) = P(X = 1) + P(X = 2) + P(X = 3)$$
$$= 0.4 + 0.3 + 0.2 = 0.9.$$

If $4 \le x < \infty$, then

$$F(x) = P(X \le x)$$
$$= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$
$$= 0.4 + 0.3 + 0.2 + 0.1 = 1.$$

Solution



Definition 4.1

Let Y denote any random variable. The **distribution function** (or **cumulative distribution function**) of Y, denoted by $F(y)$, is such that $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

Theorem 4.1

Properties of a Distribution Function[1]. If $F(y)$ is a distribution function, then

1. $\lim_{y \to -\infty} F(y) = 0$.
2. $\lim_{y \to \infty} F(y) = 1$.
3. $F(y)$ is a nondecreasing function of y.[If y_1 and y_2 are *any values such that* $y_1 < y_2$, *then* $F(y_1) \leq F(y_2)$.*]*

[1] *To be mathematically rigorous, if* $F(y)$ *is a valid distribution function, then* $F(y)$ *also must be right continuous.*

Definition 4.2

A random variable Y with distribution function $F(y)$ is said to be **continuous** if $F(y)$ is continuous, for $-\infty < y < \infty$.

Definition 4.3

Let $F(y)$ be the distribution function for a continuous random variable Y. Then $f(y)$, given by

$$f(y) = \frac{dF(y)}{dy} = F'(y)$$

wherever the derivative exists, is called the **probability density function** for the random variable Y.

Theorem 4.2

Properties of a Density Function. If $f(y)$ is a density function for a continuous random variable, then

1. $f(y) \geq 0$ for all y, $-\infty < y < \infty$.
2. $\int_{-\infty}^{\infty} f(y)dy = 1$.

Exercise 4.11

Suppose that Y possesses the density function

$$f(y) = \begin{cases} cy, & 0 \leq y \leq 2, \\ 0 & \text{elsewhere.} \end{cases}$$

a. Find the value of c that makes $f(y)$ a probability density function.
b. Find $F(y)$.
c. Graph $f(y)$ and $F(y)$.
d. Use $F(y)$ to find $P(1 \leq Y \leq 2)$.
e. Use $f(y)$ and geometry to find $P(1 \leq Y \leq 2)$.

Solution a)

$\int_{-\infty}^{0} 0dy + \int_0^2 cydy + \int_2^{\infty} 0dy = 1$
$\int_0^2 cydy = 1$
$c\int_0^2 ydy = 1$
$\int_0^2 ydy = \frac{1}{c}$
$\frac{2^2}{2} - \frac{0^2}{2} = \frac{1}{c}$
$\frac{4}{2} = \frac{1}{c}$
Therefore, $c = \frac{1}{2}$

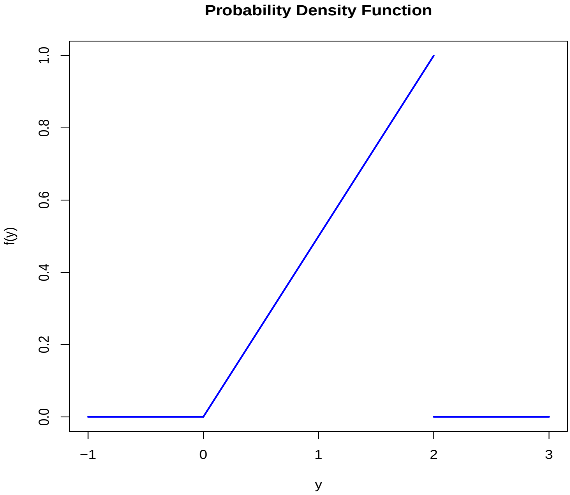
Solution b)

If $y < 0$, then $F(y) = 0$.

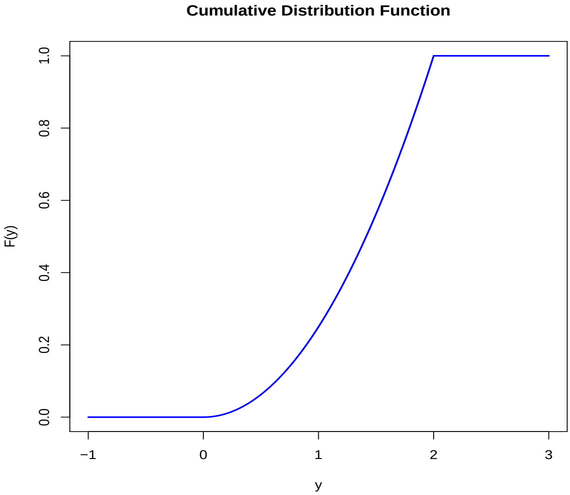
If $0 \leq y < 2$, then
$F(y) = P(Y \leq y) = \int_0^y \frac{t}{2} dt = \frac{1}{2} \int_0^y t dt = \frac{y^2}{4}$.

If $y \geq 2$, then $F(y) = 1$.

Solution c)



Solution c)


Cumulative Distribution Function
F(y)
1.0
0.8
0.6
0.4
0.2
0.0
−1
0
1
2
3
y

Solution d)

$P(1 \leq Y \leq 2) = P(Y \leq 2) - P(Y < 1)$
Since Y is a continuous random variable $P(Y < 1) = P(Y \leq 1)$

$$= P(Y \leq 2) - P(Y \leq 1)$$
$$= F(2) - F(1)$$
$$= 1 - \frac{1^2}{4} = \frac{3}{4}.$$

Exercise 4.19

Let the distribution function of a random variable Y be

$$F(y) = \begin{cases} 0, & y \leq 0, \\ \frac{y}{8}, & 0 < y < 2, \\ \frac{y^2}{16}, & 2 \leq y < 4, \\ 1, & y \geq 4. \end{cases}$$

Exercise 4.19

a. Find the density function of Y.
b. Find $P(1 \leq Y \leq 3)$.
c. Find $P(Y \geq 1.5)$.
d. Find $P(Y \geq 1 | Y \leq 3)$.

Solution a)

$$f(y) = \begin{cases} 0, & y \leq 0, \\ \frac{1}{8}, & 0 < y < 2, \\ \frac{y}{8}, & 2 \leq y < 4, \\ 0, & y \geq 4. \end{cases}$$

Solution b)

$$\begin{aligned} P(1 \leq Y \leq 3) &= P(Y \leq 3) - P(Y < 1) \\ &= P(Y \leq 3) - P(Y \leq 1) \\ &= F(3) - F(1) \\ &= \frac{3^2}{16} - \frac{1}{8} = \frac{7}{16}. \end{aligned}$$

Solution c) and d)

$P(Y \geq 1.5) = 1 - P(Y < 1.5)$
(Y is a continuous random variable)
$= 1 - P(Y \leq 1.5)$
$= 1 - \frac{1.5}{8} = \frac{13}{16}$

$$P(Y \geq 1 | Y \leq 3) = \frac{P(1 \leq Y \leq 3)}{P(Y \leq 3)} = \frac{7/16}{9/16} = \frac{7}{9}.$$

Definition 4.5

The expected value of a continuous random variable Y is

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy,$$

provided that the integral exists.

Theorem 4.4

Let $g(Y)$ be a function of Y; then the expected value of $g(Y)$ is given by

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy,$$

provided that the integral exists.

Theorem 4.5

Let c be a constant and let $g(Y)$, $g_1(y)$, $g_2(y)$, . . . , $g_k(y)$ be functions of a continuous random variable Y. Then the following results hold:

1. $E(c) = c$.
2. $E[cg(Y)] = cE[g(Y)]$.
3. $E[g_1(Y) + g_2(Y) + ... + g_k(Y)] =$
$E[g_1(Y)] + E[g_2(Y)] + ... + E[g_k(Y)]$.

Exercise 4.23

Prove Theorem 4.5.

Proof

1. $E(c) = \int_{-\infty}^{\infty} cf(y)dy = c \int_{-\infty}^{\infty} f(y)dy = c(1) = c$
2. $E[cg(Y)] = \int_{-\infty}^{\infty} cg(y)f(y)dy$
   $= c \int_{-\infty}^{\infty} g(y)f(y)dy = cE[g(Y)]$
3. $E[g_1(Y) + ... + g_k(Y)] = \int_{-\infty}^{\infty} [g_1(y) + ... + g_k(y)]f(y)dy$
   $= \int_{-\infty}^{\infty} g_1(y)f(y)dy + ... + \int_{-\infty}^{\infty} g_k(y)f(y)dy$
   $= E[g_1(Y)] + ... + E[g_k(Y)].$

Exercise 4.26

If Y is a continuous random variable with mean μ and variance σ^2 and a and b are constants, use Theorem 4.5 to prove the following:

a. $E(aY + b) = aE(Y) + b$.

b. $V(aY + b) = a^2 V(Y) = a^2\sigma^2$.

Proof a)

$$
\begin{aligned}
E(aY + b) &= E(aY) + E(b) \text{ (Theorem 4.5, part 3)} \\
&= aE(Y) + b \text{ (Theorem 4.5, parts 1 and 2).}
\end{aligned}
$$

Proof a)

(By definition of variance of $aY + b$)

$$\begin{aligned} V(aY+b) &= E\{[(aY+b)-(aE(Y)+b)]^2\} \\ &= E\{[aY+b-aE(Y)-b]^2\} \\ &= E\{[aY-aE(Y)]^2\} \\ &= E\{a^2[Y-E(Y)]^2\} \\ &= a^2E\{[Y-E(Y)]^2\} \text{ (Theorem 4.5, part 2)} \\ &= a^2V(Y) \text{ (Definition of variance of } Y\text{)} \end{aligned}$$

If $\theta_1 < \theta_2$, a random variable Y is said to have a continuous **uniform probability distribution** on the interval $(\theta_1, \ \theta_2)$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, & \theta_1 \leq y \leq \theta_2, \\ 0, & \text{elsewhere.} \end{cases}$$

Theorem 4.6

If $\theta_1 < \theta_2$ and Y is a random variable uniformly distributed on the interval $(\theta_1,\ \theta_2)$, then

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \text{ and } \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Proof

We will prove the result for $E(Y)$ and leave $V(Y)$ as an exercise.

$$
\begin{aligned}
E(Y) &= \int_{\theta_1}^{\theta_2} y \left(\frac{1}{\theta_2 - \theta_1}\right) dy \\
&= \left(\frac{1}{\theta_2 - \theta_1}\right) \int_{\theta_1}^{\theta_2} y dy \\
&= \frac{\theta_2^2 - \theta_1^2}{2(\theta_2 - \theta_1)} \\
&= \frac{(\theta_2 - \theta_1)(\theta_2 + \theta_1)}{2(\theta_2 - \theta_1)} \\
&= \frac{\theta_2 + \theta_1}{2}.
\end{aligned}
$$

A random variable Y has a uniform distribution over the interval $(\theta_1,\ \theta_2)$. Derive the variance of Y.

Example

Delta Airlines quotes a flight time of 2 hours, 5 minutes for its flights from Cincinnati to Tampa. Suppose we believe that actual flight times are uniformly distributed between 2 hours and 2 hours, 20 minutes.

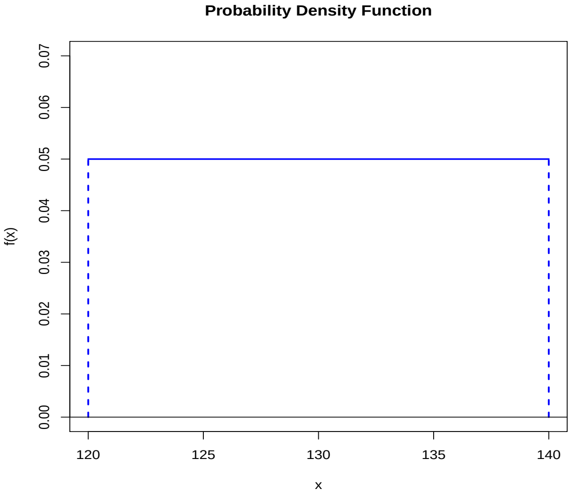
a. Show the graph of the probability density function for flight time.

b. What is the probability that the flight will be no more than 5 minutes late?

c. What is the probability that the flight will be more than 10 minutes late?

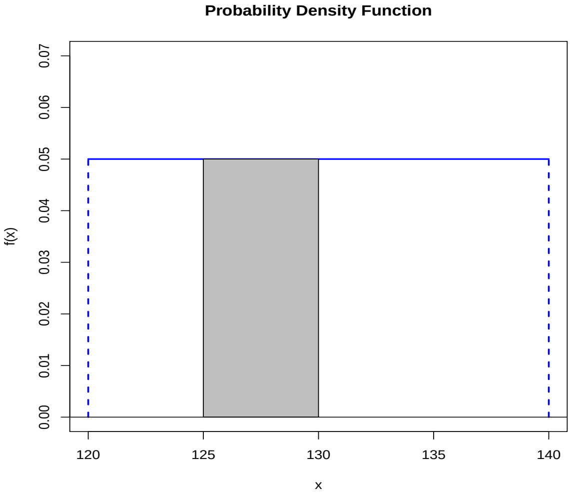
d. What is the expected flight time?

Solution a)



Probability Density Function
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
f(x)
120
125
130
135
140
x

Solution b)

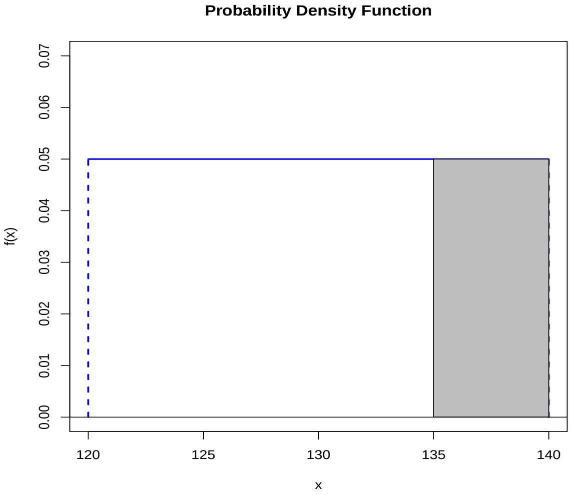

Probability Density Function
f(x)
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
120
125
130
135
140
x

Solution b)

X = flight time from Cincinnati to Tampa.

$$P(125 < X \leq 130) = \int_{125}^{130} f(x)dx$$
$$= \int_{125}^{130} \frac{1}{20}dx = \frac{5}{20} = \frac{1}{4} = 0.25$$

Solution c)



Probability Density Function
f(x)
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
120
125
130
135
140
x

Solution c)

X = flight time from Cincinnati to Tampa.

$$P(135 < X \leq 140) = \int_{135}^{140} f(x)dx$$
$$= \int_{135}^{140} \frac{1}{20} dx = \frac{5}{20} = \frac{1}{4} = 0.25$$

Solution d)

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2}$$

$$E(Y) = \frac{120 + 140}{2} = 130 \text{ minutes.}$$

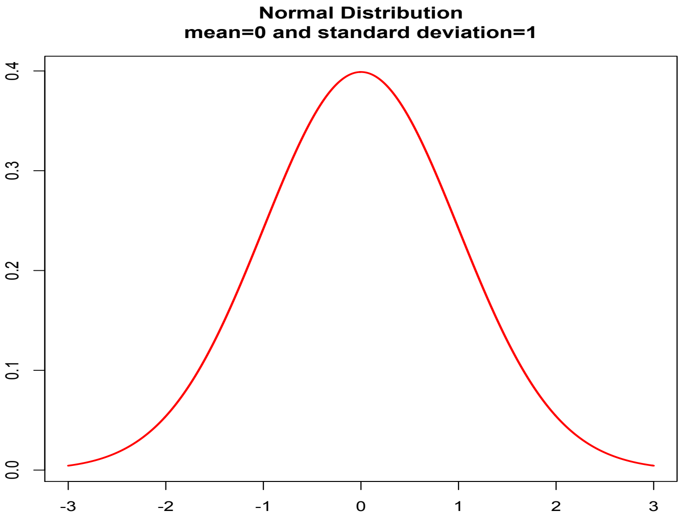
Definition 4.8

A random variable Y is said to have a **normal probability distribution** if and only if, for $\sigma > 0$ and $-\infty < \mu < \infty$, the density function of Y is

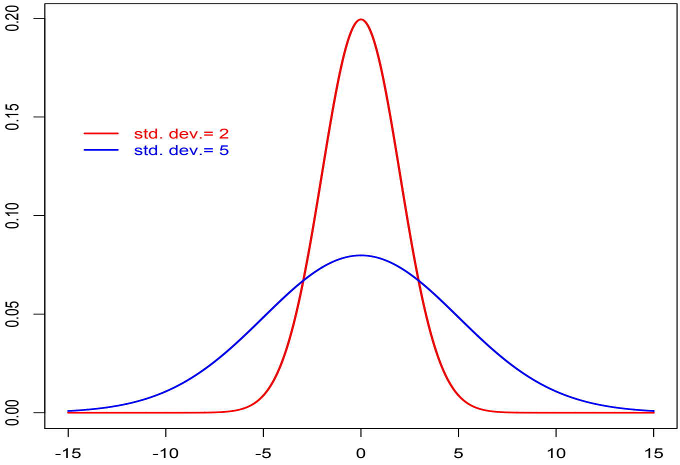
$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right], \quad -\infty < y < \infty.$$

A Normal Distribution is described by a Normal density curve. Any particular Normal distribution is completely specified by two numbers, its mean μ and standard deviation σ.
The mean of a Normal distribution is at the center of the symmetric Normal curve. The standard deviation is the distance from the center to the change-of-curvature points on either side.

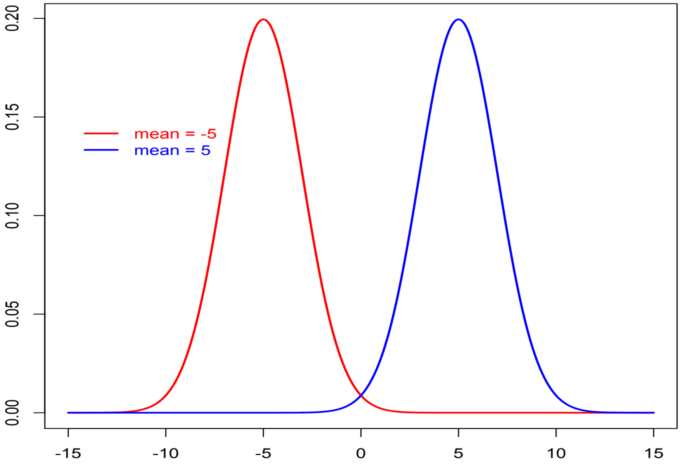
Standard Normal Distribution



Two Different Standard Deviations



Two Different Means



Theorem 4.7

If Y is a **Normally** distributed random variable with parameters μ and σ, then

$$E(Y) = \mu \text{ and } V(Y) = \sigma^2.$$

Result

If Z is a **Normally** distributed random variable with parameters $\mu = 0$ and $\sigma = 1$, then

$$E(Z) = 0 \quad \text{and} \quad V(Z) = 1.$$

Proof

By definition,

$$
\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} zf(z)dz \\
&= \int_{-\infty}^{\infty} z\left[\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\right]dz \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ze^{-z^2/2}dz \\
&= \frac{1}{\sqrt{2\pi}}\left[\int_{-\infty}^{0} ze^{-z^2/2}dz + \int_{0}^{\infty} ze^{-z^2/2}dz\right] \\
&= \frac{1}{\sqrt{2\pi}}\left[\int_{0}^{\infty}(-z)e^{-(-z)^2/2}dz + \int_{0}^{\infty} ze^{-z^2/2}dz\right] \\
&= \frac{1}{\sqrt{2\pi}}\left[-\int_{0}^{\infty} ze^{-z^2/2}dz + \int_{0}^{\infty} ze^{-z^2/2}dz\right] = 0
\end{aligned}
$$

Proof

We know that $V(Z) = E(Z^2) - [E(Z)]^2 = E(Z^2)$.

$$
\begin{aligned}
E(Z^2) &= \int_{-\infty}^{\infty} z^2 f(z) dz \\
&= \int_{-\infty}^{\infty} z^2 \left[\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right] dz \\
&= \frac{1}{\sqrt{2\pi}} lim_{N \to \infty} \int_{-N}^{N} z^2 e^{-z^2/2} dz \\
&= \frac{1}{\sqrt{2\pi}} lim_{N \to \infty} \int_{-N}^{N} z \left[z e^{-z^2/2} \right] dz
\end{aligned}
$$

Proof

Integrating by parts, we get

$$
\begin{aligned}
E(Z^2) &= \frac{1}{\sqrt{2\pi}} lim_{N\to\infty} \left\{ -ze^{-z^2/2}|_{-N}^{N} + \int_{-N}^{N} e^{-z^2/2} dz \right\} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = 1.
\end{aligned}
$$

Standard Normal Distribution

The standard Normal distribution is the Normal distribution $N(0,1)$ with mean 0 and standard deviation 1.
If a variable Y has any Normal distribution N(μ,σ) with mean μ and standard deviation σ, then the standardized variable

$$Z = \frac{Y - \mu}{\sigma}$$

has the standard Normal distribution.

Using the Normal table

Use table 4 to find the proportion of observations from a standard Normal distribution that satisfies each of the following statements. In each case, sketch a standard Normal curve and shade the area under the curve that is the answer to the question.

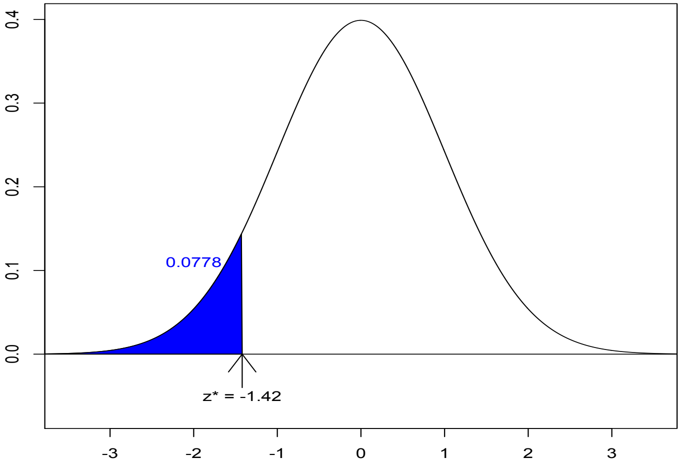
a) $Z < -1.42$

b) $Z > -1.42$

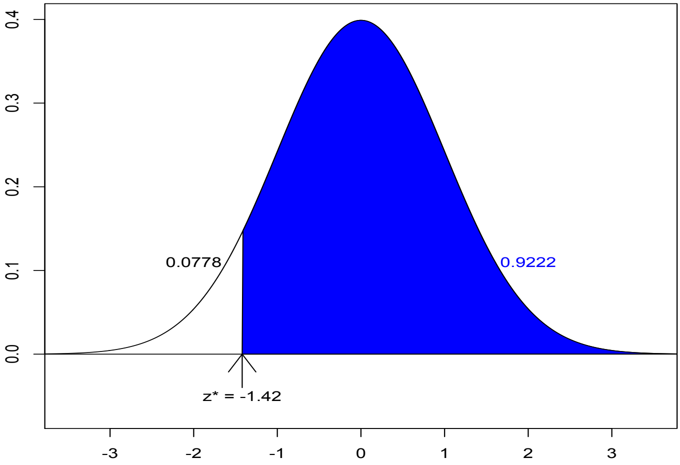
c) $Z < 2.35$

d) $-1.42 < Z < 2.35$

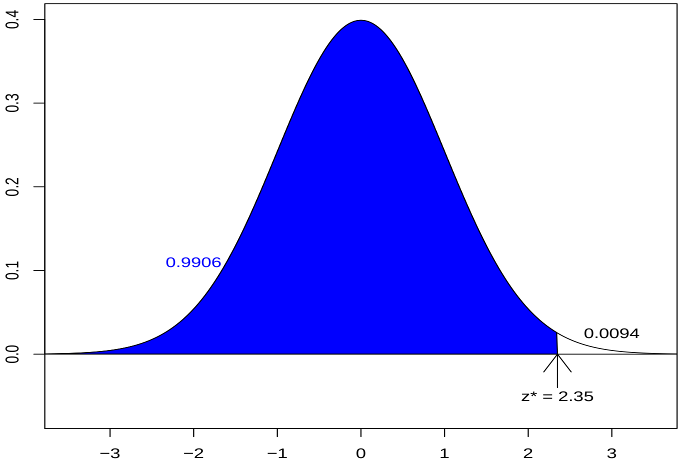
Solution a) 0.0778



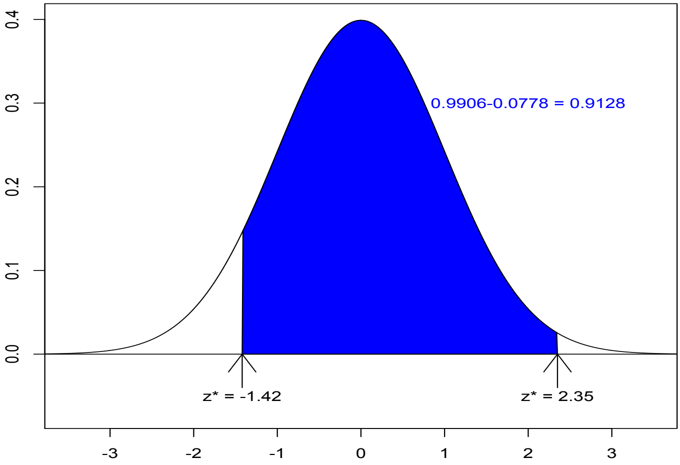
Solution b) 0.9222



Solution c) 0.9906



Solution d) 0.9906 - 0.0778 = 0.9128



The summer monsoon rains in India follow approximately a Normal distribution with mean 852 mm of rainfall and standard deviation 82 mm.

a) In the drought year 1987, 697 mm of rain fell. In what percent of all years will India have 697 mm or less of monsoon rain?

b) "Normal rainfall" means within 20% of the long-term average, or between 683 and 1022 mm. In what percent of all years is the rainfall "normal"?

Solution a)

1. State the problem. Let X be the monsoon rainfall in a given year. The variable X has the $N(852, 82)$ distribution. We want the proportion of years with $X \leq 697$.
2. Standardize. Subtract the mean, then divide by the standard deviation, to turn X into a standard Normal Z.
Hence $X \leq 697$ corresponds to $Z \leq \frac{697-852}{82} = -1.89$.
3. Use the table. From Table 4, we see that the proportion of observations less than -1.89 is 0.0294. Thus, the answer is 2.94%.

Solution b)

1. State the problem. Let X be the monsoon rainfall in a given year. The variable X has the $N(852, 82)$ distribution. We want the proportion of years with $683 < X < 1022$.
2. Standardize. Subtract the mean, then divide by the standard deviation, to turn X into a standard Normal Z.
$683 < X < 1022$ corresponds to $\frac{683-852}{82} < Z < \frac{1022-852}{82}$, or $-2.06 < Z < 2.07$.
3. Use the table. Hence, using Table 4, the area is $0.9808 - 0.0197 = 96.11\%$.

The Medical College Admission Test

Almost all medical schools in the United States require students to take the Medical College Admission Test (MCAT). The exam is composed of three multiple-choice sections (Physical Sciences, Verbal Reasoning, and Biological Sciences). The score on each section is converted to a 15-point scale so that the total score has a maximum value of 45. The total scores follow a Normal distribution, and in 2010 the mean was 25.0 with a standard deviation of 6.4. There is little change in the distribution of scores from year to year.

a) What proportion of students taking the MCAT had a score over 30?

b) What proportion had scores between 20 and 25?

Solution a)

1. State the problem. Let X be the MCAT score of a randomly selected student. The variable X has the $N(25, 6.4)$ distribution. We want the proportion of students with $X > 30$.
2. Standardize. Subtract the mean, then divide by the standard deviation, to turn X into a standard Normal Z.
Hence $X > 30$ corresponds to $Z > \frac{30-25}{6.4} = 0.78$.
3. Use the table. From Table 4, we see that the proportion of observations greater than 0.78 is 0.2177. Hence, the answer is 21.77%.

Solution b)

1. State the problem. Let X be the MCAT score of a randomly selected student. The variable X has the $N(25, 6.4)$ distribution. We want the proportion of students with $20 \leq X \leq 25$.
2. Standardize. Subtract the mean, then divide by the standard deviation, to turn X into a standard Normal Z.
$20 \leq X \leq 25$ corresponds to $\frac{20-25}{6.4} \leq Z \leq \frac{25-25}{6.4}$, or $-0.78 \leq Z \leq 0$.
3. Use the table. Using Table 4, the area is $0.5 - 0.2177 = 0.2833$, or 28.33%.

Using a table to find Normal proportions

Step 1. State the problem in terms of the observed variable X. Draw a picture that shows the proportion you want in terms of cumulative proportions.

Step 2. Standardize X to restate the problem in terms of a standard Normal variable Z.

Step 3. Use Table 4 and the fact that the total are under the curve is 1 to find the required area under the standard Normal curve.

Example

A person must score in the upper 2% of the population on an IQ test to qualify for membership in MENSA, the international high-IQ society. If IQ scores are Normally distributed with a mean of 100 and a standard deviation of 15, what score must a person have to qualify for MENSA?

1. State the problem. Let X = IQ score of a randomly selected person. We want to find the IQ score x_* with area 0.02 to its right under the Normal curve with mean $\mu = 100$ and standard deviation $\sigma = 15$. Because our table gives the areas to the right of z-values, always state the problem in terms of the area to the right of x_*.

Solution

2. Use the table. Look in the body of our table for the entry closest to 0.02. It is 0.0202. This is the entry corresponding to $z_* = 2.05$. So $z_* = 2.05$ is the standardized value with area 0.02 to its right.
3. Unstandardize to transform the solution from Z back to the original X scale. We know that the standardized value of the unknown x_* is $z_* = 2.05$. So x_* itself satisfies:
$\frac{x_* - 100}{15} = 2.05$. Solving this equation for x gives:
$x_* = 100 + (2.05)(15) = 130.75$. We see that a person must score at least 130.75 to place in the highest 2%.

Recall that the Poisson distribution is used to compute the probability of specific numbers of "events" during a particular period of time or space. In many applications, the time period or span of space is the random variable. Don't forget that a Poisson distribution has a single parameter λ, where λ may be interpreted as the mean number of events per unit "time". Consider now the random variable described by the time required for the first event to occur. Using the Poisson distribution, we find the probability of **no** events occurring in the span up to time t is given by

$$\frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Note that *probability of at least one event occurring in the span up to time* t is $1 - e^{-\lambda t}$.
We can now make use of the above and let X be the time to the first Poisson event. The probability that the length of time until the first event will be less than or equal to x is the same as the probability that at least one Poisson event will occur in x. This probability is given by $1 - e^{-\lambda x}$. As a result,

$$P(X \leq x) = 1 - e^{-\lambda x}.$$

Thus the cumulative distribution function for X is given by

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x > 0.$$

We may differentiate the cumulative distribution function above to obtain the density function

$$f(x) = \lambda e^{-\lambda x}.$$

which is the density function of the exponential distribution with $\lambda = \frac{1}{\beta}$.

Definition 4.11

A random variable Y is said to have an **exponential distribution** with parameter $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{1}{\beta} e^{-y/\beta}, & 0 \leq y < \infty, \\ 0, & \text{elsewhere.} \end{cases}$$

Theorem 4.10

If Y is an exponential random variable with parameter β, then

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

Exercise 4.89

If Y has an exponential distribution and $P(Y > 2) = 0.0821$, what is

a. $\beta = E(Y)$.

b. $P(Y \leq 1.7)$?

Solution a)

$P(Y > 2) = 0.0821$
$\int_2^\infty \frac{1}{\beta} e^{-y/\beta} dy = e^{-2/\beta}$
$e^{-2/\beta} = 0.0821$ (solving for β)
$\beta = \frac{-2}{ln(0.0821)} \approx 0.8$
$E(Y) = 0.8$

Solution b)

$$P(Y \leq 1.7) = \int_0^{1.7} \frac{1}{0.8} e^{-y/0.8} dy = 1 - e^{-1.7/0.8} = 0.8805$$

Exercise 4.91

The operator of a pumping station has observed that demand for water during early afternoon hours has an approximately exponential distribution with mean 100 cfs (cubic feet per second).
a. Find the probability that the demand will exceed 200 cfs during the early afternoon on a randomly selected day.
b. What water-pumping capacity should the station maintain during early afternoon so that the probability that demand will exceed capacity on a randomly selected day is only 0.01?

Solution

a. $P(Y > 200) = 1 - P(Y \leq 200)$
$$= 1 - [1 - e^{-200/100}] = e^{-2} \approx 0.1353.$$

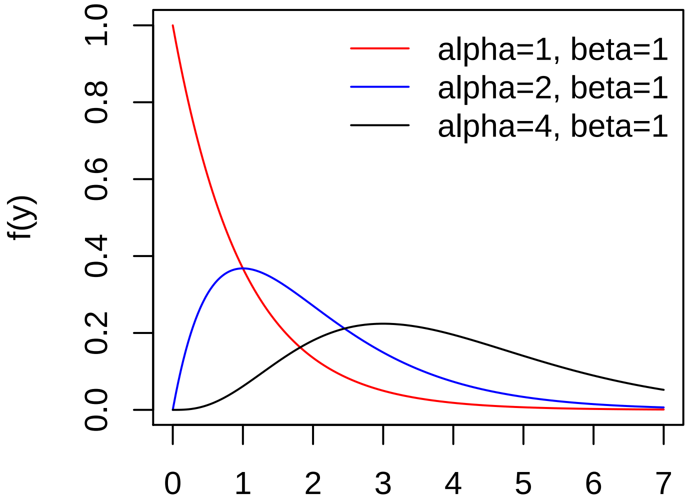
b. $P(Y \leq c*) = 0.99$
$1 - e^{-c*/100} = 0.99$
Solving for $c*$ gives us
$c* = -100 ln(0.01) = 460.517$ cfs.

Definition 4.9

A random variable Y is said to have a **Gamma distribution** with parameters $\alpha > 0$ and $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{y^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-y/\beta}, & 0 \leq y < \infty, \\ 0, & \text{elsewhere.} \end{cases}$$

where $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy$.


alpha=1, beta=1
alpha=2, beta=1
alpha=4, beta=1
f(y)
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
7

Gamma Function

The quantity $\Gamma(\alpha)$ is known as the **gamma function**. Direct integration will verify that $\Gamma(1) = 1$. Integration by parts will verify that $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for any $\alpha > 1$ and that $\Gamma(n) = (n - 1)!$, provided that n is an integer.

Gamma Function

The **gamma function** is defined by

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$$

for $\alpha > 0$.

Gamma Function

Integrating by parts with $u = y^{\alpha-1}$ and $dv = e^{-y}dy$, we obtain

$$\begin{aligned}\Gamma(\alpha) &= -e^{-y}y^{\alpha-1}|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2}dy \\ &= (\alpha-1)\int_0^\infty y^{\alpha-2}e^{-y}dy\end{aligned}$$

for $\alpha > 1$, which yields the recursion formula

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1).$$

Theorem 4.8

If Y has a Gamma distribution with parameters α and β, then

$$\mu = E(Y) = \alpha\beta \text{ and } \sigma^2 = V(Y) = \alpha\beta^2.$$

Proof

$$E(Y) = \int_0^\infty y \frac{y^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-y/\beta} dy$$
$$= \int_0^\infty \frac{y^{(\alpha+1)-1}}{\beta^\alpha \Gamma(\alpha)} e^{-y/\beta} dy$$
$$= \frac{\beta^{\alpha+1}\Gamma(\alpha+1)}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty \frac{y^{(\alpha+1)-1}}{\beta^{(\alpha+1)}\Gamma(\alpha+1)} e^{-y/\beta} dy$$

Note that the last integral equals one (we are integrating the pdf of a Gamma random variable with parameters $\alpha + 1$ and β, over its entire domain).

$$= \frac{\beta^{\alpha+1}\Gamma(\alpha+1)}{\beta^\alpha \Gamma(\alpha)}$$

Since $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, we finally have that

$$E(Y) = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta.$$

Finding $V(Y)$ is left as an exercise.

Definition 4.10

Let ν be a positive integer. A random variable Y is said to have a **chi-square distribution** with ν degrees of freedom if and only if Y is a Gamma-distributed random variable with parameters $\alpha = \frac{\nu}{2}$ and $\beta = 2$.

Exercise 4.96

Suppose that a random variable Y has a probability density function given by

$$f(y) = \begin{cases} ky^3 e^{-y/2}, & 0 < y < \infty, \\ 0, & \text{elsewhere.} \end{cases}$$

Exercise 4.96 (cont.)

a. Find the value of k that makes $f(y)$ a density function.
b. Does Y have a χ^2 distribution? If so, how many degrees of freedom?
c. What are the mean and standard deviation of Y?

Solution a)

If we compare $f(y)$ to the pdf of a Gamma random variable, it is clear that Y has a Gamma distribution with parameters $\alpha = 4$ and $\beta = 2$. Therefore,
$k = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} = \frac{1}{2^4\Gamma(4)} = \frac{1}{(16)(3!)} = \frac{1}{96}$.

Solution b)

Using the definition of a chi-square random variable and part a) $\alpha = \frac{\nu}{2} = 4$ and $\beta = 2$. Therefore, $\nu = 8$ and Y has a chi-square distribution with 8 degrees of freedom.

Solution c)

From our Table (or from Theorem 4.8),

$$E(Y) = \alpha\beta = (4)(2) = 8.$$

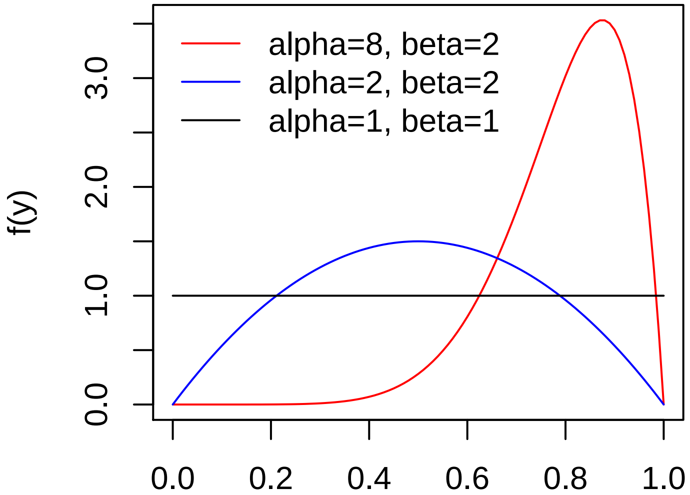
$$V(Y) = \alpha\beta^2 = (4)(2)^2 = 16.$$

$$\sigma = \sqrt{V(Y)} = \sqrt{16} = 4.$$

Definition 4.12

A random variable Y is said to have a **beta probability distribution** with parameters $\alpha > 0$ and $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, & 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$


alpha=8, beta=2
alpha=2, beta=2
alpha=1, beta=1
f(y)
0.0
1.0
2.0
3.0
0.0
0.2
0.4
0.6
0.8
1.0

Theorem 4.11

If Y is a beta-distributed random variable with parameters $\alpha > 0$ and $\beta > 0$, then

$$\mu = E(Y) = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Proof

$$E(Y) = \int_0^1 y \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1} dy$$
$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y^{(\alpha+1)-1}(1-y)^{\beta-1} dy$$
$$= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+1)} \int_0^1 \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta)} y^{(\alpha+1)-1}(1-y)^{\beta-1} dy$$

Note that the last integral equals 1 (we are integrating the pdf of a Beta random variable with parameters $\alpha + 1$ and β over its entire domain).

$$= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+1)}$$

Since $\Gamma(\alpha + \beta + 1) = (\alpha + \beta)\Gamma(\alpha + \beta)$ and $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$

$$E(Y) = \frac{\Gamma(\alpha+\beta)\alpha\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}.$$

Finding $V(Y)$ is left as an exercise.

Exercise 4.123

The relative humidity Y, when measured at a location, has a probability density function given by

$$f(y) = \begin{cases} ky^3(1-y)^2, & 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Exercise 4.123

a. Find the value of k that makes $f(y)$ a density function.
b. Find $E(Y)$.
c. Find $V(Y)$.

Solution a)

If we compare $f(y)$ to the pdf of a Beta random variable, it is clear that Y has a Beta distribution with parameters $\alpha = 4$ and $\beta = 3$, then
$k = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{\Gamma(7)}{\Gamma(4)\Gamma(3)} = \frac{6!}{3!2!} = 60.$

Solution b) and c)

b. $E(Y) = \frac{\alpha}{\alpha+\beta} = \frac{4}{7}$.

c. $V(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{12}{(49)(8)} = \frac{3}{98}$.

Exercise 4.127

Verify that if Y has a beta distribution with $\alpha = \beta = 1$, then Y has a uniform distribution over $(0, 1)$. That is, the uniform distribution over the interval $(0, 1)$ is a special case of a beta distribution.

Solution.
Easy! Right?

Exercise 4.131

Errors in measuring the time of arrival of a wave front from an acoustic source sometimes have an approximate beta distribution. Suppose that these errors, measured in microseconds, have approximately a beta distribution with $\alpha = 1$ and $\beta = 2$.
a. What is the probability that the measurement error in a randomly selected instance is less that 0.5 μs?
b. Give the mean and standard deviation of the measurement errors.

Solution

Let Y = error in measuring the time of arrival of a wave front from an acoustic source. Since Y has a Beta distribution with parameters $\alpha = 1$ and $\beta = 2$, $f(y)$ is given by

$$f(y) = \frac{\Gamma(3)}{\Gamma(1)\Gamma(2)} y^{1-1}(1-y)^{2-1} = 2(1-y), \quad 0 \leq y \leq 1.$$

Solution a)

$$P(Y < 0.5) = \int_0^{0.5} 2(1-y)dy = 2\int_0^{0.5} dy - 2\int_0^{0.5} ydy$$
$$= 2(0.5) - 2\frac{y^2}{2}\Big|_0^{0.5}$$
$$= 1 - (0.5)^2 = 1 - \frac{1}{4} = \frac{3}{4}.$$

Solution b)

Using our Table (or Theorem 4.11),

$$E(Y) = \frac{\alpha}{\alpha + \beta} = \frac{1}{3}$$

$$V(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = \frac{(1)(2)}{(3)^2(3 + 1)} = \frac{1}{18}$$

$$\sigma = \sqrt{V(Y)} = \frac{1}{\sqrt{18}}.$$

Definition 4.14

If Y is a continuous random variable, then the **moment-generating function** of Y is given by

$$M_Y(t) = E(e^{tY}).$$

The moment-generating function is said to exist if there exists a constant $b > 0$ such that $M_Y(t)$ is finite for $|t| \leq b$.

Theorem 4.12

Let Y be a random variable with density function $f(y)$ and $g(Y)$ be a function of Y. Then the moment-generating function for $g(Y)$ is

$$E[e^{tg(Y)}] = \int_{-\infty}^{\infty} e^{tg(y)} f(y) dy.$$

Example

Let $g(Y) = Y - \mu$, where Y is a Normally distributed random variable with mean μ and variance σ^2.

a. Find the moment-generating function for $g(Y)$.

b. Differentiate the moment-generating function found in part a) to find $E[g(Y)]$ and $V[g(Y)]$.

Solution a)

Let $W = Y - \mu$.
$M_W(t) = E[e^{tW}]$
$= E[e^{(Y-\mu)t}]$
$= E[e^{Yt}e^{-\mu t}]$
$= e^{-\mu t}E[e^{Yt}]$
$= e^{-\mu t}M_Y(t)$
From our Table, $M_Y(t) = e^{\mu t + (t^2\sigma^2)/2}$.

Finally, $M_W(t) = e^{(t^2\sigma^2)/2}$.

Solution b)

From part a), it is clear that $M_W(t)$ corresponds to the MGF of a Normal random variable with mean 0 and variance σ^2. Therefore, W has a Normal distribution with mean 0 and variance σ^2. Verify this by doing the following:
$E(W) = M_W^{'}(0)$ and
$V(W) = E(W^2) - [E(W)]^2 = M_W^{''}(0) - [M_W^{'}(0)]^2$.
(recall that you are taking derivatives with respect to t).

Exercise 4.139

Let Y be a Normally distributed random variable with mean μ and variance σ^2. Derive the moment-generating function of $X = -3Y + 4$. What is the distribution of X? Why?

Markov's Inequality

If X is a random variable that takes only nonnegative values, then for any value $a > 0$,

$$P[X \geq a] \leq \frac{E(X)}{a}.$$

Proof

We give a proof for the case where X is continuous with density f:

$$\begin{aligned} E[X] &= \int_0^\infty xf(x)dx \\ &= \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \\ &\geq \int_a^\infty xf(x)dx \\ &\geq \int_a^\infty af(x)dx \\ &= a\int_a^\infty f(x)dx \\ &= aP[X \geq a]. \end{aligned}$$

Tchebysheff's Theorem

Let Y be a random variable with finite mean μ and variance σ^2. Then, for any $k > 0$,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad \text{or} \quad P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof

Since $X = (Y - \mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a = k^2\sigma^2$) to obtain

$$P\{(Y - \mu)^2 \geq k^2\sigma^2\} \quad \leq \quad \frac{E[(Y - \mu)^2]}{k^2\sigma^2}$$

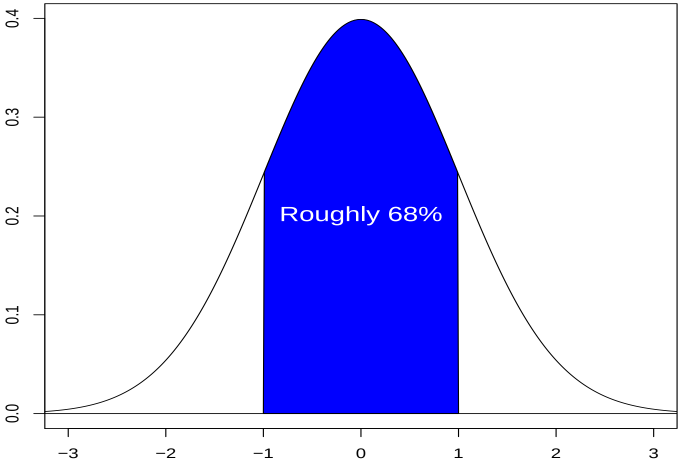
But since $(Y - \mu)^2 \geq k^2\sigma^2$ if and only if $|Y - \mu| \geq k\sigma$, the preceding is equivalent to

$$P\{|Y - \mu| \geq k\sigma\} \quad \leq \quad \frac{E[(Y - \mu)^2]}{k^2\sigma^2} = \frac{1}{k^2}$$

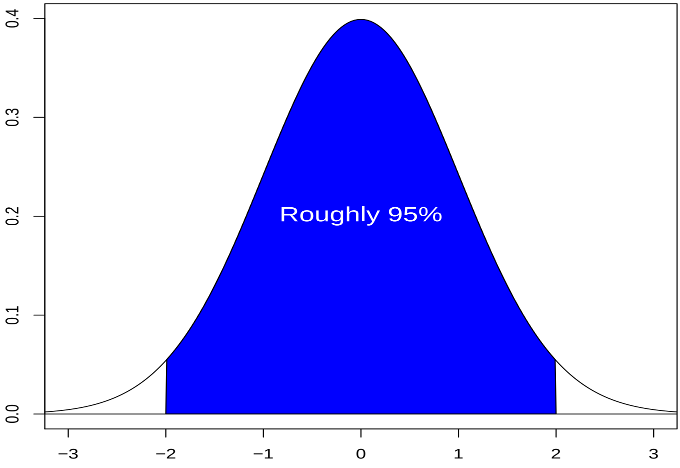
The 68-95-99.7 rule

In the Normal distribution with mean μ and standard deviation σ:
Approximately 68% of the observations fall within σ of the mean μ.
Approximately 95% of the observations fall within 2σ of μ.
Approximately 99.7% of the observations fall within 3σ of μ.
Note. The 68-95-99.7 rule is also know as the empirical rule.

Example $N(\mu = 0, \sigma = 1)$



Example $N(\mu = 0, \sigma = 1)$



Exercise 4.150

1) Find $P(|X - \mu| \leq 2\sigma)$ for an exponential random variable with mean β.
2) Compare the result in 1) to the empirical rule result.
3) Compare the result in 1) to Tchebysheff's result.

Solution 1)

First, let us find the CDF of X. By definition, if $k > 0$
$F(k) = P(X \leq k) = \int_0^k \frac{1}{\beta} e^{-x/\beta} dx = -\int_0^k -\frac{1}{\beta} e^{-x/\beta} dx$
$F(k) = -\left[e^{-k/\beta} - e^0\right] = 1 - e^{-k/\beta}$.
We want to find the following probability

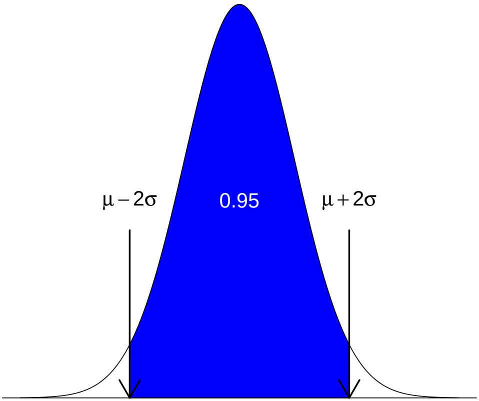
$$P(|X - \mu| \leq 2\sigma) = P(-2\sigma \leq X - \mu \leq 2\sigma)$$
$$= P(\mu - 2\sigma \leq X \leq \mu + 2\sigma).$$

Solution 1) (cont.)

Since $\mu = \beta$ and $\sigma^2 = \beta^2$, we have that

$$\begin{aligned}
P(|X - \mu| \leq 2\sigma) &= P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \\
&= P(\beta - 2\beta \leq X \leq \beta + 2\beta). \\
&= P(-\beta \leq X \leq 3\beta) \\
&= P(0 \leq X \leq 3\beta) \\
&= F(3\beta) = 1 - e^{-3\beta/\beta} \\
&= 1 - e^{-3} = 0.950213.
\end{aligned}$$

Solution 2)



Solution 3)

$$P(|X-\mu|<2\sigma)\geq 1-\frac{1}{2^2}$$

$$P(|X-\mu|<2\sigma)\geq 1-\frac{1}{4}$$

$$P(|X-\mu|<2\sigma)\geq 1-\frac{1}{4}=0.75$$