Continuous Probability Distributions: Chapter 8

Chapter Outcomes:

- General terminology and notation
- Continuous uniform distribution
- Exponential distribution
- Normal distribution

Continuous Random Variables

Such random variables take on values on the *real number line* (which is no longer countable).

They are treated differently than discrete random variables because **now** $P(X = a) = 0$ for each a (since the chance of observing the specific value of a among an infinite number of possibilities is extremely remote).

Hence, probability functions (as we know them from the discrete case) cannot be used to describe a continuous probability distribution.

With continuous random variables, we specify the probability over **intervals**, rather than individual points.

Note: For continuous r.v.'s, probability is given by the **AREA UNDER THE CURVE**. More on this later – stay tuned!

We can use the **cumulative distribution function (c.d.f.)** to describe the distribution of a continuous r.v., namely:

$$F(x) = P(X \leq x).$$

The c.d.f. satisfies the following properties:

1. $F(x)$ is defined for all real x

2. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$

3. $F(x)$ is a non-decreasing function of x

4. $P(a < X \leq b) = F(b) - F(a)$

Remark: As we indicated earlier, note that 0

$$= P(X = a)$$
$$= \lim_{\varepsilon \to 0} P(a - \varepsilon < X \leq a)$$
$$= \lim_{\varepsilon \to 0} [F(a) - F(a - \varepsilon)]$$

This leads to $\lim_{\varepsilon \to 0} F(a - \varepsilon) = F(a)$.

Similarly, we can show $\lim_{\varepsilon \to 0} F(a + \varepsilon) = F(a)$.

Together, this proves that the c.d.f. $F(x)$ is a **continuous** function of x.

Because $P(X = x) = 0$ for each point x, we also get:

$$P(a < X < b) = P(a \leq X \leq b) = P(a < X \leq b)$$
$$= P(a \leq X < b) = F(b) - F(a)$$

Probability Density Function (p.d.f.)

The c.d.f. does not give an intuitive picture of which values of x are more likely, and which are less likely. Remember that the c.d.f. "describes" the probability journey from 0 to 1 for the r.v. X.

We can use the c.d.f. to determine probabilities associated with the r.v. X.

Suppose we take a **short** interval of x-values, say $[x, x + \Delta x]$, so that the probability X lies in the interval is given by

$$P(x \leq X \leq x + \Delta x) = F(x + \Delta x) - F(x).$$

Now, we divide both sides by Δx, and consider what happens as Δx becomes small.

Definition: The **probability density function** (p.d.f.) $f(x)$ for a continuous r.v. X is the derivative of the c.d.f. given by

$$f(x) = \frac{dF(x)}{dx} = F'(x).$$

Properties of a p.d.f.:

Assume that $f(x)$ is a continuous function of x at all points for which $0 < F(x) < 1$.

1. $P(a \le X \le b) = F(b) - F(a) = \int_a^b f(x)\,dx$

2. $f(x) \ge 0$, since $F(x)$ is **non-decreasing** in x (and so its derivative would be non-negative)

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

4. $F(x) = \int_{-\infty}^{x} f(u)du$

Property 1 implies that **areas under the curve** *of* $f(x)$ **correspond to probabilities** for the r.v. X.

Also, using property 1, we clearly see that

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x)\, dx = 0.$$

Note that the p.d.f., $f(x)$ represents the **likelihood** of (small intervals around) different x-values.

To see this, note that for small Δx

$$P\left(x - \frac{\Delta x}{2} \leq X \leq x + \frac{\Delta x}{2}\right)$$
$$= F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \approx f(x)\Delta x$$

Visually, we have the following:

Thus, $f(x) \neq P(X = x)$. That is, it cannot be thought of as a probability, BUT $f(x)\Delta x$ is the **approximate** probability that X is inside an interval of length Δx centered about the value x when Δx is small.

Example:

Let X be a r.v. with c.d.f.

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \dfrac{x}{4} & \text{for } 0 < x \leq 4 \\ 1 & \text{for } x > 4 \end{cases}$$

a) Find the p.d.f. $f(x)$.

b) Solve for the 90th percentile (i.e. solve for the value, x, such that the area under the curve to the left of that value is 0.9. Maybe call this value x_{90}).

You Try: Let X be a r.v. with p.d.f.

$$f(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

be a p.d.f. for some constant c.

Find:

a) c

We can determine c by integrating the pdf and then set it equal to 1.

$$\mathbf{c\int_0^2 (4x - 2x^2)dx = c\left[2x^2 - \frac{2x^3}{3}\right]\Big|_0^2 = c\left(8 - \frac{16}{3}\right) = 1 \Rightarrow c = 3/8.}$$

b) $F(x)$

We can determine the cdf, F(x) as follows: F(x) = $\boldsymbol{\int_0^x \frac{3}{8}(4s - 2s^2)ds}$

So, F(x) = $\boldsymbol{\frac{x^2(3-x)}{4}}$**, for 0 < x < 2**

c) $P(X > 1)$

P(X > 1) = 1 – F(1) = 1 – (2/4) = 1/2

Exercise: Find the median of the distribution of X.

To find the median of the distribution of X, or the median value of X, we simply set the cdf equal to 0.5 and solve for the value, x.

So, we have $\boldsymbol{\frac{3x^2 - x^3}{4}}$ **= 0.5 =>** $\boldsymbol{3x^2 - x^3 - 2 = 0.}$

In this case, x = 1. The other two roots: x = 1$\boldsymbol{\pm\sqrt{3}}$ **are inadmissible.**

Defined Variables or Change of Variable

Sometimes our interest may actually lie in finding the p.d.f. or c.d.f. for some r.v. Y that is a function of X.

The procedure to use is as follows:

1. Write the c.d.f. of Y as a function of X.
2. Use $F_X(x)$ to find $F_Y(y)$. If you want the p.d.f. $f_Y(y)$, you simply need to differentiate $F_Y(y)$.
3. Find the range of values for the r.v. Y.

Example: Suppose X is a continuous r.v. having p.d.f.

$$f(x) = \frac{1}{4} \text{ for } 0 < x < 4$$

and c.d.f.

$$F(x) = \frac{x}{4} \text{ for } 0 \leq x \leq 4.$$

Let $Y = 1/X$.

Find the p.d.f. of Y.

You Try:

Let X be a continuous r.v. with p.d.f. given by

$$f(x) = 3x^2 \text{ for } 0 < x < 1.$$

a) Find the c.d.f. $F(x)$.

We can find F(x) as follows: F(x) = $\int_0^s 3s^2 ds$

So, F(x) = $\left[\frac{3s^3}{3}\right] |_0^x = x^3$**, for 0 < x < 1.**

b) If $Y = X^2$, determine the p.d.f. of Y.

$$F_Y(y) = P(Y \le y) = P\left(X^2 \le y\right) = P(-\sqrt{y} < X < \sqrt{y})$$

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) = y^{\frac{3}{2}}.$$

Remember that there is no support for negative values of the r.v. Y, as we know that the range of X is from 0 to 1.

We know that:

$$f_Y(y) = \left(\frac{d}{dy}\right) F_Y(y).$$

So $f_Y(y) = \left(\frac{3}{2}\right) y^{\frac{1}{2}}$**, for 0 < y < 1; 0, otherwise**

Expectation, Mean, and Variance of Continuous Distributions

Definition: If X is a continuous r.v. and g is an arbitrary function, then the expected value of $g(X)$ is given by

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx,$$

and hence the mean of X is simply

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Moreover, like with discrete random variables,

$$\sigma^2 = Var(X) = E(X^2) - [E(X)]^2$$

and

$$\sigma = sd(X) = \sqrt{Var(X)}$$

In fact, all of the earlier properties of expected value and variance still hold true even in the continuous case, including:

$$E[ag(X) + b] = aE[g(X)] + b$$

$$E[ag_1(X) + bg_2(X)] = aE[g_1(X)] + bE[g_2(X)]$$

$$Var[ag(X) + b] = a^2 Var[g(X)]$$

Example: If X is a continuous r.v. having p.d.f.

$$f(x) = \frac{1}{4} \text{ for } 0 < x < 4,$$

Calculate $\mathrm{E(X)}$ and $\mathrm{Var(X)}$.

You try: Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} k\sqrt{x} & \text{for } 0 \le x \le 1 \\ \dfrac{k}{x^4} & \text{for } x > 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Find k.

We have a piece-wise pdf here.

So, we need k $\int_0^1 \sqrt{x}dx + k\left[\frac{2x^{3/2}}{3}\right]|_0^1 + k\left[\frac{-1}{3x^3}\right]|_1^\infty$

= k$\left[\frac{2}{3} + \frac{1}{3}\right]$ **= 1. So, k = 1**

b) Find the c.d.f. $F(x)$ for all values of x.

We can find F(x) as follows:

F(x) = $\int_0^x \sqrt{s}ds = \frac{2x^{3/2}}{3}$**, for** $0 \le x \le 1$

Note: This is a piece-wise definition.

From the first part of the definition, we see that F(1) = 2/3. We will need to bring that into the next piece.

So, for the next piece: F(x) = $\frac{2}{3} + \frac{x^3-1}{3x^3}$**, for x > 1**

c) Find $P(\frac{1}{3} < X < 3)$.

We can find this probability piece-wise.

$$P\left(\frac{1}{3} < X < 3\right) = P\left(\frac{1}{3} < X < 1\right) + P(1 < X < 3).$$

$P\left(\frac{1}{3} < X < 3\right)$ **=** $\int_{1/3}^1 \sqrt{x}dx$ **+** $\int_1^3 \frac{1}{x^4}dx$ **=** $\left[\frac{2}{3} - \frac{2}{3^{5/2}}\right]$ **+** $\left[\frac{1}{3} - \frac{1}{81}\right] = 0.8594$

You could also use the definitions of F(x) that were found in part b) to solve this question. You will end up with the same answer!

You try (cont'd): Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} k\sqrt{x} & \text{for } 0 \le x \le 1 \\ \dfrac{k}{x^4} & \text{for } x > 1 \\ 0 & \text{otherwise} \end{cases}$$

d) Calculate $\mathrm{E(X)}$ and $\mathrm{Var(X)}$.

We can calculate E(X) piece-wise.

E(X) = $\int_0^1 x\sqrt{x}dx + \int_1^{\infty} \frac{x}{x^4} dx$

E(X) = $\int_0^1 x^{3/2}dx + \int_1^{\infty} \frac{1}{x^3} dx = \left[\frac{2x^{5/2}}{5}\right] |_0^1 + \left[\frac{-1}{2x^2}\right] |_1^{\infty}$

E(X) = 2/5 + 1/2 = 0.9

We can also calculate E(X²) piece-wise.

E(X²) = $\int_0^1 x^2\sqrt{x}dx + \int_1^{\infty} \frac{x^2}{x^4} dx$

E(X²) = $\int_0^1 x^{5/2}dx + \int_1^{\infty} \frac{1}{x^2} dx = \left[\frac{2x^{7/2}}{7}\right] |_0^1 + \left[\frac{-1}{x}\right] |_1^{\infty}$

E(X²) = 4/63 + 1 = 67/63

So, Var(X) = 67/63 – (0.9)² = 0.2535

Continuous Uniform Distribution (Section 8.2)

Physical setup:

X is a r.v. taking on values in the interval $[a, b]$ (it does not matter whether the interval is closed or open) with all subintervals of a fixed length being *equally likely*.

Then, we write $X \sim Uniform(a, b)$.

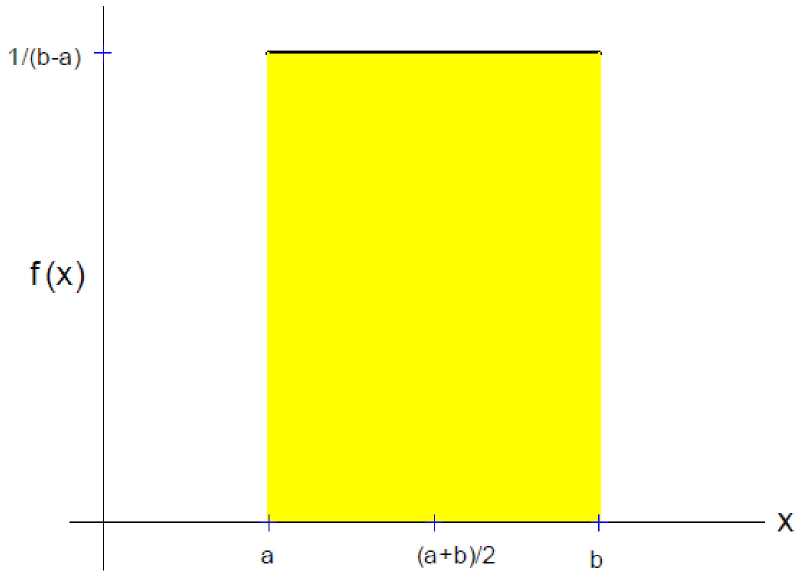
Probability density function:

Since all points are equally likely, the p.d.f. must be a constant (i.e. $f(x) = k$ for $a < x < b$ for some constant k).

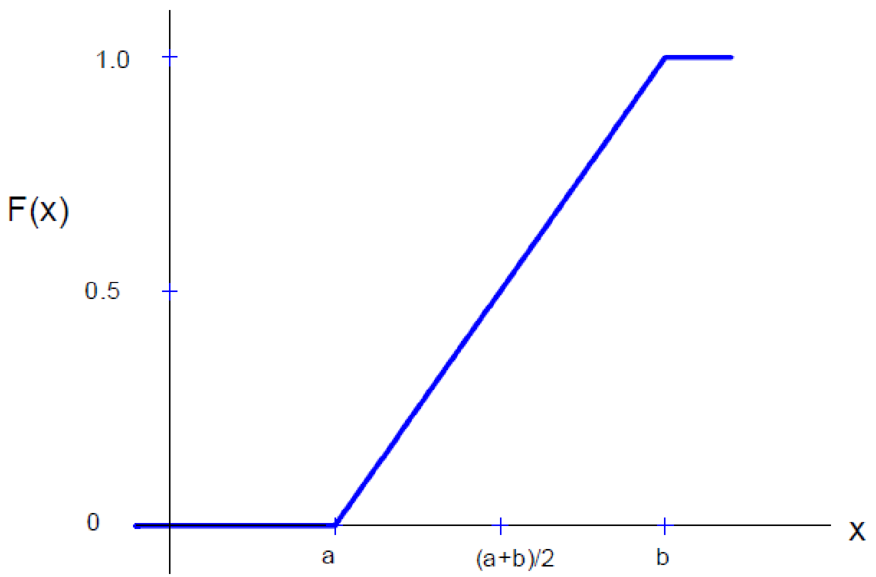
Remember that $\int_a^b f(x)dx = 1$

So, this means that k must equal ………

In addition, it follows that the **c.d.f.** is given by



1/(b-a)
f (x)
x
a
(a+b)/2
b



1.0
F(x)
0.5
0
x
a
(a+b)/2
b

Mean: E(X) = $\frac{a+b}{2}$ (the midpoint of the distribution)

Note: This distribution is symmetric about its mean.

Variance: $Var(X) = \frac{(b-a)^2}{12}$

Example: How to **transform** a r.v. with a general continuous distribution to obtain a uniformly distributed one.

Let X be a continuous r.v. with p.d.f. of the form

$$f(x) = 0.1e^{-0.1x} \text{ for } x > 0.$$

Let the **new** r.v. be given by

$$Y = e^{-0.1X}.$$

Show that Y has a uniform distribution on [0,1].

Exponential Distribution (Section 8.3)

Physical setup:

In a Poisson process which models events occurring over time, let X represent **the length of time we wait until the first event occurrence**. Then, X has an exponential distribution.

For example:

If phone calls to a fire station follow a Poisson process, then the ***length of time between*** consecutive phone calls to the station follows an exponential distribution.

The p.d.f. and c.d.f. of *X* are given by:

Alternate Form: For $\theta > 0$,

$$f(x) = \frac{1}{\theta} e^{-x/\theta} \text{ for } x > 0$$

and

$$F(x) = 1 - e^{-x/\theta} \text{ for } x \geq 0$$

Short-hand notation:

For this alternate form, we write:

$X \sim Exponential(\theta)$.

PDF and CDF of an Exponential(θ) Distribution

Figure 8.9: Graph of the probability density function of a $Exponential\ (\theta)$ random variable

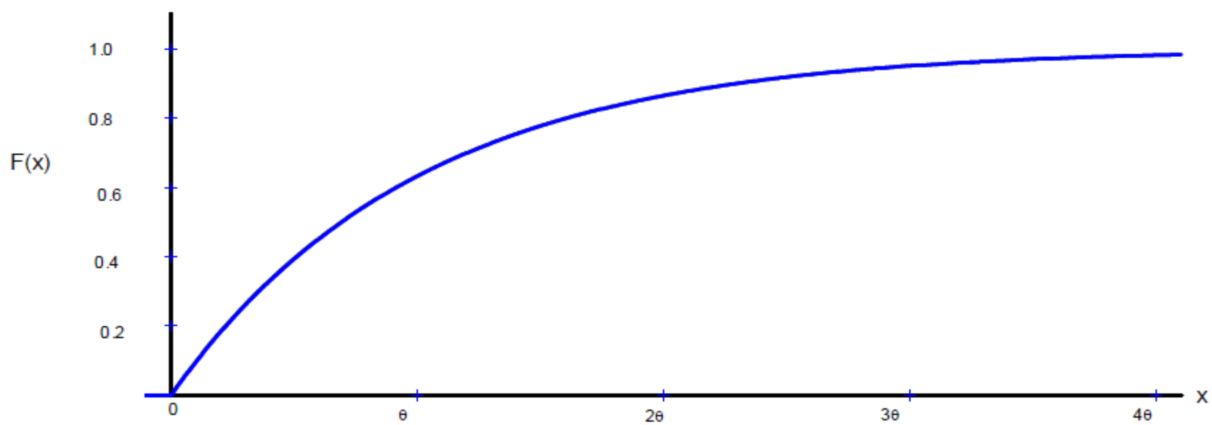


Figure 8.10: Cumulative distribution function for a $Exponential\ (\theta)$ random variable

Exponential pdf's with Varying Values of Theta

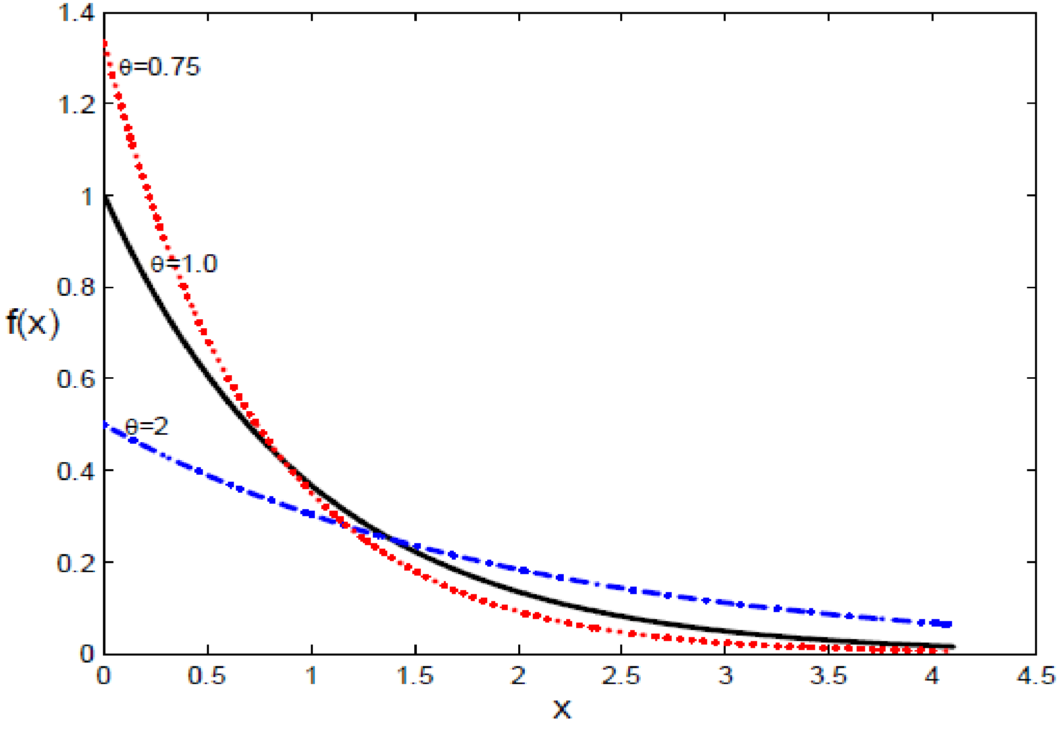


Figure 8.11: Exponential probability density functions for different values of θ

Mean and Variance:

Finding the mean and variance directly involves the use of **integration by parts**.

To avoid this integration, we can make use of the properties of the **Gamma function**.

Definition: The function

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

is called the **Gamma function** of α where $\alpha > 0$.

Note that

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

Properties of the Gamma function we will use:

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

2. $\Gamma(\alpha) = (\alpha - 1)!$ If α is a positive integer

3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Returning to the exponential distribution.

If $X \sim \text{Exponential}(\theta)$, then:

Mean

Variance

Keep in mind:

$\boldsymbol{\lambda}$ is the average **rate** of occurrence in a Poisson process

$\boldsymbol{\theta = 1/\lambda}$ is the average waiting **time** for an occurrence

Example: The amount of time in hours that a computer survives before breaking down is exponentially distributed with a mean of 100 hours. What is the probability that:

a) a computer will function between 50 and 150 hours before breaking down?

b) it will function for fewer than 100 hours?

c) if a computer survives more than 100 hours, what is the probability it survives an additional 50 hours?

Part c) of the previous example illustrates the well-known **"memoryless property"** of the exponential distribution:

$$\boldsymbol{P(X > b + c | X > b) = P(X > c)}$$

Interpretation: Given that you have waited b units of time for the next event, the probability you wait an additional c units of time **does not depend** on b and only depends on c.

You Try:
Suppose that the length of a phone call in minutes is an exponential random variable with a mean of 10 minutes. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait

Let X represent the length of a phone call (in minutes).
In this case, X~Exponential(θ = 10)

a) more than 10 minutes.

We want P(X > 10) = 1 – F(10) = $e^{-1} = \frac{1}{e}$ **= 0.3679**

b) between 10 and 20 minutes.

We want P(10 < X < 20) = F(20) – F(10) = e^{-1} **-** e^{-2} **= 0.2325**

c) more than an additional 7 minutes, given you have been waiting for more than 10 minutes.

We want P(X > 17 | X > 10) = $\frac{P(X>17)}{P(X>10)} = \frac{e^{-17/10}}{e^{-10/10}} = e^{-7/10}$ **= 0.4966**

A Method for Computer Generation of Random Variables (Section 8.4)

Theorem:

If F is an arbitrary cumulative distribution function and U~Uniform(0, 1), then the random variable defined by $X = F^{-1}(U)$ has a cumulative distribution function F(x).

A Method for Computer Generation of Random Exponential Example

Example: Exponential Random Number Generator.

If U is a Uniform random variable in (0, 1), we want a value of X from an Exponential(θ) distribution.

If X~Exp(θ), we know that F(x) = $1 - e^{-x/\theta}$

We know that U~Uniform(0, 1), so, we have:

U = F(X) = $1 - e^{-X/\theta}$.

We simply need to isolate for X.

This gives us $e^{-X/\theta}$ = 1 – U => $-X/\theta$ = log(1 – U)

So, X = - θ*log(1-U)

Note: log = natural logarithm, or ln.

A Method for Computer Generation of Random Exponential Example

Example: Exponential Random Number Generator.

If U is a Uniform random variable in (0, 1), let's generate a value of X from an Exponential(5) distribution.

If X~Exp(5), we know that F(x) = $1 - e^{-x/5}$

We know that U~Uniform(0, 1).

Let's generate the value of U using R: runif(1) = 0.902066
Note: runif(**number**) will generate **number** values from a (0,1) Uniform distribution.

So, we have 0.902066 = F(x) = $1 - e^{-x/5}$.

We simply need to isolate for the value, x.

This gives us $e^{-x/5}$ = 1 – 0.902066 => $-x/5$ = log(1 – 0.902066)

So, x = - 5*log(1-0.902066) = **11.617307**
Note: log = natural logarithm, or ln.

Normal Distribution (Section 8.5)

Physical setup:

A r.v. X defined on $(-\infty,\infty)$ has a *normal distribution* if it has p.d.f. of the form

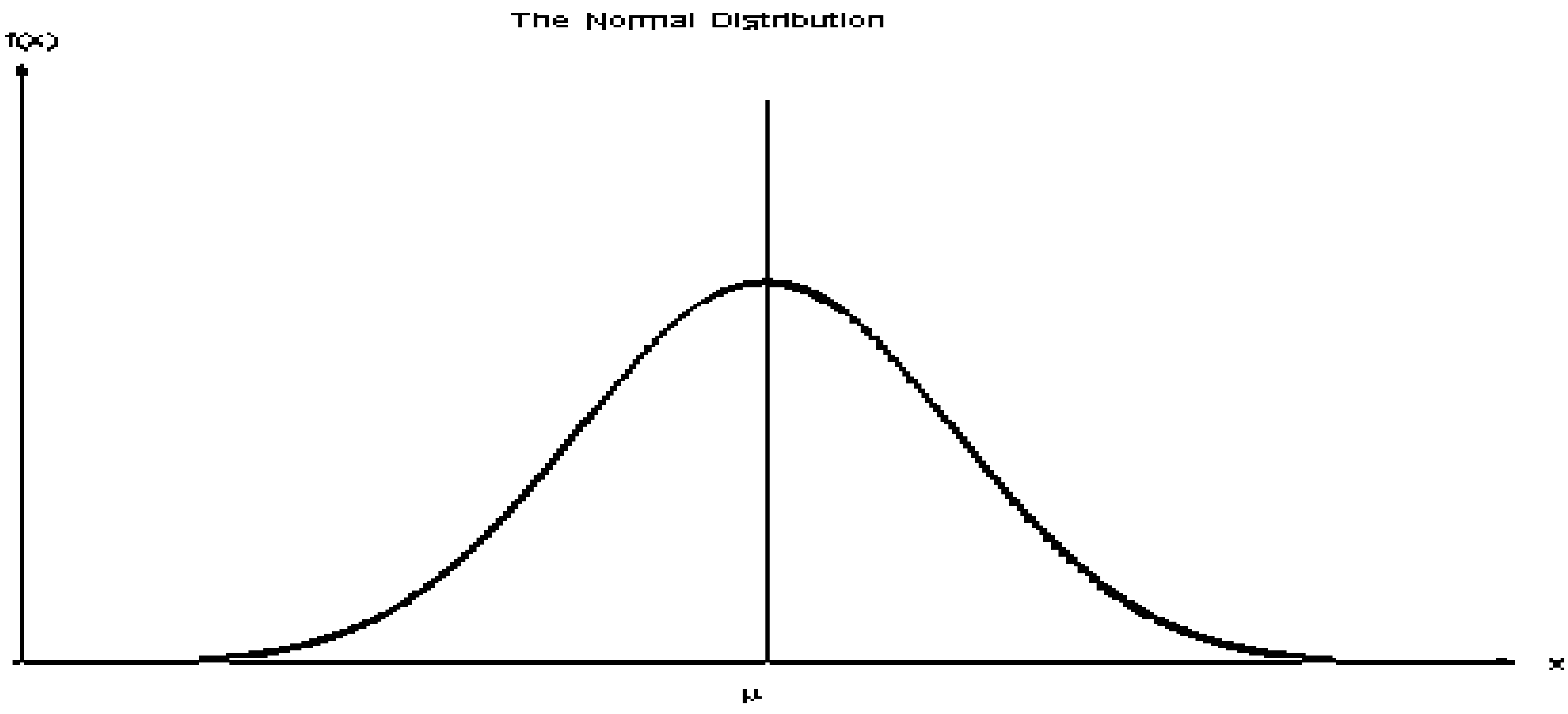
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad \text{for } -\infty < x < \infty$$

where $-\infty < \mu < \infty$ and $\sigma > 0$ are parameters.

- It turns out that $\mu = E(X)$ and $\sigma^2 = Var(X)$
- We write $X \sim N(\mu, \sigma^2)$

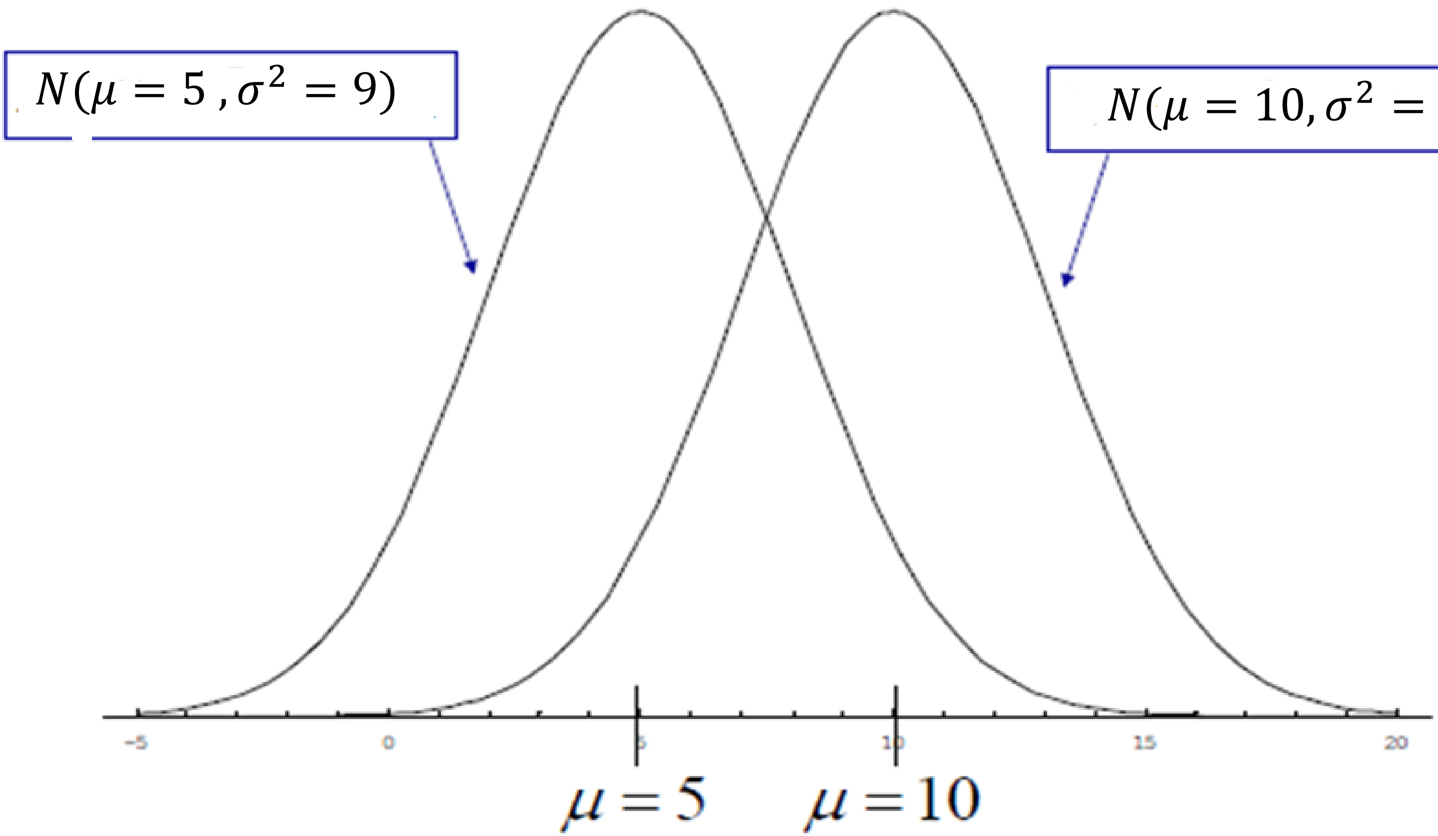
<u>or</u>

$$\boldsymbol{X \sim G(\mu, \sigma)}$$



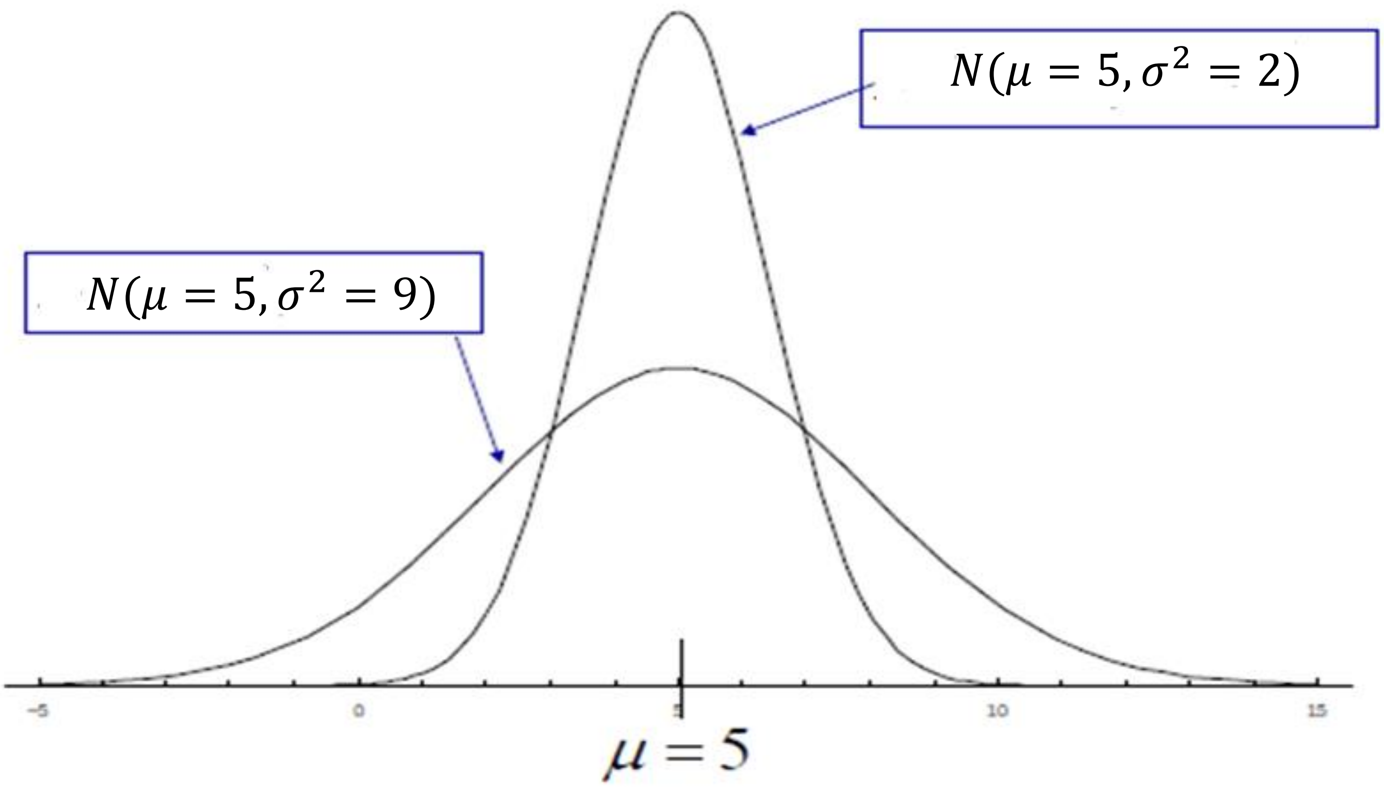
The Effect of the Mean μ

The mean shifts the distribution along the *x*-axis (i.e. μ is a **location** parameter).



The Effect of the Variance σ^2

The variance either stretches out or pulls in the distribution (i.e. σ^2 is a **scale** parameter).



For example:

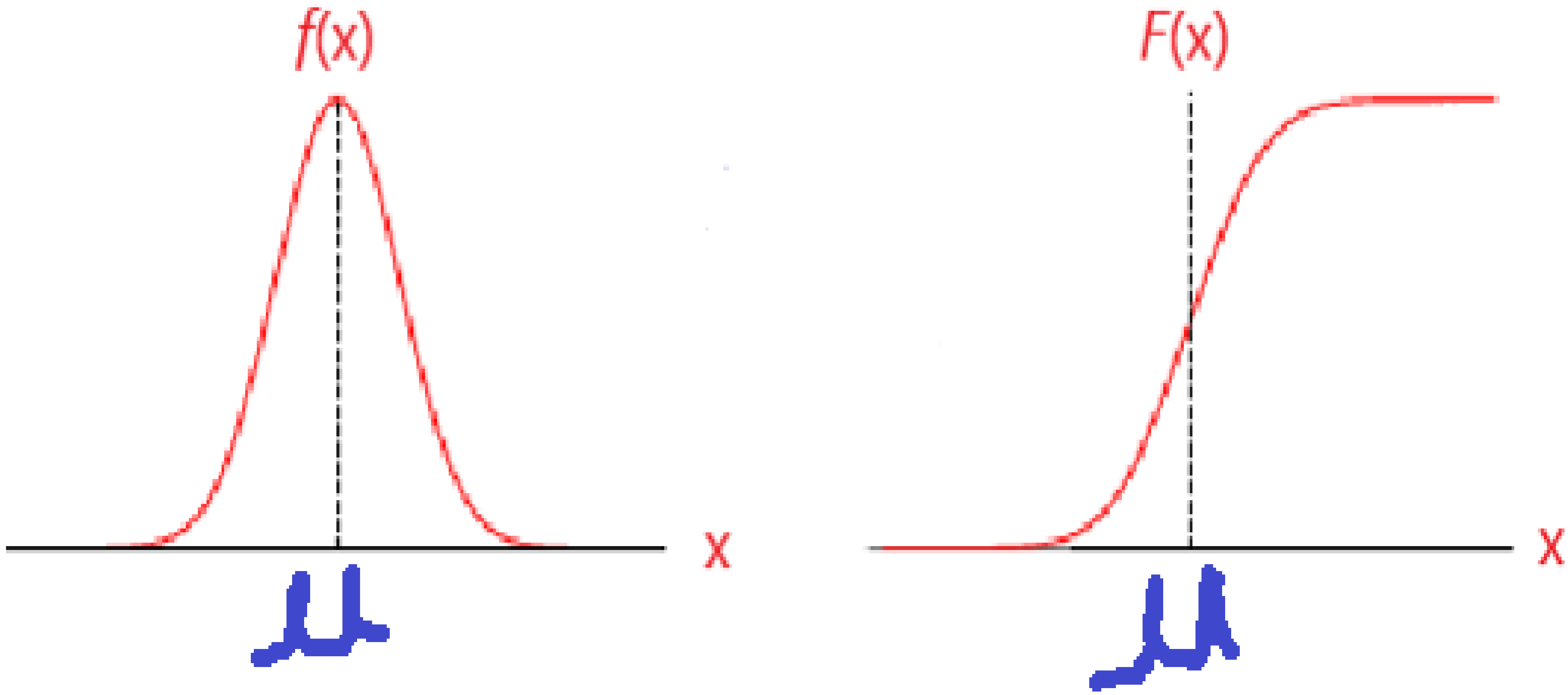
The heights or weights of males (or females) in large populations tend to follow normal distributions.

The c.d.f. of a $N(\mu, \sigma^2)$ r.v. is given by

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2} dy$$

However, this integral does not yield a simple mathematical expression, and so numerical methods have to be used in order to evaluate it.

We can make use of Standard Normal probability tables (in the Course Notes)



Numerical methods need to be used, and prior to computers, statisticians produced **tables of probabilities** of $F(x)$ by numerical integration, using mechanical calculators.

Finding Normal Probabilities via a $N(0,1)$ Table

Fortunately, only one table is needed and is based on the **standard normal distribution** (i.e. $Z \sim N(0,1)$) and the following result:

Theorem: Let $X \sim N(\mu, \sigma^2)$ and define

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

Then, $Z \sim N(0,1)$ and

$$\boldsymbol{P(X \leq x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right)}.$$

Note: The r.v. Z is a **unitless** quantity, and its p.d.f.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \text{ for } -\infty < z < \infty$$

can be shown to satisfy $\int_{-\infty}^{\infty} f_Z(z)dz = 1.$

The observed value, z, that you obtain (also called a z-score) gives you the number of standard deviations away from the mean that your observation is.

If the value is **negative**, then the observation is to the **left of the mean**.

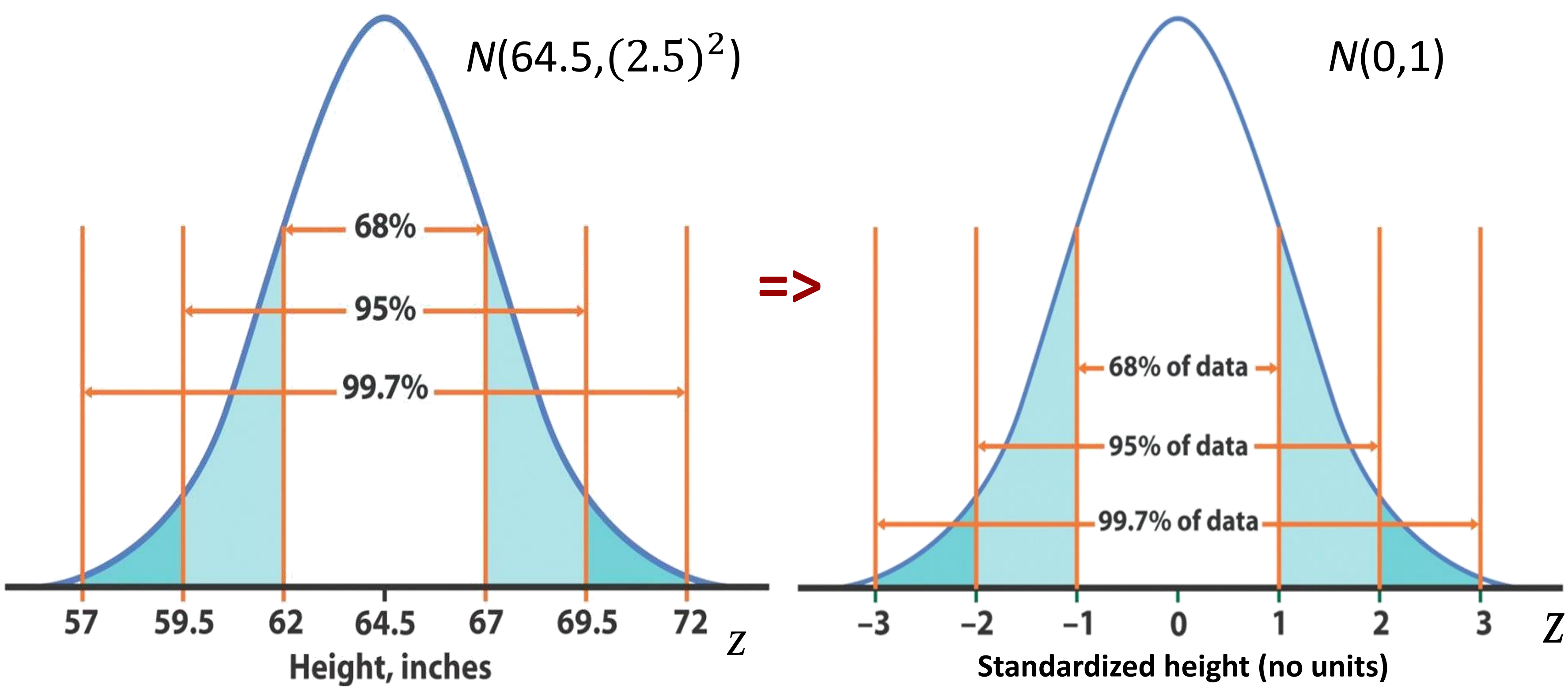
If the value is **positive**, then the observation is to the **right of the mean**.

If the **value is 0**, then the observation is **equal to the mean**.

Note the median equals the mean in this case, as the distribution is **symmetric** about its mean of 0.

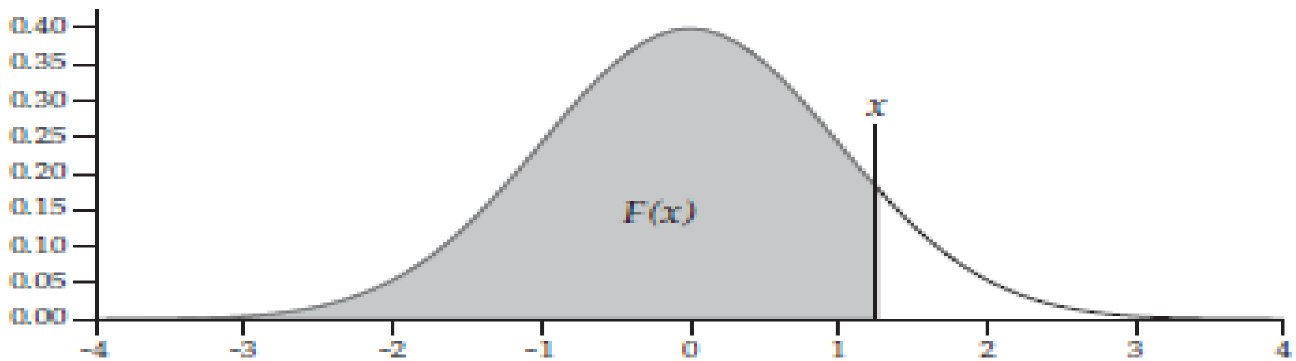
68 – 95 – 99.7 Rule

Consider the distribution of heights of young women aged 18 to 24, denoted by the r.v. X.
The distribution is approximately normal with mean 64.5 inches and standard deviation 2.5 inches.
In other words, $X \sim N(64.5, (2.5)^2)$.



Probability Table for a $N(0,1)$ Distribution

Probabilities for Standard Normal $N(0,1)$ Distribution



This table gives the values of $F(x)$ for $x \geq 0$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56750	0.57142	0.57534
0.2	0.57926	0.58317	0.58706	0.59095	0.59484	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189

You try: Let $Z \sim N(0,1)$.
Using the Normal table in the Course Notes, find the following probabilities:

a) $P(Z < 2.11)$

P(Z < 2.11) = 0.98257

b) $P(Z > 1.06)$

P(Z > 1.06) = 1 – P(Z < 1.06) = 1 – 0.85543
So, P(Z > 1.06) = 0.14457

c) $P(Z < -1.06)$

P(Z < -1.06) = P(Z > 1.06) (by symmetry)
From part b), P(Z > 1.06) = 1 – P(Z < 1.06) = 1 – 0.85543
So, P(Z < -1.06) = 0.14457

d) $P(-1.06 < Z < 2.11)$

P(-1.06 < Z < 2.11) = P(Z < 2.11) – P(Z < -1.06) = P(Z < 2.11) – [1 – P(Z < 1.06)]
So P(-1.06 < Z < 2.11) = 0.98257 – [1 – 0.85543] = 0.98257 – 0.14457 = 0.838

Example: Let $Z \sim N(0,1)$.

a) Find a number c such that $P(Z < c) = 0.85$.

b) Find a number d such that $P(Z > d) = 0.90$.

c) Find a number b such that $P(-b < Z < b) = 0.95$.

Example: Suppose that $X \sim N(10, 2)$.

Calculate $P(|X - 10| \leq 3)$.

You Try:

Now, consider a r.v. X such that $X \sim N(-7, 14)$.

Calculate $P(|X + 7| \geq 8)$.

Let Y = X + 7. In this case, Y ~ N(0, 14).

We subtract off the mean of 0, and divide through by $\sqrt{\mathbf{14}}$.

This gives $\boldsymbol{P\left(|Z| \geq \frac{8}{\sqrt{14}}\right) = P(|Z| \geq}$ 2.14).

$\boldsymbol{P(|Z| \geq}$ 2.14) = P(Z > 2.14) + P(Z < -2.14)

= 1 – P(Z < 2.14) + P(Z > 2.14) (by symmetry)

= 1 – P(Z < 2.14) + [1 – P(Z < 2.14)]

= 2*[1 – P(Z < 2.14)] = 2*[1 – 0.98382] = 0.03236

Example: Suppose that a certain mechanical component produced by a company has a width that is normally distributed with mean $\mu = 2600$ and standard deviation $\sigma = 0.6$.

a) What proportion of the components have a width outside the range of 2599 to 2601?

b) If the company needs to be able to guarantee to its purchaser that no more than 1 in 1000 of its components have a width outside the range of 2599 to 2601, by how much does the value of σ need to be reduced?

You Try:

The intelligence quotient (IQ) score, as measured by the Stanford-Binet IQ test, is normally distributed in a certain population of children. The mean IQ score is 100 points, and the standard deviation is 16 points.

Let X represent IQ score for this population of children. X ~ N(100, 256).

a) What percentage of children in the population have IQ scores of 140 or more?

We want P(X ≥ 140) = 1 – P(X < 140)

= 1 - $\boldsymbol{P\left(Z < \frac{140-100}{16}\right) = 1 - P(Z < 2.50)}$

$\boldsymbol{= 1 - 0.99379 = 0.00621}$

b) What percentage of children in the population have IQ scores between 80 and 120?

We want P(80 < X < 120) = P(X < 120) – P(X < 80)

= $\boldsymbol{P\left(Z < \frac{120-100}{16}\right) - P\left(Z < \frac{80-100}{16}\right) = P(Z <}$ **1.25) – P(Z < -1.25)**

= 2*P(Z < 1.25) – 1 = 2*(0.89435) – 1 = 0.7887

You Try (cont'd):

The intelligence quotient (IQ) score, as measured by the Stanford-Binet IQ test, is normally distributed in a certain population of children. The mean IQ score is 100 points, and the standard deviation is 16 points.

c) Solve for the 95th percentile (i.e. the IQ level such that 95% of children have an IQ below it).

Start with determining the 95th percentile of the Z distribution, where Z ~ N(0, 1).

Go to the table below the Standard Normal probability table, and we see, with p = 0.95, we have z = 1.6449.

We are looking for the value, x, such that P(X < x) = 0.95.

This corresponds to $P(Z < \frac{x-100}{16})$ = 0.95. The corresponding value, z, is 1.6449.

So, we set $\frac{x-100}{16}$ = 1.6449 and solve for x.

This gives us x = 100 + 16*1.6449 = 126.3184.