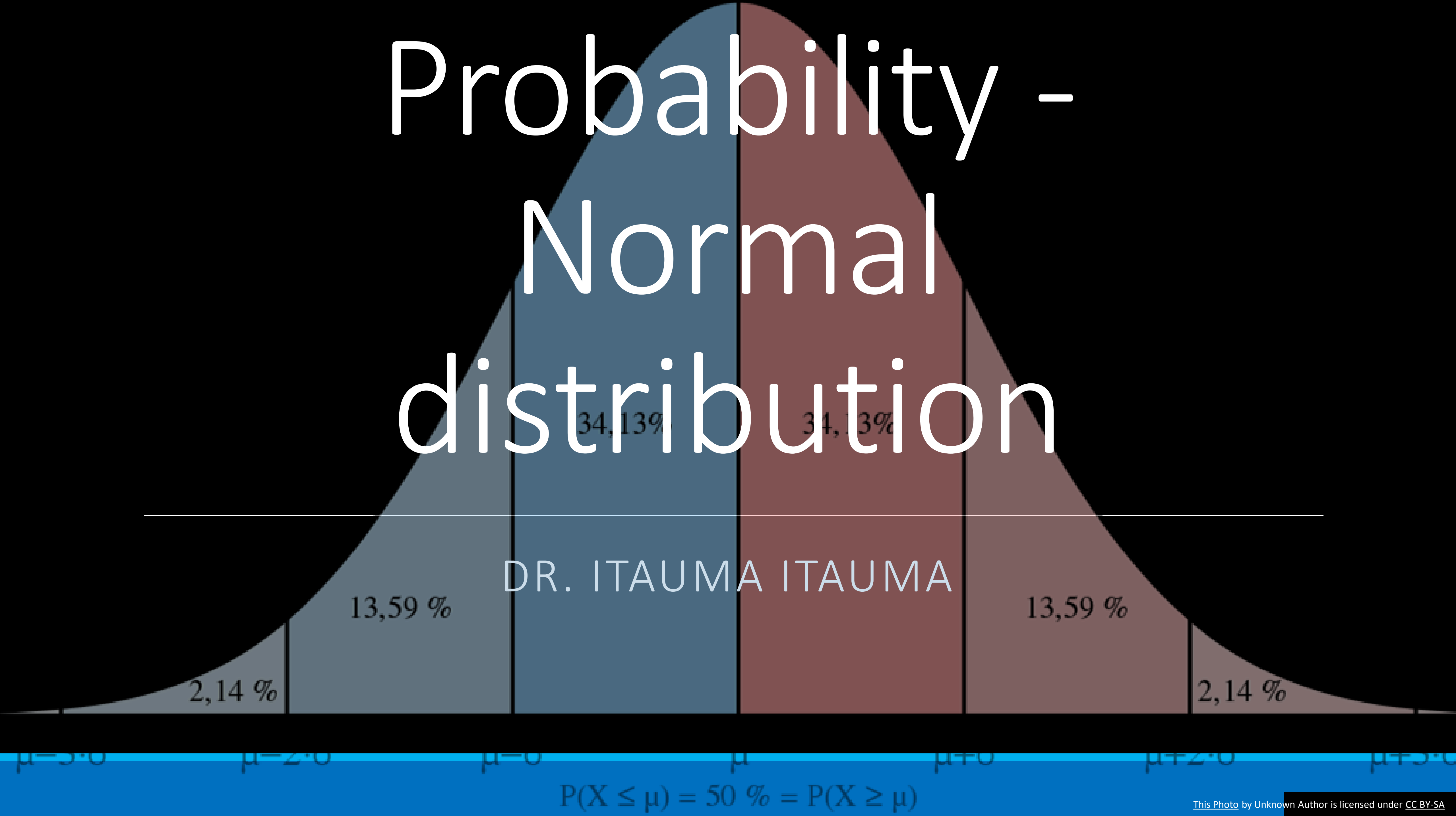
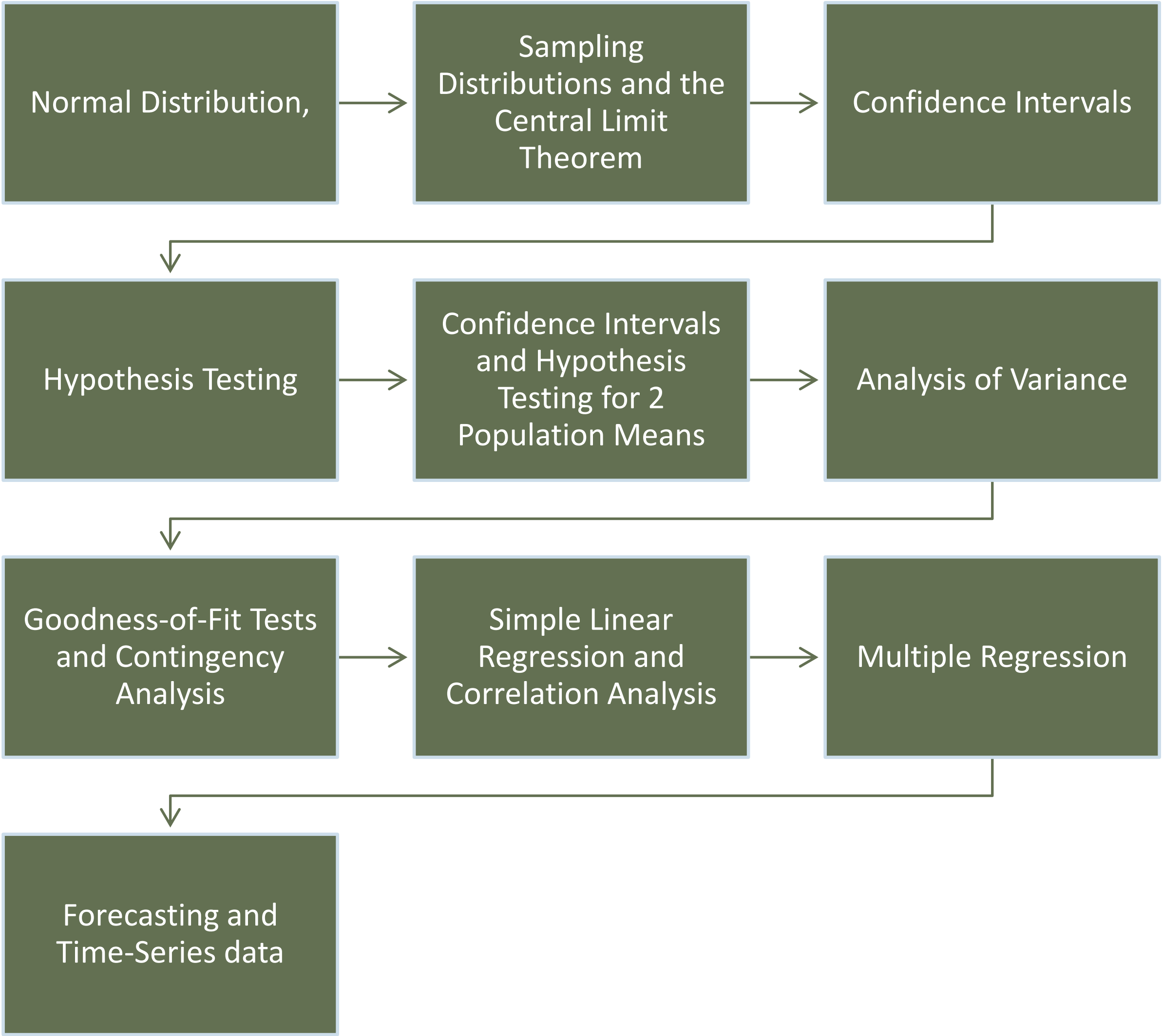
Probability - Normal distribution

DR. ITAUMA ITAUMA





Course Learning Objectives

Learning Objectives

This week we will review:

- Normal distribution
- Using Minitab for Normal Calculations

Continuous Random Variables

Continuous random variables take on any value within a specific range.

For example:

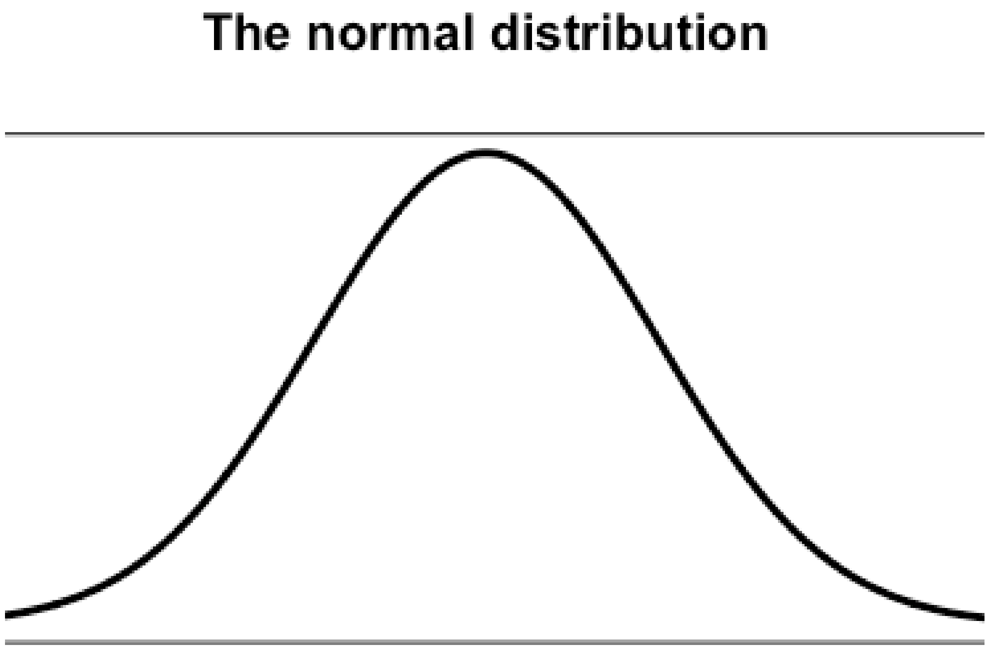
- Amount of time it takes a customer to check out at an online retailer.
- Distance a supplier travels to deliver goods.

Distributions for continuous random variables are defined by smooth functions/curves, and probabilities are represented by areas under the curve.

Intro to the Normal Distribution

The <u>normal distribution</u> is a commonly used continuous distribution.

- The normal distribution is often called the “bell-shaped curve” because of its shape:



The Normal Distribution (cont.)

A normal distribution is defined by two parameters: the mean μ and the variance σ^2.

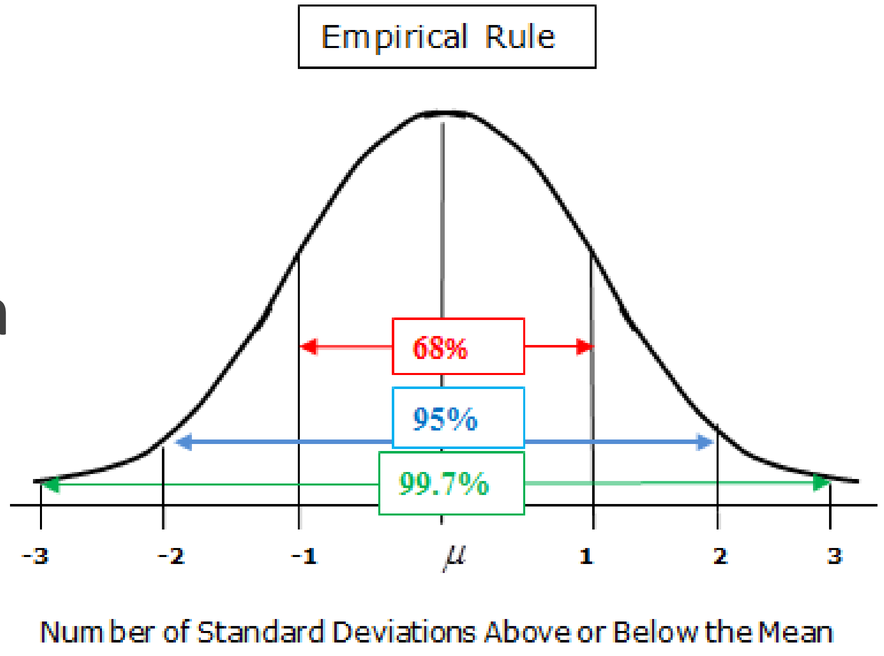
If the random variable X follows a normal distribution, it is denoted in shorthand as $X \sim \mathrm{N}(\mu, \sigma^2)$.

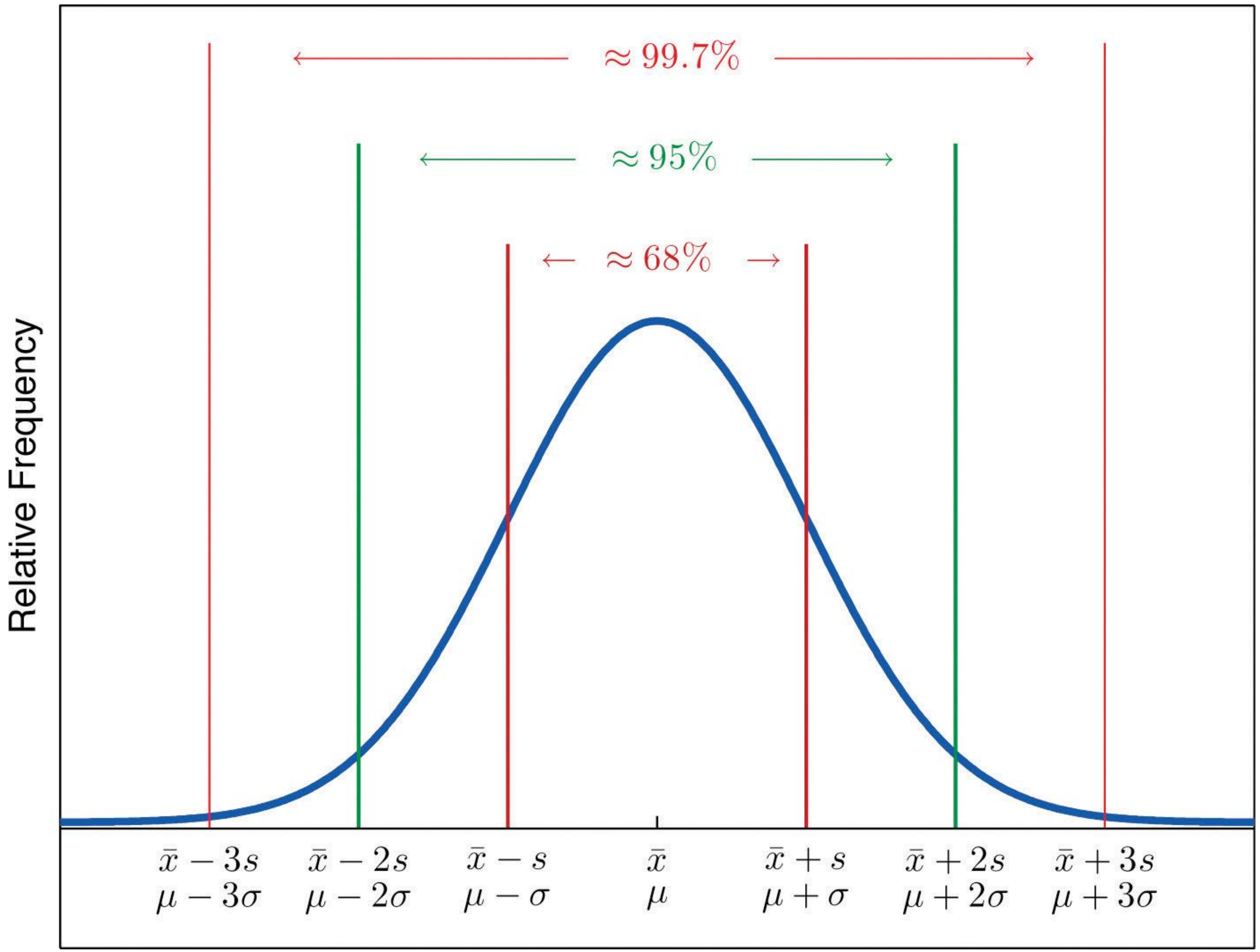
All normal distributions have the same bell shape and are centered at a mean of μ and have a spread based on the variance of σ^2.

Empirical Rule

Probabilities for a normal distribution can be roughly calculated from the <u>empirical rule</u> (sometimes called the “68-95-99.7” rule):

- 68% of a normal distribution lies between -1 and +1 standard deviations away from the mean.
- 95% lies between -2 and +2 standard deviations.
- 99.7% lies between -3 and +3 standard deviations.





Empirical Rule Curve

Empirical Rule for Data

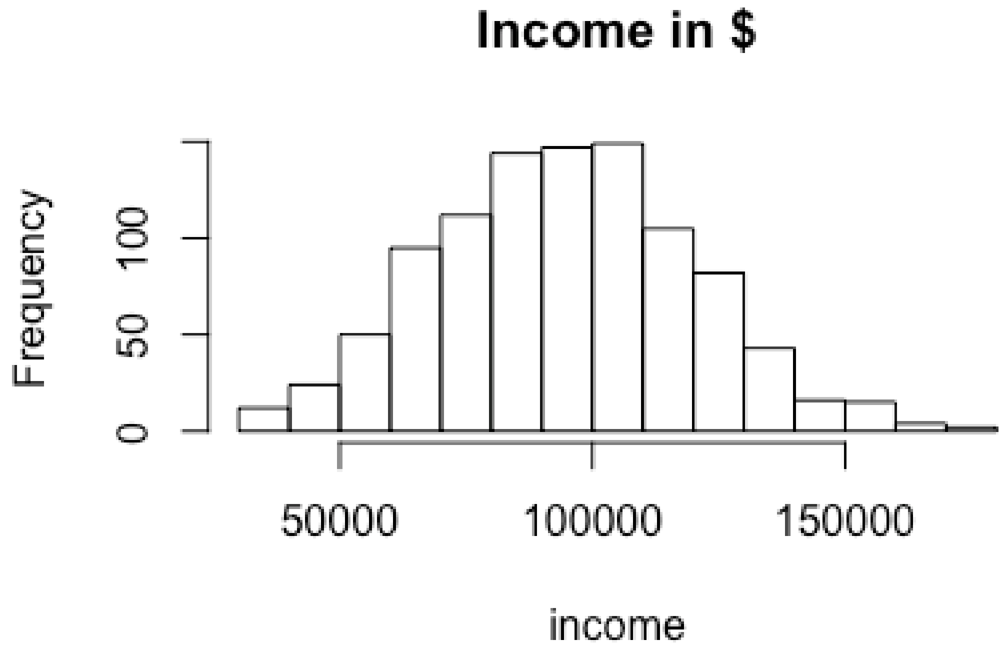
The empirical rule also holds for data whose histogram looks roughly normal.

These data have:

$$\bar{x} = 94500$$

$$s = 25300$$

So about 95% of
the data lie between
94500 – 2(25300) = \$43,900 and
94500 + 2(25300) = \$145,100.



Z-score

Often times it is easier to convert a value into a *z-score* by subtracting off the mean and dividing by standard deviation (also called standardizing) .

For example if X ~ N(μ = 140, σ^2 = 30^2), then the value x = 200 becomes a *z*-score of:

$$z = \frac{x - \mu}{\sigma} = \frac{200 - 140}{30} = 2$$

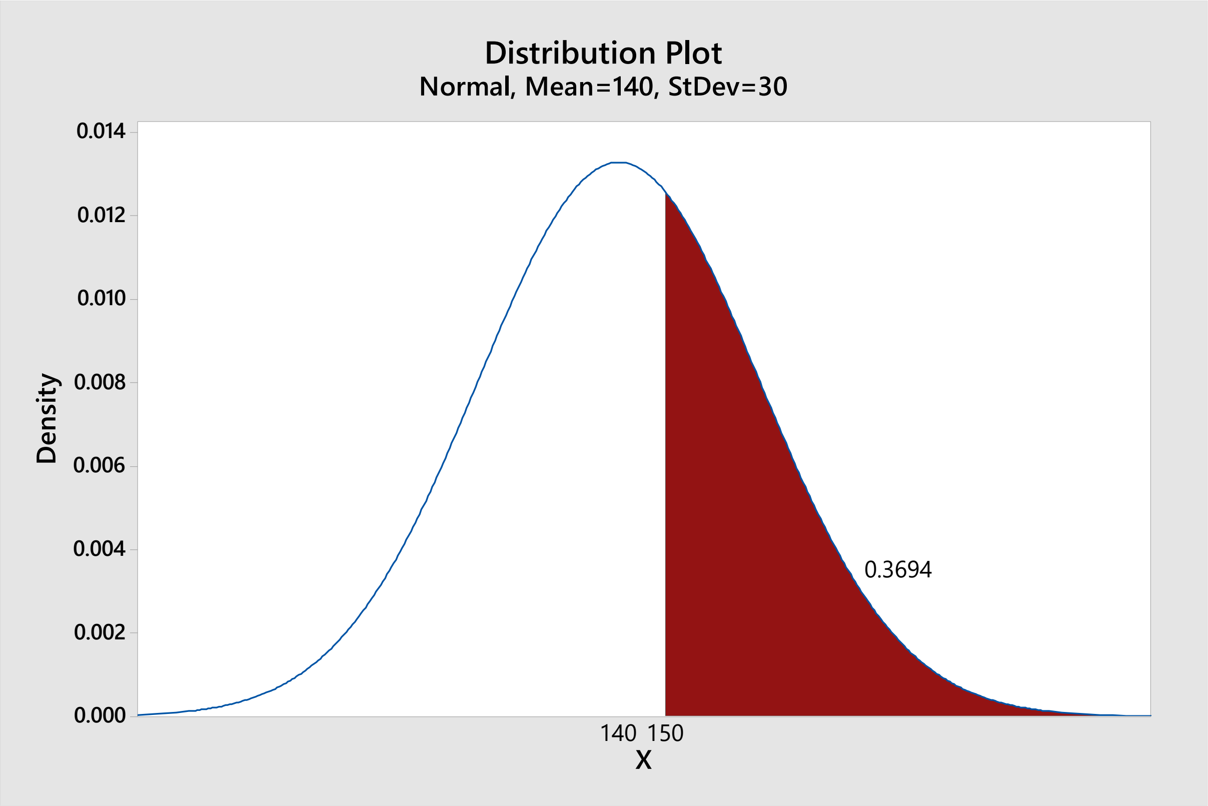
The *z*-score can be interpreted as the number of standard deviation away from the mean.

Using Minitab for Normal Calculations

The empirical rule only works if your values of interest happen to lie at exactly 1, 2, or 3 standard deviations.

If you want to calculate probabilities at other values, you need to use Minitab.

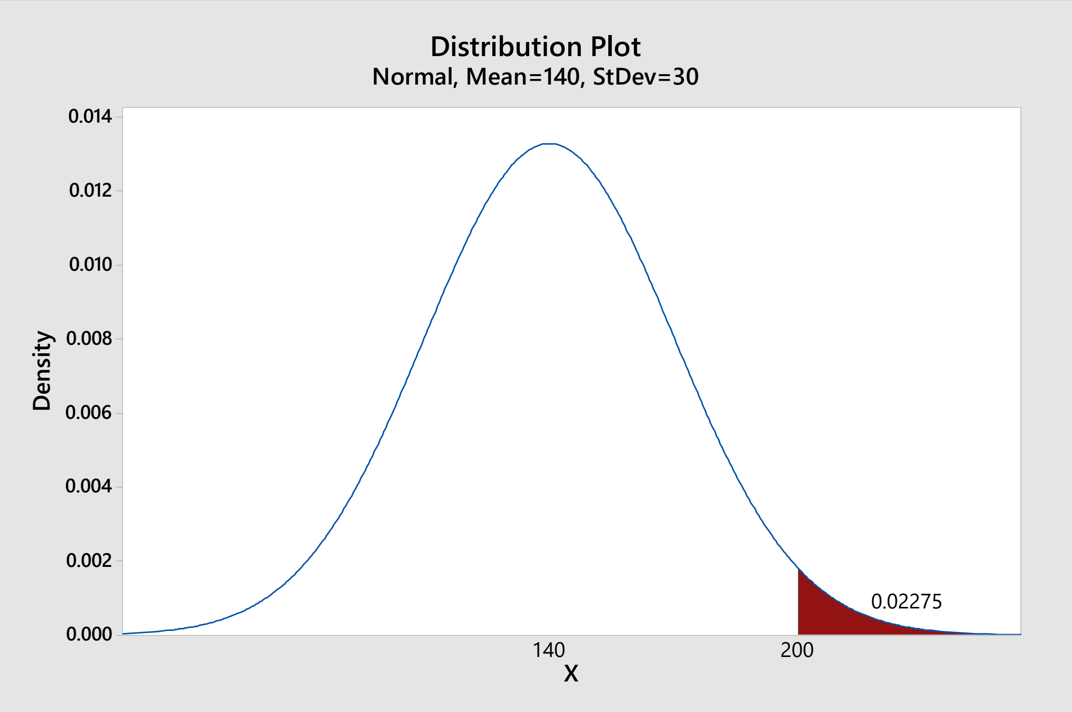
If $X \sim N(\mu = 140, \sigma^2 = 30^2)$, find $P(X > 150)$.



Calculating normal probabilities

Assume that salaries of recent MBA graduates are normally distributed with mean \$140K and standard deviation \$30K.

1. What is the probability of a recent MBA graduate getting paid more than \$200K?



Answer:

$P(X > 200) = P(Z > 2)$

Using the Empirical Rule 2

Assume that salaries of recent MBA graduates are normally distributed with mean \$140K and standard deviation \$30K.

2. Getting paid between \$110K and \$200K?

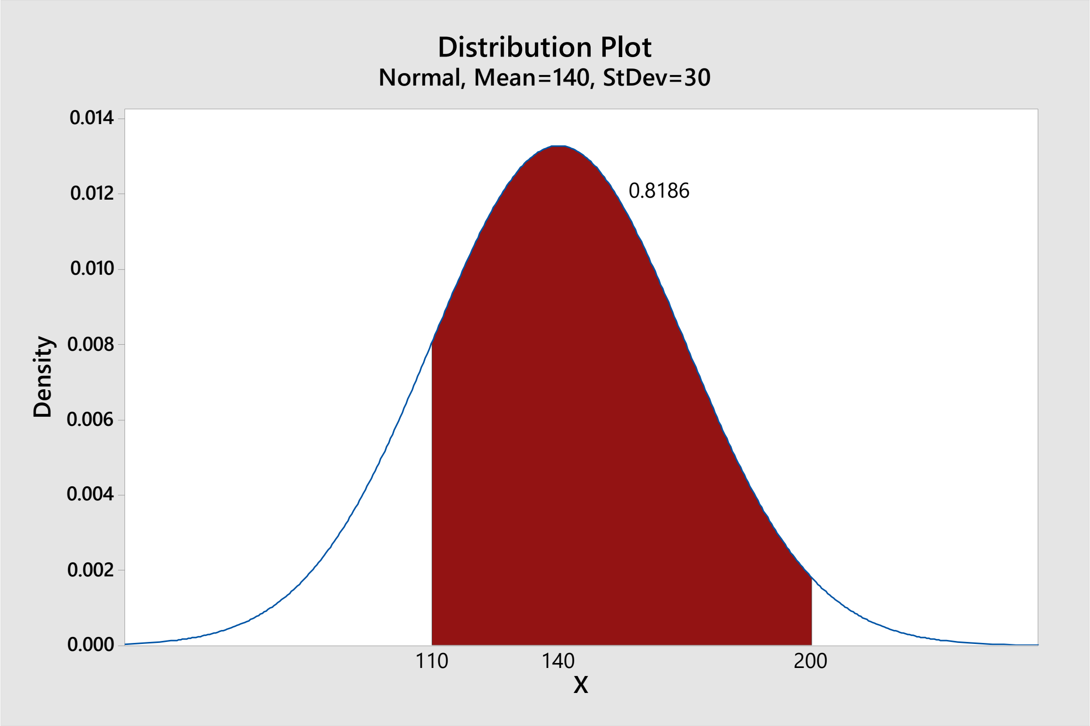
<u>Answer:</u>

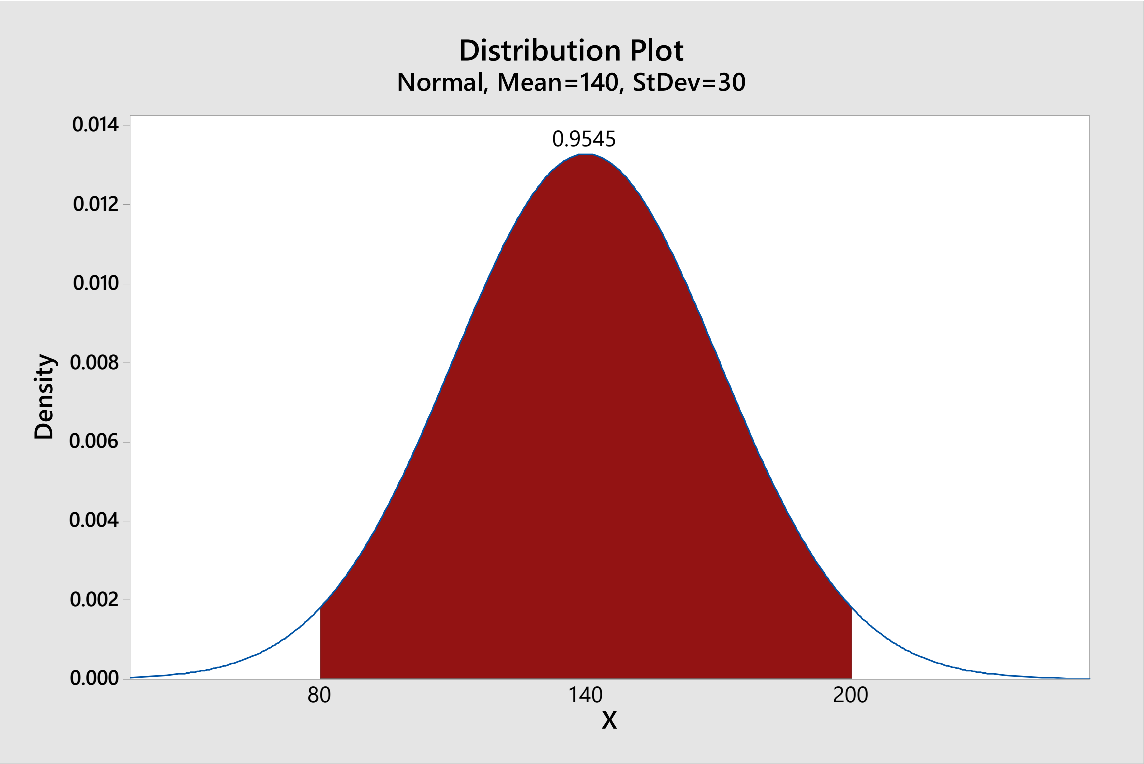
$P(110 < X < 200) = P(-1 < Z < 2)$

$= P(Z < 2) – P(Z < -1)$

$= 0.975 – 0.16$

$= 0.815$





Using the Empirical Rule 3

Assume that salaries of recent MBA graduates are normally distributed with mean \$140K and standard deviation \$30K.

3. What is the middle 95% of salaries for recent MBAs?

Answer:

The middle 95% of the distribution lies between $\mu - 2\sigma$ and $\mu + 2\sigma$, which correspond to salaries of 140 – 2(30) = 80 and 140 + 2(30) = 200 (in thousands of \$).