Statistics for Managers Using Microsoft Excel
7th Edition

Chapter 6

The Normal Distribution & Other Continuous Distributions

Learning Objectives

In this chapter, you learn:

- To compute probabilities from the normal distribution
- How to use the normal distribution to solve business problems
- To use the normal probability plot to determine whether a set of data is approximately normally distributed
- To compute probabilities from the uniform distribution
- To compute probabilities from the exponential distribution

Continuous Probability Distributions

- A continuous random variable is a variable that can assume any value on a continuum (can assume an uncountable number of values)
  - thickness of an item
  - time required to complete a task
  - temperature of a solution
  - height, in inches
- These can potentially take on any value depending only on the ability to precisely and accurately measure

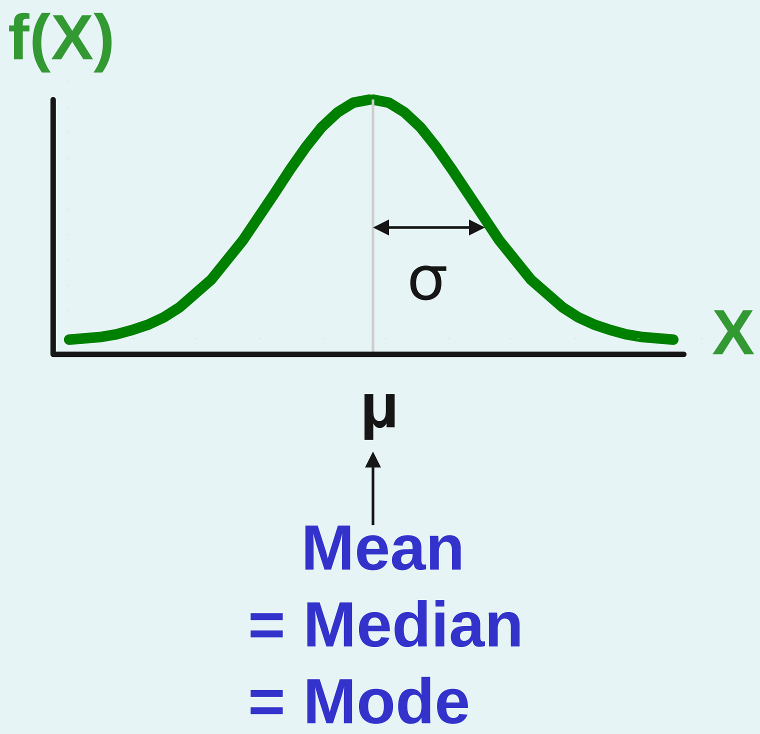
The Normal Distribution

- 'Bell Shaped'
- Symmetrical
- Mean, Median and Mode are Equal

Location is determined by the mean, μ

Spread is determined by the standard deviation, σ

The random variable has an infinite theoretical range: + ∞ to − ∞



The Normal Distribution Density Function

- The formula for the normal probability density function is

$$f(X) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{(X-\mu)}{\sigma}\right)^2}$$

Where e = the mathematical constant approximated by 2.71828

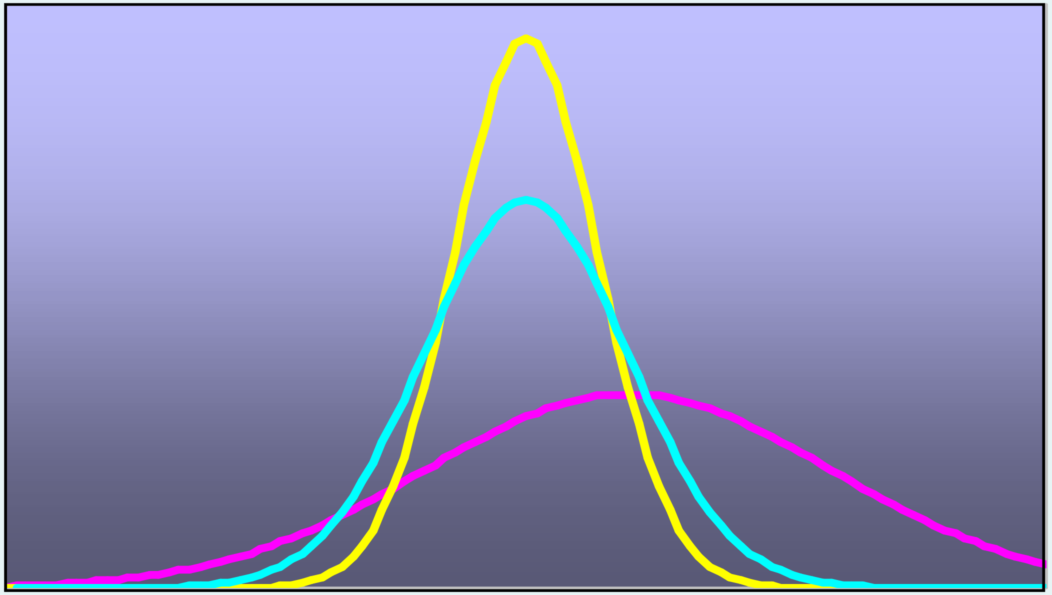
π = the mathematical constant approximated by 3.14159

μ = the population mean

σ = the population standard deviation

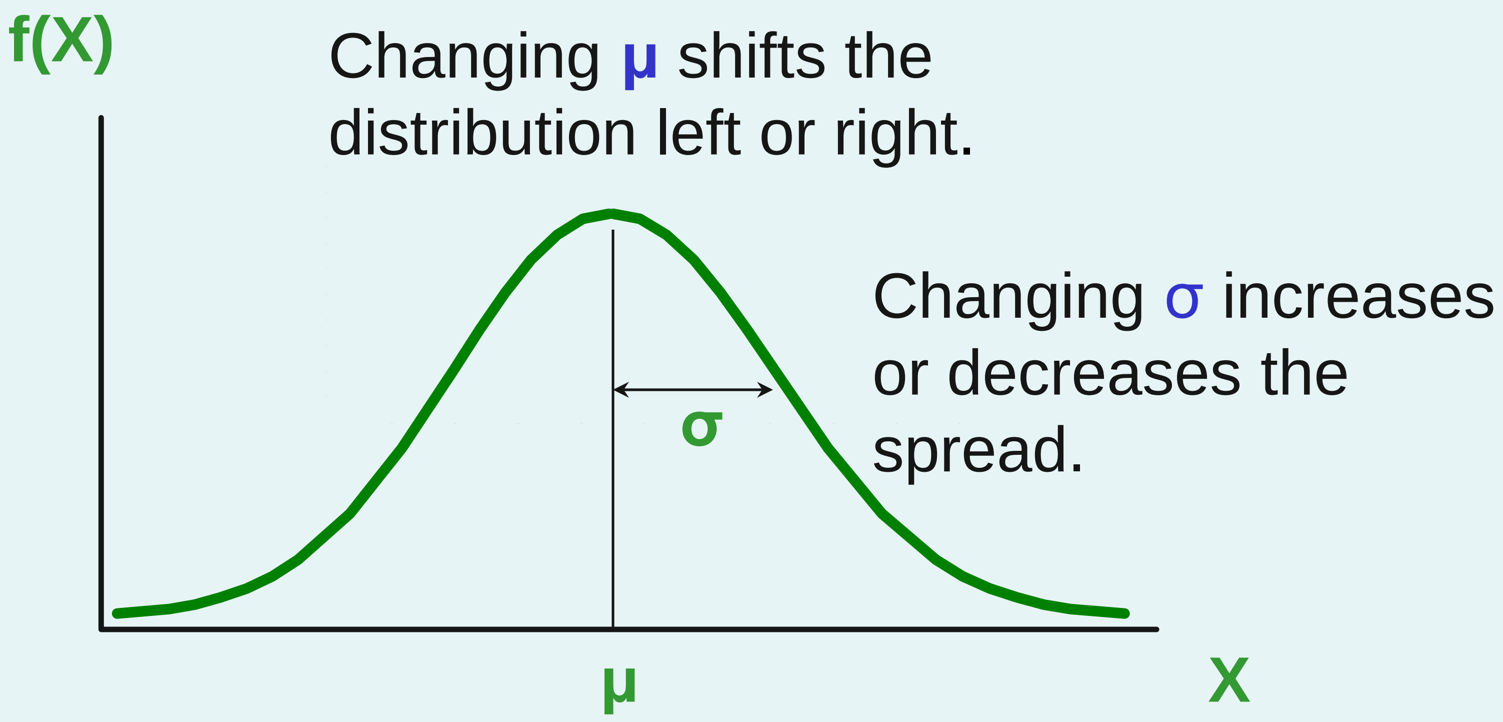
X = any value of the continuous variable

Many Normal Distributions



By varying the parameters μ and σ, we obtain different normal distributions

The Normal Distribution Shape



The Standardized Normal

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standardized normal distribution (Z)
- Need to transform X units into Z units
- The standardized normal distribution (Z) has a mean of 0 and a standard deviation of 1

Translation to the Standardized Normal Distribution

- Translate from X to the standardized normal (the "Z" distribution) by subtracting the mean of X and dividing by its standard deviation:

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

The Z distribution always has mean = 0 and standard deviation = 1

The Standardized Normal Probability Density Function

- The formula for the standardized normal probability density function is

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-(1/2)Z^2}$$

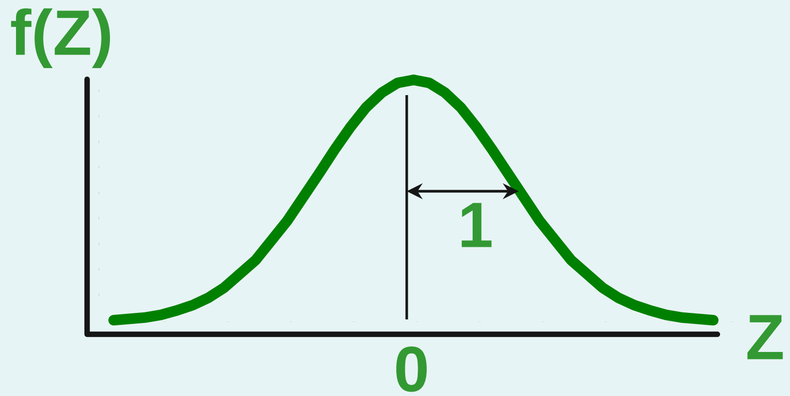
Where e = the mathematical constant approximated by 2.71828

π = the mathematical constant approximated by 3.14159

Z = any value of the standardized normal distribution

The Standardized Normal Distribution

- Also known as the “Z” distribution
- Mean is 0
- Standard Deviation is 1



Values above the mean have positive Z-values, values below the mean have negative Z-values

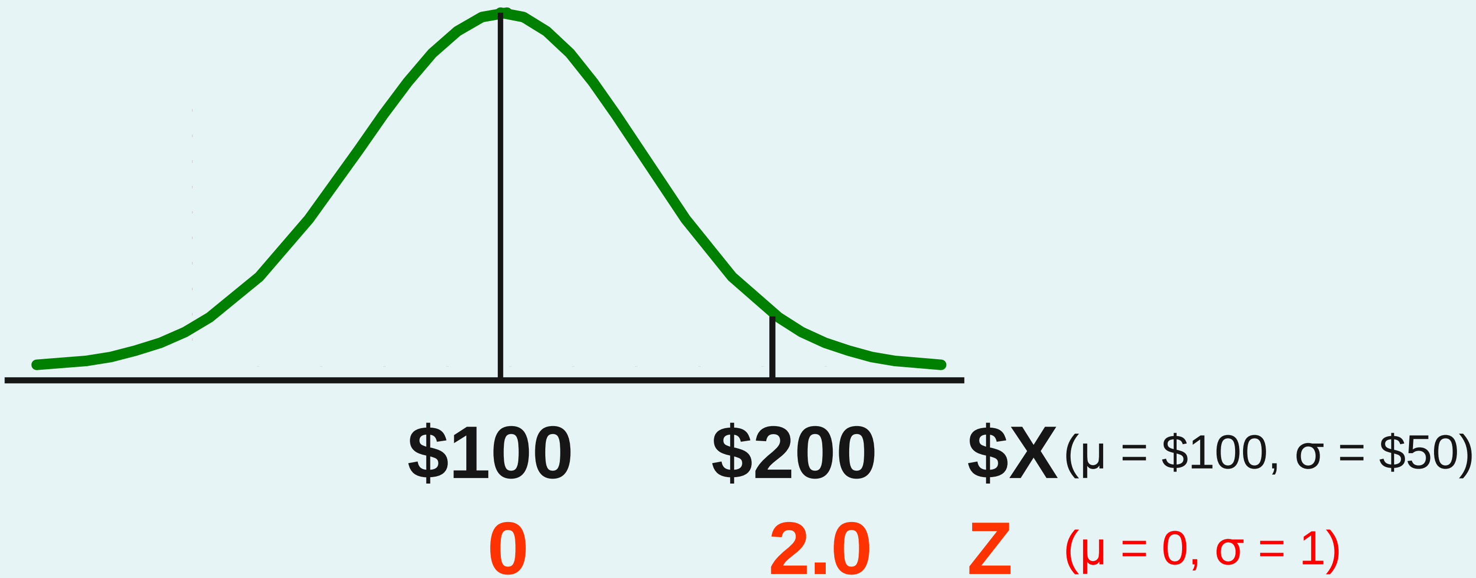
Example

- If X is distributed normally with mean of \$100 and standard deviation of \$50, the Z value for X = \$200 is

$$Z = \frac{X - \mu}{\sigma} = \frac{\$200 - \$100}{\$50} = 2.0$$

- This says that X = \$200 is two standard deviations (2 increments of \$50 units) above the mean of \$100.

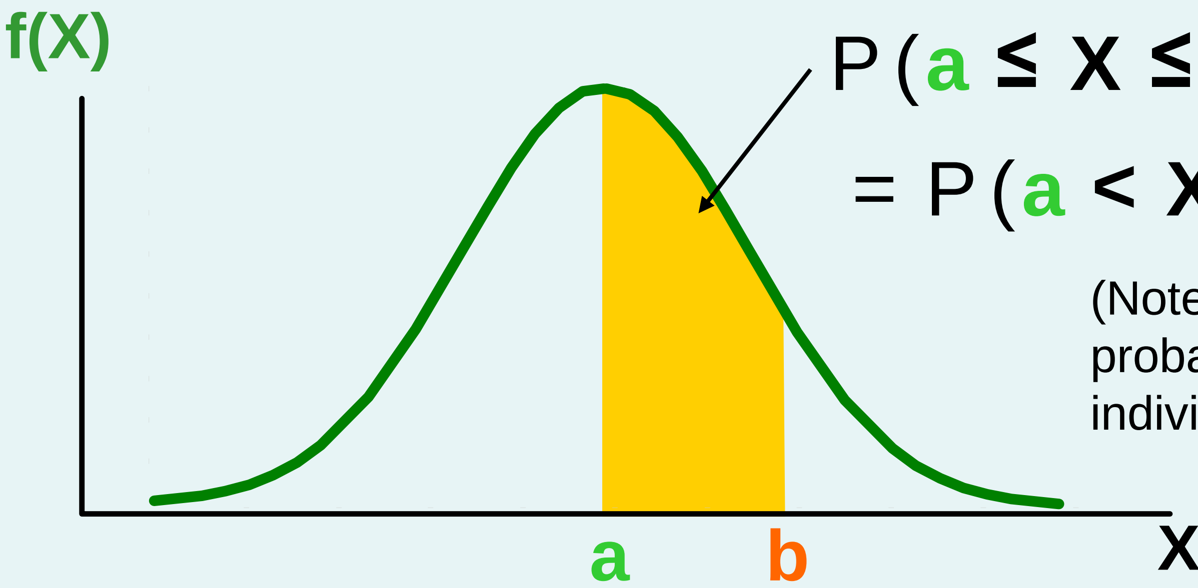
Comparing X and Z units



Note that the shape of the distribution is the same, only the scale has changed. We can express the problem in the original units (X in dollars) or in standardized units (Z)

Finding Normal Probabilities

Probability is measured by the area under the curve



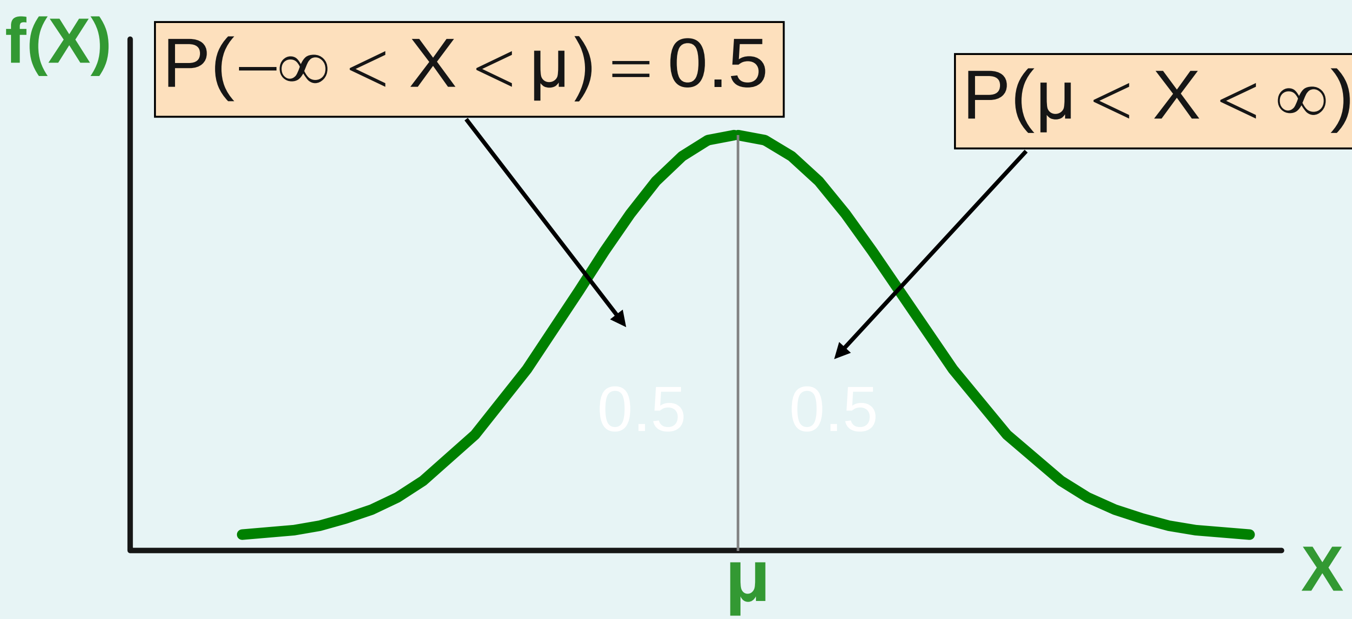
$$P(a \leq X \leq b)$$

$$= P(a < X < b)$$

(Note that the probability of any individual value is zero)

Probability as Area Under the Curve

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below



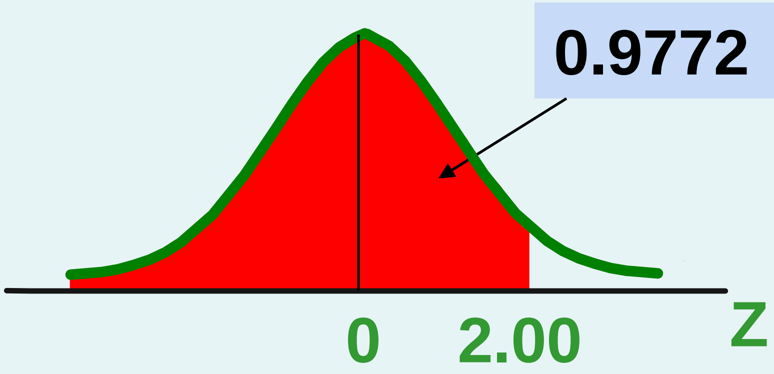
$$P(-\infty < X < \infty) = 1.0$$

The Standardized Normal Table

- The Cumulative Standardized Normal table in the textbook (Appendix table E.2) gives the probability less than a desired value of Z (i.e., from negative infinity to Z)

Example:

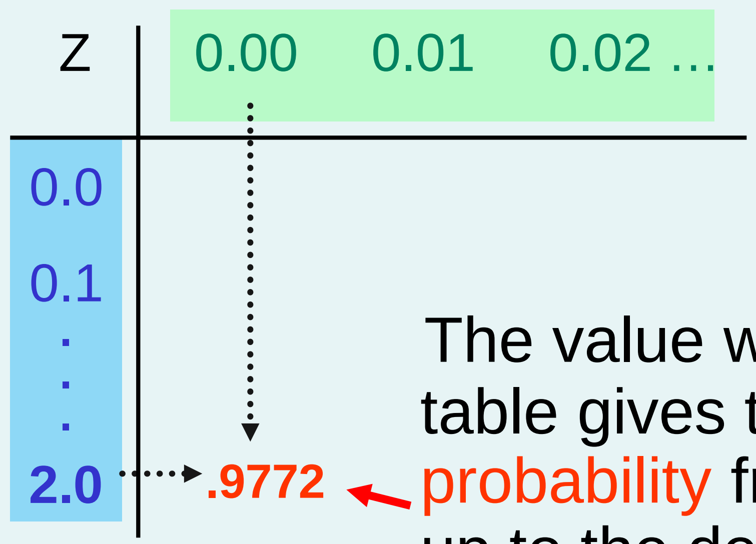
P(Z < 2.00) = 0.9772



The Standardized Normal Table

(continued)

The column gives the value of Z to the second decimal point



The row shows the value of Z to the first decimal point

The value within the table gives the probability from Z = – ∞ up to the desired Z value

P(Z < 2.00) = 0.9772

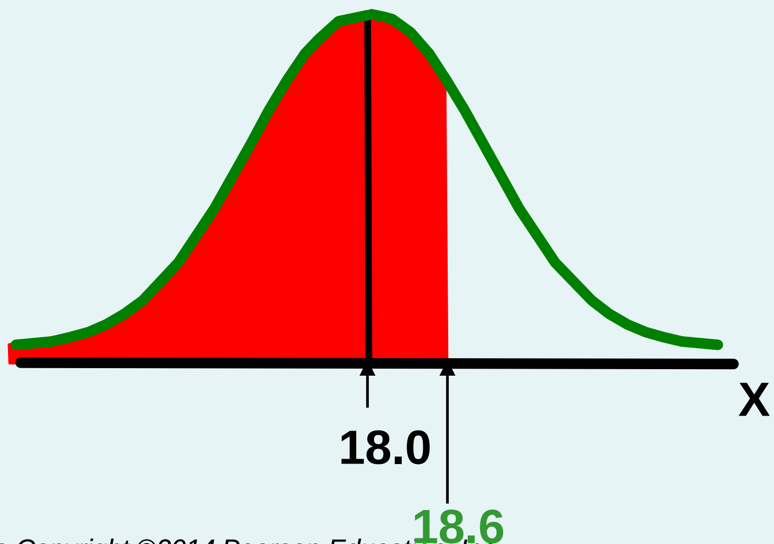
General Procedure for Finding Normal Probabilities

To find $P(a < X < b)$ when X is distributed normally:

- Draw the normal curve for the problem in terms of X
- Translate X-values to Z-values
- Use the Standardized Normal Table

Finding Normal Probabilities

- Let X represent the time it takes (in seconds) to download an image file from the internet.
- Suppose X is normal with a mean of18.0 seconds and a standard deviation of 5.0 seconds. Find P(X < 18.6)

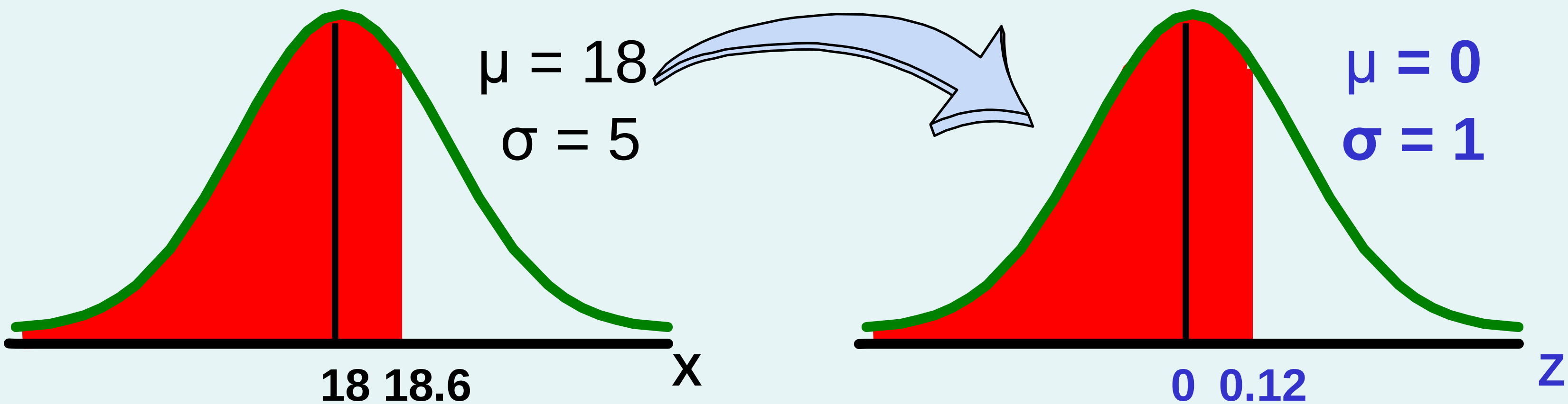


Finding Normal Probabilities

(continued)

- Let X represent the time it takes, in seconds to download an image file from the internet.
- Suppose X is normal with a mean of 18.0 seconds and a standard deviation of 5.0 seconds. Find P(X < 18.6)

$$Z = \frac{X - \mu}{\sigma} = \frac{18.6 - 18.0}{5.0} = 0.12$$



P(X < 18.6)

P(Z < 0.12)

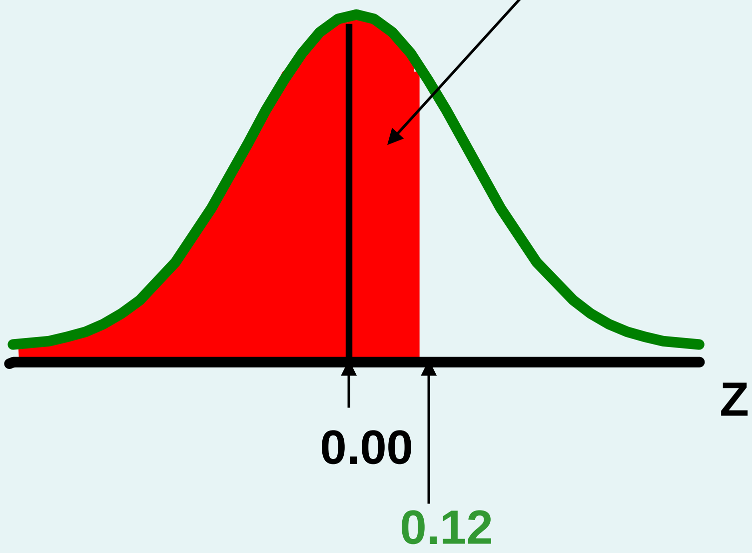
Solution: Finding P(Z < 0.12)

Standardized Normal Probability Table (Portion)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

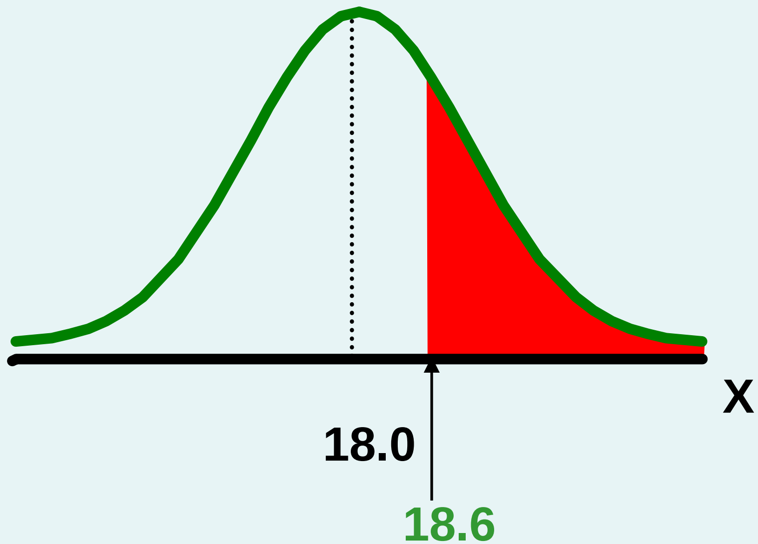
P(X < 18.6)
= P(Z < 0.12)

0.5478



Finding Normal Upper Tail Probabilities

- Suppose X is normal with mean 18.0 and standard deviation 5.0.
- Now Find P(X > 18.6)



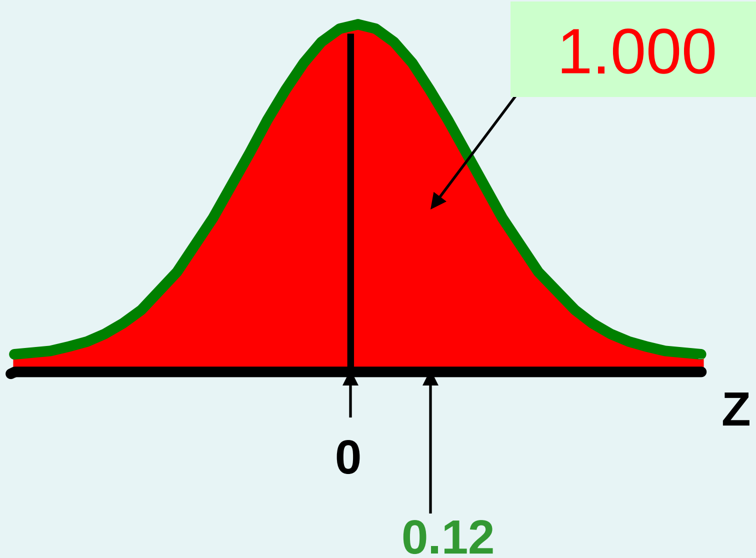
Finding Normal Upper Tail Probabilities

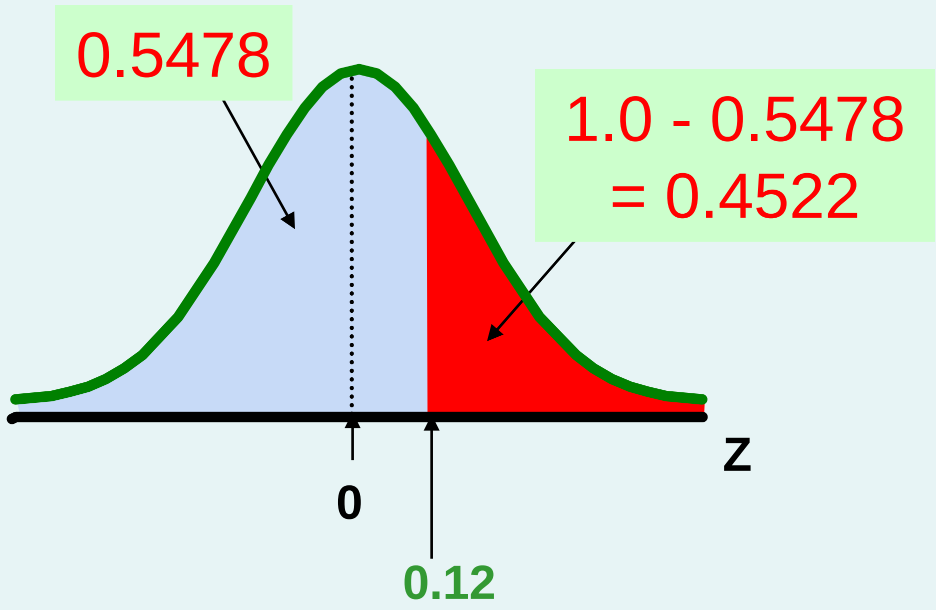
(continued)

- Now Find P(X > 18.6)…

$$P(X > 18.6) = P(Z > 0.12) = 1.0 - P(Z \le 0.12)$$

$$= 1.0 - 0.5478 = 0.4522$$





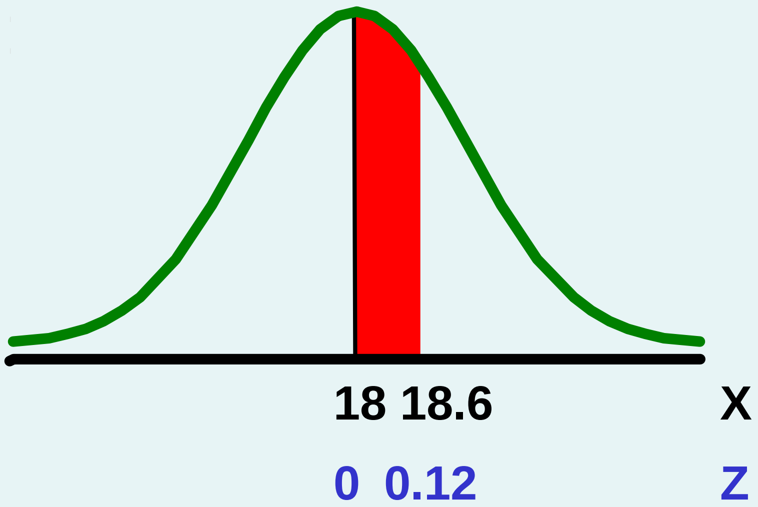
Finding a Normal Probability Between Two Values

- Suppose X is normal with mean 18.0 and standard deviation 5.0. Find P(18 < X < 18.6)

Calculate Z-values:

$$Z = \frac{X - \mu}{\sigma} = \frac{18 - 18}{5} = 0$$

$$Z = \frac{X - \mu}{\sigma} = \frac{18.6 - 18}{5} = 0.12$$



P(18 < X < 18.6)

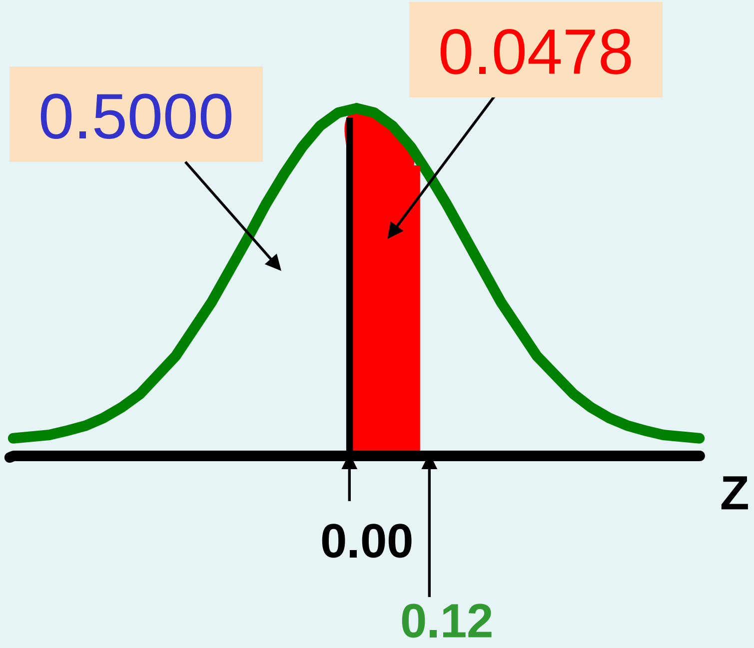
= P(0 < Z < 0.12)

Solution: Finding P(0 < Z < 0.12)

Standardized Normal Probability Table (Portion)

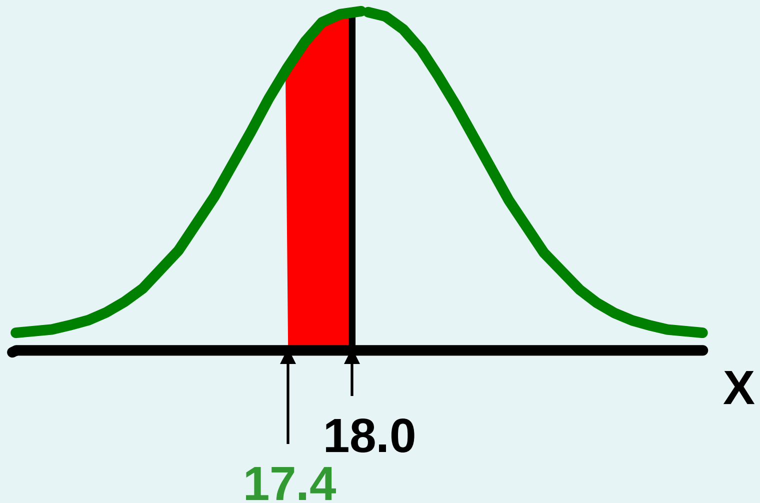
Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

P(18 < X < 18.6)
= P(0 < Z < 0.12)
= P(Z < 0.12) – P(Z ≤ 0)
= 0.5478 - 0.5000 = 0.0478



Probabilities in the Lower Tail

- Suppose X is normal with mean 18.0 and standard deviation 5.0.
- Now Find P(17.4 < X < 18)



Probabilities in the Lower Tail

(continued)

Now Find P(17.4 < X < 18)…

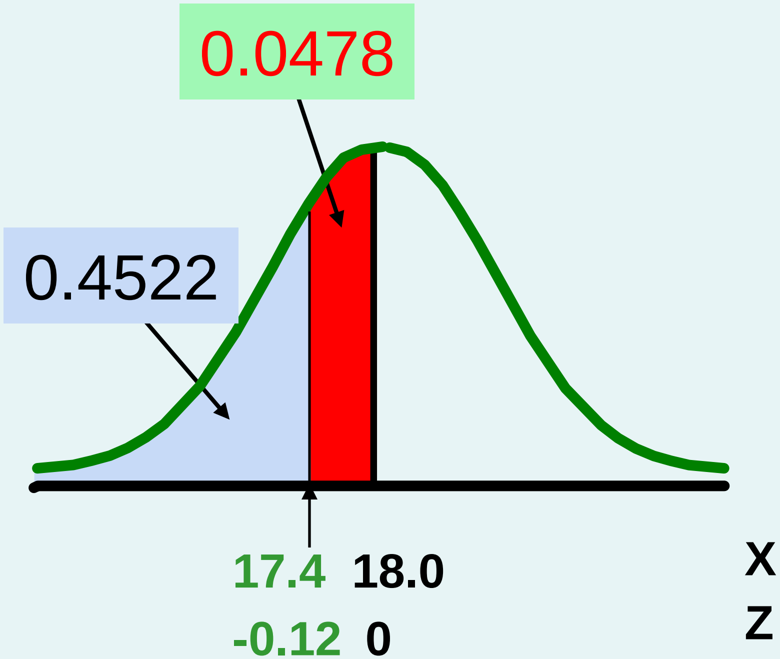
$P(17.4 < X < 18)$

$= P(-0.12 < Z < 0)$

$= P(Z < 0) - P(Z \leq -0.12)$

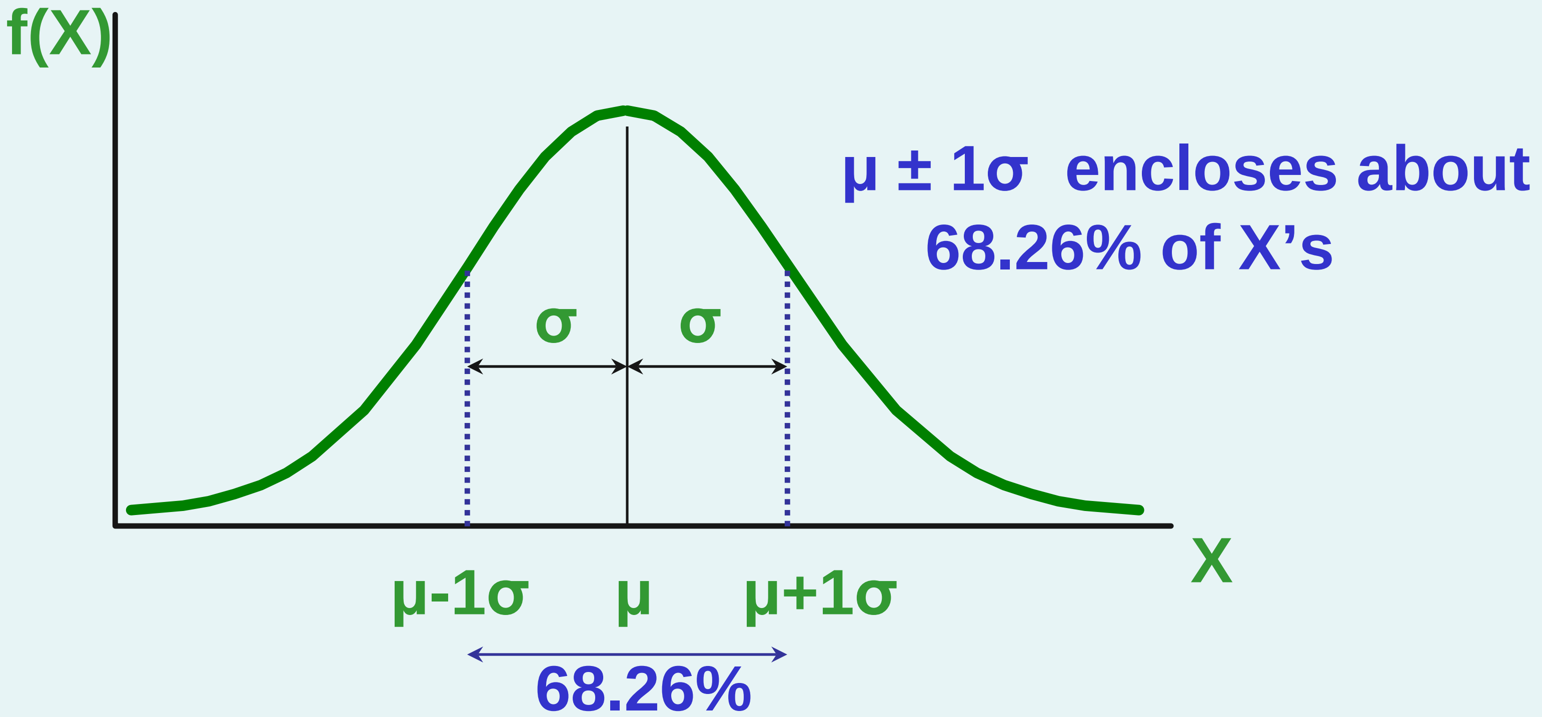
$= 0.5000 - 0.4522 = 0.0478$

The Normal distribution is symmetric, so this probability is the same as $P(0 < Z < 0.12)$



Empirical Rules

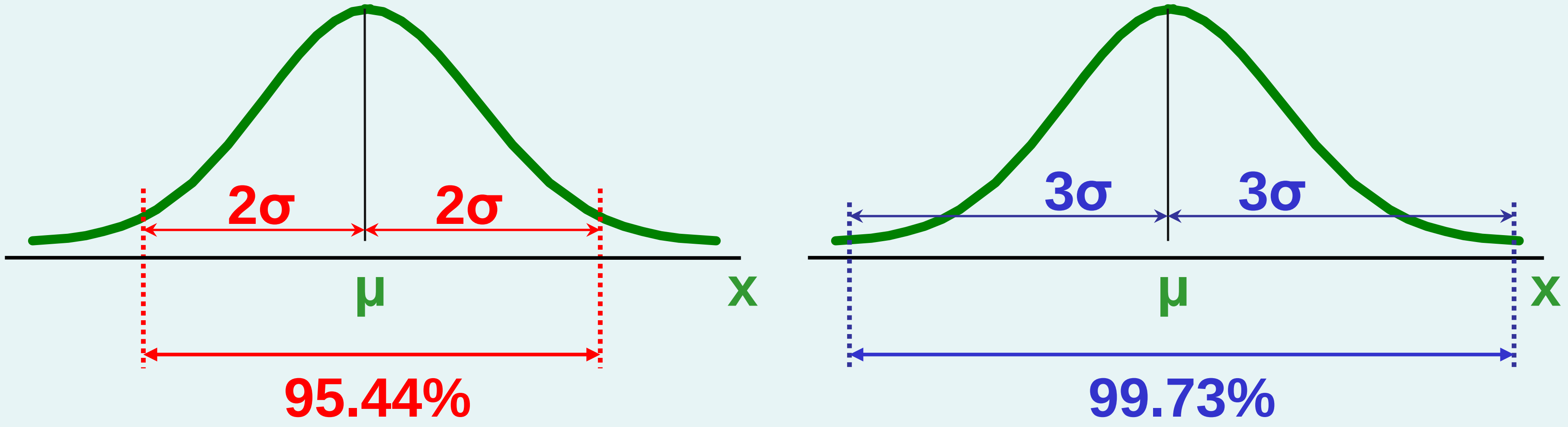
What can we say about the distribution of values around the mean? For any normal distribution:



The Empirical Rule

(continued)

- μ ± 2σ covers about 95% of X's
- μ ± 3σ covers about 99.7% of X's



Given a Normal Probability Find the X Value

- Steps to find the X value for a known probability:
  1. Find the Z value for the known probability
  2. Convert to X units using the formula:

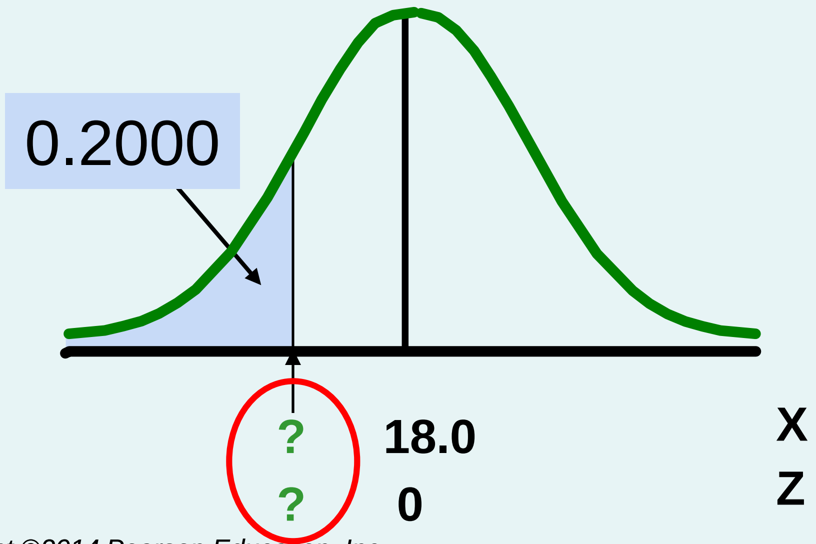
$$X = \mu + Z\sigma$$

Finding the X value for a Known Probability

(continued)

Example:

- Let X represent the time it takes (in seconds) to download an image file from the internet.
- Suppose X is normal with mean 18.0 and standard deviation 5.0
- Find X such that 20% of download times are less than X.



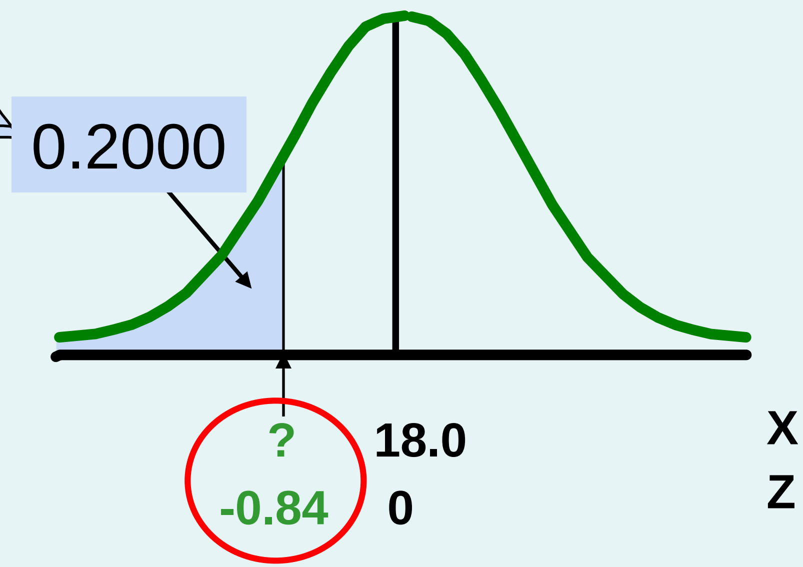
Find the Z value for 20% in the Lower Tail

1. Find the Z value for the known probability

Standardized Normal Probability Table (Portion)

Z	…	.03	.04	.05
-0.9	…	.1762	.1736	.1711
-0.8	…	.2033	.2005	.1977
-0.7	…	.2327	.2296	.2266

- 20% area in the lower tail is consistent with a Z value of -0.84



Finding the X value

2. Convert to X units using the formula:

$$
\begin{aligned}
X &= \mu + Z\sigma \\
&= 18.0 + (-0.84)5.0 \\
&= 13.8
\end{aligned}
$$

So 20% of the values from a distribution with mean 18.0 and standard deviation 5.0 are less than 13.80

Evaluating Normality

- Not all continuous distributions are normal
- It is important to evaluate how well the data set is approximated by a normal distribution.
- Normally distributed data should approximate the theoretical normal distribution:
  - The normal distribution is bell shaped (symmetrical) where the mean is equal to the median.
  - The empirical rule applies to the normal distribution.
  - The interquartile range of a normal distribution is 1.33 standard deviations.

Evaluating Normality

(continued)

Comparing data characteristics to theoretical properties

- Construct charts or graphs
  - For small- or moderate-sized data sets, construct a stem-and-leaf display or a boxplot to check for symmetry
  - For large data sets, does the histogram or polygon appear bell-shaped?
- Compute descriptive summary measures
  - Do the mean, median and mode have similar values?
  - Is the interquartile range approximately 1.33σ?
  - Is the range approximately 6σ?

Evaluating Normality

(continued)

Comparing data characteristics to theoretical properties

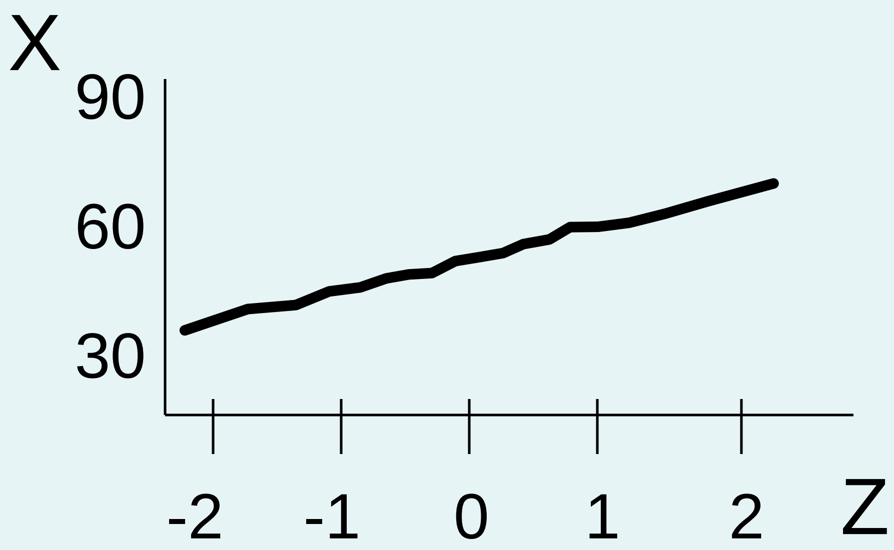
- Observe the distribution of the data set
  - Do approximately 2/3 of the observations lie within mean ±1 standard deviation?
  - Do approximately 80% of the observations lie within mean ±1.28 standard deviations?
  - Do approximately 95% of the observations lie within mean ±2 standard deviations?
- Evaluate normal probability plot
  - Is the normal probability plot approximately linear (i.e. a straight line) with positive slope?

Constructing A Normal Probability Plot

- Normal probability plot
  - Arrange data into ordered array
  - Find corresponding standardized normal quantile values (Z)
  - Plot the pairs of points with observed data values (X) on the vertical axis and the standardized normal quantile values (Z) on the horizontal axis
  - Evaluate the plot for evidence of linearity

The Normal Probability Plot Interpretation

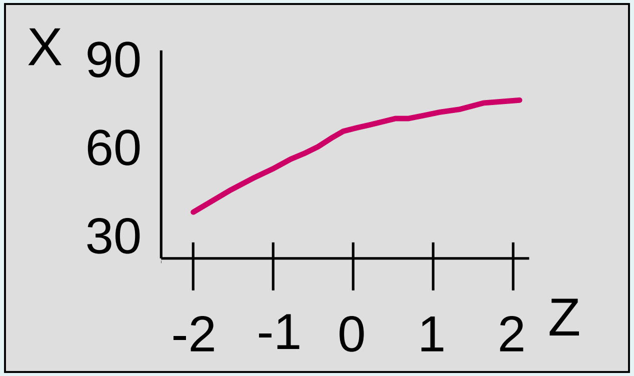
A normal probability plot for data from a normal distribution will be approximately linear:



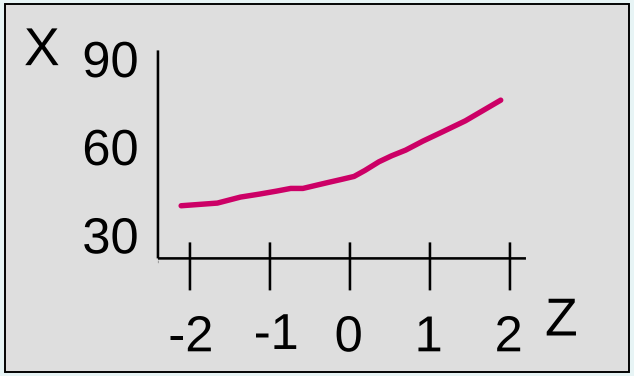
Normal Probability Plot Interpretation

(continued)

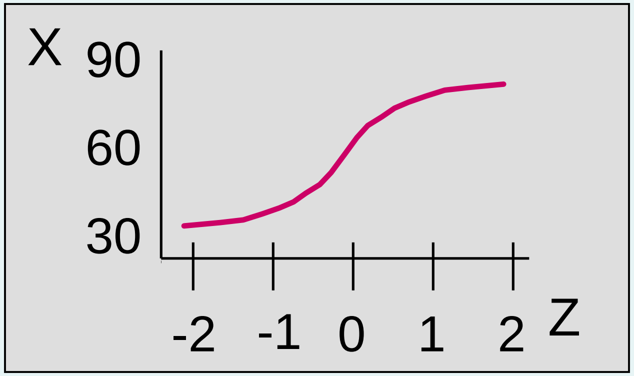
Left-Skewed



Right-Skewed



Rectangular

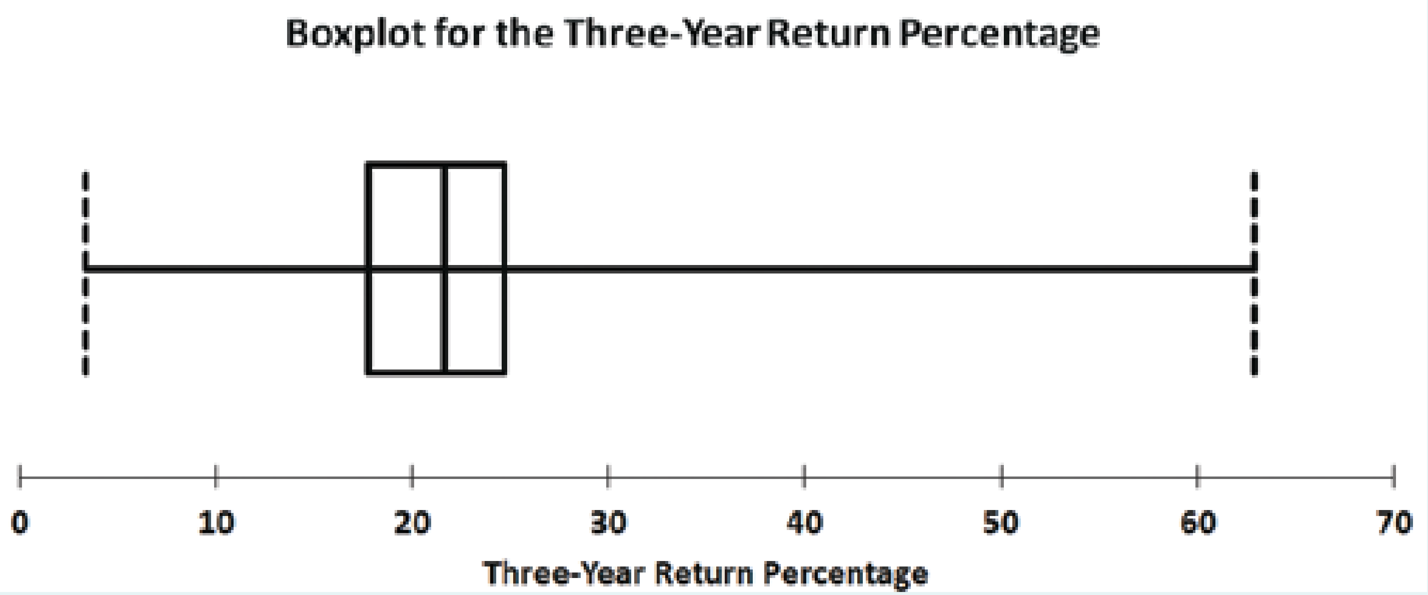


Nonlinear plots indicate a deviation from normality

Evaluating Normality
An Example: Bond Funds Returns

Five-Number Summary	
Minimum	3.39
First quartile	17.76
Median	21.65
Third quartile	24.74
Maximum	62.91

The boxplot is skewed to the right. (The normal distribution is symmetric.)



Evaluating Normality
An Example: Bond Funds Returns
(continued)

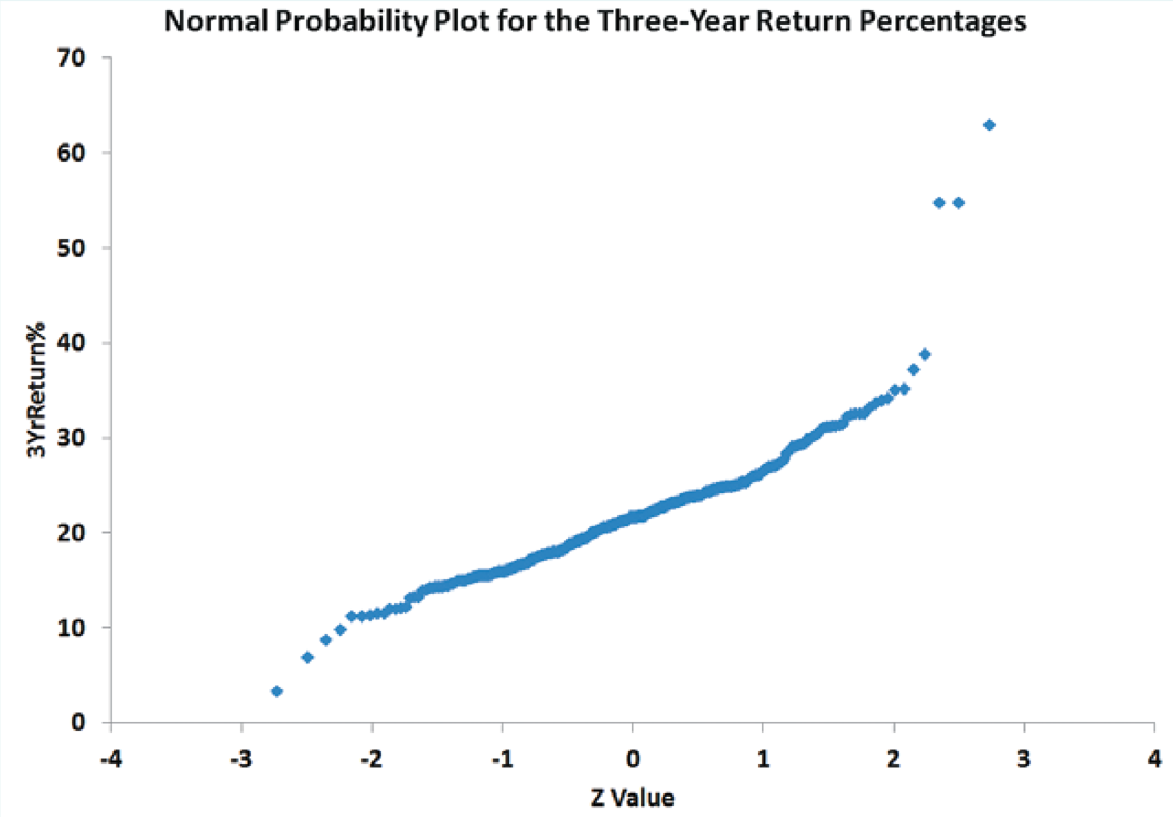
Descriptive Statistics

	3YrReturn%
Mean	21.84
Median	21.65
Mode	21.74
Minimum	3.39
Maximum	62.91
Range	59.52
Variance	41.2968
Standard Deviation	6.4263
Coeff. of Variation	29.43%
Skewness	1.6976
Kurtosis	8.4670
Count	318
Standard Error	0.3604

- The mean (21.84) is approximately the same as the median (21.65). (In a normal distribution the mean and median are equal.)
- The interquartile range of 6.98 is approximately 1.09 standard deviations. (In a normal distribution the interquartile range is 1.33 standard deviations.)
- The range of 59.52 is equal to 9.26 standard deviations. (In a normal distribution the range is 6 standard deviations.)
- 77.04% of the observations are within 1 standard deviation of the mean. (In a normal distribution this percentage is 68.26%.
- 86.79% of the observations are within 1.28 standard deviations of the mean. (In a normal distribution this percentage is 80%.)
- 96.86% of the observations are within 2 standard deviations of the mean. (In a normal distribution this percentage is 95.44%.)
- The skewness statistic is 1.698 and the kurtosis statistic is 8.467. (In a normal distribution, each of these statistics equals zero.)

Evaluating Normality
An Example: Bond Funds Returns

(continued)



Plot is not a straight line and shows the distribution is skewed to the right. (The normal distribution appears as a straight line.)

Evaluating Normality
An Example: Mutual Funds Returns

(continued)

- Conclusions
  - The returns are right-skewed
  - The returns have more values concentrated around the mean than expected
  - The range is larger than expected
  - Normal probability plot is not a straight line
  - Overall, this data set greatly differs from the theoretical properties of the normal distribution

The Uniform Distribution

- The uniform distribution is a probability distribution that has equal probabilities for all possible outcomes of the random variable

- Also called a rectangular distribution

The Uniform Distribution

(continued)

The Continuous Uniform Distribution:

$$f(X) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le X \le b \\ 0 & \text{otherwise} \end{cases}$$

where

f(X) = value of the density function at any X value

a = minimum value of X

b = maximum value of X

Properties of the Uniform Distribution

- The mean of a uniform distribution is

$$\mu = \frac{a+b}{2}$$

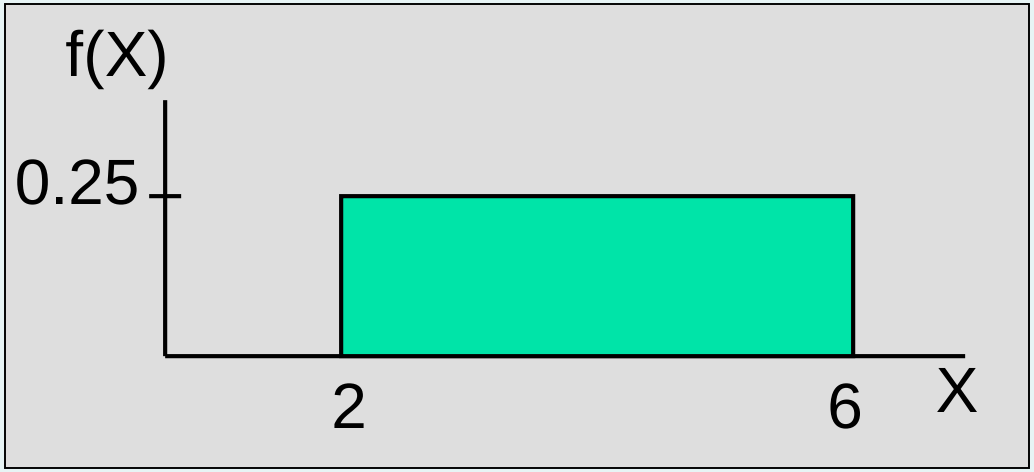
- The standard deviation is

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

Uniform Distribution Example

Example: Uniform probability distribution over the range $2 \le X \le 6$:

$$f(X) = \frac{1}{6 - 2} = 0.25 \quad \text{for } 2 \le X \le 6$$



$$\mu = \frac{a+b}{2} = \frac{2+6}{2} = 4$$

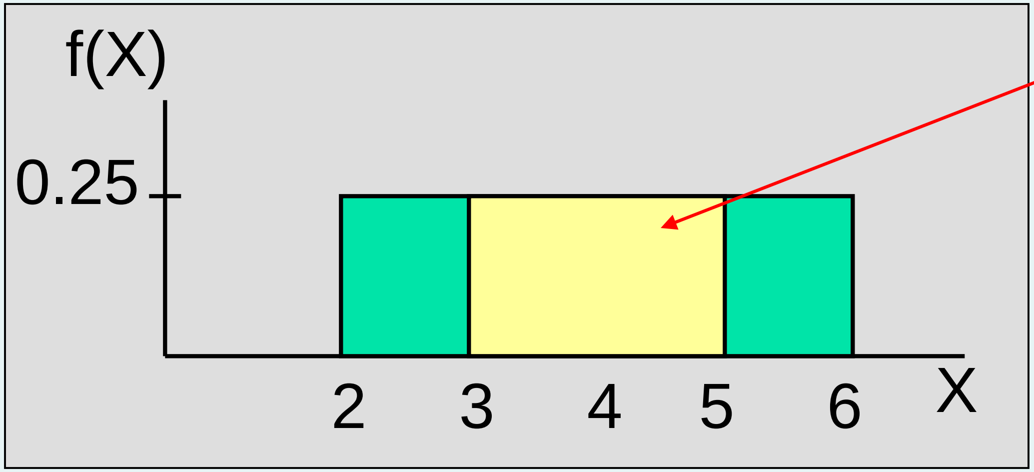
$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(6-2)^2}{12}} = 1.1547$$

Uniform Distribution Example

(continued)

Example: Using the uniform probability distribution to find $P(3 \le X \le 5)$:

$$P(3 \le X \le 5) = (\text{Base})(\text{Height}) = (2)(0.25) = 0.5$$



The Exponential Distribution

- Often used to model the length of time between two occurrences of an event (the time between arrivals)

  - Examples:
    - Time between trucks arriving at an unloading dock
    - Time between transactions at an ATM Machine
    - Time between phone calls to the main operator

The Exponential Distribution

- Defined by a single parameter, its mean λ (lambda)
- The probability that an arrival time is less than some specified time X is

$$P(\text{arrival time} < X) = 1 - e^{-\lambda X}$$

where e = mathematical constant approximated by 2.71828

λ = the population mean number of arrivals per unit

X = any value of the continuous variable where $0 < X < \infty$

Exponential Distribution Example

Example: Customers arrive at the service counter at the rate of 15 per hour. What is the probability that the arrival time between consecutive customers is less than three minutes?

- The mean number of arrivals per hour is 15, so $\lambda = 15$
- Three minutes is 0.05 hours
- P(arrival time < .05) $= 1 - e^{-\lambda X} = 1 - e^{-(15)(0.05)} =$ **0.5276**
- So there is a 52.76% chance that the arrival time between successive customers is less than three minutes

The Exponential Distribution In Excel

Calculating the probability that an exponential distribution with an mean of 20 is less than 0.1

	A	B
1	Exponential Probability	
2		
3	Data	
4	Mean	20
5	X Value	0.1
6		
7	Results	
8	P(<=X)	0.8647

=EXPON.DIST(B5, B4, TRUE)

Chapter Summary

In this chapter we discussed

- Key continuous distributions
  - normal, uniform, exponential
- Finding probabilities using formulas and tables
- Recognizing when to apply different distributions
- Applying distributions to decision problems

Statistics for Managers Using Microsoft Excel
7th Edition

On Line Topic

Normal Approximation To The Binomial

Learning Objectives

In this topic, you learn:

- When it is appropriate to use the normal distribution to approximate binomial probabilities
- How to use the normal distribution to approximate binomial probabilities

Normal Approximation to the Binomial Distribution

- The binomial distribution is a discrete distribution, but the normal is continuous
- To use the normal to approximate the binomial, accuracy is improved if you use a correction for continuity adjustment
- Example:
  - X is discrete in a binomial distribution, so $P(X = 4)$ can be approximated with a continuous normal distribution by finding

$$P(3.5 < X < 4.5)$$

Normal Approximation to the Binomial Distribution

(continued)

- The closer π is to 0.5, the better the normal approximation to the binomial
- The larger the sample size n, the better the normal approximation to the binomial
- General rule:
  - The normal distribution can be used to approximate the binomial distribution if

$$n\pi \geq 5$$

and

$$n(1 - \pi) \geq 5$$

Normal Approximation to the Binomial Distribution

(continued)

- The mean and standard deviation of the binomial distribution are

$$\mu = n\pi$$

$$\sigma = \sqrt{n\pi(1-\pi)}$$

- Transform binomial to normal using the formula:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

Using the Normal Approximation to the Binomial Distribution

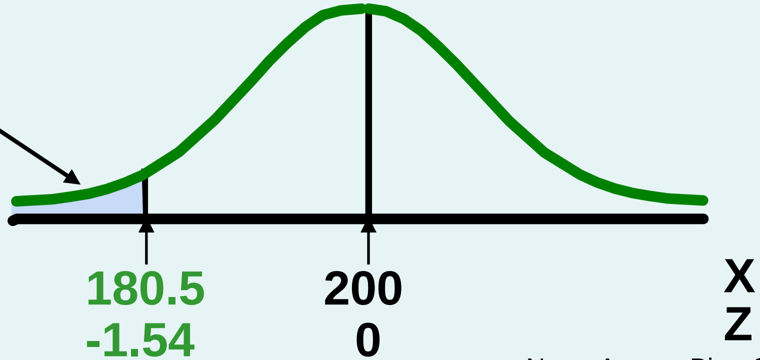
- If n = 1000 and π = 0.2, what is P(X ≤ 180)?
- Approximate P(X ≤ 180) using a continuity correction adjustment:

P(X ≤ 180.5)

- Transform to standardized normal:

$$Z = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}} = \frac{180.5 - (1000)(0.2)}{\sqrt{(1000)(0.2)(1-0.2)}} = -1.54$$

- So P(Z ≤ -1.54) = 0.0618



Topic Summary

In this topic we discussed:

- When it is appropriate to use the normal distribution to approximate binomial probabilities
- How to use the normal distribution to approximate binomial probabilities