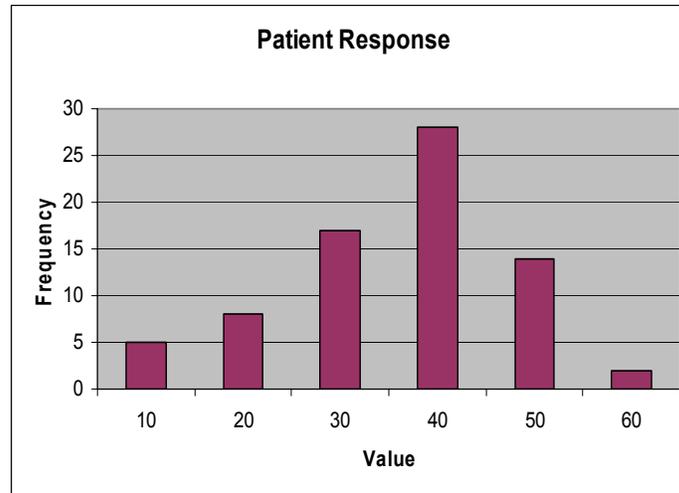

Chapter 4

Using Probability and Probability Distributions

Chapter Goals

- How to use models to make decisions

Why Model?



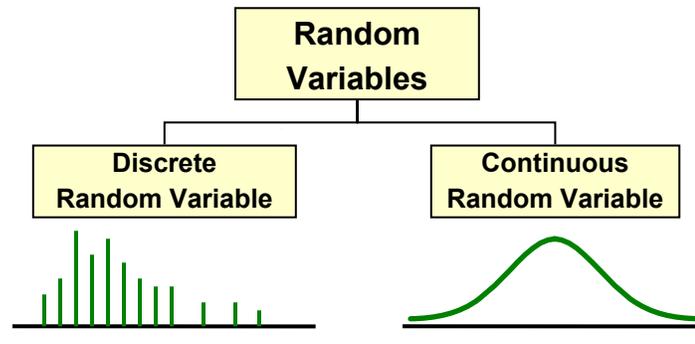
Example

Suppose we wish to compare two drugs, Drug A and Drug B, for relieving arthritis pain. Subjects suitable for the study are randomly assigned one of the two drugs. Results of the study are summarized in the following model for the time to relief for the two drugs.

Introduction to Probability Distributions

■ Random Variable

- Represents a possible numerical value from a random event



Discrete Random Variables

- Can only assume a countable number of values



Example:

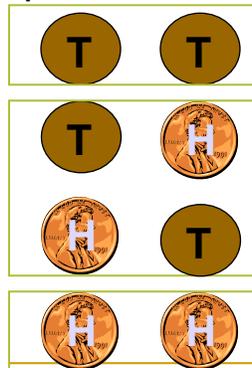
- **Roll a die twice. Let X be the random variable representing the number of times 4 comes up.**

Then, X takes can be

Discrete Probability Distribution

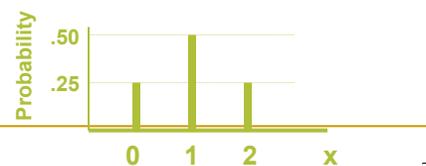
Experiment: Toss 2 Coins. Let $X = \#$ heads.

4 possible outcomes



Probability Distribution

| $X=x$ | $P(X=x)$ |
|-------|----------|
| 0 | |
| 1 | |
| 2 | |



Fall 2006 – Fundamentals of Business Statistics

7

Discrete Probability Distribution

The **probability mass function** of a discrete variable is a graph, table, or formula that specifies the proportion associated with each possible value the variable can take. The mass function $p(X = x)$ (or just $p(x)$) has the following properties:

- All values of the discrete function $p(x)$ must be between 0 and 1, both inclusive, and
- if you add up all values, they should sum to 1.

Fall 2006 – Fundamentals of Business Statistics

8

Example

- Let X represent the number of books in a backpack for students enrolled at KSU. The probability mass function for X is given below:

| | | | |
|----------|-----|-----|-----|
| X | 2 | 3 | 4 |
| $P(X=x)$ | 0.5 | 0.3 | 0.2 |

YDI 6.11

Consider the following discrete mass function:

| | | | | |
|----------|-----|---|-----|-----|
| X | -1 | 2 | 3 | 4 |
| $P(X=x)$ | 0.1 | | 0.6 | 0.2 |

- Complete the table
- $P(X < 3.1) =$
- $P(X \geq -1.1) =$
- $P(2 < X < 3) =$

Discrete Random Variable Summary Measures

- **Expected Value** of a discrete distribution (Average)

$$E(X) = \sum x_i P(x_i)$$

- **Example:** Toss 2 coins,
 $X = \# \text{ of heads}$

| x | P(x) |
|---|------|
| 0 | .25 |
| 1 | .50 |
| 2 | .25 |

Discrete Random Variable Summary Measures

(continued)

- **Standard Deviation** of a discrete distribution

$$\sigma_x = \sqrt{\sum \{x - E(x)\}^2 P(x)}$$

where:

$E(X)$ = Expected value of the random variable

Discrete Random Variable

Summary Measures

- **Example:** Toss 2 coins, $x = \#$ heads, *(continued)*
compute standard deviation (recall $E(x) = 1$)

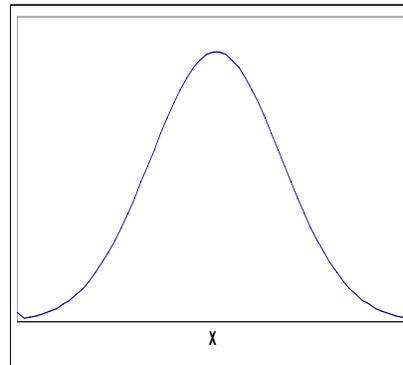
$$\sigma_x = \sqrt{\sum \{x - E(x)\}^2 P(x)}$$

$$\sigma_x = \sqrt{(0-1)^2(.25) + (1-1)^2(.50) + (2-1)^2(.25)} = \sqrt{.50} = .707$$

Possible number of heads
= 0, 1, or 2

Continuous Variables

A density function is a (nonnegative) function or curve that describes the overall shape of a distribution. The total area under the entire curve is equal to **one**, and proportions are measured as areas under the density function.



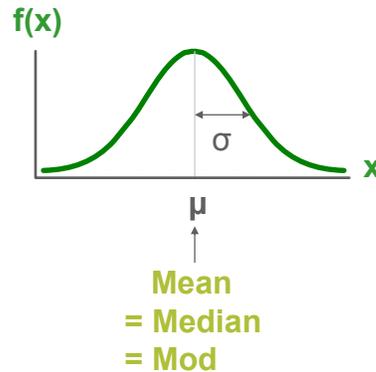
The Normal Distribution

■ $X \sim N(\mu, \sigma^2)$ means that the variable X is normally distributed with mean μ and variance σ^2 (or standard deviation σ).

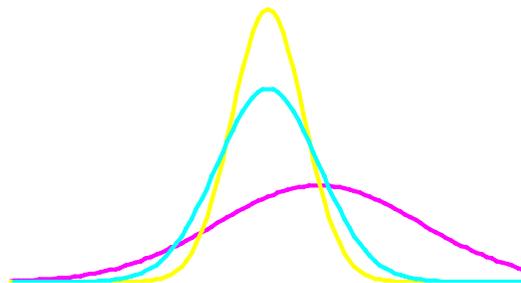
■ If $X \sim N(\mu, \sigma^2)$, then the standardized normal variable $Z = (X - \mu) / \sigma \sim N(0, 1)$. Z is called the **standard normal**.

The random variable has an infinite theoretical range:

$+\infty$ to $-\infty$



Many Normal Distributions



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

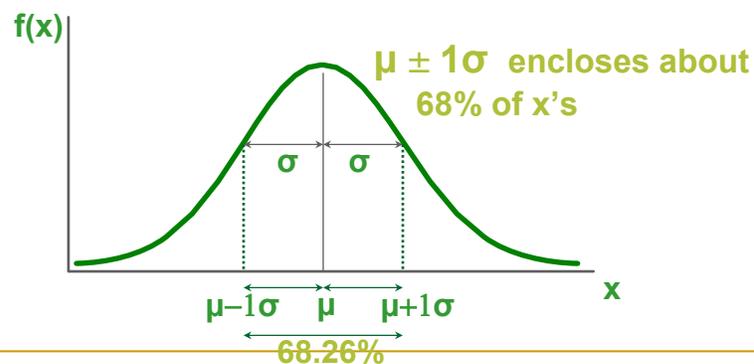
Properties of the Normal Distribution

- Symmetric about the mean μ .
- Bell-shaped
- Mean = Median = Mode
- Approximately 68% of the area under the curve is within ± 1 standard deviation of the mean.
- Approximately 95% of the area under the curve is within ± 2 standard deviation of the mean.
- Approximately 99.7% of the area under the curve is within ± 3 standard deviation of the mean.

Note: Any normal distribution $N(\mu, \sigma^2)$ can be transformed to a standard normal distribution $N(0,1)$

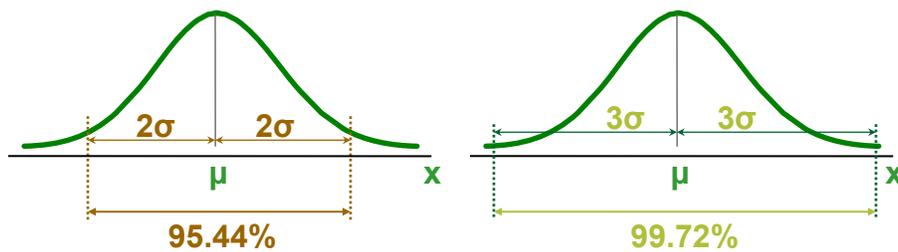
Empirical Rules

What can we say about the distribution of values around the mean? There are some general rules:



The Empirical Rule

- $\mu \pm 2\sigma$ covers about **95%** of x 's *(continued)*
- $\mu \pm 3\sigma$ covers about **99.7%** of x 's



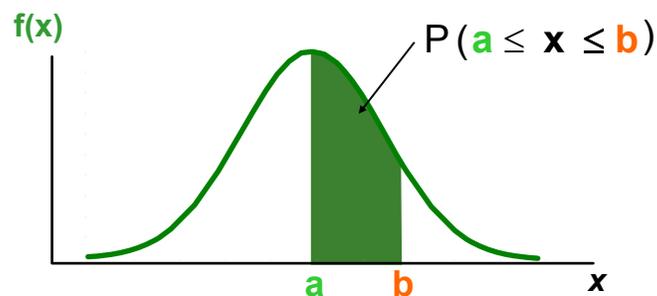
Translation to the Standard Normal Distribution

- Translate from x to the standard normal (the “ z ” distribution) by **subtracting the mean** of x and **dividing by its standard deviation**:

$$z = \frac{x - \mu}{\sigma}$$

Finding Normal Probabilities

Probability is measured by the area under the curve



Example

Let the variable X represent IQ scores of 12-year-olds. Suppose $X \sim N(100, 256)$. Jessica is a 12-year-old and has an IQ score of 132. What proportion of 12-year-olds have IQ scores less than Jessica's score of 132?

YDI 6.1

- Find the area under a standard normal distribution between $z = 0$ and $z = 1.22$.
- Find the area under a standard normal distribution to the left of $z = -2.55$.
- Find the area under a standard normal distribution between $z = -1.22$ and $z = 1.22$.

YDI 6.2

Consider the previous IQ Scores example, where $X \sim N(100, 256)$.

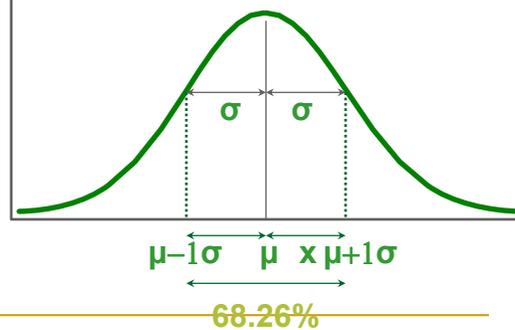
- What proportion of the 12-year-olds have IQ scores below 84?
- What proportion of the 12-year-olds have IQ scores 84 or more?
- What proportion of the 12-year-olds have IQ scores between 84 and 116?

Empirical Rules

$\mu \pm 1\sigma$ covers about 68% of x's

$\mu \pm 2\sigma$ covers about 95% of x's

$\mu \pm 3\sigma$ covers about 99.7% of x's



Example

Suppose cholesterol measures for healthy individuals have a normal distribution. Kyle's standardized cholesterol measure was $z = -2$. Using the 68-95-99 rule, what percentile does Kyle's measure represent?

Lee's standardized cholesterol measure was $z = 3.2$. Does Lee's cholesterol seem unusually high?

YDI 6.4

Different species of pine trees are grown at a Christmas-tree farm. It is known that the length of needles on a Species A pine tree follows a normal distribution. About 68% of such needles have lengths centered around the mean between 5.9 and 6.9 inches.

1. What are the mean and standard deviation of the model for Species A pine needle lengths?
2. A 5.2-inch pine needle is found that looks like a Species A needle but is somewhat shorter than expected. Is it likely that this needle is from a Species A pine tree?

YDI 6.6

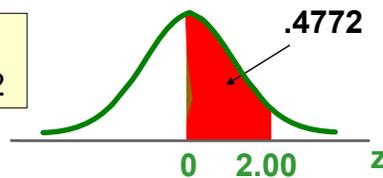
The finishing times for swimmers performing the 100-meter butterfly are normally distributed with a mean of 55 seconds and a standard deviation of 5 seconds.

1. The sponsors decide to give certificates to all those swimmers who finish in under 49 seconds. If there are 50 swimmers entered in the 100-meter butterfly, approximately how many certificates will be needed?
2. What time must a swimmer finish to be in the top fastest 2% of the distribution of finishing times?

The Standard Normal Table

- The Standard Normal table in the textbook (Appendix D) gives the probability from the mean (zero) up to a desired value for z

Example:
 $P(0 < z < 2.00) = .4772$



The Standard Normal Table

The **column** gives the value of z to the second decimal point *(continued)*

The **row** shows the value of z to the first decimal point

| z | 0.00 | 0.01 | 0.02 | ... |
|-----|-------|------|------|-----|
| 0.1 | | | | |
| 0.2 | | | | |
| ⋮ | | | | |
| ⋮ | | | | |
| 2.0 | .4772 | | | |

The value within the table gives the **probability** from $z = 0$ up to the desired z value

$$P(0 < z < 2.00) = .4772$$

General Procedure for Finding Probabilities

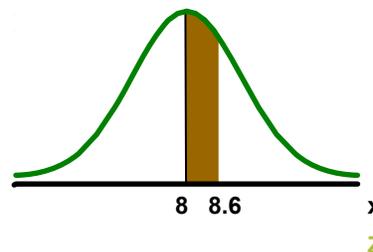
To find $P(a < x < b)$ when x is distributed normally:

- Draw the normal curve for the problem in terms of x
- Translate x -values to z -values
- Use the Standard Normal Table

Z Table example

- Suppose x is normal with mean 8.0 and standard deviation 5.0. Find $P(8 < x < 8.6)$

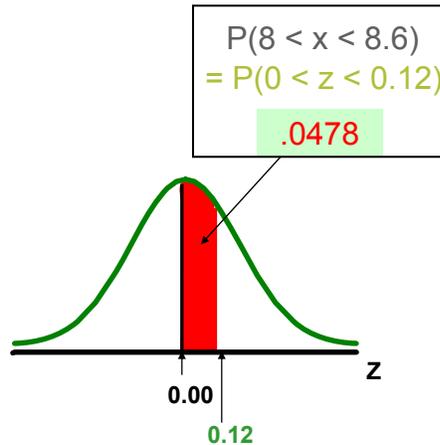
Calculate z -values:



Solution: Finding $P(0 < z < 0.12)$

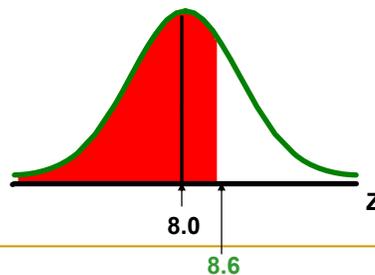
Standard Normal Probability
Table (Portion)

| z | .00 | .01 | .02 |
|-----|-------|-------|-------|
| 0.0 | .0000 | .0040 | .0080 |
| 0.1 | .0398 | .0438 | .0478 |
| 0.2 | .0793 | .0832 | .0871 |
| 0.3 | .1179 | .1217 | .1255 |



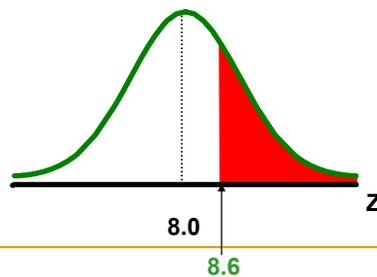
Finding Normal Probabilities

- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(x < 8.6)$



Upper Tail Probabilities

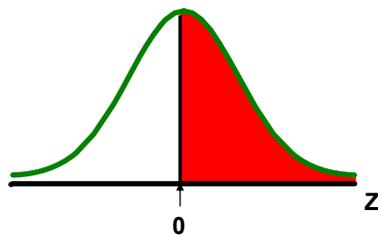
- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(x > 8.6)$



Upper Tail Probabilities

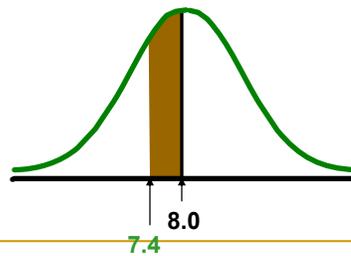
(continued)

- Now Find $P(x > 8.6)$...



Lower Tail Probabilities

- Suppose x is normal with mean 8.0 and standard deviation 5.0.
- Now Find $P(7.4 < x < 8)$



Quiz Problems

- The distribution of resistance for a certain type of resistors is known to be normal; 2.5% of these resistors have a resistance exceeding 15 ohms and 2.5% have resistance smaller than 8 ohms. What is the standard deviation of this resistance distribution?

Quiz Problems

- For a population under study, the variable weight follows a normal distribution. If the variance in weights is 100 and the 20th percentile corresponds to 114 pounds, what is the mean weight?