Business Statistics: A Decision-Making Approach

6th Edition

Chapter 4

Using Probability and Probability Distributions

Chapter Goals

After completing this chapter, you should be able to:

- Explain three approaches to assessing probabilities
- Apply common rules of probability
- Use Bayes' Theorem for conditional probabilities
- Distinguish between discrete and continuous probability distributions
- Compute the expected value and standard deviation for a discrete probability distribution

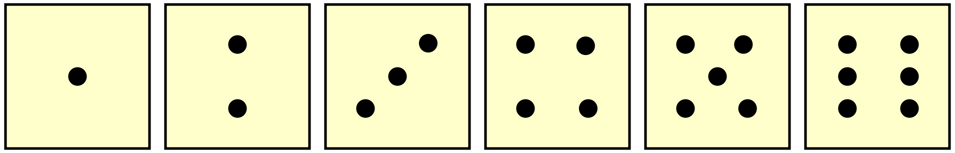
Important Terms

- Probability – the chance that an uncertain event will occur (always between 0 and 1)
- Experiment – a process of obtaining outcomes for uncertain events
- Elementary Event – the most basic outcome possible from a simple experiment
- Sample Space – the collection of all possible elementary outcomes

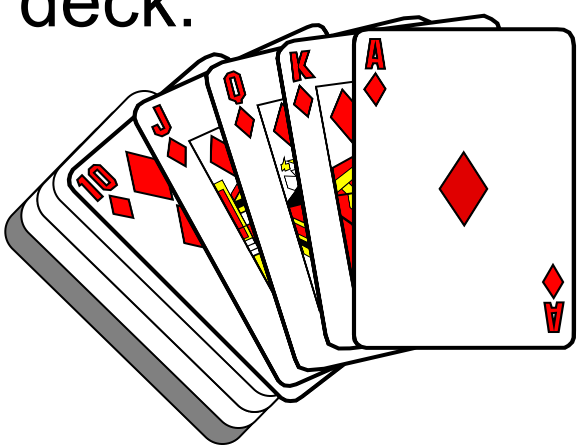
Sample Space

The Sample Space is the collection of all possible outcomes

e.g. All 6 faces of a die:



e.g. All 52 cards of a bridge deck:



Events

- Elementary event – An outcome from a sample space with one characteristic
  - Example: A red card from a deck of cards
- Event – May involve two or more outcomes simultaneously
  - Example: An ace that is also red from a deck of cards

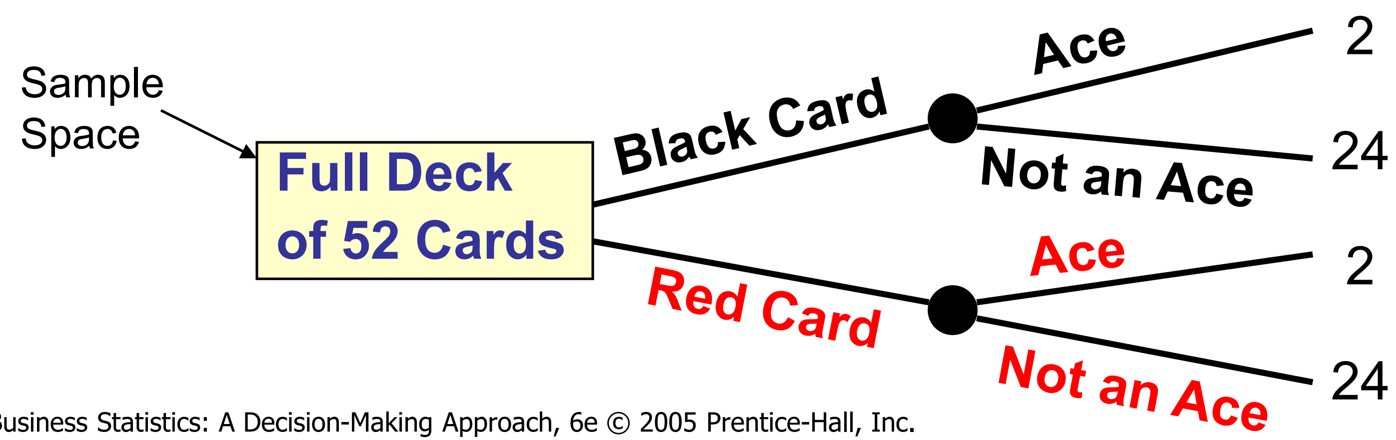
Visualizing Events

- Contingency Tables

	Ace	Not Ace	Total
Black	2	24	26
Red	2	24	26
Total	4	48	52

Sample Space

- Tree Diagrams



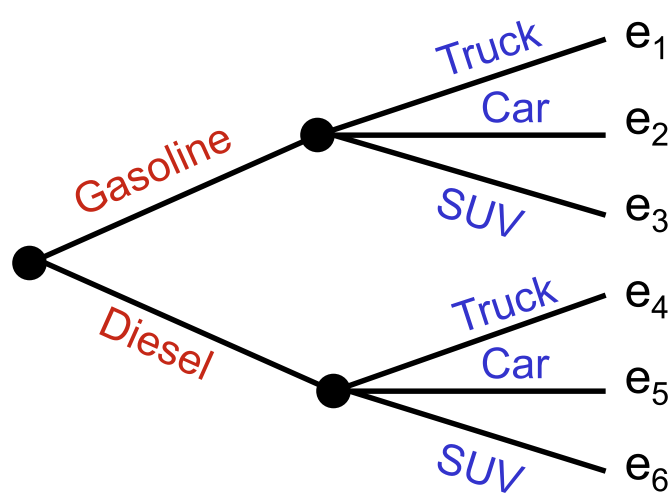
Elementary Events

- A automobile consultant records fuel type and vehicle type for a sample of vehicles

  2 Fuel types: Gasoline, Diesel

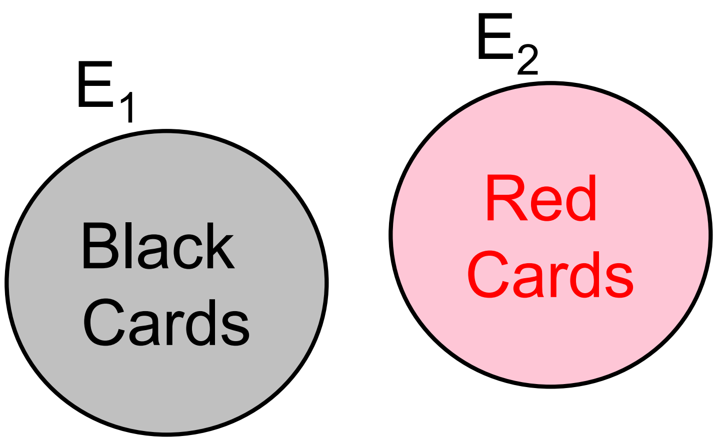
  3 Vehicle types: Truck, Car, SUV

6 possible elementary events:	
e_1	Gasoline, Truck
e_2	Gasoline, Car
e_3	Gasoline, SUV
e_4	Diesel, Truck
e_5	Diesel, Car
e_6	Diesel, SUV



Probability Concepts

- **Mutually Exclusive Events**
  - If E_1 occurs, then E_2 cannot occur
  - E_1 and E_2 have no common elements



A card cannot be Black and Red at the same time.

Probability Concepts

- **Independent and Dependent Events**
  - Independent: Occurrence of one does not influence the probability of occurrence of the other
  - Dependent: Occurrence of one affects the probability of the other

Independent vs. Dependent Events

- **Independent Events**

E_1 = heads on one flip of fair coin

E_2 = heads on second flip of same coin

Result of second flip does not depend on the result of the first flip.

- **Dependent Events**

E_1 = rain forecasted on the news

E_2 = take umbrella to work

Probability of the second event is affected by the occurrence of the first event

Assigning Probability

- Classical Probability Assessment

$$P(E_i) = \frac{\text{Number of ways } E_i \text{ can occur}}{\text{Total number of elementary events}}$$

- Relative Frequency of Occurrence

$$\text{Relative Freq. of } E_i = \frac{\text{Number of times } E_i \text{ occurs}}{N}$$

- Subjective Probability Assessment

An opinion or judgment by a decision maker about the likelihood of an event

Rules of Probability

Rules for Possible Values and Sum

- **Individual Values**

  $0 \leq P(e_i) \leq 1$

  For any event e_i

- **Sum of All Values**

  $$\sum_{i=1}^{k} P(e_i) = 1$$

where:

k = Number of elementary events in the sample space

e_i = i^{th} elementary event

Addition Rule for Elementary Events

- The probability of an event E_i is equal to the sum of the probabilities of the elementary events forming E_i.
- That is, if:

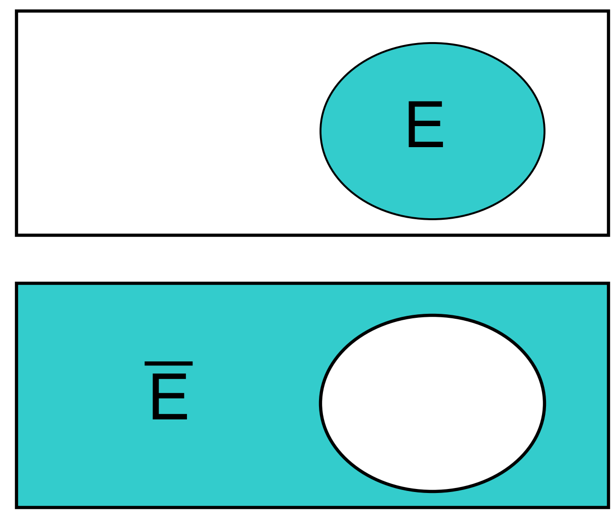
$$E_i = \{e_1, e_2, e_3\}$$

then:

$$P(E_i) = P(e_1) + P(e_2) + P(e_3)$$

Complement Rule

- The complement of an event E is the collection of all possible elementary events **not** contained in event E. The complement of event E is represented by $\overline{E}$.



- Complement Rule:

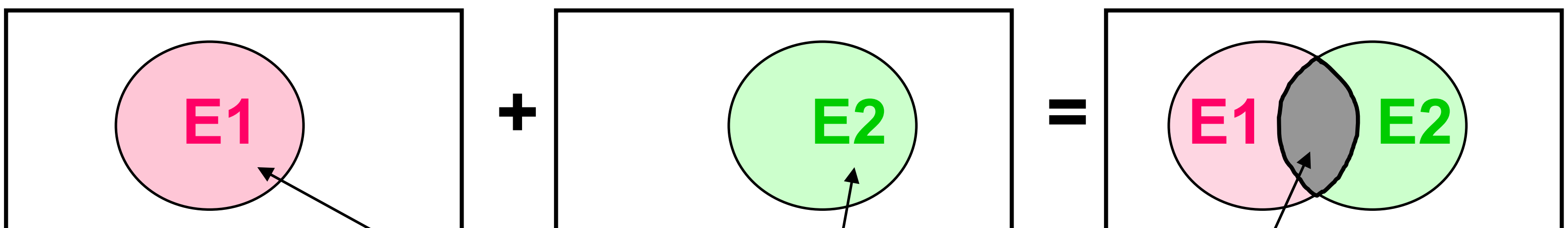
$$P(\overline{E}) = 1 - P(E)$$

Or, $$P(E) + P(\overline{E}) = 1$$

Addition Rule for Two Events

- **Addition Rule:**

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$



$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

Don't count common elements twice!

Addition Rule Example

P(Red or Ace) = P(Red) +P(Ace) - P(Red and Ace)

= 26/52 + 4/52 - 2/52 = 28/52

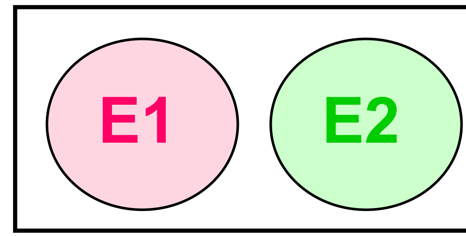
Don't count the two red aces twice!

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Addition Rule for Mutually Exclusive Events

- If E1 and E2 are mutually exclusive, then

P(E1 and E2) = 0



So

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

= 0 if mutually exclusive

$$= P(E_1) + P(E_2)$$

Conditional Probability

- Conditional probability for any two events E_1 , E_2:

$$P(E_1 | E_2) = \frac{P(E_1 \text{ and } E_2)}{P(E_2)}$$

where $P(E_2) > 0$

Conditional Probability Example

- Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both.

- What is the probability that a car has a CD player, given that it has AC ?

i.e., we want to find P(CD | AC)

Conditional Probability Example

(continued)

- Of the cars on a used car lot, **70%** have air conditioning (AC) and **40%** have a CD player (CD). **20%** of the cars have both.

	CD	No CD	Total
AC	.2	.5	.7
No AC	.2	.1	.3
Total	.4	.6	1.0

$$P(CD \mid AC) = \frac{P(CD \text{ and } AC)}{P(AC)} = \frac{.2}{.7} = .2857$$

Conditional Probability Example

(continued)

- Given AC, we only consider the top row (70% of the cars). Of these, 20% have a CD player. 20% of 70% is about 28.57%.

	CD	No CD	Total
AC	.2	.5	.7
No AC	.2	.1	.3
Total	.4	.6	1.0

$$P(CD\,|\,AC) = \frac{P(CD \text{ and } AC)}{P(AC)} = \frac{.2}{.7} = .2857$$

For Independent Events:

- Conditional probability for independent events E_1 , E_2:

$$P(E_1 \mid E_2) = P(E_1)$$

where $P(E_2) > 0$

$$P(E_2 \mid E_1) = P(E_2)$$

where $P(E_1) > 0$

Multiplication Rules

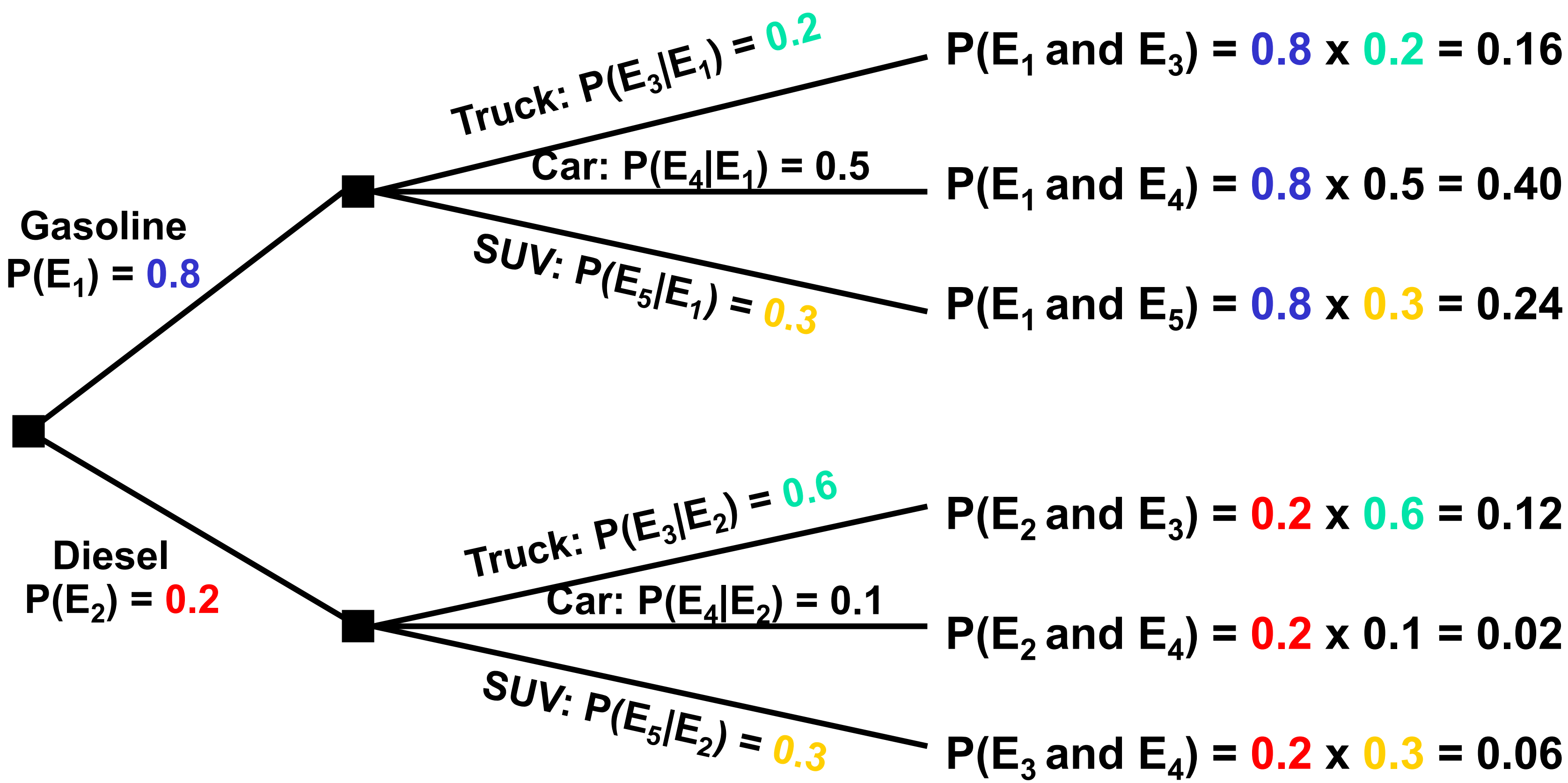
- Multiplication rule for two events E_1 and E_2:

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2 | E_1)$$

Note: If E_1 and E_2 are independent, then $P(E_2 | E_1) = P(E_2)$ and the multiplication rule simplifies to

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2)$$

Tree Diagram Example



Bayes' Theorem

$$P(E_i \mid B) = \frac{P(E_i)P(B \mid E_i)}{P(E_1)P(B \mid E_1) + P(E_2)P(B \mid E_2) + \ldots + P(E_k)P(B \mid E_k)}$$

- where:

  E_i = i^{th} event of interest of the k possible events

  B = new event that might impact $P(E_i)$

  Events E_1 to E_k are mutually exclusive and collectively exhaustive

Bayes' Theorem Example

- A drilling company has estimated a 40% chance of striking oil for their new well.
- A detailed test has been scheduled for more information. Historically, 60% of successful wells have had detailed tests, and 20% of unsuccessful wells have had detailed tests.
- Given that this well has been scheduled for a detailed test, what is the probability that the well will be successful?

Bayes' Theorem Example

(continued)

- Let S = successful well and U = unsuccessful well
- P(S) = .4 , P(U) = .6 (prior probabilities)
- Define the detailed test event as D
- Conditional probabilities:

$$P(D|S) = .6 \qquad P(D|U) = .2$$

- Revised probabilities

Event	Prior Prob.	Conditional Prob.	Joint Prob.	Revised Prob.
S (successful)	.4	.6	.4*.6 = .24	.24/.36 = .67
U (unsuccessful)	.6	.2	.6*.2 = .12	.12/.36 = .33

Sum = .36

Bayes’ Theorem Example

(continued)

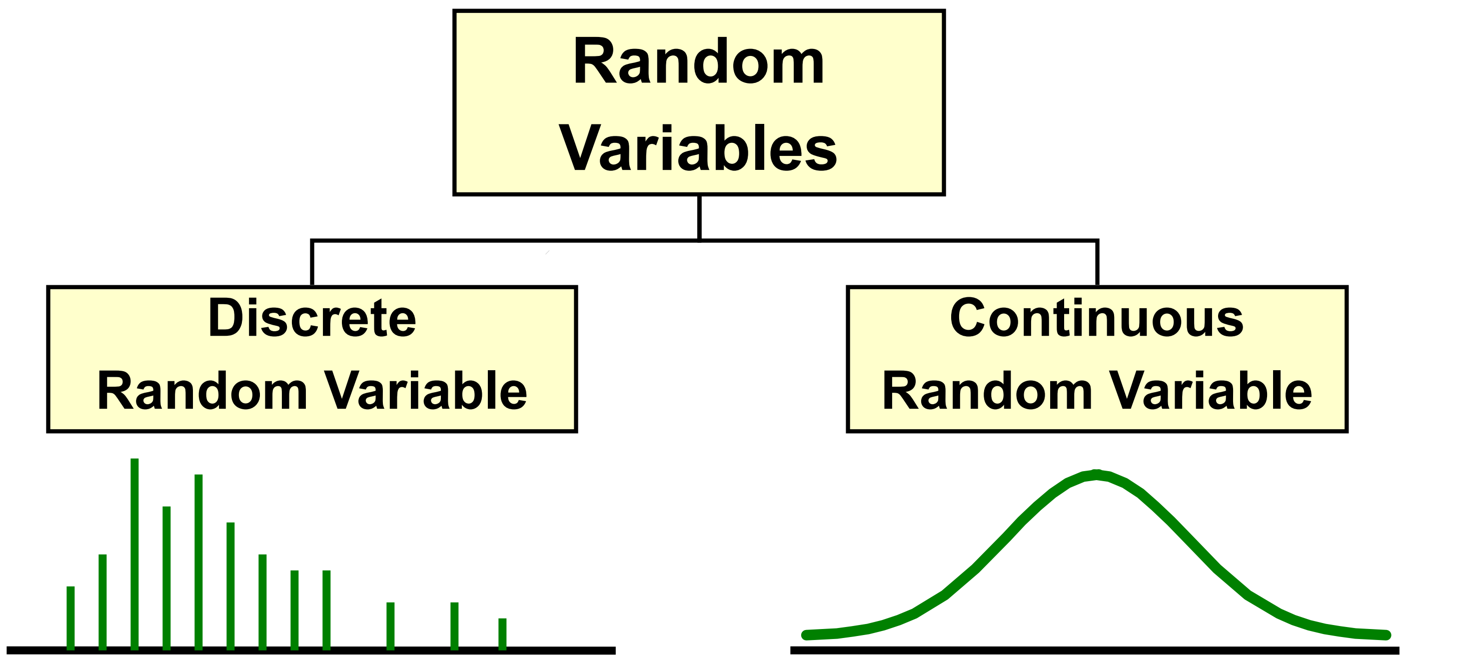
- Given the detailed test, the revised probability of a successful well has risen to .67 from the original estimate of .4

Event	Prior Prob.	Conditional Prob.	Joint Prob.	Revised Prob.
S (successful)	.4	.6	.4*.6 = .24	.24/.36 = .67
U (unsuccessful)	.6	.2	.6*.2 = .12	.12/.36 = .33

Sum = .36

Introduction to Probability Distributions

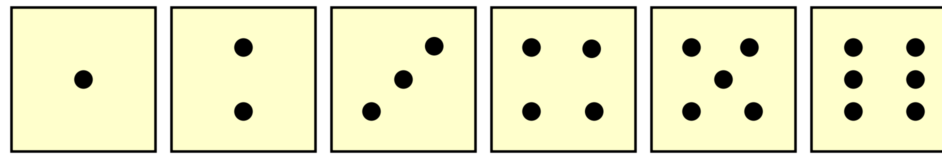
- **Random Variable**
  - Represents a possible numerical value from a random event



Discrete Random Variables

- Can only assume a countable number of values

Examples:



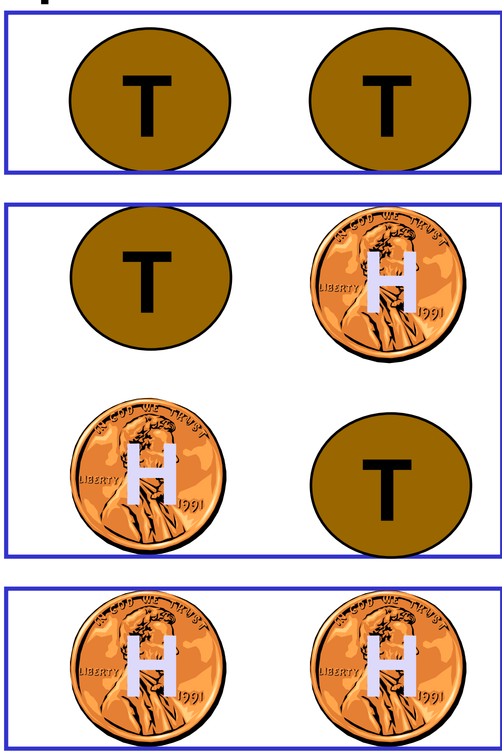
- **Roll a die twice**
  Let x be the number of times 4 comes up
  (then x could be 0, 1, or 2 times)

- **Toss a coin 5 times.**
  Let x be the number of heads
  (then x = 0, 1, 2, 3, 4, or 5)

Discrete Probability Distribution

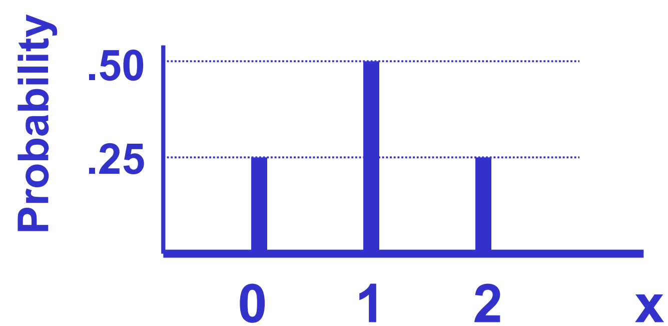
Experiment: Toss 2 Coins. Let x = # heads.

4 possible outcomes



Probability Distribution

x Value	Probability
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25



Discrete Probability Distribution

- A list of all possible [x_i , $P(x_i)$] pairs

  x_i = Value of Random Variable (Outcome)

  $P(x_i)$ = Probability Associated with Value

- x_i's are mutually exclusive
  (no overlap)
- x_i's are collectively exhaustive
  (nothing left out)
- $0 \le P(x_i) \le 1$ for each x_i
- $\Sigma P(x_i) = 1$

Discrete Random Variable Summary Measures

- Expected Value of a discrete distribution
  (Weighted Average)

$$E(x) = \Sigma x_i P(x_i)$$

- Example: Toss 2 coins,
  x = # of heads,
  compute expected value of x:

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

$E(x) = (0 \times .25) + (1 \times .50) + (2 \times .25)$
$= 1.0$

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

x = Values of the random variable

P(x) = Probability of the random variable having the value of x

Discrete Random Variable Summary Measures

(continued)

- Example: Toss 2 coins, x = # heads, 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

Two Discrete Random Variables

- Expected value of the sum of two discrete random variables:

$$E(x + y) = E(x) + E(y)$$
$$= \Sigma\, x\, P(x) + \Sigma\, y\, P(y)$$

(The expected value of the sum of two random variables is the sum of the two expected values)

Covariance

- **Covariance** between two discrete random variables:

$$\sigma_{xy} = \Sigma\,[x_i - E(x)][y_j - E(y)]P(x_i y_j)$$

where:

x_i = possible values of the x discrete random variable

y_j = possible values of the y discrete random variable

$P(x_i\ ,y_j)$ = joint probability of the values of x_i and y_j occurring

Interpreting Covariance

- **Covariance** between two discrete random variables:

$\sigma_{xy} > 0 \longrightarrow$ x and y tend to move in the same direction

$\sigma_{xy} < 0 \longrightarrow$ x and y tend to move in opposite directions

$\sigma_{xy} = 0 \longrightarrow$ x and y do not move closely together

Correlation Coefficient

- **The Correlation Coefficient** shows the strength of the linear association between two variables

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

where:

ρ = correlation coefficient ("rho")
σ_{xy} = covariance between x and y
σ_x = standard deviation of variable x
σ_y = standard deviation of variable y

Interpreting the Correlation Coefficient

- **The Correlation Coefficient always falls between -1 and +1**

$\rho = 0 \longrightarrow$ x and y are not linearly related.

The farther ρ is from zero, the stronger the linear relationship:

$\rho = +1 \longrightarrow$ x and y have a perfect positive linear relationship

$\rho = -1 \longrightarrow$ x and y have a perfect negative linear relationship

Chapter Summary

- Described approaches to assessing probabilities
- Developed common rules of probability
- Used Bayes' Theorem for conditional probabilities
- Distinguished between discrete and continuous probability distributions
- Examined discrete probability distributions and their summary measures

SAS Exercise

- Based on the Ketchup data
  - What is the covariance between the prices of the different brands?
  - What is the correlations between the prices of the different brands?
  - What is your inference based on the above findings?
  - What is the correlation between the promotions of the different brands? What is your inference based on these findings?