Statistics for Managers Using Microsoft® Excel 5th Edition

Chapter 4

Basic Probability

Learning Objectives

In this chapter, you will learn:

- Basic probability concepts
- Conditional probability

Definitions

- **Probability**: the chance that an uncertain event will occur (always between 0 and 1)
- **Event**: Each possible type of occurrence or outcome
- **Simple Event**: an event that can be described by a single characteristic
- **Sample Space**: the collection of all possible events

Types of Probability

There are three approaches to assessing the probability of an uncertain event:

1. ***a priori* classical probability**: the probability of an event is based on prior knowledge of the process involved.

2. **empirical classical probability**: the probability of an event is based on observed data.

3. **subjective probability**: the probability of an event is determined by an individual, based on that person's past experience, personal opinion, and/or analysis of a particular situation.

Calculating Probability

1. ***a priori* classical probability**

$$\text{Probability of Occurrence} = \frac{X}{T} = \frac{\text{number of ways the event can occur}}{\text{total number of possible outcomes}}$$

2. **empirical classical probability**

$$\text{Probability of Occurrence} = \frac{\text{number of favorable outcomes observed}}{\text{total number of outcomes observed}}$$

These equations assume all outcomes are equally likely.

Example of *a priori* classical probability

Find the probability of selecting a face card (Jack, Queen, or King) from a standard deck of 52 cards.

$$\text{Probability of Face Card} = \frac{X}{T} = \frac{\text{number of face cards}}{\text{total number of cards}}$$

$$\frac{X}{T} = \frac{12 \text{ face cards}}{52 \text{ total cards}} = \frac{3}{13}$$

Example of empirical classical probability

Find the probability of selecting a male taking statistics from the population described in the following table:

	Taking Stats	Not Taking Stats	Total
Male	84	145	229
Female	76	134	210
Total	160	279	439

$$\text{Probability of Male Taking Stats} = \frac{\text{number of males taking stats}}{\text{total number of people}} = \frac{84}{439} = 0.191$$

Class Exercise

PH Grade ASSIST **4.5** For each of the following, indicate whether the type of probability involved is an example of *a priori* classical probability, empirical classical probability, or subjective probability.

a. The next toss of a fair coin will land on heads.

b. Italy will win soccer's World Cup the next time the competition is held.

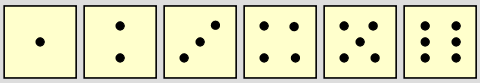
c. The sum of the faces of two dice will be seven.

d. The train taking a commuter to work will be more than 10 minutes late.

Examples of Sample Space

The Sample Space is the collection of all possible events

ex. All 6 faces of a die:



ex. All 52 cards in a deck of cards

ex. All possible outcomes when having a child:
Boy or Girl

Events in Sample Space

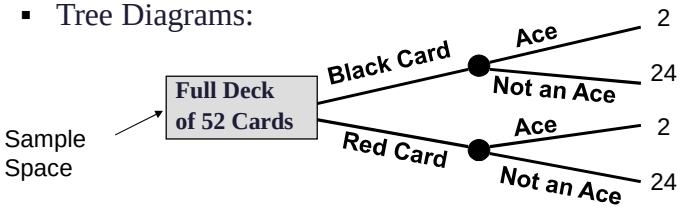
- Simple event
  - An outcome from a sample space with one characteristic
  - ex. A red card from a deck of cards
- Complement of an event A (denoted A')
  - All outcomes that are not part of event A
  - ex. All cards that are not diamonds
- Joint event
  - Involves two or more characteristics simultaneously
  - ex. An ace that is also red from a deck of cards

Visualizing Events in Sample Space

- Contingency Tables:

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

- Tree Diagrams:



Definitions Simple vs. Joint Probability

- Simple (Marginal) Probability refers to the probability of a simple event.
  - ex. P(King)
- Joint Probability refers to the probability of an occurrence of two or more events.
  - ex. P(King and Spade)

Definitions
Mutually Exclusive Events

- **Mutually exclusive events** are events that cannot occur together (simultaneously).
- example:
  - A = queen of diamonds; B = queen of clubs
  - Events A and B are mutually exclusive if only one card is selected
- example:
  - B = having a boy; G = having a girl
  - Events B and G are mutually exclusive if only one child is born

Definitions
Collectively Exhaustive Events

- **Collectively exhaustive events**
  - One of the events must occur
  - The set of events covers the entire sample space
- example:
  - A = aces; B = black cards; C = diamonds; D = hearts
  - Events A, B, C and D are **collectively exhaustive** (but not mutually exclusive – a selected ace may also be a heart)
  - Events B, C and D are **collectively exhaustive** and also mutually exclusive

Class Exercise

4.6 For each of the following, state whether the events created are mutually exclusive and collectively exhaustive. If they are not mutually exclusive and collectively exhaustive, either reword the categories to make them mutually exclusive and collectively exhaustive or explain why that would not be useful.

a. Registered voters in the United States were asked whether they registered as Republicans or Democrats.

b. Each respondent was classified by the type of car he or she drives: American, European, Japanese, or none.

c. People were asked, "Do you currently live in (i) an apartment or (ii) a house?"

d. A product was classified as defective or not defective.

Computing Joint and Marginal Probabilities

- The probability of a joint event, A and B:

$$P(\text{A and B}) = \frac{\text{number of outcomes satisfying A and B}}{\text{total number of elementary outcomes}}$$

- Computing a marginal (or simple) probability:

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \cdots + P(A \text{ and } B_k)$$

  - Where B_1, B_2, …, B_k are k mutually exclusive and collectively exhaustive events

Example: Joint Probability

P(Red and Ace)

$$= \frac{\text{number of cards that are red and ace}}{\text{total number of cards}} = \frac{2}{52}$$

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

Example: Marginal (Simple) Probability

P(Ace)

$$= P(\text{Ace and Red}) + P(\text{Ace and Black}) = \frac{2}{52} + \frac{2}{52} = \frac{4}{52}$$

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

Joint Probability Using a Contingency Table

Event	Event B_1	Event B_2	Total
A_1	$P(A_1 \text{ and } B_1)$	$P(A_1 \text{ and } B_2)$	$P(A_1)$
A_2	$P(A_2 \text{ and } B_1)$	$P(A_2 \text{ and } B_2)$	$P(A_2)$
Total	$P(B_1)$	$P(B_2)$	1

Joint Probabilities

Marginal (Simple) Probabilities

Class Exercise

4.8 According to an Ipsos poll, the perception of unfairness in the U.S. tax code is spread fairly evenly across income groups, age groups, and education levels. In an April 2006 survey of 1,005 adults, Ipsos reported that almost 60% of all people said the code is unfair, while slightly more than 60% of those making more than $50,000 viewed the code as unfair ("People Cry Unfairness," *The Cincinnati Enquirer*, April 16, 2006, p. A8). Suppose that the following contingency table represents the specific breakdown of responses:

U.S. TAX CODE	INCOME LEVEL: Less Than $50,000	More Than $50,000	Total
Fair	225	180	405
Unfair	280	320	600
Total	505	500	1,005

a. Give an example of a simple event.
b. Give an example of a joint event.
c. What is the complement of "tax code is fair"?
d. Why is "tax code is fair *and* makes less than $50,000" a joint event?

4.9 Referring to the contingency table in Problem 4.8, if a respondent is selected at random, what is the probability that he or she
a. thinks the tax code is unfair?
b. thinks the tax code is unfair and makes less than $50,000?
c. thinks the tax code is unfair or makes less than $50,000?
d. Explain the difference in the results in (b) and (c).

Class Exercise

SELF Test **4.10** A yield improvement study at a semiconductor manufacturing facility provided defect data for a sample of 450 wafers. The following table presents a summary of the responses to two questions: "Was a particle found on the die that produced the wafer?" and "Is the wafer good or bad?"

QUALITY OF WAFER	CONDITION OF DIE		
	No Particles	Particles	Totals
Good	320	14	334
Bad	80	36	116
Totals	400	50	450

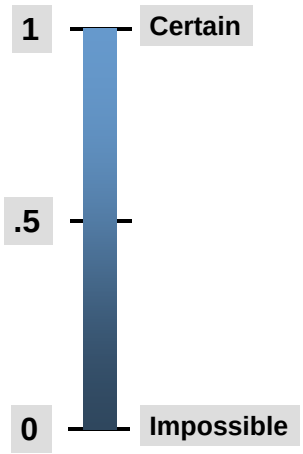
Source: Extracted from S. W. Hall, "Analysis of Defectivity of Semiconductor Wafers by Contingency Table," Proceedings Institute of Environmental Sciences, Vol. 1, 1994, pp. 177–183.

a. Give an example of a simple event.
b. Give an example of a joint event.
c. What is the complement of a good wafer?
d. Why is a "good wafer" and a die "with particles" a joint event?

4.11 Referring to the contingency table in Problem 4.10, if a wafer is selected at random, what is the probability that
a. it was produced from a die with no particles?
b. it is a bad wafer and was produced from a die with no particles?
c. it is a bad wafer or was produced from a die with particles?
d. Explain the difference in the results in (b) and (c).

Probability Summary So Far

- Probability is the numerical measure of the likelihood that an event will occur.
- The probability of any event must be between 0 and 1, inclusively
  - $0 \le P(A) \le 1$ for any event A.
- The sum of the probabilities of all mutually exclusive and collectively exhaustive events is 1.
  - $P(A) + P(B) + P(C) = 1$
  - A, B, and C are mutually exclusive and collectively exhaustive



General Addition Rule

General Addition Rule:

P(A or B) = P(A) + P(B) - P(A and B)

If A and B are mutually exclusive, then

P(A and B) = 0, so the rule can be simplified:

P(A or B) = P(A) + P(B)

for mutually exclusive events A and B

General Addition Rule Example

Find the probability of selecting a male or a statistics student from the population described in the following table:

	Taking Stats	Not Taking Stats	Total
Male	84	145	229
Female	76	134	210
Total	160	279	439

P(Male or Stat) = P(M) + P(S) – P(M or S)
= 229/439 + 160/439 – 84/439 = 305/439

Conditional Probability

- A conditional probability is the probability of one event, given that another event has occurred:

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)} \Rightarrow$$ The conditional probability of A given that B has occurred

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)} \Rightarrow$$ The conditional probability of B given that A has occurred

Where P(A and B) = joint probability of A and B
P(A) = marginal probability of A
P(B) = marginal probability of B

Computing Conditional Probability

- 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 ?
- We want to find P(CD | AC).

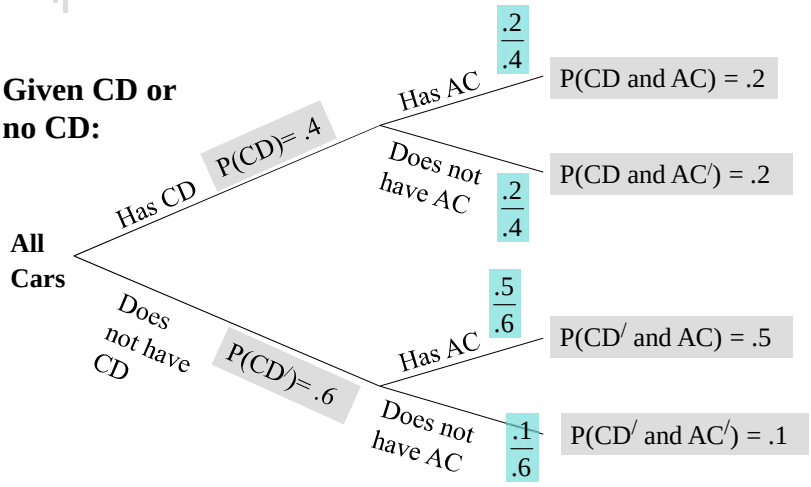
Computing Conditional Probability

	CD	No CD	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

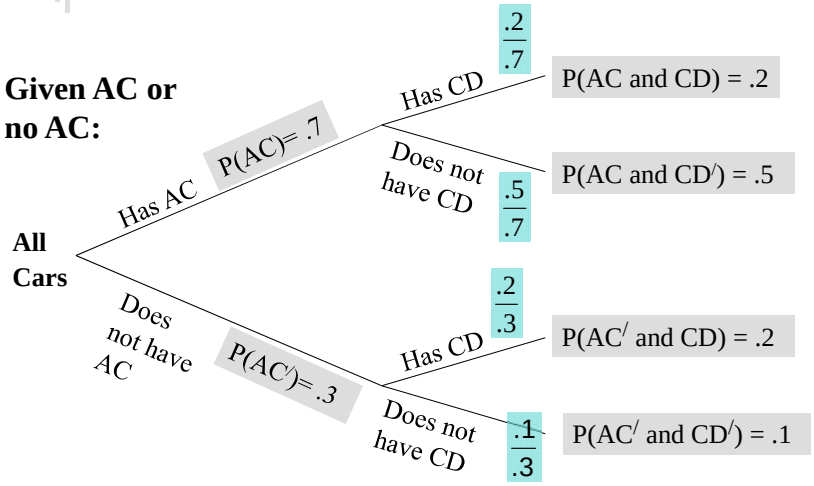
$$P(CD \mid AC) = \frac{P(CD \text{ and } AC)}{P(AC)} = \frac{.2}{.7} = .2857$$

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%.

Computing Conditional Probability: Decision Trees



Computing Conditional Probability: Decision Trees



Statistical Independence

- Two events are **independent** if and only if:

$$P(A \mid B) = P(A)$$

- Events A and B are independent when the probability of one event is not affected by the other event

Class Exercise

PH Grade ASSIST **4.16** Given the following contingency table:

	B	B'
A	10	20
A'	20	40

What is the probability of
a. $A \mid B$?
b. $A \mid B'$?
c. $A' \mid B'$?
d. Are events A and B statistically independent?

4.17 Given the following contingency table:

	B	B'
A	10	30
A'	25	35

What is the probability of
a. $A \mid B$?
b. $A' \mid B'$?
c. $A \mid B'$?
d. Are events A and B statistically independent?

Multiplication Rules

- Multiplication rule for two events A and B:

$$P(A \text{ and } B) = P(A \mid B)\,P(B)$$

- If A and B are independent, then $P(A \mid B) = P(A)$

and the multiplication rule simplifies to:

$$P(A \text{ and } B) = P(A)\,P(B)$$

Class Exercise

PH Grade ASSIST **4.19** If $P(A) = 0.7$, $P(B) = 0.6$, and A and B are statistically independent, find $P(A \text{ and } B)$.

PH Grade ASSIST **4.20** If $P(A) = 0.3$, $P(B) = 0.4$, and if $P(A \text{ and } B) = 0.2$, are A and B statistically independent?

Class Exercise

4.21 Where people turn to for news is different for various age groups. Suppose that a study conducted on this issue (extracted from P. Johnson, "Young People Turn to the Web for News," *USA Today*, March 23, 2006, p. 9D) was based on 200 respondents who were between the ages of 36 and 50 and 200 respondents who were over age 50. Of the 200 respondents who were between the ages of 36 and 50, 82 got their news primarily from newspapers. Of the 200 respondents who were over age 50, 104 got their news primarily from newspapers.

a. Given that a respondent is over age 50, what then is the probability that he or she gets news primarily from newspapers?

b. Given that a respondent gets news primarily from newspapers, what is the probability that he or she is over age 50?

c. Explain the difference in the results in (a) and (b).

d. Are the two events, whether the respondent is over age 50 and whether he or she gets news primarily from newspapers, statistically independent?

Class Exercise

SELF Test **4.22** A yield improvement study at a semiconductor manufacturing facility provided defect data for a sample of 450 wafers. The following table presents a summary of the responses to two questions: "Were particles found on the die that produced the wafer?" and "Is the wafer good or bad?"

QUALITY OF WAFER	CONDITION OF DIE		
	No Particles	Particles	Totals
Good	320	14	334
Bad	80	36	116
Totals	400	50	450

Source: Extracted from S. W. Hall, "Analysis of Defectivity of Semiconductor Wafers by Contingency Table," Proceedings Institute of Environmental Sciences, Vol. 1, 1994, pp. 177–183.

a. Suppose you know that a wafer is bad. What is the probability that it was produced from a die that had particles?
b. Suppose you know that a wafer is good. What is the probability that it was produced from a die that had particles?
c. Are the two events, a good wafer and a die with no particles, statistically independent? Explain.