# ECON 2110 Fall 2025

Chapter 6: Probability

Instructor: Dennis McCornac



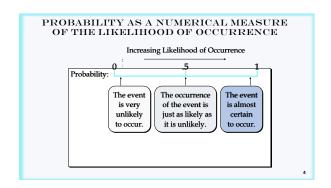
### INTRODUCTION

Although descriptive statistics consist of a set of useful graphical and numerical methods, we are particularly interested in developing statistical inference about a population from a sample.

Our primary objective in this and the following two chapters is to develop the probability-based tools that are at the basis of statistical inference.

,

# INTRODUCTION TO PROBABILITY Probability is a numerical measure of the likelihood that an event will occur. Probability values are always assigned on a scale from 0 to 1. A probability near zero indicates an event is quite unlikely to occur. A probability near one indicates an event is almost certain to occur.



### STATISTICAL EXPERIMENTS

In statistics, the notion of an experiment differs somewhat from that of an experiment in the physical sciences.

In statistical experiments, probability determines

Even though the experiment is repeated in exactly the same way, an entirely different outcome may occur.

For this reason, statistical experiments are sometimes called *random experiments*.

### RANDOM EXPERIMENT

A random experiment is an action or process that leads to one of several possible outcomes.

Examples:

Illustration 1. Experiment: Flip a coin.

Outcomes: Heads and tails

Illustration 2. Experiment: Record marks on a statistics test (out of 100).

Outcomes: Numbers between 0 and 100 Illustration 3. Experiment: Record grade on a statistics test.

Outcomes: A, B, C, D, and F

The list of possible outcomes of an experiment must be:

Exhaustive: All possible outcomes must be included

Mutually Exclusive: Two outcomes cannot occur at the same time

### SAMPLE SPACE AND PROBABILITY REQUIREMENTS

A sample space of a random experiment is a list of all possible outcomes of the experiment.

Using set notation, we represent the sample space and its k outcomes as:

$$S = \{O_1, O_2, O_3, \dots, O_k \,\}$$

When we begin the task of assigning probabilities to the outcomes, we must obey two probability requirements:

- 1. The probability of any outcome must lie between 0 and 1. That is,  $0 \le P(O_i) \le 1$  for each i
- 2. The sum of the probabilities of all the outcomes in a sample space must be 1.

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

### AN EXPERIMENT AND ITS SAMPLE SPACE

Experiment

Toss a coin Inspection of a part

Conduct a sales call Purchase, no purchase

Roll a die Play a football game Experiment Outcomes

Head, tail Defective, non-defective

1, 2, 3, 4, 5, 6 Win, lose, tie

### THREE APPROACHES TO ASSIGNING PROBABILITIES

The  ${\it classical\ approach\ }$  is used by mathematicians to determine probability based on games of chance, and it consists of equally likely outcomes.

The relative frequency approach is applied to experimentation or historical data, and it defines probability as the long-run relative frequency with which an outcome  $% \left( 1\right) =\left( 1\right) \left( 1\right) \left$ 

The subjective approach assigns probabilities based on a subjective judgment, and it is used when the classical approach is not reasonable and there are no available experimentations or historical data to apply the relative frequency approach.

### INTERPRETING PROBABILITY

No matter what method was used to assign probability, we always interpret it using the relative frequency approach for an infinite number of experiments

Subjective approach: an investor assumed that there is a 65% probability that a particular stock's price will increase over the next month.

Interpretation:

if we had an infinite number of stocks with the exact same economic and market characteristics as the one the investor will buy, 65% of them will increase in price over the next month.

the probability that a balanced die lands on the number 5 is 1/6. Classical approach: Interpretation: the proportion of times that a 5 is observed on a balanced die thrown an infinite number of times.

### CLASSICAL METHOD EXAMPLE

### Example: Rolling a Die

If an experiment has n possible outcomes, the classical method would assign a probability of 1/n to each outcome.

Experiment: Rolling a die

Sample Space:  $S = \{1, 2, 3, 4, 5, 6\}$ 

Probabilities: Each sample point has a 1/6 chance of occurring

RELATIVE FREQUENCY METHOD EXAMPLE

Example: Lucas Tool Rental

- Lucas Tool Rental would like to assign probabilities to the number of car polishers it rents each day, and office records show the following frequencies of daily rentals for the last 40 days.
- Each probability assignment is given by dividing the frequency (number of days) by the total frequency (total number of days).

Number of Polishers Rented	Number of Days	Probability	
0 1 2 3 4	4 6 18 10 <u>2</u> 40	.10 .15 .45 .25 <u>.05</u> 1.00	4/40

### SUBJECTIVE METHOD

- When economic conditions or a company's circumstances change rapidly it might be inappropriate to assign probabilities based solely on historical data
- We can use any data available as well as our experience and intuition, but ultimately a probability value should express our <u>degree of belief</u> that the experimental outcome will occur.
- The best probability estimates often are obtained by combining the estimates from the classical or relative frequency approach with the subjective estimate.

### DEFINING EVENTS

An individual outcome of a sample space is called a simple event. An **event** is a collection or set of one or more simple events in a sample space.

### Example:

We can define the event to achieve a grade of A as the set of numbers that lie between 80 and 100, inclusive.
Using set notation, we have:

 $A = \{80, 81, 82, ..., 100\}$ 

# PROBABILITY OF AN EVENT

The probability of an event is the sum of the probabilities of the simple events that constitute the event.

Example

Suppose that for a grade example, we employed the relative frequency approach to assign probabilities to the simple events as follows: P(A)=.20, P(B)=.30, P(C)=.25, P(D)=.15, P(F)=.10

The probability of the event, passing the course, is:

P(Pass the course) = P(A) + P(B) + P(C) + P(D)= 0.2 + 0.3 + 0.25 + 0.1 = 0.9

15

### SOME BASIC RELATIONSHIPS OF PROBABILITY

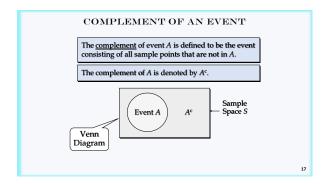
There are some <u>basic probability relationships</u> that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

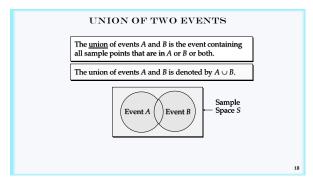
Complement of an Event

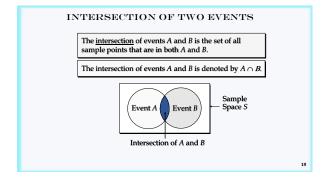
Union of Two Events

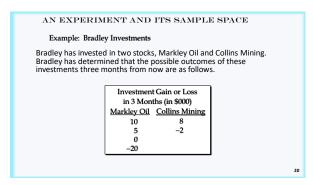
Intersection of Two Events

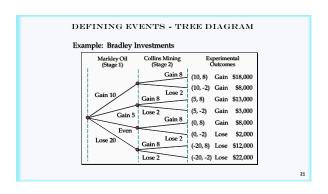
Mutually Exclusive Events

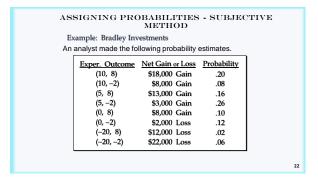












EVENTS AND THEIR PROBABILITES

Example: Bradley Investments

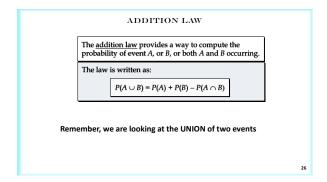
Event M = Markley Oil Profitable  $M = \{(10, 8), (10, -2), (5, 8), (5, -2)\}$  P(M) = P(10, 8) + P(10, -2) + P(5, 8) + P(5, -2) = .20 + .08 + .16 + .26Event C = Collins Mining Profitable  $C = \{(10, 8), (5, 8), (0, 8), (-20, 8)\}$  P(C) = P(10, 8) + P(5, 8) + P(0, 8) + P(-20, 8) = .20 + .16 + .10 + .02 = 0.48

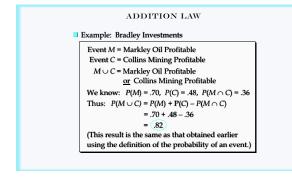
UNION OF TWO EVENTS

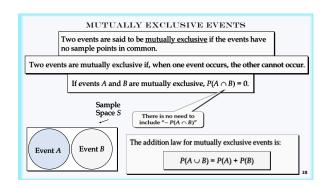
Example: Bradley Investments

Event M = Markley Oil ProfitableEvent C = Collins Mining Profitable  $M \cup C = Markley Oil Profitable$ or Collins Mining Profitable (or both)  $M \cup C = \{(10, 8), (10, -2), (5, 8), (5, -2), (0, 8), (-20, 8)\}$   $P(M \cup C) = P(10, 8) + P(10, -2) + P(5, 8) + P(5, -2)$  + P(0, 8) + P(-20, 8) = .20 + .08 + .16 + .26 + .10 + .02 = .82

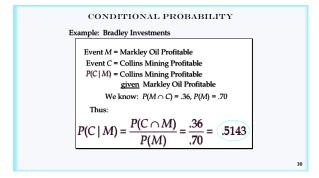
# INTERSECTION OF TWO EVENTS Example: Bradley Investments Event M = Markley Oil ProfitableEvent C = Collins Mining Profitable $M \cap C = Markley Oil Profitable$ and Collins Mining Profitable $M \cap C = \{(10, 8), (5, 8)\}$ $P(M \cap C) = P(10, 8) + P(5, 8)$ = .20 + .16 = .36







# The probability of an event given that another event has occurred is called a <u>conditional probability</u>. The conditional probability of <u>A given B</u> is denoted by $P(A \mid B)$ . A conditional probability is computed as follows: $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$



MULTIPLICATION LAW

The <u>multiplication law</u> provides a way to compute the probability of the intersection of two events.

The law is written as:  $P(A \cap B) = P(A)P(B \mid A)$ or  $P(A \cap B) = P(B)P(A \mid B)$ 

MULTIPLICATION LAW

Example: Bradley Investments

Event M = Markley Oil ProfitableEvent C = Collins Mining Profitable  $M \cap C = \text{Markley Oil Profitable}$ and Collins Mining Profitable

We know: P(M) = .70, P(C|M) = .5143Thus:  $P(M \cap C) = P(M)P(C|M)$  = (.70)(.5143) = (.36)(This result is the same as that obtained earlier using the definition of the probability of an event.)

A financial analyst has determined that there is a 22% probability that a mutual fund will outperform the market over a 1-year period provided that it outperformed the market the previous year. If only 15% of mutual funds outperform the market during any year, what is the probability that a mutual fund will outperform the market 2 years in a row?

Let A = mutual fund outperforms the market in the first year

B = mutual fund outperforms the market in the second year

 $P(A \text{ and } B) = P(A)P(B \mid A) = (.15)(.22) = .033$ 

JOINT PROBABILITY TABLE Collins Mining
Profitable (C) Not Profitable (C) Markley Oil Profitable (M) .70 Not Profitable (M<sup>c</sup>) .12 .18 .30 1.00 .52 Joint Probabilities (appear in the body of the table) Marginal Probabilities (appear in the margins of the table)

To investigate how often families eat at home, Harris Interactive surveyed 496 adults living with children under the age of 18 (USA Today, January 3, 2007). The survey results are shown in the following table.

Number of	Number of
Family Meals	Survey
Week	Responses
0	11
1	11
2	30
3	36
4	36
5	119
6	114
7	120

For a randomly selected family with children under the age of 18, compute the following.

- a. The probability the family eats no meals at home during the week
  b. The probability the family eats at least four meals at home during the week
  c. The probability the family eats two or fewer meals at home during the week

	Number of Family Meals Week	Number of Survey Responses
a. The probability the family eats no meals at home during the week	0	11
b. The probability the family eats at least four meals at home during the week	1	11
c. The probability the family eats two or fewer meals at home during the week	2	30
	3	36
Solution	4	36
11	5	119
. a. $P(\text{no meals}) = \frac{11}{496} = .0222$	6 7 or more	114 139
b. $P(\text{at least four meals}) = P(4) + P(5) + P(6) + P(7)$ 36 , 119 , 114 , 139	or more)	
$= \frac{36}{496} + \frac{119}{496} + \frac{114}{496} + \frac{139}{496} = .8226$		
c. $P(\text{two or fewer meals}) = P(2) + P(1) + P(0)$		
30 11 11		
$= \frac{30}{496} + \frac{11}{496} + \frac{11}{496} = .1048$		

# DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

Suppose that a potential investor examined the relationship between how well the mutual fund performs and which university awarded the manager's MBA. After the analysis, the following table of joint probabilities, was developed:

MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET
Top 20	.11	.29
Non top 20	.06	.54

Analyze these probabilities and interpret the results.

# $\begin{array}{c} \textbf{EXAMPLE} - \textbf{DETERMINANTS} \ \textbf{OF} \ \textbf{SUCCESS} \ \textbf{AMONG} \\ \textbf{MUTUAL} \ \textbf{FUND} \ \textbf{MANAGERS} \end{array}$

Let us represent the events as follows:

- $A_1$  = Fund manager graduated from a top-20 MBA program
- $A_2$  = Fund manager did not graduate from a top-20 MBA program
- $B_1$  = Fund outperforms the market
- $B_2$  = Fund does not outperform the market

Thus, the joint probabilities are:

- $P(A_1 \text{ and } B_1) = .11$
- $P(A_2 \text{ and } B_1) = .06$
- $P(A_1 \text{ and } B_2) = .29$
- $P(A_2 \text{ and } B_2) = .54$

38

### MARGINAL PROBABILITIES

Marginal probabilities are calculated in the margins of the table, and they are computed by adding across rows or down columns.

In Example 6.1, adding across the two rows and column produces the marginal probabilities:

First row:  $P(A_1) = P(A_1 \text{ and } B_1) + P(A_1 \text{ and } B_2) = .11 + .29 = .40$ Second row:  $P(A_1) = P(A_1 \text{ and } B_1) + P(A_1 \text{ and } B_2) = .06 + .54 = .60$ 

Second row:  $P(A_2) = P(A_2 \ and \ B_1) + P(A_2 \ and \ B_2) = .06 + .54 = .60$ First column:  $P(B_1) = P(A_1 \ and \ B_1) + P(A_2 \ and \ B_1) = .11 + .06 = .17$ 

Second column:  $P(B_2) = P(A_1 \text{ and } B_2) + P(A_2 \text{ and } B_2) = .29 + .54 = .83$ 

We can visualize marginal and joint probabilities as follows:

		B1	B2	
	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### CONDITIONAL PROBABILITY ONCE AGAIN

- The **conditional probability** is the probability of one event given the occurrence of another related event.
- The probability of event A given event B is:

• 
$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

• The probability of event B given event A is:

• 
$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

 For example, what is the probability that a fund managed by a graduate of a top-20 MBA program (event A<sub>1</sub>) will outperform the market (event B<sub>2</sub>)?

### CONDITIONAL PROBABILITY ONCE AGAIN

• The conditional probability that a fund managed by a graduate of a top-20 MBA program (event  $A_1$ ) will outperform the market (event  $B_1$ ) is:

• 
$$P(B_1|A_1) = \frac{P(A_1 \text{ and } B_1)}{P(A_1)} = \frac{.11}{.40} = .275$$

B2 B2

	PROGRAM	OUTPERFORMS MARKET	OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### EXAMPLE - DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

We can reformulate the conditional probability we need to calculate as follows:

The probability that a randomly selected mutual fund is managed by a graduate of a top-20 MBA program (event  $A_1$ ), **given the fact that** the fund did not outperform the market (event

$$P(A_1|B_2) = \frac{P(A_1 \text{ and } B_2)}{P(B_2)} = \frac{.29}{.83} = .349$$
B1 B2

	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### EXAMPLE - DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

Suppose that we select one mutual fund at random and discover that it did not outperform the market.

What is the probability that a graduate of a top-20 MBA program manages it?

		B1	B2	
	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \ and \ B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### CONDITIONAL PROBABILITY - ANOTHER EXAMPLE

**Example:** Promotion status of police officers over the past two years

	Men	Women	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

### JOINT PROBABILITY TABLE

**Example:** Promotion status of police officers over the past two years

	Men (M)	Women (W)	Total
Promoted (A)	288/1200 = 0.24	0.03	0.27
Not Promoted (A <sup>c</sup> )	0.56	0.17	0.73
Total	0.80	0.20	1.00

### CONDITIONAL PROBABILITY -ANOTHER EXAMPLE

Example: Promotion status of police officers over the past two years

Event A = An officer is promoted
Event M = The promoted officer is a man
P(A | M) = An officer is promoted given that the officer is a man

$$P(A|M) = \frac{P(A \cap M)}{P(M)}$$

From the table we know:

 $P(A \cap M) = 0.24$  P(M) = 0.8 P(A|M) = 0.24/0.8 = 0.3

### INDEPENDENCE

One of the objectives of calculating conditional probability is to determine whether two events are related.

In particular, we would like to know whether they are independent events.

Two events A and B are said to be independent if:

P(A|B) = P(A)

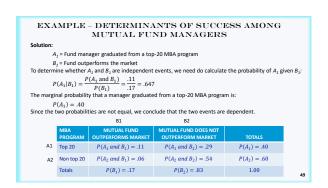
P(B|A)=P(B)

In other words, whether one event happens has no influence on the likelihood that the other event happens.

EXAMPLE - DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

Determine whether the event that the manager graduated from a top-20 MBA program and the event the fund outperforms the market are independent events.

 $P(A_1 \text{ and } B_1) = .11$  $P(A_1 \text{ and } B_2) = .29$ A1 Top 20 A2 Non top 20  $P(A_2 \text{ and } B_1) = .06$  $P(A_2 \text{ and } B_2) = .54$  $P(A_2) = .60$  $P(B_1) = .17$  $P(B_2) = .83$ 1.00



### COMPLEMENT RULE AND UNION OF TWO EVENTS ONCE AGAIN

The **complement** of event A is the event that occurs when event A does not occur. The complement of event A is denoted by  $A^{C}$ .

The complement rule defined here derives from the fact that the probability of an event and the probability of the event's complement must sum to 1.

 $P(A^{\mathcal{C}}) = 1 - P(A)$ for any event A.

The union of events A and B is the event that occurs when either A or B or both

It is denoted as:

A or B

52

# $\begin{array}{c} {\bf EXAMPLE-DETERMINANTS\ OF\ SUCCESS\ AMONG}\\ {\bf MUTUAL\ FUND\ MANAGERS} \end{array}$ Determine the probability that a randomly selected fund outperforms the market, or the manager graduated from a top-20 MBA program. → We want to compute the probability of the union of two events:

		B1	B2	
	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### EXAMPLE - DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

**Solution:** The union  $A_1$  or  $B_1$  occurs when any of the following joint events occurs: A1 and B1

- A2 and B1

• A1 and B2  $P(A_1 \text{ or } B_1) = .11 + .06 + .29 = .46$ 

Notice that the only joint probability not representing an event that is part of the union is the joint

Which is the probability that the union does not occur. Thus, the probability of the union can also be

 $P(A_1 \text{ or } B_1) = 1 - P(A_2 \text{ and } B_2) = 1 - .54 = .46$ 

	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### MULTIPLICATION RULE ONCE AGAIN

The multiplication rule is used to calculate the joint probability of two events. It is based on the formula for conditional probability:

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

We derive the multiplication rule by solving for P(A and B):

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

Recollect that, if A and B are independent events, P(B|A) = P(B). It follows that the joint probability of two independent events is simply the product of the probabilities of the two events:

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

## EXAMPLES - SELECTING TWO STUDENTS

### • Example - Without Replacement

A graduate statistics course has seven male and three female students. The professor wants to select two students at random.

What is the probability that the two students chosen are female?

Let us define the following events:

A. the first chosen student is female.

B. the second chosen student is also female.

We need: P(A and B) = P(A)P(B|A) $P(A) = \frac{3}{10} \text{ and } P(B|A) = \frac{2}{9}$ 

Thus:  $P(A \text{ and } B) = \frac{3}{10} \frac{2}{9} = \frac{6}{90} = .067$ 

### EXAMPLES - SELECTING TWO STUDENTS

### • Example - With Replacement

The professor needs to select one student at random as a substitute for the next two classes. What is the probability that the students selected for the two classes are both females?

This time the same student can be selected for both classes.

Therefore, A and B are now independent events.

We need: P(A and B) = P(A)P(B)With:  $P(A) = \frac{3}{10} \text{ and } P(B) = \frac{3}{10}$ 

 $P(A \text{ and } B) = \frac{3}{10} \frac{3}{10} = \frac{9}{100} = .09$ Thus:

ADDITION RULE

The addition rule calculates the probability of the union of two events. That is, the probability that event A, or event B, or both occur is:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

This is because:

 $P(A_1) = P(A_1 \text{ and } B_1) + P(A_1 \text{ and } B_2)$   $P(B_1) = P(A_1 \text{ and } B_1) + P(A_2 \text{ and } B_1)$ 

Plugging  $P(A_1)$  and  $P(B_2)$  into the addition rule:  $P(A_1 \text{ or } B_1) = P(A_1 \text{ and } B_1) + P(A_1 \text{ and } B_2) + P(A_1 \text{ and } B_1) + P(A_2 \text{ and } B_1) - P(A_1 \text{ and } B_2)$ 

If we now simplify  $P(A_1 \text{ or } B_1)$ , we obtain the sum of the joint probabilities as seen in the example on the

 $P(A_1 \text{ or } B_1) = P(A_1 \text{ and } B_2) + P(A_1 \text{ and } B_1) + P(A_2 \text{ and } B_1)$ 

### EXAMPLE - DETERMINANTS OF SUCCESS AMONG MUTUAL FUND MANAGERS

The union  $A_1$  or  $B_1$  occurs when any of the following joint events occurs:

- A1 and B1
- A2 and B1
- A1 and B2

 $P(A_1 \text{ or } B_1) = .11 + .06 + .29 = .46$ 

Another way of getting the same answer is:

 $P(A_1 \text{ or } B_1) = P(A_1) + P(B_1) - P(A_1 \text{ and } B_1)$ 

	MBA PROGRAM	MUTUAL FUND OUTPERFORMS MARKET	MUTUAL FUND DOES NOT OUTPERFORM MARKET	TOTALS
A1	Top 20	$P(A_1 \text{ and } B_1) = .11$	$P(A_1 \text{ and } B_2) = .29$	$P(A_1) = .40$
A2	Non top 20	$P(A_2 \text{ and } B_1) = .06$	$P(A_2 \text{ and } B_2) = .54$	$P(A_2) = .60$
	Totals	$P(B_1) = .17$	$P(B_2) = .83$	1.00

### EXAMPLE - APPLYING THE ADDITION RULE

In a large city, two newspapers are published, the Sun and the Post.

The circulation departments report that 22% of the city's households have a subscription to the Sun and 35% subscribe to the Post. A survey reveals that 6% of all households subscribe to both newspapers.

What proportion of the city's households subscribe to at least one newspaper?

Solution: We can reformulate the question as: "what is the probability of selecting a household at random that subscribes to the Sun or the Post or both?"

P(Sun or Post) = P(Sun) + P(Post) - P(Sun and Post)

P(Sun or Post) = .22 + .35 - .06 = .51

Interpret: There is a 51% probability that a randomly selected household subscribes to

one or the other or both papers.

Discrimination in the workplace is illegal, and companies that discriminate are often sued. The female instructors at a large university recently lodged a complaint about the most recent round of promotions from assistant professor to associate professor. An analysis of the relationship between gender and promotion produced the following joint probabilities.

	Promoted	Not Promoted
Female	.03	.12
Male	.17	.68

a. What is the rate of promotion among female assistant professors?

b. What is the rate of promotion among male assistant professors?

c. Is it reasonable to accuse the university of gender bias?

Discrimination in the workplace is illegal, and companies that discriminate are often sued. The female instructors at a large university recently lodged a complaint about the most recent round of promotions from assistant professor to associate professor. An analysis of the relationship between gender and promotion produced the following joint probabilities.

	Promoted	Not Promoted
Female	.03	.12
Male	.17	.68

a. What is the rate of promotion among female assistant professors?

b. What is the rate of promotion among male assistant professors?

c. Is it reasonable to accuse the university of gender bias?

a.  $P(promoted \mid female) = \frac{P(promoted and female)}{P(female)} = \frac{.03}{.03 + .12} = .20$ 

b.  $P(promoted \mid male) = \frac{P(promoted and male)}{P(male)} = \frac{.17}{.17 + .68} = .20$ 

c. No, because promotion and gender are independent events.

The following table lists the joint probabilities associated with gender and cigarette smoking.

Gender	Never Gender Smoked		Current Smoker
Male	.23	.05	.20
Female	.29	.16	.07

- a. Find the probability that a male is a current smoker.
- b. Compute the probability that a former smoker is female.

Gender	Never Smoked		Current Smoker
Male	.23	.05	.20
Female	.29	.16	.07

- a. Find the probability that a male is a current smoker.
- b. Compute the probability that a former smoker is female.

- a. P(Current smoker | Male) = .20/(.23 + .05 + .20) = .20/.48 = .417
- b. (Female | Former smoker) = .16/(.05 + .16) = .16/.21 = .762

The generally accepted definitions of Generation X and Millennials is that the former was born between 1965 and 1980 and the latter born after 1981. Baby Boomers are defined as people born between 1946 and 1964. An analysis conducted by the Pew Research Center produced the following table of joint probabilities relating marital status of the three groups defined have.

groups defined here.				
	Marital Status	Millennial	Gen X	Boomer
	Single, never married	0.195	0.058	0.030
	Married	0.089	0.223	0.201
	Living with not married	0.030	0.025	0.009
	Divorced, separated,	0.017	0.054	0.070

- a. Find the probability that a Millennial is married.
- b. Compute the probability that a Baby Boomer is single, never married.
- c. Suppose that one person is selected at random. What is the probability that that person is married?
- d. What is the probability that someone who is living with a partner but not married is a Generation Xer?

Marital Status	Millennial	Gen X	Boomer	
Single, never married	0.195	0.058	0.030	
Married	0.089	0.223	0.201	
Living with not married	0.030	0.025	0.009	
Divorced, separated,	0.017	0.054	0.070	

- Find the probability that a Millennial is married.
   Compute the probability that a Baby Boomer is single, never married.
   Suppose that one person is selected at random. What is the probability that that person is married?
- d. What is the probability that someone who is living with a partner but not married is a Generation Xer?

- Solutions:
  a. P(Married | Millennia) = .089/.331 = .269
  b. P(Single, never married | Baby boomer) = .030/.310 = .0968
  c. P(Married) = .089 + .223 + .201 = .513
  d. P(Generation X | Living with not married) = .025/.064 = .391

The costs of medical care in North America are increasing faster than inflation, and with the baby boom generation soon to need health care, it becomes imperative that countries find ways to reduce both costs and demand. The following table lists the joint probabilities associated with smoking and lung disease among 60- to 65-year-old men.

	He is a Smoker	He is a Nonsmoker
He has lung disease	.12	.03
He does not have lung disease	.19	.66

- a. He is a smoker.
- b. He does not have lung disease.
- c. He has lung disease given that he is a smoker.
- d. He has lung disease given that he does not smoke.

osts of medical care in North America are increasing faster than inflation, and with the baby boom generation soon to need health care, it becomes imperative that countries find ways to reduce both costs and demand. The following table lists the joint probabilities associated with smoking and lung disease among 60-to 65-year-old men.

	He is a Smoker	He is a Nonsmokei
He has lung disease	.12	.03
He does not have lung disease	.19	.66
	ted at rando	n. What is the
events?	ted at rando	m. What is the
vents? . He is a smoker.	ted at rando	m. What is the

- d.  $P(\text{He has lung disease} \mid \text{he does not smoke})$   $= \frac{P(\text{he has lung disease and he does not smoke})}{P(\text{he does not smoke})} = \frac{.03}{.69} = .043$

A financial analyst has determined that there is a 22% probability that a mutual fund will outperform the market over a 1-year period provided that it outperformed the market the previous year. If only 15% of mutual funds outperform the market during any year, what is the probability that a mutual fund will outperform the market 2 years in a row?

Let A = mutual fund outperforms the market in the first year

B = mutual fund outperforms the market in the second year

 $P(A \text{ and } B) = P(A)P(B \mid A) = (.15)(.22) = .033$ 

PROBABILITY TREES

In a **probability tree**, the events in an experiment are represented by branches, which are lines linked to each other.

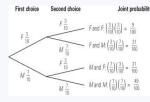
We calculate the joint probabilities by multiplying the probabilities on the linked branches.

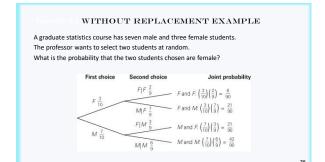
Parallel branches from the same node are mutually exclusive and can be added together.

### EXAMPLE - WITH REPLACEMENT EXAMPLE

The professor needs to select one student at random as a substitute for the next two classes.

What is the probability that the students selected for the two classes are both females?





### EXAMPLE - PROBABILITY OF PASSING THE BAR EXAM

Suppose that the pass rate for law school graduates taking the bar exam for the first time is 72%.

Candidates who fail the first exam may take it again later. Of those who fail their first test, 88% pass on their second attempt.

 $Find \ the \ probability \ that \ a \ randomly \ selected \ law \ school \ graduate \ becomes \ a \ lawyer.$ 

Assume that candidates cannot take the exam more than twice.

EXAMPLE - PROBABILITY OF PASSING THE BAR

First exam Second exam Joint probability

Pass .72

Pass|Fail .88 Fail and Pass .1281.888 = .2464

Fail .28 Fail .12 Fail and Fail .(28)(.12) = .0336

The included probability tree describes the experiment.  $P(\text{Becoming Lawyer}) = P(\text{Pass}) + P(\text{Fail})P(\text{Pass}|\text{Fail}) = .72 + .28 \cdot .88$  = .9664

### BAYES'S LAW

Conditional probability measures the **likelihood** that event *B* occurs given that a possible cause of the event *A* has occurred.

We calculated the probability that a mutual fund outperformed the market (the effect) given that the fund manager graduated from a top-20 MBA program (the possible cause)

**Bayes's Law** is used when we witness event *B*, and we need to compute the probability of one of its possible causes (event *A*).

$$P(B|A) \Rightarrow P(A|B)$$

If we observe that a mutual fund outperformed the market, what is the probability that it was managed by a graduate from a top-20 MBA program?

### BAYES' LAW - TERMINOLOGY

The probabilities P(A) and  $P(A^C)$  are called **prior probabilities** because they are determined prior to determining whether event B has taken place.

The conditional probability P(B|A) is called a likelihood probability.

The conditional probability P(A | B) is called a **posterior probability** (or **revised probability**), because the prior probability is revised after event B has taken place.

We use Bayes' rule to determine posterior probabilities (or revised probabilities).



The production of the bound of the production o

### EXAMPLE - BUILDING A PROBABILITY TREE

• Should an MBA Applicant Take a Preparatory Course?

Suppose that a survey of MBA students reveals that among GMAT scores of at least 650, 52% took a preparatory course, whereas among GMAT scores of less than 650, only 23% took a preparatory course.

An applicant to an MBA program knows that the probability of scoring at least 650 is 10%, and he is willing to purchase the prep course only if the probability of achieving at least 650 at least doubles, or 20%.

### SHOULD AN MBA APPLICANT TAKE A PREPARATORY COURSE? We can begin by defining the events A and B as follows: A = GMAT score is 650 or more B = Took preparatory course Then, we can define the provided probabilities as: P(A) = .10Joint probability Preparatory course $P(A^c) = .90$ B|A=.52 A and B: |.100|.521 = .052P(B|A) = .52Using the complement and multiplications rules $B|A^{C}|23$ $A^{C}$ and B|1.90((.23) = .207we can complete the probability tree: $B^{\circ}A^{\circ}$ .77 $A^{\circ}$ and $B^{\circ}$ : (.90)(.77) = .693

### SOLVING A PROBABILITY TREE WITH BAYES'S LAW

We need to calculate the probability that an MBA applicant scores at least 650 (event A), given that he has taken a preparatory course (event B).

Using the conditional probability formula, we can write:

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

The probability tree provides: P(A and B) = .052

From the probability tree we can also derive the marginal probability P(B) as:

$$P(B) = P(A \text{ and } B) + P(A^{C} \text{ and } B) = .052 + .207 = .259$$

Thus, the probability that an MBA applicant scores at least 650, given that he has taken a preparatory course is:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{.052}{.259} = .201 \text{ which is slightly greater than 20\%}.$$

BAYES'S LAW FORMULA

For those who prefer an algebraic approach rather than a probability tree, Bayes's Law can be expressed as a formula:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_k)P(B|A_k)}$$

where:

B is the given event,

 $A_1, A_2, ..., A_k$  are the events with known prior probabilities  $P(A_1), P(A_2), ..., P(A_k)$ ,

 $P(B|A_1)$ ,  $P(B|A_2)$ ,...,  $P(B|A_k)$  are the likelihood probabilities,

 $P(A_i|B)$ , with i=1,2,...,k are the posterior probabilities we seek.

EXAMPLE - SOLVING WITH BAYES'S LAW

• Should an MBA Applicant Take a Preparatory Course?

We define the events as follows:

 $A_1$  = GMAT score is 650 or more

A<sub>2</sub> = GMAT score is less than 650 B = Student took preparatory course

The provided prior probabilities are:

 $P(A_1) = .10$  (probability student scored 650 or more)

 $P(A_2) = 1 - .10 = .90$  (complement probability – probability student scored less than 650)

The provided conditional (likelihood) probabilities are:

 $P(B|A_1)$  = .52 (probability student took prep course among those scoring 650 or more)

 $P(B|A_2)$  = .23 (probability student took prep course among those scoring less than 650)

The Bayes's Law formula yields the probability a student scores 650 or more after taking the prep course:

 $P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} = \frac{.10(.52)}{.10(.52) + .90(.23)} = \frac{.052}{.052 + .207} = .201$ 

### ADDITIONAL EXAMPLE - NARCOTICS

In a particular pain clinic, 10% of patients are prescribed narcotic pain killers. Overall, five percent of the clinic's patients are addicted to narcotics (including pain killers and illegal substances). Out of all the people prescribed pain pills, 8% are addicts. If a patient is an addict, what is the probability that they will be prescribed pain pills?

### Step 1: Figure out what your event "A" is from the question.

The event that happens first (A) is being prescribed pain pills. That's given as 10%.

### Step 2: Figure out what your event "B" is from the question.

Event B is being an addict. That's given as 5%.

<u>Step 3</u>: Figure out what the probability of event B (Step 2) given event A (Step 1). In other words, find what  $(B \mid A)$  is. We want to know "Given that people are prescribed pain pills, what's the probability they are an addict?" That is given in the question as 8%.

# Step 4: Insert your answers from Steps 1, 2 and 3 into the formula and solve. $P(A \mid B) = P(B \mid A) * P(A) / P(B) = (0.08 * 0.1)/0.05 = 0.16$

The probability of an addict being prescribed pain pills is 0.16 (16%).

### IDENTIFYING THE CORRECT METHOD

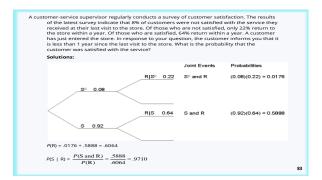
The key issue is to determine whether joint probabilities are provided or are required:

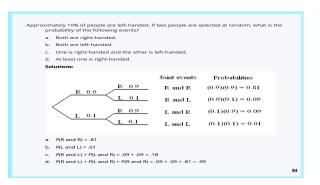
### · Joint probabilities are provided

- 1. Compute marginal probabilities by adding across rows and down columns.
- $2. \quad \text{Use the joint and marginal probabilities to compute conditional probabilities}.$
- 3. Determine whether the events described by the table are independent.
- 4. Apply the addition rule to compute the probability that either of two events

### Joint probabilities are required

- 1. Apply probability rules or build a probability tree.
- 2. Use the multiplication rule to calculate the probability of intersections.
- 3. Apply addition and complement rules for mutually exclusive events.
- 4. Compute posterior probability using the Bayes's Law.





A study undertaken by the Miami-Dade Supervisor of Elections revealed that 44% of registered voters are Democrats, 37% are Republicans, and 19% are others. If two registered voters are selected at random, what is the probability that both of them have the same party affiliation?

### Solutions

		Joint Events	Probabilities
	D 0.44	DD	(0.44)(0.44) = .1936
D 0.44	R 0.37	_ DR	(0.44)(0.37) = .1628
	0 0.19	DO	(0.44)(0.19) = .0836
	D 0.44	RD	(0.37)(0.44) = .1628
R 0.37	R 0.37	_ RR	(0.37)(0.37) = .1369
	0 0.19	RO	(0.37)(0.19) = .0703
	D 0.44	OD	(0.19)(0.44) = .0836
O 0.19	R 0.37	OR	(0.19)(0.37) = .0703
	.0 0.19	00	(0.19)(0.19) = .0361

P(Same party affiliation) = P(DD) + P(RR) + P(OO) = .1936 + .1369 + .0361.3666

How does level of affluence affect health care? To address one dimension of the problem, a group of heart attack victims was selected, fach was categorized as a low, medium, or more than the problem of the medium-income group, and 30% are in the high-income group. Furthermore, an analysis of heart attack victims reveals that 12% of low-income group, 9% or netium that a survivor of a heart attack is in the low-income group. Furthermore, an analysis of heart attack with the probability of the problem of the probability of the probability of the probability of the probability of the probabilities.

Solutions:

Define events,  $A_1 = Low-income$  earner,  $A_2 = medium-income$  earner,  $A_3 = high-income$  earner,  $B = die of a heart attack, <math>B^c = survive$  a heart attack.

Probabilities:  $A_1 = 0.21$   $B^c | A_1 = 0.88$   $A_2 = 0.49$   $B^c | A_2 = 0.91$   $A_3 = 0.30$   $B^c | A_2 = 0.91$   $A_3 = 0.30$   $B^c | A_3 = 0.30$ 

The most recent U.S. Census reveals that 17.6% of Americans live in the Northeast, 21.3% live in the Midwest, 23.5% live in the West, and 37.6% live in the South. A survey asked American adults whether they approve of unions. From the survey, the probabilities below were generated.

	P(Approve unions)
Northeast	.69
Midwest	.65
West	.60
South	.58

Find the probability that a randomly selected person would not approve of unions.

The most recent U.S. Census reveals that 17.6% of Americans live in the Northeast, 21.3% live in the Midwest, 23.5% live in the West, and 37.6% live in the South. A survey asked American adults whether they approve of unions. From the survey, the probabilities below were generated.

	P(Approve unions)
Northeast	.69
Midwest	.65
West	.60
South	58

Find the probability that a randomly selected person would not approve of unions.

### Solutions

 $\label{eq:proposed} $$P(Don't approve) = P(NE)P(Don't approve \mid NE) + P(MW)P(Don't approve \mid MW) + P(West)P(Don't approve \mid West) + P(South)P(Don't approve \mid South) = (.176)(.31) + (.213)(.35) + (.235)(.40) + (.376)(.42) = .0546 + .0746 + .0940 + .1579 = .3810$