3.5: Contingency Tables and Probability Trees

A **contingency table** provides a way of portraying data that can facilitate calculating probabilities. The table helps in determining conditional probabilities quite easily. The table displays sample values in relation to two different variables that may be dependent or contingent on one another. Later on, we will use contingency tables again, but in another manner.

Exercise 3.5.1

Suppose a study of speeding violations and drivers who use cell phones produced the following fictional data:

	Speeding violation in the last year	No speeding violation in the last year	Total
Uses cell phone while driving	25	280	305
Does not use cell phone while driving	45	405	450
Total	70	685	755

Table 3.5.1

The total number of people in the sample is 755. The row totals are 305 and 450. The column totals are 70 and 685. Notice that 305 + 450 = 755 and 70 + 685 = 755.

Calculate the following probabilities using the table.

a. Find *P*(Driver is a cell phone user).
b. Find *P*(driver had no violation in the last year).
c. Find *P*(Driver had no violation in the last year AND was a cell phone user).
d. Find *P*(Driver is a cell phone user OR driver had no violation in the last year).
e. Find *P*(Driver is a cell phone user GIVEN driver had a violation in the last year).
f. Find *P*(Driver had no violation last year GIVEN driver was not a cell phone user)

Answer

a. $\frac{\text{number of cell phone users}}{\text{total number in study}} = \frac{305}{755}$
b. $\frac{\text{number that had no violation}}{\text{total number in study}} = \frac{685}{755}$
c. $\frac{280}{755}$
d. $\left(\frac{305}{755} + \frac{685}{755}\right) - \frac{280}{755} = \frac{710}{755}$
e. $\frac{25}{70}$ (The sample space is reduced to the number of drivers who had a violation.)
f. $\frac{405}{450}$ (The sample space is reduced to the number of drivers who were not cell phone users.)

▼

Solution

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

Try It 3.5.1

Table 3.5.2 shows the number of athletes who stretch before exercising and how many had injuries within the past year.

	Injury in last year	No injury in last year	Total
Stretches	55	295	350
Does not stretch	231	219	450
Total	286	514	800

Table 3.5.2

a. What is *P*(athlete stretches before exercising)?
b. What is *P*(athlete stretches before exercising||no injury in the last year)?

▼

Solution

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

Exercise 3.5.2

Table ❓ (\PageIndex{3}\) shows a random sample of 100 hikers and the areas of hiking they prefer.

Sex	The Coastline	Near Lakes and Streams	On Mountain Peaks	Total
Women	18	16		45
Men			14	55
Total		41		

Table ❓ 3.5.3Hiking Area Preference

a. Complete the table.

b. Are the events "being a woman" and "preferring the coastline" independent events?

Let $F =$ being a woman and let $C =$ preferring the coastline.
1. Find $P(F \cap C)$.
2. Find $P(F)P(C)$

Are these two numbers the same? If they are, then F and C are independent. If they are not, then F and C are not independent.

c. Find the probability that a person is a man given that the person prefers hiking near lakes and streams. Let $M =$ being a man, and let $L =$ prefers hiking near lakes and streams.

1. What word tells you this is a conditional?
2. Fill in the blanks and calculate the probability: P (______ | ______) = ______.
3. Is the sample space for this problem all 100 hikers? If not, what is it?

d. Find the probability that a person is a woman or prefers hiking on mountain peaks. Let $F =$ being a woman, and let $P =$ prefers mountain peaks.

1. Find $P(F)$.
2. Find $P(P)$.
3. Find $P(F \cap P)$.
4. Find $P(F \cup P)$.

Answer

a.

Sex	The coastline	Near lakes and streams	On mountain peaks	Total
Women	18	16	**11**	45
Men	**16**	**25**	14	55
Total	**34**	41	**25**	**100**

Table 3.5.4Hiking Area Preference

b.

1. $P(F \cap C) = \frac{18}{100} = 0.18$
2. $P(F)P(C) = \left(\frac{45}{100}\right)\left(\frac{34}{100}\right) = (0.45)(0.34) = 0.153$

$P(F \cap C) \neq P(F)P(C)$, so the events F and C are not independent.

c.

1. The word 'given' tells you that this is a conditional.
2. $P(M \mid L) = \frac{25}{41}$
3. No, the sample space for this problem is the 41 hikers who prefer lakes and streams.

d.

1. $P(F) = \frac{45}{100}$
2. $P(P) = \frac{25}{100}$

3. $P(F \cap P) = \dfrac{11}{100}$
4. $P(F \cup P) = \dfrac{45}{100} + \dfrac{25}{100} - \dfrac{11}{100} = \dfrac{59}{100}$

▼

Solution

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

Try It 3.5.2

Table 3.5.5 shows a random sample of 200 cyclists and the routes they prefer. Let O = older and H = hilly path.

Age Group	Lake Path	Hilly Path	Wooded Path	Total
Younger	45	38	27	110
Older	26	52	12	90
Total	71	90	39	200

Table 3.5.5

a. Out of the older group, what is the probability that the cyclist prefers a hilly path?
b. Are the events "being older" and "preferring the hilly path" independent events?

Exercise 3.5.3

Muddy Mouse lives in a cage with three doors. If Muddy goes out the first door, the probability that he gets caught by Alissa the cat is $\dfrac{1}{5}$ and the probability he is not caught is $\dfrac{4}{5}$. If he goes out the second door, the probability he gets caught by Alissa is $\dfrac{1}{4}$ and the probability he is not caught is $\dfrac{3}{4}$. The probability that Alissa catches Muddy coming out of the third door is $\dfrac{1}{2}$ and the probability she does not catch Muddy is $\dfrac{1}{2}$. It is equally likely that Muddy will choose any of the three doors so the probability of choosing each door is $\dfrac{1}{3}$.

Caught or Not	Door One	Door Two	Door Three	Total
Caught	\(\dfrac{1}{15}\)$\frac{1}{15}$	\(\dfrac{1}{12}\)	1\(\dfrac{1}{6}\)$\frac{1}{6}$	-
Not Caught	\(\dfrac{4}{15}\)	\(\dfrac{3}{12}\)	\(\dfrac{1}{6}\)	-
Total	-	-	-	1

Table 3.5.6Door Choice

- The first entry $\dfrac{1}{15} = \left(\dfrac{1}{5}\right)\left(\dfrac{1}{3}\right)$ is $P($ Door One $\cap$ Caught)
- The entry $\dfrac{4}{15} = \left(\dfrac{4}{5}\right)\left(\dfrac{1}{3}\right)$ is P (Door One $\cap$ Not Caught)

Verify the remaining entries.

Problem

a. Complete the probability contingency table. Calculate the entries for the totals. Verify that the lower-right corner entry is 1 .

b. What is the probability that Alissa does not catch Muddy?

c. What is the probability that Muddy chooses Door One OR Door Two given that Muddy is caught by Alissa?

Answer

a.

Caught or Not	Door One	Door Two	Door Three	Total
Caught	\(\dfrac{1}{15}\)$\frac{1}{15}$	\(\dfrac{1}{12}\)	1\(\dfrac{1}{6}\)$\frac{1}{6}$	\(\dfrac{19}{60}\)
Not Caught	\(\dfrac{4}{15}\)	\(\dfrac{3}{12}\)	\(\dfrac{1}{6}\)	\(\dfrac{41}{60}\)
Total	\(\dfrac{5}{15}\)	\(\dfrac{4}{12}\)	\(\dfrac{2}{6}\)	1

Table 3.5.7Door Choice

b. \(\dfrac{41}{60}\)

c. \(\dfrac{9}{19}\)

▼

Solution

a.

Caught or Not	Door One	Door Two	Door Three	Total
Caught	115115" role="presentation" style="position: relative;"> $\frac{1}{15}$	112112" role="presentation" style="position: relative;"> $\frac{1}{12}$	1616" role="presentation" style="position: relative;"> $\frac{1}{6}$	19601960" role="presentation" style="position: relative;"> $\frac{19}{60}$
Not Caught	415415" role="presentation" style="position: relative;"> $\frac{4}{15}$	312312" role="presentation" style="position: relative;"> $\frac{3}{12}$	1616" role="presentation" style="position: relative;"> $\frac{1}{6}$	41604160" role="presentation" style="position: relative;"> $\frac{41}{60}$
Total	515515" role="presentation" style="position: relative;"> $\frac{5}{15}$	412412" role="presentation" style="position: relative;"> $\frac{4}{12}$	2626" role="presentation" style="position: relative;"> $\frac{2}{6}$	1

Table 3.9 Door Choice

b. 41604160" role="presentation" style="position: relative;"> $\frac{41}{60}$

c. 919919" role="presentation" style="position: relative;"> $\frac{9}{19}$

▼

Solution

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

Try It 3.5.3

Fred's preference to drink coffee or iced tea depends on the season. When it is summer, Fred's preference to drink coffee is $\frac{1}{4}$, and his preference to drink iced tea is $\frac{3}{4}$. When it is rainy season, Fred's preference to drink coffee is $\frac{1}{2}$, and his preference to drink iced tea is $\frac{1}{2}$. When it is winter, Fred's preference to drink coffee is $\frac{1}{5}$, and his preference to drink iced tea is $\frac{1}{5}$. A day randomly selected in a year has a probability of $\frac{1}{3}$ for each season. The event here is selection of a day randomly in a year.

Season	**Summer**	**Rainy**	**Winter**	**Total**
Coffee	\(\dfrac{1}{12}\)	\(\dfrac{1}{6}\)	\(\dfrac{4}{15}\)	__
Iced tea	\(\dfrac{1}{4}\)	\(\dfrac{1}{6}\)	\(\dfrac{1}{15}\)	__
Total	__	__	__	1

Table 3.5.8

Table 3.10

- The first entry is $\left(\frac{1}{4}\right)\left(\frac{1}{3}\right)=\frac{1}{12}$.
- The second entry is $\left(\frac{3}{4}\right)\left(\frac{1}{3}\right)=\frac{1}{4}$.

Verify the remaining entries.
a. Complete the probability contingency table. Calculate the entries for the totals. Verify that the lower-right corner entry is 1 .
b. What is the probability that Fred will drink iced tea?
c. What is the probability that the day is in summer or rainy season given that Fred drinks iced tea?

$P(AORB)$

Exercise 3.5.4

Table 3.5.9 contains the number of crimes per 100,000 inhabitants in the United States over the span of several years.

Year	Robbery	Burglary	Vandalism	Vehicle	Total
1	145.7	732.1	29.7	314.7	
2	133.1	717.7	29.1	259.2	
3	119.3	701	27.7	239.1	
4	113.7	702.2	26.8	229.6	
Total					

Table 3.5.9 United States Crime Index Rates Per 100,000 Inhabitants

TOTAL each column and each row. Total data = 4,520.7

a. Find *P* (Year 2 AND Robbery).
b. Find *P* (Year 3 AND Burglary).
c. Find *P* (Year 3 OR Burglary).
d. Find *P* (Year 4 | Vandalism).
e. Find *P* (Vehicle | Year 1).

Answer

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

▼

Solution

a. 0.0294, b. 0.1551, c. 0.7165, d. 0.2365, e. 0.2575

Try It 3.5.4

Table 3.5.10 relates the weights and heights of a group of individuals participating in an observational study.

Weight/height	Tall	Medium	Short	Totals
Overweight	18	28	14	
Typical Weight Range	20	51	28	
Underweight	12	25	9	
Totals				

Table 3.5.10

a. Find the total for each row and column
b. Find the probability that a randomly chosen individual from this group is Tall.
c. Find the probability that a randomly chosen individual from this group is Overweight and Tall.
d. Find the probability that a randomly chosen individual from this group is Tall given that the individual is Overweight.
e. Find the probability that a randomly chosen individual from this group is Overweight given that the individual is Tall.
f. Find the probability a randomly chosen individual from this group is Tall and Underweight.
g. Are the events Overweight and Tall independent?