

3.5: Contingency Tables and Probability Trees

Contingency Tables

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.

Example 3.5.1

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

Table 3.5.1

	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

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 $P(\text{Driver is a cell phone user})$.

Answer

$$\text{a. } \frac{\text{number of cell phone users}}{\text{total number in study}} = \frac{305}{755}$$

- b. Find $P(\text{Driver had no violation in the last year})$.

Answer

$$\text{b. } \frac{\text{number that had no violation}}{\text{total number in study}} = \frac{685}{755}$$

- c. Find $P(\text{Driver had no violation in the last year} \cap \text{was a cell phone user})$.

Answer

$$\text{c. } \frac{280}{755}$$

- d. Find $P(\text{Driver is a cell phone user} \cup \text{driver had no violation in the last year})$.

Answer

$$\text{d. } \left(\frac{305}{755} + \frac{685}{755} \right) - \frac{280}{755} = \frac{710}{755}$$

- e. Find $P(\text{Driver is a cell phone user} \mid \text{driver had a violation in the last year})$.

Answer

$$\text{e. } \frac{25}{70} \text{ (The sample space is reduced to the number of drivers who had a violation.)}$$

f. Find $P(\text{Driver had no violation last year} \mid \text{driver was not a cell phone user})$

Answer

f. $\frac{405}{450}$ (The sample space is reduced to the number of drivers who were not cell phone users.)

Exercise 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

1. What is $P(\text{athlete stretches before exercising})$?
2. What is $P(\text{athlete stretches before exercising} \mid \text{no injury in the last year})$?

Example 3.5.2

Table 3.5.3 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
Female	18	16	—	45
Male	—	—	14	55
Total	—	41	—	—

Table 3.5.3 Hiking Area Preference

a. Complete the table.

Answer

a.

Table 3.5.4 Hiking Area Preference

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

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

Let F = being female 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.

Answer

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. Find the probability that a person is male given that the person prefers hiking near lakes and streams. Let M = being male, 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?

Answer

c.

1. The word 'given' tells you that this is a conditional.

$$2. P(M|L) = \frac{25}{41}$$

3. No, the sample space for this problem is the 41 hikers who prefer lakes and streams.

d. Find the probability that a person is female or prefers hiking on mountain peaks. Let F = being female, 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

d.

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

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

$$3. P(F \cap P) = \frac{11}{100}$$

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

Exercise 3.5.2

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

Table 3.5.5

Gender	Lake path	Hilly path	Wooded path	Total
Female	45	38	27	110
Male	26	52	12	90
Total	71	90	39	200

- a. Out of the males, what is the probability that the cyclist prefers a hilly path?
- b. Are the events “being male” and “preferring the hilly path” independent events?

Example 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 $\frac{1}{15}$ and the probability he is not caught is $\frac{4}{15}$. If he goes out the second door, the probability he gets caught by Alissa is $\frac{1}{12}$ and the probability he is not caught is $\frac{3}{12}$. The probability that Alissa catches Muddy coming out of the third door is $\frac{1}{6}$ and the probability she does not catch Muddy is $\frac{1}{6}$. It is equally likely that Muddy will choose any of the three doors so the probability of choosing each door is $\frac{1}{3}$.

Table 3.5.6 Door Choice

Caught or not	Door one	Door two	Door three	Total
Caught	$\frac{1}{15}$	$\frac{1}{12}$	$\frac{1}{6}$	_____
Not caught	$\frac{4}{15}$	$\frac{3}{12}$	$\frac{1}{6}$	_____
Total	_____	_____	_____	1

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

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.

Answer

a.

Table 3.5.7 Door Choice

Caught or not	Door one	Door two	Door three	Total
Caught	$\frac{1}{15}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{19}{60}$
Not caught	$\frac{4}{15}$	$\frac{3}{12}$	$\frac{1}{6}$	$\frac{41}{60}$
Total	$\frac{5}{15}$	$\frac{4}{12}$	$\frac{2}{6}$	1

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

Answer

b. $\frac{41}{60}$

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

Answer

c. $\frac{9}{19}$

Example 3.5.4

Table 3.5.8 contains the number of crimes per 100,000 inhabitants from 2008 to 2011 in the U.S.

Table 3.5.8 United States Crime Index Rates Per 100,000 Inhabitants 2008–2011

Year	Robbery	Burglary	Rape	Vehicle	Total
2008	145.7	732.1	29.7	314.7	
2009	133.1	717.7	29.1	259.2	

Year	Robbery	Burglary	Rape	Vehicle	Total
2010	119.3	701	27.7	239.1	
2011	113.7	702.2	26.8	229.6	
Total					

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

1. Find $P(2009 \cap \text{Robbery})$.
2. Find $P(2010 \cap \text{Burglary})$.
3. Find $P(2010 \cup \text{Burglary})$.
4. Find $P(2011|\text{Rape})$.
5. Find $P(\text{Vehicle}|2008)$.

Answer

1. 0.0294
2. 0.1551
3. 0.7165
4. 0.2365
5. 0.2575

Exercise 3.5.3

Table 3.5.9 relates the weights and heights of a group of individuals participating in an observational study.

Table 3.5.9

Weight/height	Tall	Medium	Short	Totals
Obese	18	28	14	
Normal	20	51	28	
Underweight	12	25	9	
Totals				

1. Find the total for each row and column
2. Find the probability that a randomly chosen individual from this group is Tall.
3. Find the probability that a randomly chosen individual from this group is Obese and Tall.
4. Find the probability that a randomly chosen individual from this group is Tall given that the individual is Obese.
5. Find the probability that a randomly chosen individual from this group is Obese given that the individual is Tall.
6. Find the probability a randomly chosen individual from this group is Tall and Underweight.
7. Are the events Obese and Tall independent?

Tree Diagrams

Sometimes, when the probability problems are complex, it can be helpful to graph the situation. Tree diagrams can be used to visualize and solve conditional probabilities.

Tree Diagrams

A **tree diagram** is a special type of graph used to determine the outcomes of an experiment. It consists of "branches" that are labeled with either frequencies or probabilities. Tree diagrams can make some probability problems easier to visualize and solve. The following example illustrates how to use a tree diagram.

Example 3.5.5

In an urn, there are 11 balls. Three balls are red (R) and eight balls are blue (B). Draw two balls, one at a time, **with replacement**. "With replacement" means that you put the first ball back in the urn before you select the second ball. The tree diagram using frequencies that show all the possible outcomes follows.

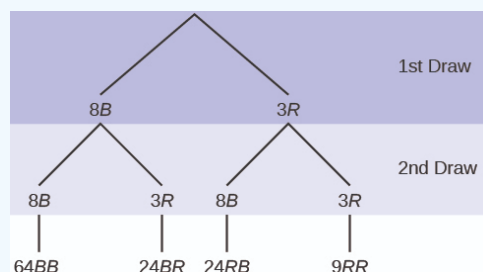


Figure 3.5.1 Total = $64 + 24 + 24 + 9 = 121$

The first set of branches represents the first draw. The second set of branches represents the second draw. Each of the outcomes is distinct. In fact, we can list each red ball as R_1 , R_2 , and R_3 and each blue ball as B_1 , B_2 , B_3 , B_4 , B_5 , B_6 , B_7 , and B_8 . Then the nine RR outcomes can be written as:

R_1R_1 ; R_1R_2 ; R_1R_3 ; R_2R_1 ; R_2R_2 ; R_2R_3 ; R_3R_1 ; R_3R_2 ; R_3R_3

The other outcomes are similar.

There are a total of 11 balls in the urn. Draw two balls, one at a time, with replacement. There are $11(11) = 121$ outcomes, the size of the **sample space**.

a. List the 24 BR outcomes: B_1R_1 , B_1R_2 , B_1R_3 , ...

Answer

a. B_1R_1 ; B_1R_2 ; B_1R_3 ; B_2R_1 ; B_2R_2 ; B_2R_3 ; B_3R_1 ; B_3R_2 ; B_3R_3 ; B_4R_1 ; B_4R_2 ; B_4R_3 ; B_5R_1 ; B_5R_2 ; B_5R_3 ; B_6R_1 ; B_6R_2 ; B_6R_3 ; B_7R_1 ; B_7R_2 ; B_7R_3 ; B_8R_1 ; B_8R_2 ; B_8R_3

b. Using the tree diagram, calculate $P(RR)$.

Answer

$$b. P(RR) = \left(\frac{3}{11}\right) \left(\frac{3}{11}\right) = \frac{9}{121}$$

c. Using the tree diagram, calculate $P(RB \cup BR)$.

Answer

$$c. P(RB \cup BR) = \left(\frac{3}{11}\right) \left(\frac{8}{11}\right) + \left(\frac{8}{11}\right) \left(\frac{3}{11}\right) = \frac{48}{121}$$

d. Using the tree diagram, calculate $P(R \text{ on 1st draw} \cap B \text{ on 2nd draw})$.

Answer

$$d. P(R \text{ on 1st draw} \cap B \text{ on 2nd draw}) = \left(\frac{3}{11}\right) \left(\frac{8}{11}\right) = \frac{24}{121}$$

e. Using the tree diagram, calculate $P(R \text{ on 2nd draw} | B \text{ on 1st draw})$.

Answer

$$e. P(R \text{ on 2nd draw} | B \text{ on 1st draw}) = P(R \text{ on 2nd} | B \text{ on 1st}) = \frac{24}{88} = \frac{3}{11}$$

This problem is a conditional one. The sample space has been reduced to those outcomes that already have a blue on the first draw. There are $24 + 64 = 88$ possible outcomes (24 BR and 64 BB). Twenty-four of the 88 possible outcomes are BR. $\frac{24}{88} = \frac{3}{11}$.

f. Using the tree diagram, calculate $P(BB)$.

Answer

$$f. P(BB) = \frac{64}{121}$$

g. Using the tree diagram, calculate $P(B \text{ on the 2nd draw} | R \text{ on the first draw})$.

Answer

$$g. P(B \text{ on 2nd draw} | R \text{ on 1st draw}) = \frac{8}{11}$$

There are $9 + 24$ outcomes that have R on the first draw (9 RR and 24 RB). The sample space is then $9 + 24 = 33$. 24 of the 33 outcomes have B on the second draw. The probability is then $\frac{24}{33}$.

Exercise 3.5.4

In a standard deck, there are 52 cards. 12 cards are face cards (event F) and 40 cards are not face cards (event N). Draw two cards, one at a time, with replacement. All possible outcomes are shown in the tree diagram as frequencies. Using the tree diagram, calculate $P(FF)$.

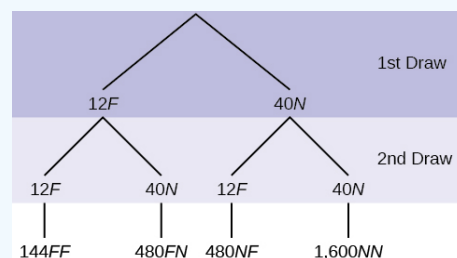


Figure 3.5.2

Example 3.5.6

An urn has three red marbles and eight blue marbles in it. Draw two marbles, one at a time, this time without replacement, from the urn. "**Without replacement**" means that you do not put the first ball back before you select the second marble. Following is a tree diagram for this situation. The branches are labeled with probabilities instead of frequencies. The numbers at the ends of the branches are calculated by multiplying the numbers on the two corresponding branches, for example, $\left(\frac{3}{11}\right)\left(\frac{2}{10}\right) = \frac{6}{110}$.

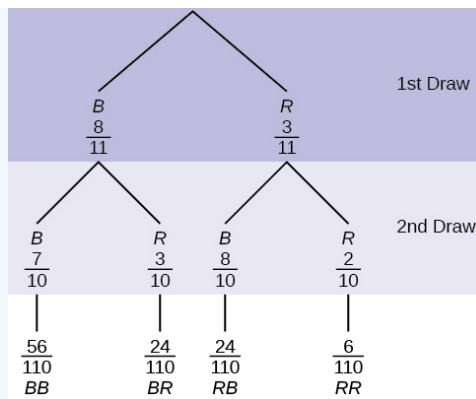


Figure 3.5.3

$$\text{Total} = \frac{56+24+24+6}{110} = \frac{110}{110} = 1$$

NOTE

If you draw a red on the first draw from the three red possibilities, there are two red marbles left to draw on the second draw. You do not put back or replace the first marble after you have drawn it. You draw **without replacement**, so that on the second draw there are ten marbles left in the urn.

Calculate the following probabilities using the tree diagram.

a. $P(RR) = \underline{\hspace{2cm}}$

Answer

a. $P(RR) = \left(\frac{3}{11}\right) \left(\frac{2}{10}\right) = \frac{6}{110}$

b. Fill in the blanks:

$$P(RB \cup BR) = \left(\frac{3}{11}\right) \left(\frac{8}{10}\right) + (\underline{\hspace{1cm}})(\underline{\hspace{1cm}}) = \frac{48}{110}$$

Answer

b. $P(RB \cup BR) = \left(\frac{3}{11}\right) \left(\frac{8}{10}\right) + \left(\frac{8}{11}\right) \left(\frac{3}{10}\right) = \frac{48}{110}$

c. $P(R \text{ on 2nd} | B \text{ on 1st}) =$

Answer

c. $P(R \text{ on 2nd} | B \text{ on 1st}) = \frac{3}{10}$

d. Fill in the blanks.

$$P(R \text{ on 1st} \cap B \text{ on 2nd}) = (\underline{\hspace{1cm}})(\underline{\hspace{1cm}}) = \frac{24}{110}$$

Answer

d. $P(R \text{ on 1st} \cap B \text{ on 2nd}) = \left(\frac{3}{11}\right) \left(\frac{8}{10}\right) = \frac{24}{110}$

e. Find $P(BB)$.

Answer

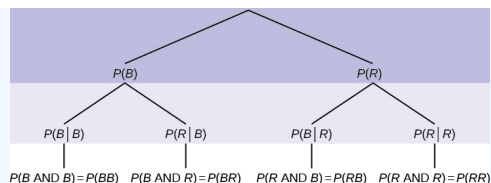
e. $P(BB) = \left(\frac{8}{11}\right) \left(\frac{7}{10}\right)$

f. Find $P(B \text{ on 2nd} | R \text{ on 1st})$.

Answer

f. Using the tree diagram, $P(B \text{ on 2nd} | R \text{ on 1st}) = P(R|B) = \frac{8}{10}$.

If we are using probabilities, we can label the tree in the following general way.



- $P(R|R)$ here means $P(R \text{ on 2nd} | R \text{ on 1st})$
- $P(B|R)$ here means $P(B \text{ on 2nd} | R \text{ on 1st})$
- $P(R|B)$ here means $P(R \text{ on 2nd} | B \text{ on 1st})$
- $P(B|B)$ here means $P(B \text{ on 2nd} | B \text{ on 1st})$

Exercise 3.5.5

In a standard deck, there are 52 cards. Twelve cards are face cards (F) and 40 cards are not face cards (N). Draw two cards, one at a time, without replacement. The tree diagram is labeled with all possible probabilities.

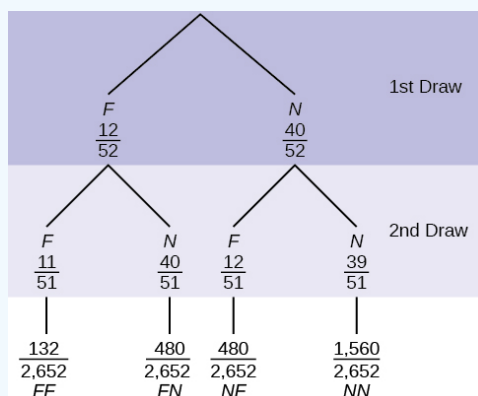
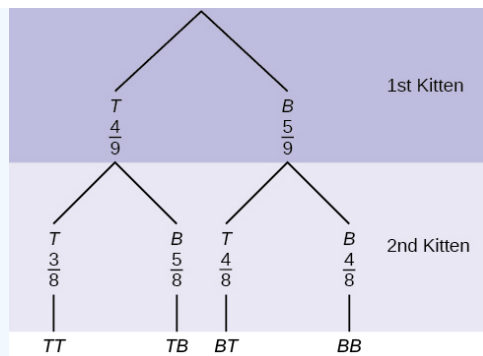


Figure 3.5.4

1. Find $P(FN \cup NF)$.
2. Find $P(N|F)$.
3. Find $P(\text{at most one face card})$.
Hint: "At most one face card" means zero or one face card.
4. Find $P(\text{at least one face card})$.
Hint: "At least one face card" means one or two face cards.

Example 3.5.7

A litter of kittens available for adoption at the Humane Society has four tabby kittens and five black kittens. A family comes in and randomly selects two kittens (without replacement) for adoption.



1. What is the probability that both kittens are tabby?

a. $\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$ b. $\left(\frac{4}{9}\right)\left(\frac{4}{9}\right)$ c. $\left(\frac{4}{9}\right)\left(\frac{3}{8}\right)$ d. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right)$

2. What is the probability that one kitten of each coloring is selected?

a. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right)$ b. $\left(\frac{4}{9}\right)\left(\frac{5}{8}\right)$ c. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right) + \left(\frac{5}{9}\right)\left(\frac{4}{9}\right)$ d. $\left(\frac{4}{9}\right)\left(\frac{5}{8}\right) + \left(\frac{5}{9}\right)\left(\frac{4}{8}\right)$

3. What is the probability that a tabby is chosen as the second kitten when a black kitten was chosen as the first?

4. What is the probability of choosing two kittens of the same color?

Answer

1. c, 2. d, 3. $\frac{4}{8}$, 4. $\frac{32}{72}$

Exercise 3.5.6

Suppose there are four red balls and three yellow balls in a box. Two balls are drawn from the box without replacement. What is the probability that one ball of each coloring is selected?