Chapter 4

Probability Concepts

Section 4.1

Probability Basics

EXAMPLE 4.3 Probability for Equally Likely Outcomes

Dice When two balanced dice are rolled, 36 equally likely outcomes are possible, as depicted in Fig. 4.1.

a. Find the probability that we obtain 1 and 3.

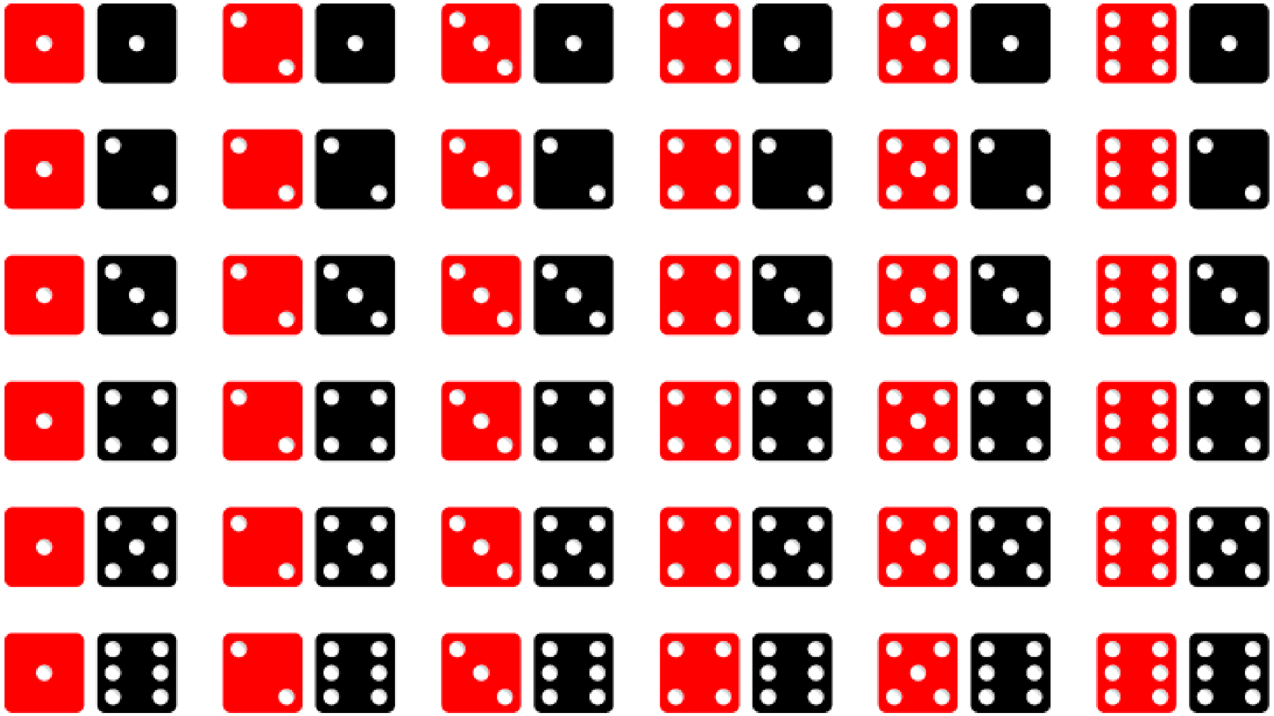
b. Find the probability that 2 is included in the sample.

c. Find the probability that the sum of the dice is 11.

d. Find the probability that doubles are rolled; that is, both dice come up the same number.

Figure 4.1

Possible outcomes for rolling a pair of dice



Definition 4.1

Probability for Equally Likely Outcomes (f/N Rule)

Suppose an experiment has N possible outcomes, all equally likely. An event that can occur in f ways has probability f/N of occurring:

Number of ways event can occur

$$\text{Probability of an event} = \frac{f}{N}.$$

Total number of possible outcomes

Example: The Nature of Probability

Example: Consider an experiment in which we roll two six-sided fair dice and record the number of 3s face up. The only possible outcomes are zero 3s, one 3, or two 3s. Here are the results after 100 rolls, and after 1000 rolls:

100 Rolls

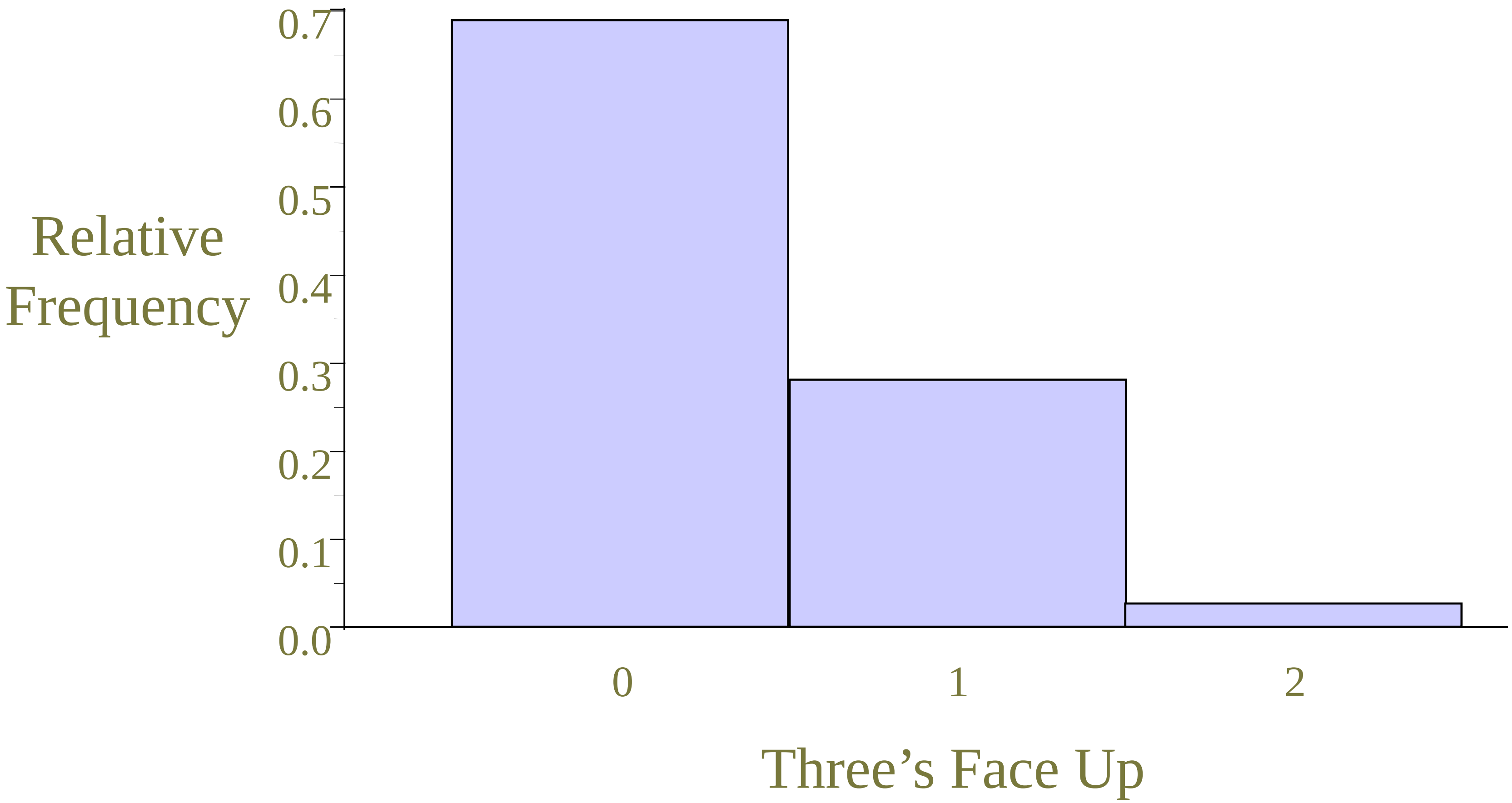
Outcome	Frequency
0	80
1	19
2	1

1000 Rolls

Outcome	Frequency
0	690
1	282
2	28

Using a Histogram

We can express these results (from the 1000 rolls) in terms of relative frequencies and display the results using a histogram:



Continuing the Experiment

If we continue this experiment for several thousand more rolls:

1. The frequencies will have approximately a 25:10:1 ratio in totals
2. The relative frequencies will settle down

Note: We can simulate many probability experiments:

- Use random number tables
- Use a computer to randomly generate number values representing the various experimental outcomes
- Key to either method is to maintain the probabilities

Definitions

A probability is a measure of the likelihood that an event in the future will happen.

- It can only assume a value between 0 and 1.
- A value near zero means the event is not likely to happen. A value near one means it is likely.

Probability of Events

Probability that an Event Will Occur: The relative frequency with which that event can be expected to occur

- The probability of an event may be obtained in three different ways:
  - Empirically
  - Theoretically
  - Subjectively

Definitions *continued*

1. The classical definition applies when there are *n* equally likely outcomes.

2. The empirical definition applies when the number of times the event happens is divided by the number of observations.

3. Subjective probability is based on whatever information is available.

Experimental or Empirical Probability

Experimental or Empirical Probability:

1. The observed relative frequency with which an event occurs
2. Prime notation is used to denote empirical probabilities:

$$P'(\mathrm{A}) = \frac{n(\mathrm{A})}{n}$$

3. n(A): number of times the event A has occurred
4. n: number of times the experiment is attempted

➔ Question: What happens to the observed relative frequency as n increases?

Example

Throughout his teaching career Professor Weiss has awarded 186 A's out of 1,200 students. What is the probability that a student in her section this semester will receive an A?

This is an example of the empirical definition of probability.

To find the probability a selected student earned an A:

$$P(A) = \frac{186}{1200} = 0.155$$

Subjective Probability

1. Suppose the sample space elements are not equally likely, and empirical probabilities cannot be used
2. Only method available for assigning probabilities may be personal judgment
3. These probability assignments are called *subjective probabilities*
4. Personal judgment of the probability is expressed by comparing the likelihood among the various outcomes

Frequentist interpretation of probability

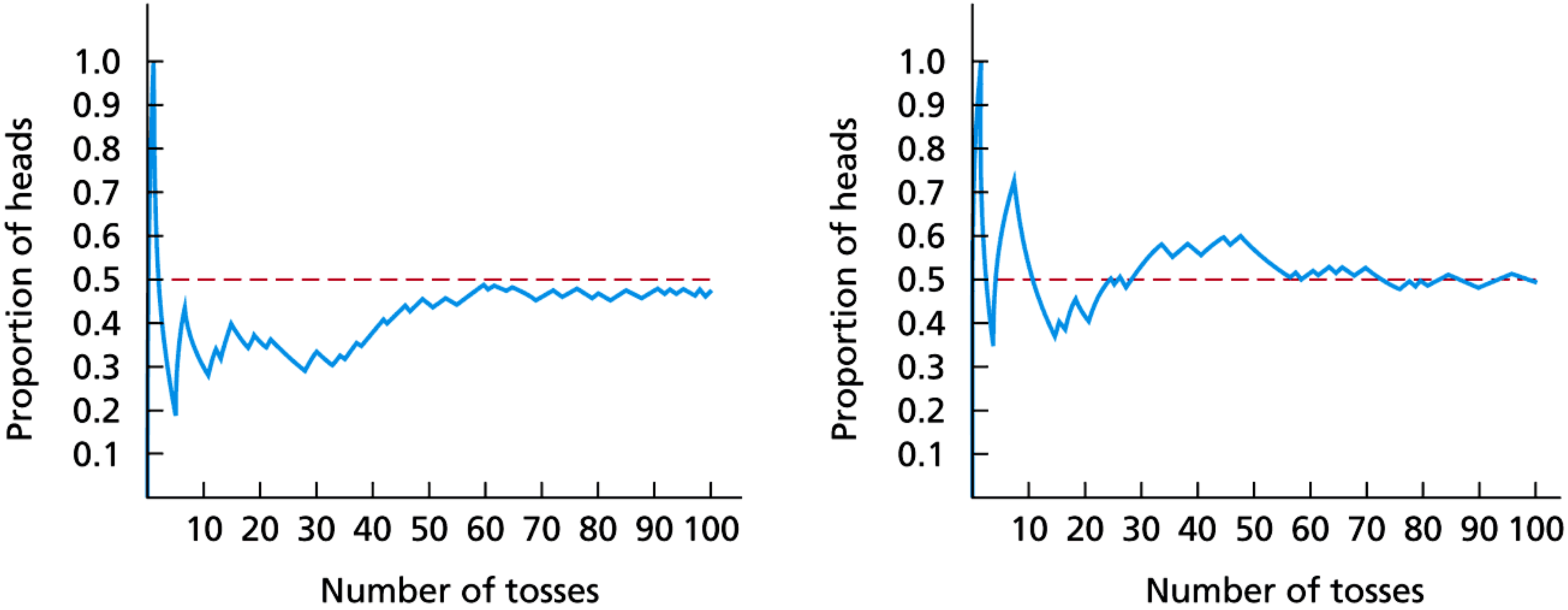
Used a computer to perform two simulations of tossing a balanced coin 100 times. The results are displayed in Fig. 4.2.

Each graph shows the number of tosses of the coin versus the proportion of heads. Both graphs seem to corroborate the frequentist interpretation.

The **frequentist interpretation of probability** construes the probability of an event to be the proportion of times it occurs in a large number of repetitions of the experiment.

Figure 4.2

Two computer simulations of tossing a balanced coin 100 times



Long-Run Behavior

To illustrate the ***long-run behavior***:

1. Consider an experiment in which we toss the coin several times and record the number of heads
2. A trial is a set of 10 tosses
3. Graph the relative frequency and cumulative relative frequency of occurrence of a head
4. A cumulative graph demonstrates the idea of long-run behavior
5. This cumulative graph suggests a stabilizing, or settling down, effect on the observed cumulative probability
6. This stabilizing effect, or long-term average value, is often referred to as the law of large numbers

Example

Example: Consider tossing a fair coin. Define the event H as the occurrence of a head. What is the probability of the event H, $P(H)$?

1. In a single toss of the coin, there are two possible outcomes

2. Since the coin is fair, each outcome (side) should have an equally likely chance of occurring

3. Intuitively, $P(H) = 1/2$ (the expected relative frequency)

Notes:

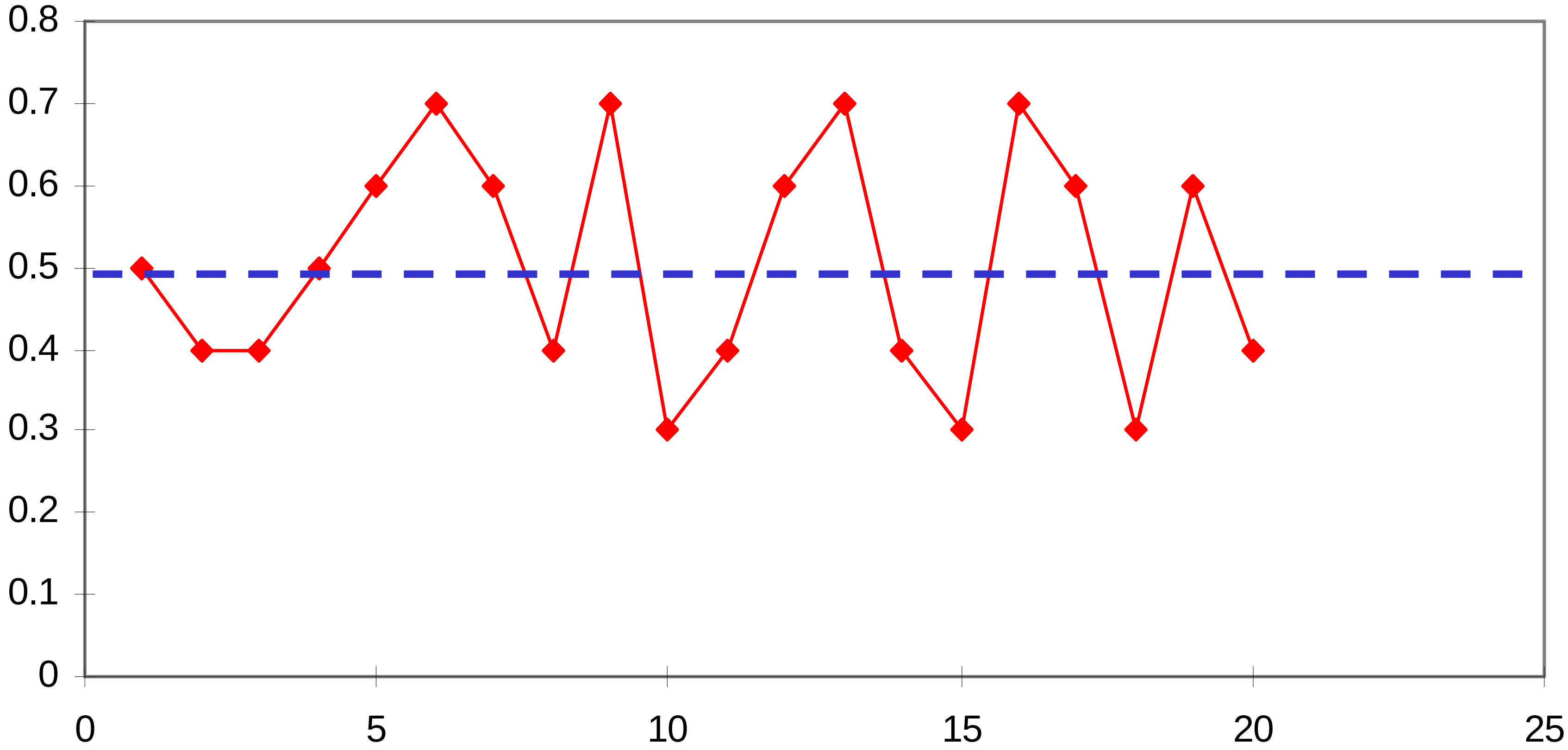
- This does not mean exactly one head will occur in every two tosses of the coin
- In the long run, the proportion of times that a head will occur is approximately 1/2

Experiment

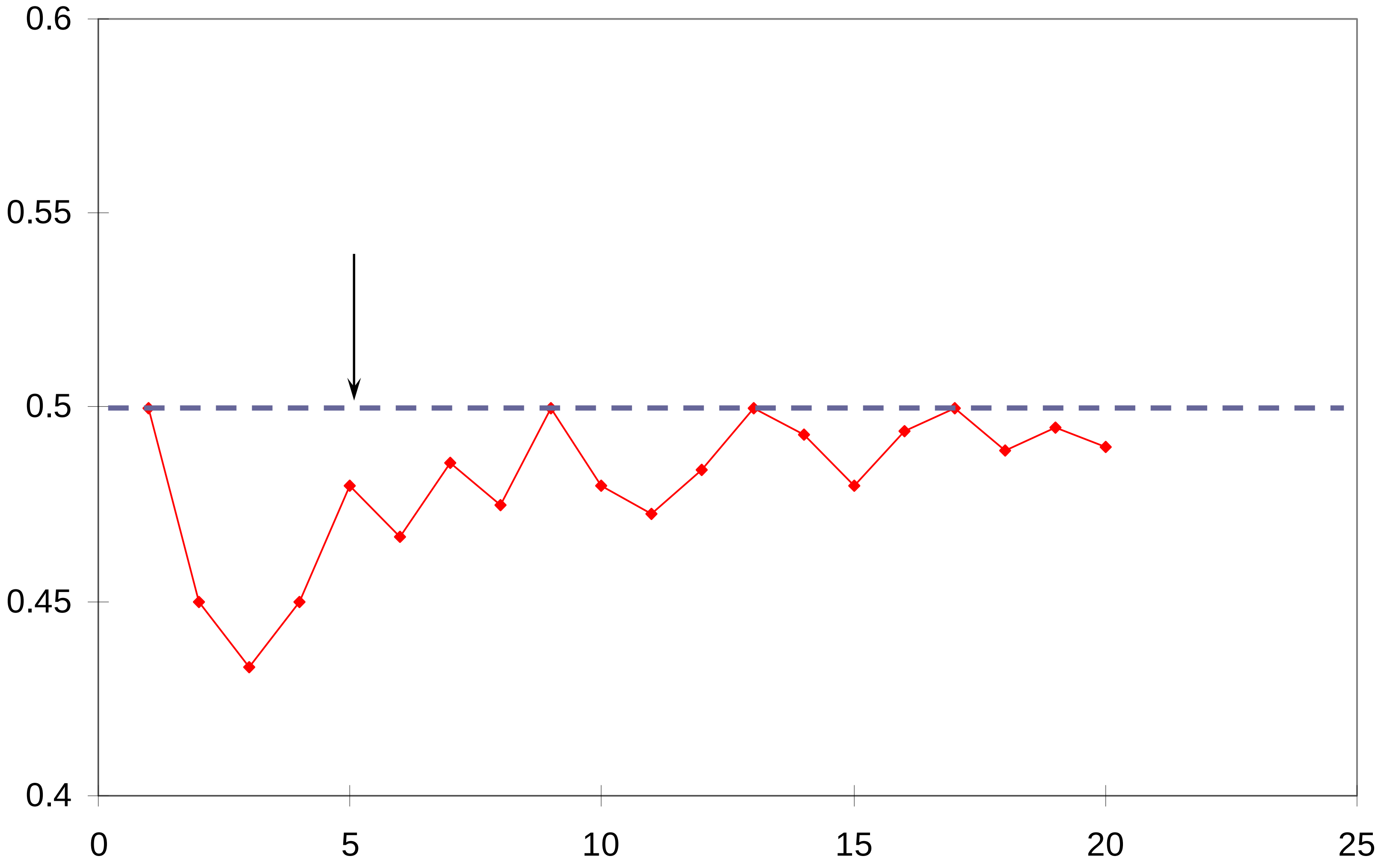
Experimental results of tossing a coin 10 times each trial:

Trial	Number of Heads Observed	Relative Frequency	Cumulative Relative Frequency
1	5	5/10	5/10 = 0.5000
2	4	4/10	9/20 = 0.4500
3	4	4/10	13/30 = 0.4333
4	5	5/10	18/40 = 0.4500
5	6	6/10	24/50 = 0.4800
6	7	7/10	28/60 = 0.4667
7	6	6/10	34/70 = 0.4857
8	4	4/10	38/80 = 0.4750
9	7	7/10	45/90 = 0.5000
10	3	3/10	48/100 = 0.4800
11	4	4/10	52/110 = 0.4727
12	6	6/10	58/120 = 0.4838
13	7	7/10	65/130 = 0.5000
14	4	4/10	69/140 = 0.4929
15	3	3/10	72/150 = 0.4800
16	7	7/10	79/160 = 0.4938
17	6	6/10	85/170 = 0.5000
18	3	3/10	88/180 = 0.4889
19	6	6/10	94/190 = 0.4947
20	4	4/10	98/200 = 0.4900

Relative Frequency



Cumulative Relative Frequency



Law of Large Numbers

Law of Large Numbers: If the number of times an experiment is repeated is increased, the ratio of the number of successful occurrences to the number of trials will tend to approach the theoretical probability of the outcome for an individual trial

– ***Interpretation:*** The law of large numbers says: the larger the number of experimental trials n, the closer the empirical probability $P'(\mathrm{A})$ is expected to be to the true probability $P(\mathrm{A})$

In symbols: As $n \to \infty, \quad P'(\mathrm{A}) \to P(\mathrm{A})$

Key Fact 4.1

Basic Properties of Probabilities

Property 1: The probability of an event is always between 0 and 1, inclusive.

Property 2: The probability of an event that cannot occur is 0. (An event that cannot occur is called an **impossible event.**)

Property 3: The probability of an event that must occur is 1. (An event that must occur is called a **certain event.**)

EXAMPLE 4.4 Basic Properties of Probabilities

Dice Let's return to Example 4.3, in which two balanced dice are rolled. Determine the probability that

a. the sum of the dice is 1.

b. the sum of the dice is 12 or less.

Solution for example:

Solution:

a. Figure 4.1 on page 147 shows that the sum of the dice must be 2 or more. Thus the probability that the sum of the dice is 1 equals $f/N = 0/36 = 0$.

Interpretation Getting a sum of 1 when two balanced dice are rolled is ***impossible*** and hence has probability 0.

b. From Fig. 4.1, the sum of the dice must be 12 or less. Thus the probability of that event equals $f/N = 36/36 = 1$.

Interpretation Getting a sum of 12 or less when two balanced dice are rolled is ***certain*** and hence has probability 1.

Section 4.2

Events

Simple Sample Spaces

We need to talk about data collection and experimentation more precisely

- With many activities, like tossing a coin, rolling a die, selecting a card, there is uncertainty as to what will happen

- We will study and characterize this uncertainty

Definition 4.2

Sample Space and Event

Sample space: The collection of all possible outcomes for an experiment.

Event: A collection of outcomes for the experiment, that is, any subset of the sample space. An event **occurs** if and only if the outcome of the experiment is a member of the event.

Definitions

- An experiment is the observation of some activity or the act of taking some measurement.
- An outcome is the particular result of an experiment.
- An event is the collection of one or more outcomes of an experiment.

Example

- A fair die is rolled once.
- The experiment is rolling the die.
- The possible outcomes are the numbers 1, 2, 3, 4, 5, and 6.
- An event is the occurrence of an even number. That is, we collect the outcomes 2, 4, and 6.

Experiment & Outcome

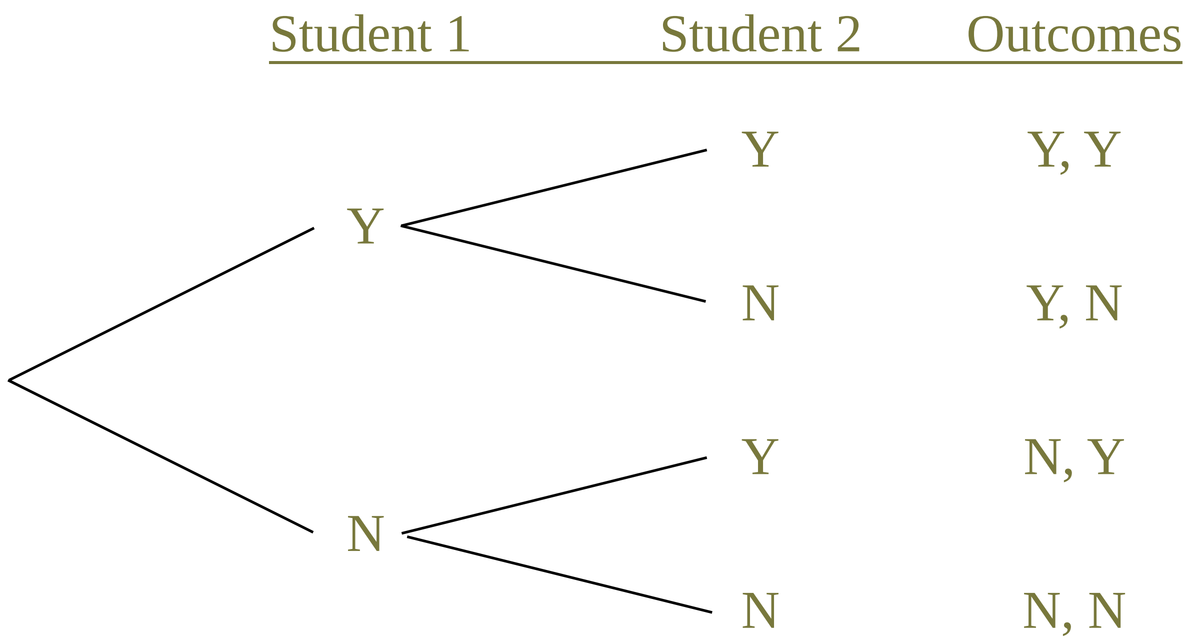
Experiment: Any process that yields a result or an observation

Outcome: A particular result of an experiment

Example: Suppose we select two students at random and ask each if they have a car on campus:

1. A list of possible outcomes: (Y, Y), (Y, N), (N, Y), (N, N)
2. This is called ordered pair notation
3. The outcomes may be displayed using a tree diagram

Tree Diagram



1. This diagram consists of four branches: 2 first generation branches and 4 second generation branches

2. Each branch shows a possible outcome

Sample Space & Event

Sample Space: The set of all possible outcomes of an experiment. The sample space is typically called S and may take any number of forms: a list, a tree diagram, a lattice grid system, etc. The individual outcomes in a sample space are called *sample points*. $n(S)$ is the number of sample points in the sample space.

Event: any subset of the sample space. If A is an event, then $n(A)$ is the number of sample points that belong to A

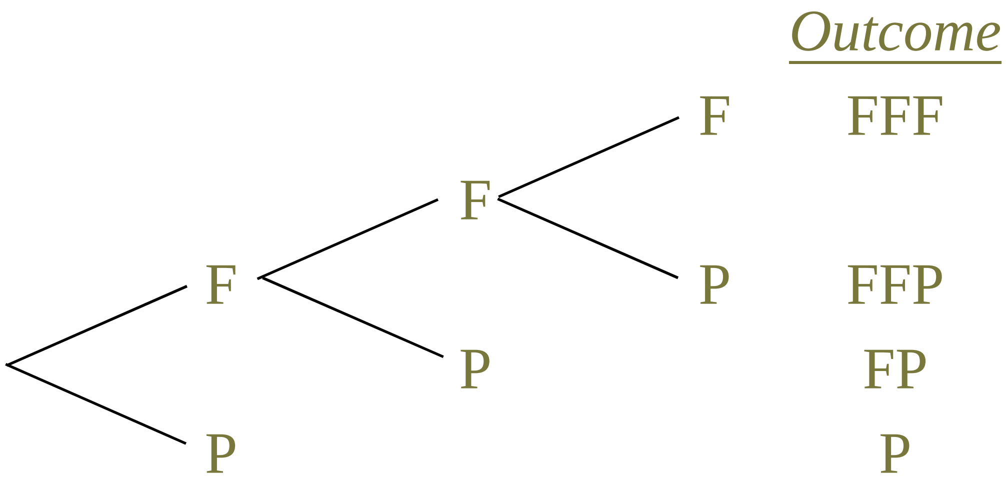
Example: For the student car example above:

$$S = \{ (Y, Y),\ (Y, N),\ (N, Y),\ (N, N) \}$$

$$n(S) = 4$$

Example

Example: An experiment consists of selecting electronic parts from an assembly line and testing each to see if it passes inspection (P) or fails (F). The experiment terminates as soon as one acceptable part is found or after three parts are tested. Construct the sample space:



S = { FFF, FFP, FP, P }

Example

<u>Example</u>: The 1200 students at a local university have been cross tabulated according to resident and college status:

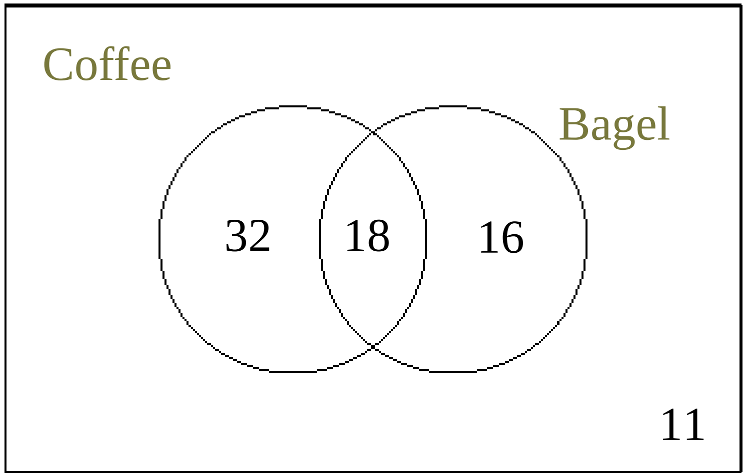
	Arts and Sciences	Business
Resident	600	280
Nonresident	175	145

– The experiment consists of selecting one student at random from the entire student body, what is the sample space?

– $n(S) = 1200$ (i.e. 600+175+280+145)

Example

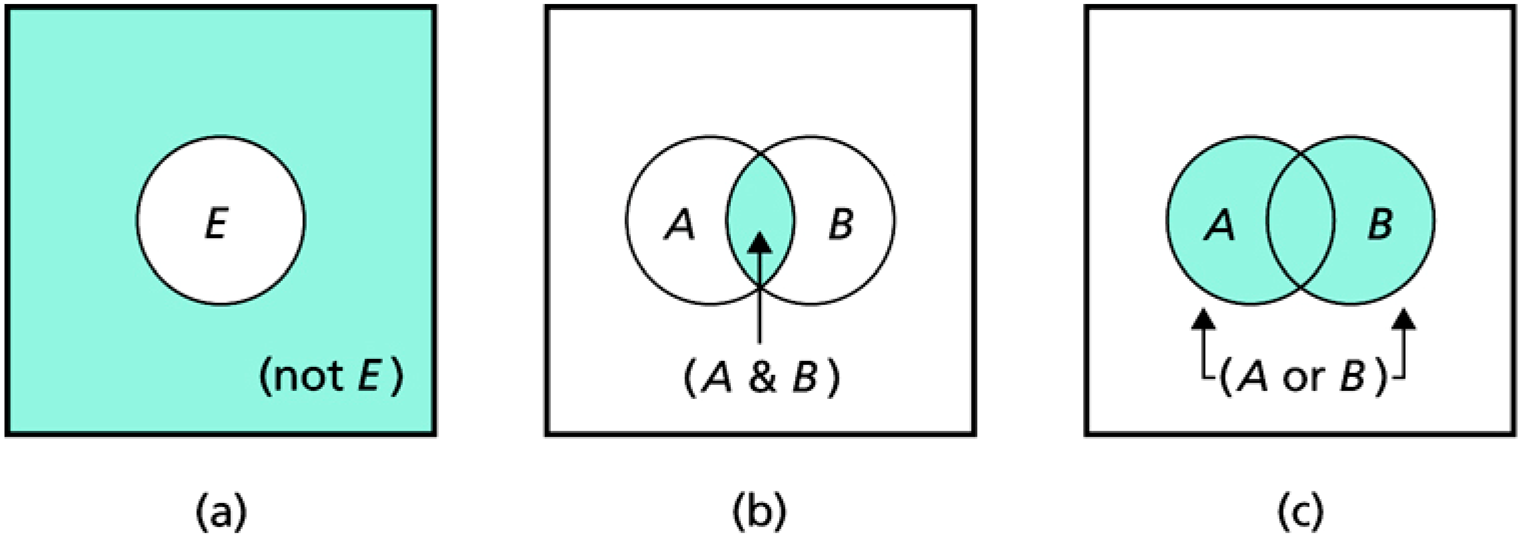
<u>Example</u>: On the way to work, some employees at a certain company stop for a bagel and/or a cup of coffee. The accompanying Venn diagram summarizes the behavior of the employees for a randomly selected work day:



– The experiment consists of selecting one employee at random
– $n(S) = 77$

Figure 4.9

Venn diagrams for (a) event (not E), (b) event (A & B), and (c) event (A or B)



Examples

- ✓ **Example:** An experiment consists of two trials. The first is tossing a penny and observing a head or a tail; the second is rolling a die and observing a 1, 2, 3, 4, 5, or 6. Construct the sample space:

S = { H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6 }

- ✓ Example: Three voters are randomly selected and asked if they favor an increase in property taxes for road construction in the county. Construct the sample space:

S = { NNN, NNY, NYN, NYY, YNN, YNY, YYN, YYY}

Definition 4.3

Relationships Among Events

(not *E*): The event "*E* does not occur"

(*A* & *B*): The event "both *A* and *B* occur"

(*A* or *B*): The event "either *A* or *B* or both occur"

Definition 4.4

Mutually Exclusive Events

Two or more events are **mutually exclusive events** if no two of them have outcomes in common.

Mutually Exclusive Events

Events are mutually exclusive if the occurrence of any one event means that none of the others can occur at the same time.

Events are independent if the occurrence of one event does not affect the occurrence of another.

Figure 4.14

(a) Two mutually exclusive events;
(b) two non–mutually exclusive events

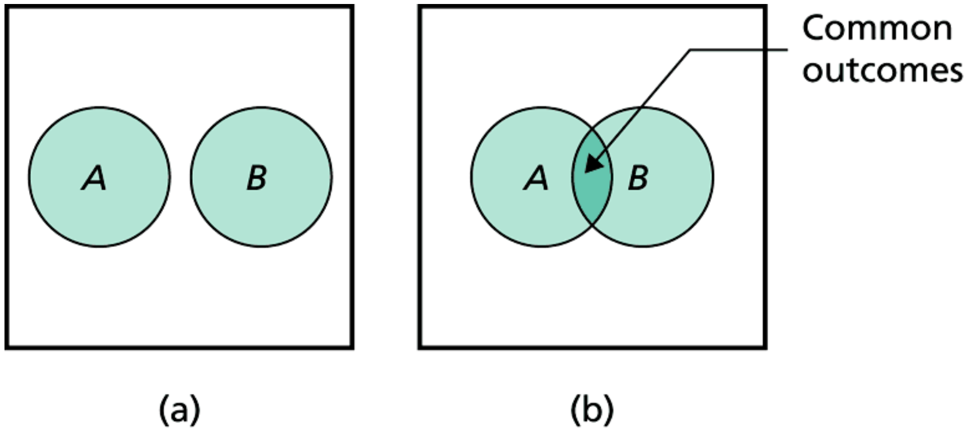
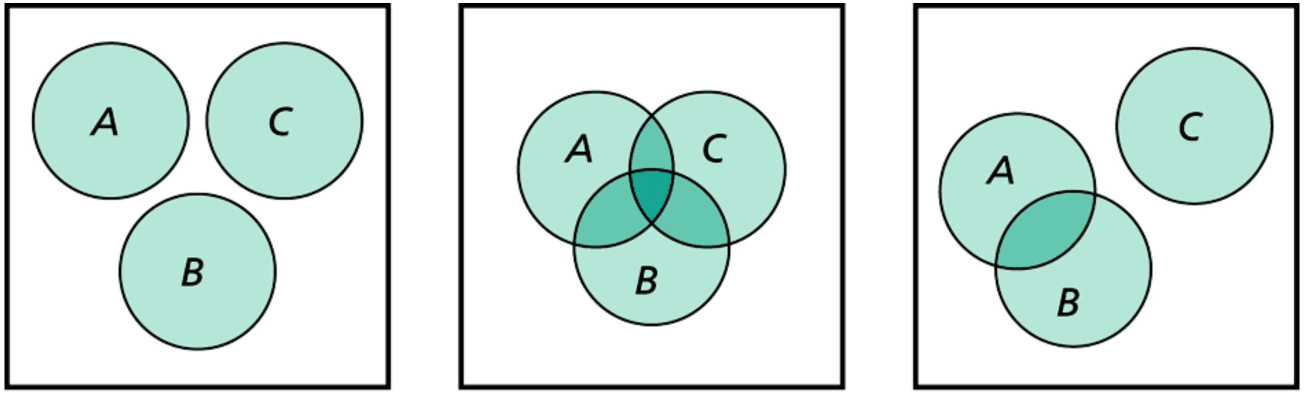


Figure 4.15

(a) Three mutually exclusive events;
(b) three non–mutually exclusive events;
(c) three non–mutually exclusive events



(a) (b) (c)

Collectively Exhaustive Events

Events are collectively exhaustive if at least one of the events must occur when an experiment is conducted.

- i.e. faulty switches and working switches.
- i.e. $x < 50$, $50 \leq x < 100$, and $100 \leq x$.

Notes

1. The outcomes in a sample space can never overlap

2. All possible outcomes must be represented

3. These two characteristics are called *mutually exclusive* and *collectively exhaustive* (all inclusive)

Section 4.3

Some Rules of Probability

Rules of Probability

Consider the concept of probability and relate it to the sample space

Recall: the probability of an event is the relative frequency with which the event could be expected to occur, the long-term average

Equally Likely Events

1. In a sample space, suppose all sample points are equally likely to occur

2. The probability of an event A is the ratio of the number of sample points in A to the number of sample points in S

3. In symbols: $P(\mathrm{A}) = \dfrac{n(\mathrm{A})}{n(S)}$

4. This formula gives a theoretical probability value of event A's occurrence

5. The use of this formula requires the existence of a sample space in which each outcome is equally likely

Example

✓ **Example:** A fair coin is tossed 5 times, and a head (H) or a tail (T) is recorded each time. What is the probability of:

A = {exactly one head in 5 tosses}, and

B = {exactly 5 heads}?

- The outcomes consist of a sequence of 5 Hs and Ts
- A typical outcome includes a mixture of Hs and Ts, like: HHTTH
- There are 32 possible outcomes, all equally likely

1. A = {HTTTT, THTTT, TTHTT, TTTHT, TTTTH}

$$P(\mathrm{A}) = \frac{n(\mathrm{A})}{n(S)} = \frac{5}{32}$$

2. B = {HHHHH}

$$P(\mathrm{B}) = \frac{n(\mathrm{B})}{n(S)} = \frac{1}{32}$$

Basic Probability Ideas

1. Probability represents a relative frequency
2. P(A) is the ratio of the number of times an event can be expected to occur divided by the number of trials
3. The numerator of the probability ratio must be a positive number or zero
4. The denominator of the probability ratio must be a positive number (greater than zero)
5. The number of times an event can be expected to occur in n trials is always less than or equal to the total number of trials, n

Properties

1. The probability of any event A is between 0 and 1:

$$0 \leq P(\mathrm{A}) \leq 1$$

2. The sum of the probabilities of all outcomes in the sample space is 1:

$$\sum_{\text{all outcomes}} P(\mathrm{A}) = 1$$

Notes:

- The probability is zero if the event cannot occur
- The probability is one if the event occurs every time (a sure thing)

Mutually Exclusive Events & the Addition Rule

Compound Events: formed by combining several simple events:

- The probability that either event A or event B will occur: $P(\text{A or B})$
- The probability that both events A and B will occur: $P(\text{A and B})$
- The probability that event A will occur given that event B has occurred: $P(A \mid B)$

Mutually Exclusive Events

Mutually Exclusive Events: Events defined in such a way that the occurrence of one event precludes the occurrence of any of the other events. (In short, if one of them happens, the others cannot happen.)

Notes:

- Complementary events are also mutually exclusive
- Mutually exclusive events are not necessarily complementary

Example

Example: The following table summarizes visitors to a local amusement park:

	All-Day Pass	Half-Day Pass	*Total*
Male	1200	800	2000
Female	900	700	1600
Total	2100	1500	3600

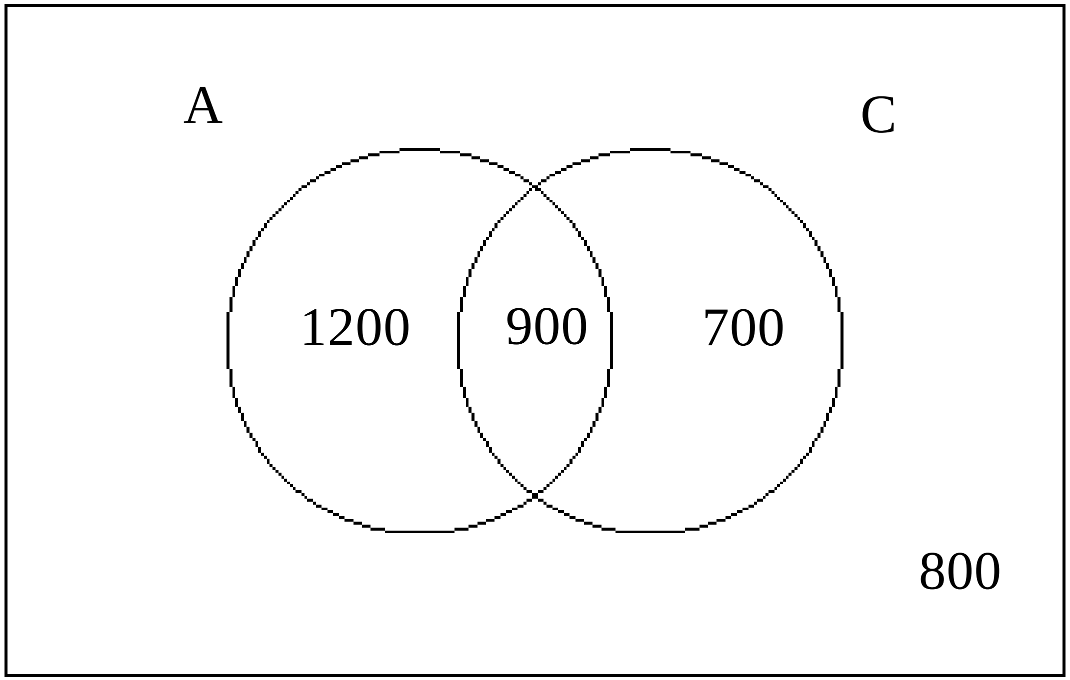
One visitor from this group is selected at random:

1) Define the event A as "the visitor purchased an all-day pass"
2) Define the event B as "the visitor selected purchased a half-day pass"
3) Define the event C as "the visitor selected is female"

Solutions

1) The events A and B are mutually exclusive

2) The events A and C are *not* mutually exclusive. The intersection of A and C can be seen in the table above or in the Venn diagram below:

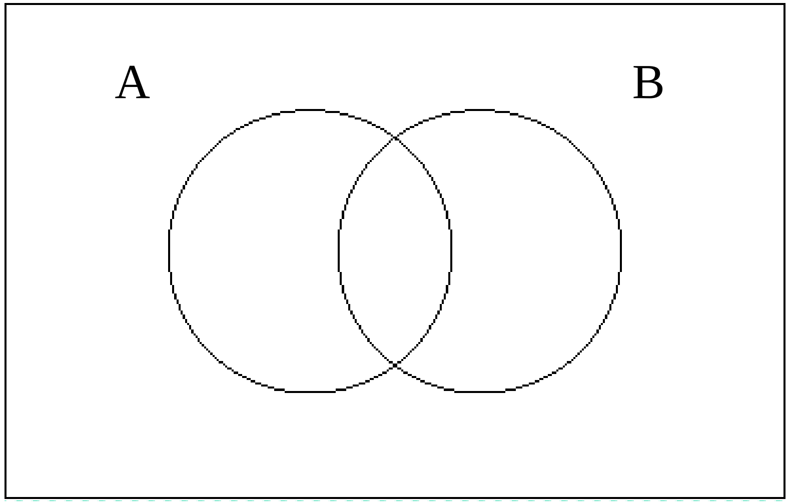


General Addition Rule

General Addition Rule: Let A and B be two events defined in a sample space *S*:

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

- Illustration:



Note: If two events A and B are mutually exclusive: $P(\text{A and B}) = 0$

Formula 4.3

The General Addition Rule

If A and B are any two events, then

$$P(A \text{ or } B) = P(A) + P(B) - P(A \,\&\, B).$$

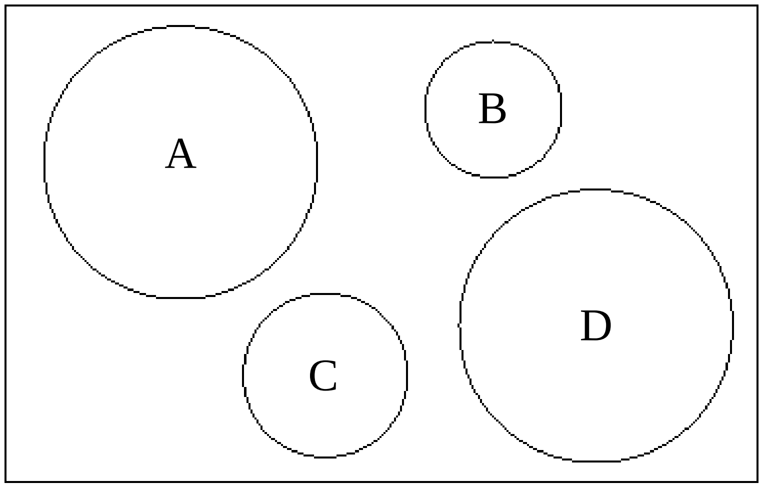
Special Addition Rule

Special Addition Rule: Let A and B be two events defined in a sample space. If A and B are *mutually exclusive* events, then:

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

- This can be expanded to consider *more than two* mutually exclusive events:

$$P(\text{A or B or C}\ldots) = P(\text{A}) + P(\text{B}) + \ldots + P(\text{E})$$



Formula 4.1

The Special Addition Rule

If event A and event B are mutually exclusive, then

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

More generally, if events A, B, C, ... are mutually exclusive, then

$$P(A \text{ or } B \text{ or } C \text{ or } \cdots) = P(A) + P(B) + P(C) + \cdots.$$

Example

- **<u>Example:</u>** On the way to work Bob's personal judgment is that he is four times more likely to get caught in a traffic jam (TJ) than have an easy commute (EC). What values should be assigned to *P*(TJ) and *P*(EC)?

$$P(\text{TJ}) = 4 \cdot P(\text{EC})$$

$$P(\text{TJ}) + P(\text{EC}) = 1$$

$$4 \cdot P(\text{EC}) + P(\text{EC}) = 1$$

$$5 \cdot P(\text{EC}) = 1$$

$$P(\text{EC}) = \frac{1}{5}$$

$$P(\text{TJ}) = 4 \cdot P(\text{EC}) = 4\left(\frac{1}{5}\right) = \frac{4}{5}$$

Odds

Odds: another way of expressing probabilities

- If the *odds in favor of an event* A are *a* to *b*, then:

  1. The odds against A are *b* to *a*

  2. The probability of event A is: $P(\mathrm{A}) = \dfrac{a}{a+b}$

  3. The probability that event A will not occur is

$$P(\text{A does not occur}) = \frac{b}{a+b}$$

Example

✓ **Example:** The odds in favor of you passing a course are 11 to 3. Find the probability you will pass and the probability you will fail.

- Using the preceding notation: $a = 11$ and $b = 3$:

$$P(\text{pass}) = \frac{11}{11+3} = \frac{11}{14} \qquad P(\text{fail}) = \frac{3}{11+3} = \frac{3}{14}$$

Complement of An Event

Complement of an Event: The set of all sample points in the sample space that do not belong to event A. The complement of event A is denoted by $\neg A, or \sim A, or\ \bar{A}\ , or\ (not\ A)$ (read "A complement").

Formula 4.2

The Complementation Rule

For any event E,

$$P(E) = 1 - P(\text{not } E).$$

Example

Examples:

1. The complement of the event “success” is “failure”
2. The complement of the event “rain” is “no rain”
3. The complement of the event “at least 3 patients recover” out of 5 patients is “2 or fewer recover”

Notes:

- $P(\text{A}) + P(\overline{\text{A}}) = 1$ for any event A
- $P(\overline{\text{A}}) = 1 - P(\text{A})$
- Every event A has a complementary event $\overline{\text{A}}$
- Complementary probabilities are very useful when the question asks for the probability of “at least one.”

Example

Example: A fair coin is tossed 5 times, and a head (H) or a tail (T) is recorded each time. What is the probability of

1) B = {at most 3 heads in 5 tosses}

2) A = {at least one head in 5 tosses}

Solution: (1) $P(\text{B}) = 1 - P(\overline{\text{B}})$

$$= 1 - P(4 \text{ or } 5 \text{ heads})$$

$$= 1 - (P(4 \text{ heads}) + P(5 \text{ heads}))$$

$$= 1 - \left(\frac{5}{32} + \frac{1}{32}\right) = 1 - \frac{6}{32} = \frac{26}{32} = \frac{13}{16}$$

(2) $P(\text{A}) = 1 - P(\overline{\text{A}})$

$$= 1 - P(0 \text{ heads in } 5 \text{ tosses})$$

$$= 1 - \frac{1}{32} = \frac{31}{32}$$

Example

Example: All employees at a certain company are classified as only one of the following: manager (A), service (B), sales (C), or staff (D). It is known that $P(\text{A}) = 0.15$, $P(\text{B}) = 0.40$, $P(\text{C}) = 0.25$, and $P(\text{D}) = 0.20$

$$P(\overline{\text{A}}) = 1 - P(\text{A}) = 1 - 0.15 = 0.85$$

$P(\text{A and B}) = 0$ i.e. A and B are mutually exclusive

$$P(\text{B or C}) = P(\text{B}) + P(\text{C}) = 0.40 + 0.25 = 0.65$$

$$P(\text{A or B or C}) = P(\text{A}) + P(\text{B}) + P(\text{C})$$
$$= 0.15 + 0.40 + 0.25 = 0.80$$

Example

Example: A consumer is selected at random. The probability the consumer has tried a snack food (F) is 0.5, tried a new soft drink (D) is 0.6, and tried both the snack food and the soft drink is 0.2: Then we have

(1). $P(\text{Tried the snack food or the soft drink})$

$=P(F\ or\ D) = P(F) + P(D) - P(F\ and\ D)$

$=0.5+0.6-0.2 = 0.9$

(2). $P(\text{Not tried the snack food})$

$=P(\bar{F}) = 1 - P(F)$

$=1-0.5 = 0.5$

(3). $P(\text{Tried neither the snack food nor the soft drink})$

$=P(\overline{F\ or\ D}) = 1 - P(F\ or\ D)$

$=1-0.9 = 0.1$

(4). $P(\text{Tried only the soft drink})$

$=P(D) - P(F\ and\ D) = 0.6-0.2 = 0.4$

Section 4.4

Contingency Tables; Joint and Marginal Probabilities

Bivariate Data

Bivariate Data: Consists of the values of two different response variables that are obtained from the same population of interest

Three combinations of variable types:

1. Both variables are qualitative (*attribute*). We will cover here!

Later stages for these two:

2. One variable is qualitative (*attribute*) and the other is quantitative (*numerical*)
3. Both variables are quantitative (*both numerical*)

Example

Example: A local automobile dealer classifies purchases by number of doors and transmission type. The table below gives the number of each classification.

	Manual Transmission	Automatic Transmission
2-door	75	155
4-door	85	170

If one customer is selected at random, find the probability that:

1) The selected individual purchased a car with automatic transmission
2) The selected individual purchased a 2-door car

Solutions

	Manual Transmission	Automatic Transmission
2-door	75	155
4-door	85	170

1) $P(\text{Automatic Transmission})$

$$= \frac{155+170}{75+85+155+170} = \frac{325}{485} = \frac{65}{97}$$

2) $P(2\text{-door})$

$$= \frac{75+155}{75+85+155+170} = \frac{230}{485} = \frac{46}{97}$$

Two Qualitative Variables

When bivariate data results from two qualitative (attribute or categorical) variables, the data is often arranged on a cross-tabulation or contingency table

➔ Example: A survey was conducted to investigate the relationship between preferences for television, radio, or newspaper for national news, and gender. The results are given in the table below:

	TV	Radio	NP
Male	280	175	305
Female	115	275	170

Marginal Totals

This table may be extended to display the ***marginal totals***. The total of the marginal totals is the ***grand total*****:**

	TV	Radio	NP	Row Totals
Male	280	175	305	760
Female	115	275	170	560
Col. Totals	395	450	475	1320

Note: Contingency tables often show percentages (relative frequencies). These percentages are based on the entire sample or on the subsample (row or column) classifications.

Percentages Based on the Grand Total

- The previous contingency table may be converted to percentages of the grand total by dividing each frequency by the grand total and multiplying by 100

For example, 175 becomes 13.3%, i.e.

$$\frac{175}{1320} \times 100\% = 13.3\%$$

	TV	Radio	NP	Row Totals
Male	21.2	**13.3**	23.1	57.6
Female	8.7	20.8	12.9	42.4
Col. Totals	29.9	34.1	36.0	100.0

Percentages Based on Row Totals and Column Totals

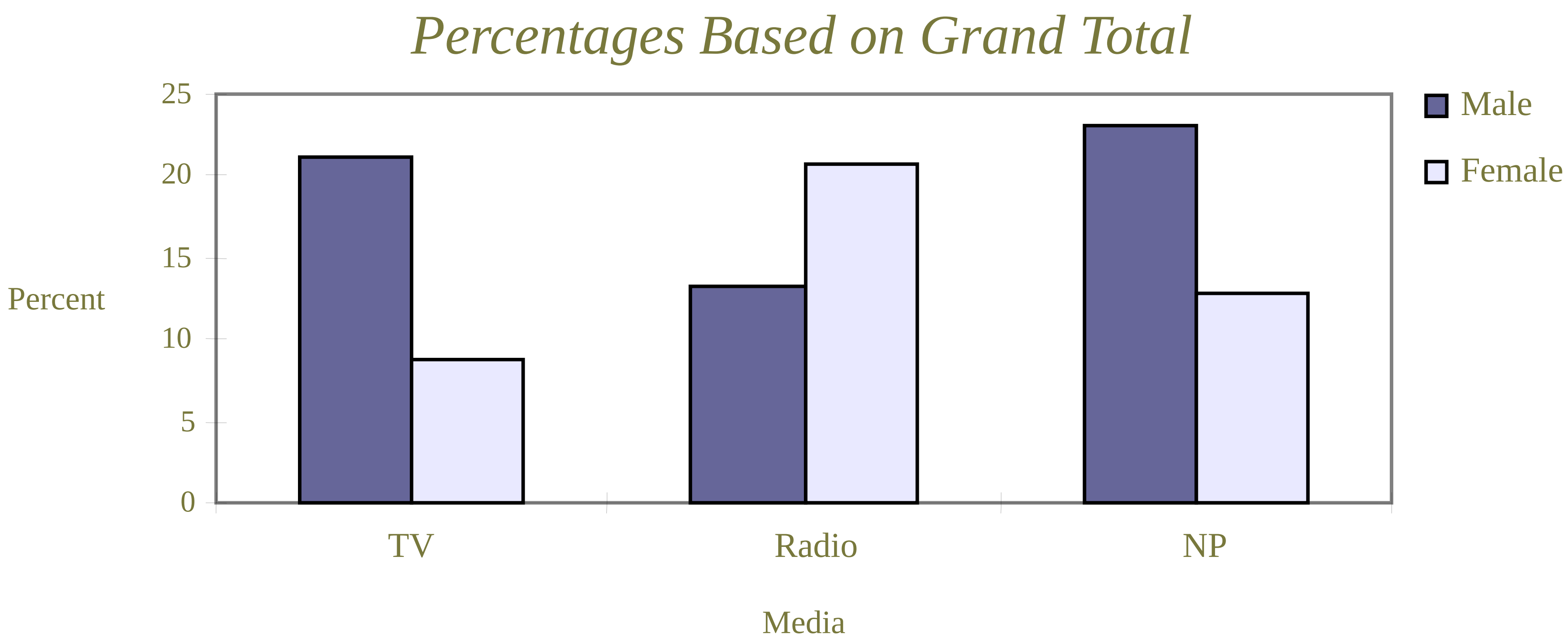
- The entries in a contingency table may also be expressed as percentages of the row (column) totals by dividing each row (column) entry by that row's (column's) total and multiplying by 100. The entries in the contingency table below are expressed as percentages of the column totals:

	TV	Radio	NP	Row Totals
Male	70.9	38.9	64.2	57.6
Female	29.1	61.1	35.8	42.4
Col. Totals	100.0	100.0	100.0	100.0

Note: These statistics may also be displayed in a side-by-side bar graph

Illustration

- These same statistics (numerical values describing sample results) can be shown in a (side-by-side) bar graph:



EXAMPLE 4.14 Introducing Contingency Tables

Age and Rank of Faculty Data about two variables—age and rank—of the faculty members at a university yielded the contingency table shown in Table 4.6. Discuss and interpret the numbers in the table.

Table 4.6 Contingency table for age and rank of faculty members

		Rank				
		Full professor R_1	Associate professor R_2	Assistant professor R_3	Instructor R_4	**Total**
Age (yr)	Under 30 A_1	2	3	57	6	68
	30–39 A_2	52	170	163	17	402
	40–49 A_3	156	125	61	6	348
	50–59 A_4	145	68	36	4	253
	60 & over A_5	75	15	3	0	93
	Total	430	381	320	33	1164

EXAMPLE 4.15 Joint and Marginal Probabilities

Age and Rank of Faculty Refer to Example 4.14. Suppose that a faculty member is selected at random.

a. Identify the events represented by the subscripted letters that label the rows and columns of the contingency table shown in Table 4.6.

b. Identify the events represented by the cells of the contingency table.

c. Determine the probabilities of the events discussed in parts (a) and (b).

d. Summarize the results of part (c) in a table.

e. Discuss the relationship among the probabilities in the table obtained in part (d).

Table 4.7 Joint probability distribution corresponding to Table 4.6

		Rank				
		Full professor R_1	Associate professor R_2	Assistant professor R_3	Instructor R_4	$P(A_i)$
Age (yr)	Under 30 A_1	0.002	0.003	0.049	0.005	0.058
	30–39 A_2	0.045	0.146	0.140	0.015	0.345
	40–49 A_3	0.134	0.107	0.052	0.005	0.299
	50–59 A_4	0.125	0.058	0.031	0.003	0.217
	60 & over A_5	0.064	0.013	0.003	0.000	0.080
	$P(R_j)$	0.369	0.327	0.275	0.028	1.000

More on Bivariate Data

One Qualitative & One Quantitative Variable

1. When bivariate data results from one qualitative and one quantitative variable, the quantitative values are viewed as separate samples

2. Each set is identified by levels of the qualitative variable
3. Each sample is described using summary statistics, and the results are displayed for side-by-side comparison

4. Statistics for comparison: measures of central tendency, measures of variation, 5-number summary

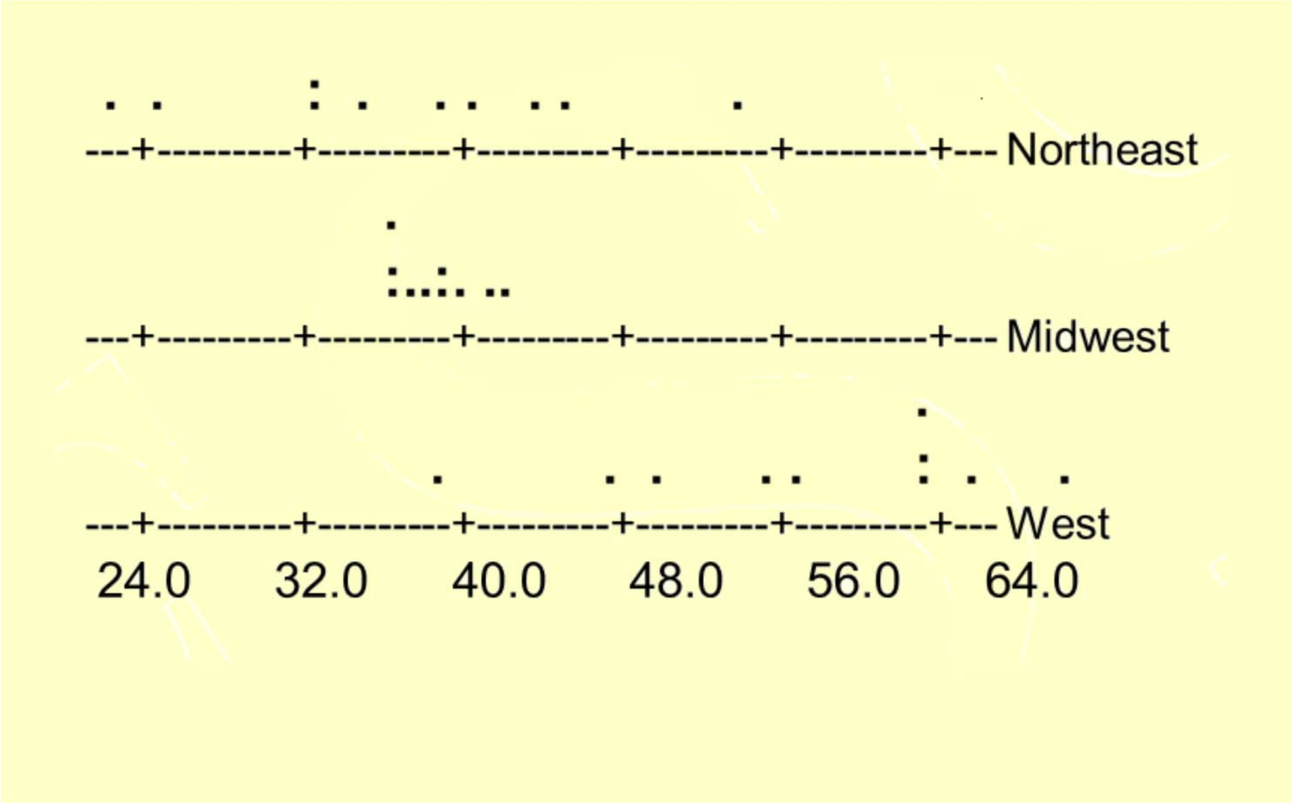
5. Graphs for comparison: dotplot, boxplot

Example: A random sample of households from three different parts of the country was obtained and their electric bill for June was recorded. The data is given in the table below:

Northeast		Midwest		West	
23.75	40.50	34.38	34.35	54.54	65.60
33.65	31.25	39.15	37.12	59.78	45.12
42.55	50.60	36.71	34.39	60.35	61.53
37.70	31.55	35.12	35.80	52.79	47.37
38.85	21.25	37.24	40.01	59.64	37.40

- The part of the country is a qualitative variable with three levels of response. The electric bill is a quantitative variable. The electric bills may be compared with numerical and graphical techniques.

Comparison Using Dotplots



- The electric bills in the Northeast tend to be more spread out than those in the Midwest. The bills in the West tend to be higher than both those in the Northeast and Midwest.

Comparison Using Box-and-Whisker Plots

The Monthly Electric Bill