STA 315
Probability & Statistics

Dr. Mouhamad Ibrahim

mouhamaad.ibrahim@gmail.com

Chapter 1
Introduction to statistics and data analysis

Chapter 1: Introduction to statistics and data analysis

– An overview of statistical terminology

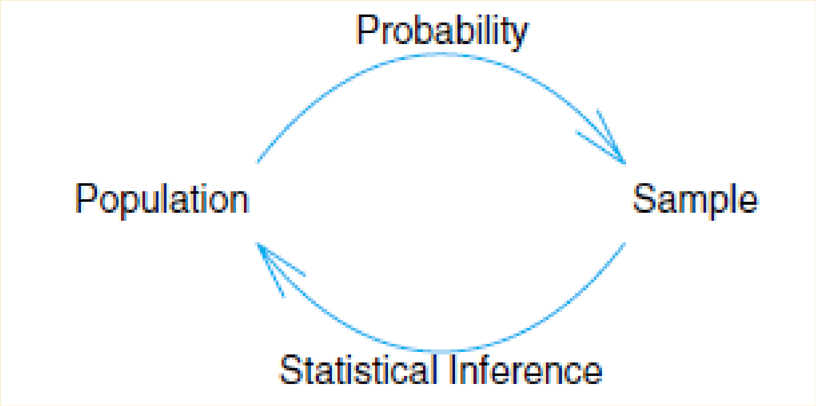
- We introduce the concepts of discrete and continuous data
- We introduce the techniques for summarizing data
- **Exploring Univariate Data**
  - Types of data
  - Mean and Median
  - Standard Deviation and Variance
  - Range, IQR and Finding Outliers
  - Graphs and Describing Distributions

1.1 Statistical Inference, Samples, Populations, and the Role of Probability

- Use of Scientific Data
  - Gathering of data is nothing new, distinction between **collection** of scientific information and **inferential statistics**
  - Inferential statistics: **uncertainty** and **variation** to improve the **quality** of the process
    - using techniques that allow us to go beyond merely reporting data to drawing conclusions (or inferences) about the scientific system.
  - **Sources of variation**
    - Ex: measure sulfur monoxide in the air during pollution studies
- Variability in Scientific Data
  - Information is gathered in the form of **samples**, or collections of **observations**
  - Samples are collected from **populations**
    - **Experimental design: factors**
    - **Observational study: factor levels** can not be preselected

1.1 Statistical Inference, Samples, Populations, and the Role of Probability

- The Role of Probability
  - The sample along with inferential statistics allows us to draw conclusions about the population, with inferential statistics making clear use of elements of probability
  - Elements in probability allow us to draw conclusions about characteristics of hypothetical data taken from the population, based on known features of the population



1.2 Sampling Procedures; Collection of Data

- Simple Random Sampling

- Experimental Design

Table 1.1: Data Set for Example 1.2

No Nitrogen	Nitrogen
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46

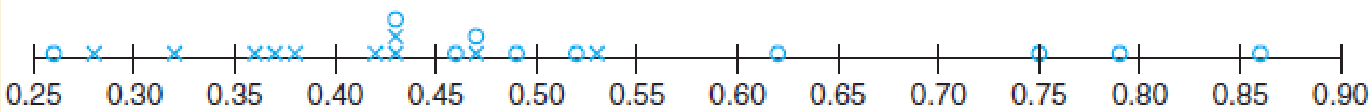


Figure 1.1: A dot plot of stem weight data.

1.3 Measures of Location: The Sample Mean and Median

- Measures of location are designed to provide the analyst with some quantitative values of where the center, or some other location, of data is located
  - In Example 1.2, it appears as if the center of the nitrogen sample clearly exceeds that of the no-nitrogen sample

- Sample Mean
  - The mean is simply a numerical average

Definition 1.1: Suppose that the observations in a sample are $x_1, x_2, \ldots, x_n$. The **sample mean**, denoted by $\bar{x}$, is

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

1.3 Measures of Location: The Sample Mean and Median

- Sample Median
  - The purpose of the sample median is to reflect the central tendency of the sample in such a way that it is uninfluenced by extreme values or outliers

Definition 1.2: Given that the observations in a sample are $x_1, x_2, \ldots, x_n$, arranged in **increasing order** of magnitude, the sample median is

$$\tilde{x} = \begin{cases} x_{(n+1)/2}, & \text{if } n \text{ is odd,} \\ \frac{1}{2}(x_{n/2} + x_{n/2+1}), & \text{if } n \text{ is even.} \end{cases}$$

As an example, suppose the data set is the following: 1.7, 2.2, 3.9, 3.11, and 14.7. The sample mean and median are, respectively,

$$\bar{x} = 5.12, \quad \tilde{x} = 3.9.$$

Clearly, the mean is influenced considerably by the presence of the extreme observation, 14.7, whereas the median places emphasis on the true "center" of the data

Example 1.2 ?

1.3 Measures of Location: The Sample Mean and Median

- Other Measures of Locations
  - Trimmed means
    - A trimmed mean is computed by “trimming away” a certain percent of both the largest and the smallest set of values
    - For example, the 10% trimmed mean is found by eliminating the largest 10% and smallest 10% and computing the average of the remaining values.

Example 1.2 ?

1.3 Measures of Location: The Sample Mean and Median

- **Exercises**

1.1 The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint.

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

Assume that the measurements are a simple random sample.

(a) What is the sample size for the above sample?

(b) Calculate the sample mean for these data.

(c) Calculate the sample median.

(d) Plot the data by way of a dot plot.

(e) Compute the 20% trimmed mean for the above data set.

(f) Is the sample mean for these data more or less descriptive as a center of location than the trimmed mean?

1.3 Measures of Location: The Sample Mean and Median

- **Exercises**

1.5 Twenty adult males between the ages of 30 and 40 participated in a study to evaluate the effect of a specific health regimen involving diet and exercise on the blood cholesterol. Ten were randomly selected to be a control group, and ten others were assigned to take part in the regimen as the treatment group for a period of 6 months. The following data show the reduction in cholesterol experienced for the time period for the 20 subjects:

Control group:	7	3	−4	14	2
	5	22	−7	9	5
Treatment group:	−6	5	9	4	4
	12	37	5	3	3

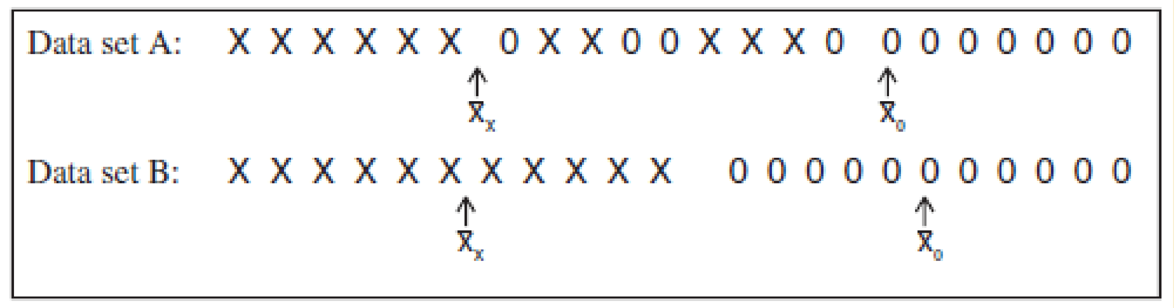
(a) Do a dot plot of the data for both groups on the same graph.

(b) Compute the mean, median, and 10% trimmed mean for both groups.

(c) Explain why the difference in means suggests one conclusion about the effect of the regimen, while the difference in medians or trimmed means suggests a different conclusion.

1.4 Measures of Variability

- Sample variability plays an important role in data analysis
  - Example, contrast the two data sets below;
    - Each contains two samples and the difference in the means is roughly the same for the two samples, but data set B seems to provide a much sharper contrast between the two populations.
    - If the purpose of such an experiment is to detect differences between the two populations, the task is accomplished in the case of data set B.
    - However, in data set A it is not clear that there is a distinction between the two populations.



Example 1.2 ?

1.4 Measures of Variability

- Sample range $X_{max} - X_{min}$
- Sample standard deviation

Definition 1.3: The **sample variance**, denoted by s^2, is given by

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}.$$

The **sample standard deviation**, denoted by s, is the positive square root of s^2, that is,

$$s = \sqrt{s^2}.$$

Example 1.4: In an example discussed extensively in Chapter 10, an engineer is interested in testing the "bias" in a pH meter. Data are collected on the meter by measuring the pH of a neutral substance (pH = 7.0). A sample of size 10 is taken, with results given by

7.07 7.00 7.10 6.97 7.00 7.03 7.01 7.01 6.98 7.08.

1.4 Measures of Variability

- **Which Variability Measure Is More Important ?**

- Population parameters
  - Two important parameters are the population mean and the population variance.
  - The sample variance plays an explicit role in the statistical methods used to draw inferences about the population variance.
  - The sample standard deviation has an important role along with the sample mean in inferences that are made about the population mean.
  - In general, the variance is considered more in inferential theory, while the standard deviation is used more in applications.

1.4 Measures of Variability

- **Exercises**

1.7 Consider the drying time data for Exercise 1.1 on page 13. Compute the sample variance and sample standard deviation.

1.11 Consider the data in Exercise 1.5 on page 13. Compute the sample variance and the sample standard deviation for both control and treatment groups.

1.5 Discrete and Continuous Data

- **Discrete** or **continuous**, depending on the area of application
  - For example,
    - A chemical engineer may be interested in conducting an experiment that will lead to conditions where yield is maximized. Here, of course, the yield may be in percent or grams/pound, measured on a continuum.

    - On the other hand, a toxicologist conducting a combination drug experiment may encounter data that are binary in nature (i.e., the patient either responds or does not).
      - Often the measure that is used in the analysis is the **sample proportion.**
      - If there are n units involved in the data and **x** is defined as the number that fall into category 1, then **n − x** fall into category 2.
      - **x/n** is the **sample proportion** in category 1, and **1 − x/n** is the sample proportion in category 2.

1.5 Discrete and Continuous Data

- **Exercises**

1.17 A study of the effects of smoking on sleep patterns is conducted. The measure observed is the time, in minutes, that it takes to fall asleep. These data are obtained:

Smokers:	69.3	56.0	22.1	47.6
	53.2	48.1	52.7	34.4
	60.2	43.8	23.2	13.8
Nonsmokers:	28.6	25.1	26.4	34.9
	29.8	28.4	38.5	30.2
	30.6	31.8	41.6	21.1
	36.0	37.9	13.9	

(a) Find the sample mean for each group.

(b) Find the sample standard deviation for each group.

(c) Make a dot plot of the data sets A and B on the same line.

(d) Comment on what kind of impact smoking appears to have on the time required to fall asleep.

1.5 Discrete and Continuous Data

- **Qxercises**

1.21 The lengths of power failures, in minutes, are recorded in the following table.

22	18	135	15	90	78	69	98	102
83	55	28	121	120	13	22	124	112
70	66	74	89	103	24	21	112	21
40	98	87	132	115	21	28	43	37
50	96	118	158	74	78	83	93	95

(a) Find the sample mean and sample median of the power-failure times.

(b) Find the sample standard deviation of the power-failure times.

Chapter 2
Probability

Chapter 2: Probability

- Definition of a sample space, events, complements
- Intersection, exclusive events, union of events, Venn diagrams
- Probability, additive rules, multiplicative rules, conditional probability, Bayes' rule

- Will learn how to compute probability
  - Counting Techniques, Combinations and Permutations
  - Sets and Venn Diagrams
  - Basic Probability Models
  - General Probability Rules

2.1 Sample Space

- Study of statistics: presentation and interpretation of **chance outcomes**

2.1: The set of all possible outcomes of a statistical experiment is called the **sample space** and is represented by the symbol S.

Each outcome in a sample space is called an **element** or a **member** of the sample space, or simply a **sample point**. If the sample space has a finite number of elements, we may *list* the members separated by commas and enclosed in braces. Thus, the sample space S, of possible outcomes when a coin is flipped, may be written

$$S = \{H, T\},$$

where H and T correspond to heads and tails, respectively.

2.1 Sample Space

- Study of statistics: presentation and interpretation of **chance outcomes**

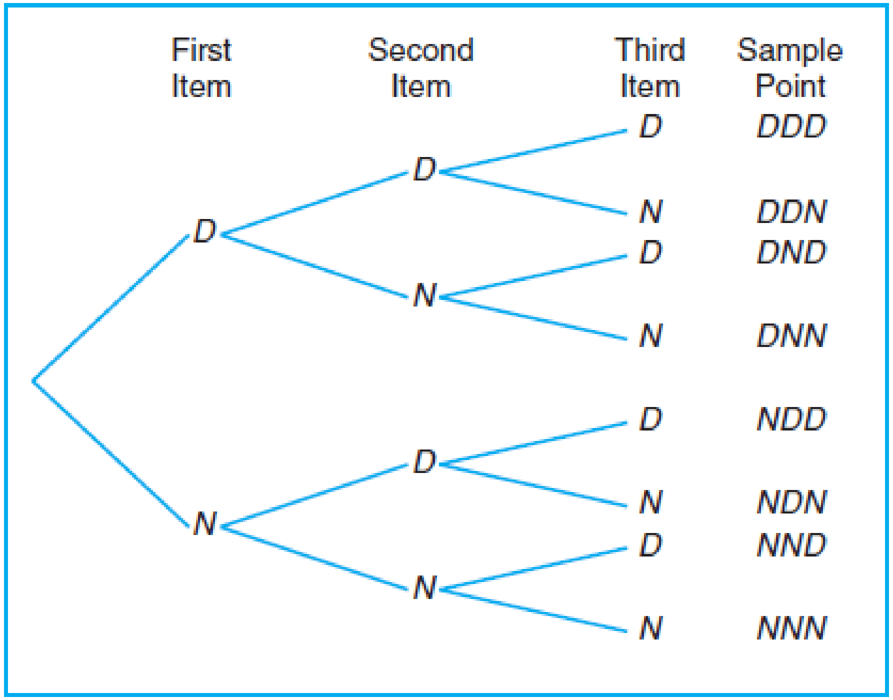


Figure 2.2: Tree diagram for Example 2.3.

2.2 Events

An **event** is a subset of a sample space.

Consider the experiment of tossing a die. If we are interested in the number that shows on the top face, the sample space is

$$S_1 = \{1, 2, 3, 4, 5, 6\}.$$

For any given experiment, we may be interested in the occurrence of certain **events** rather than in the occurrence of a specific element in the sample space. For instance, we may be interested in the event A that the outcome when a die is tossed is divisible by 3. This will occur if the outcome is an element of the subset $A = \{3, 6\}$

The **complement** of an event A with respect to S is the subset of all elements of S that are not in A. We denote the complement of A by the symbol A'.

Example 2.5: Let R be the event that a red card is selected from an ordinary deck of 52 playing cards, and let S be the entire deck. Then R' is the event that the card selected from the deck is not a red card but a black card.

2.2 Events

The **intersection** of two events A and B, denoted by the symbol $A \cap B$, is the event containing all elements that are common to A and B.

Example 2.7: Let E be the event that a person selected at random in a classroom is majoring in engineering, and let F be the event that the person is female. Then $E \cap F$ is the event of all female engineering students in the classroom.

Two events A and B are **mutually exclusive**, or **disjoint**, if $A \cap B = \phi$, that is, if A and B have no elements in common.

The **union** of the two events A and B, denoted by the symbol $A \cup B$, is the event containing all the elements that belong to A or B or both.

Example 2.10: Let $A = \{a, b, c\}$ and $B = \{b, c, d, e\}$; then $A \cup B = \{a, b, c, d, e\}$.

Example 2.12: If $M = \{x \mid 3 < x < 9\}$ and $N = \{y \mid 5 < y < 12\}$, then

$$M \cup N = \{z \mid 3 < z < 12\}.$$

2.2 Events

The relationship between events and the corresponding sample space can be illustrated graphically by means of **Venn diagrams**. In a Venn diagram we let the sample space be a rectangle and represent events by circles drawn inside the rectangle. Thus, in Figure 2.3, we see that

$$A \cap B = \text{ regions 1 and 2,}$$
$$B \cap C = \text{ regions 1 and 3,}$$

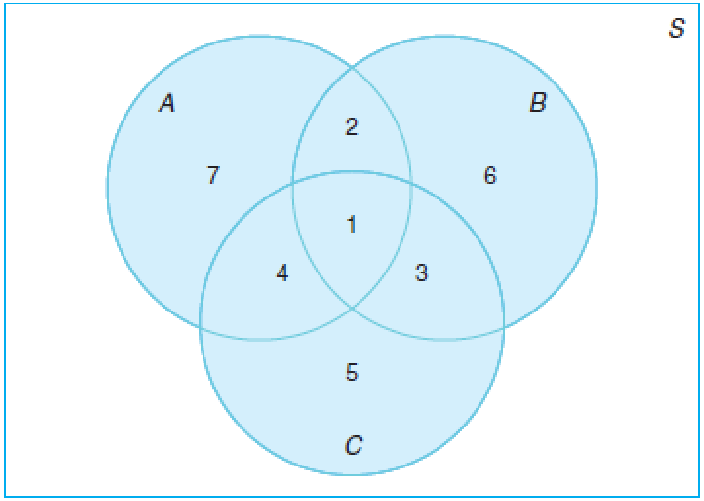


Figure 2.3: Events represented by various regions.

$$A \cup C = \text{ regions 1, 2, 3, 4, 5, and 7,}$$
$$B' \cap A = \text{ regions 4 and 7,}$$
$$A \cap B \cap C = \text{ region 1,}$$
$$(A \cup B) \cap C' = \text{ regions 2, 6, and 7,}$$

2.2 Events

- **Exercises**

2.1 List the elements of each of the following sample spaces:

(a) the set of integers between 1 and 50 divisible by 8;

(b) the set $S = \{x \mid x^2 + 4x - 5 = 0\}$;

(c) the set of outcomes when a coin is tossed until a tail or three heads appear;

(d) the set $S = \{x \mid x \text{ is a continent}\}$;

(e) the set $S = \{x \mid 2x - 4 \geq 0 \text{ and } x < 1\}$.

2.5 An experiment consists of tossing a die and then flipping a coin once if the number on the die is even. If the number on the die is odd, the coin is flipped twice. Using the notation $4H$, for example, to denote the outcome that the die comes up 4 and then the coin comes up heads, and $3HT$ to denote the outcome that the die comes up 3 followed by a head and then a tail on the coin, construct a tree diagram to show the 18 elements of the sample space S.

2.2 Events

- **Exercises**

2.9 For the sample space of Exercise 2.5,

(a) list the elements corresponding to the event A that a number less than 3 occurs on the die;

(b) list the elements corresponding to the event B that two tails occur;

(c) list the elements corresponding to the event A';

(d) list the elements corresponding to the event $A' \cap B$;

(e) list the elements corresponding to the event $A \cup B$.

2.2 Events

- **Exercises**

2.12 Exercise and diet are being studied as possible substitutes for medication to lower blood pressure. Three groups of subjects will be used to study the effect of exercise. Group 1 is sedentary, while group 2 walks and group 3 swims for 1 hour a day. Half of each of the three exercise groups will be on a salt-free diet. An additional group of subjects will not exercise or restrict their salt, but will take the standard medication. Use Z for sedentary, W for walker, S for swimmer, Y for salt, N for no salt, M for medication, and F for medication free.

(a) Show all of the elements of the sample space S.

(b) Given that A is the set of nonmedicated subjects and B is the set of walkers, list the elements of $A \cup B$.

(c) List the elements of $A \cap B$.

2.3 Counting Sample Points

- **Multiplication rule**

Rule 2.1: If an operation can be performed in n_1 ways, and if for each of these ways a second operation can be performed in n_2 ways, then the two operations can be performed together in $n_1 n_2$ ways.

Example 2.13: How many sample points are there in the sample space when a pair of dice is thrown once?

***Solution*:** The first die can land face-up in any one of $n_1 = 6$ ways. For each of these 6 ways, the second die can also land face-up in $n_2 = 6$ ways. Therefore, the pair of dice can land in $n_1 n_2 = (6)(6) = 36$ possible ways.

Generalized multiplication rule

Rule 2.2: If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, and for each of the first two a third operation can be performed in n_3 ways, and so forth, then the sequence of k operations can be performed in $n_1 n_2 \cdots n_k$ ways.

2.3 Counting Sample Points

A **permutation** is an arrangement of all or part of a set of objects.

The number of permutations of n objects is $n!$.

The number of permutations of n distinct objects taken r at a time is

$$_nP_r = \frac{n!}{(n-r)!}.$$

For any non-negative integer n, $n!$, called "n factorial," is defined as

$$n! = n(n-1)\cdots(2)(1),$$

with special case $0! = 1$.

2.3 Counting Sample Points

Example 2.18: In one year, three awards (research, teaching, and service) will be given to a class of 25 graduate students in a statistics department. If each student can receive at most one award, how many possible selections are there?

***Solution*:** Since the awards are distinguishable, it is a permutation problem. The total number of sample points is

$$_{25}P_3 = \frac{25!}{(25-3)!} = \frac{25!}{22!} = (25)(24)(23) = 13,800.$$

ple 2.19: A president and a treasurer are to be chosen from a student club consisting of 50 people. How many different choices of officers are possible if

(a) there are no restrictions;

(b) A will serve only if he is president;

(c) B and C will serve together or not at all;

(d) D and E will not serve together?

***Solution*:** (a) The total number of choices of officers, without any restrictions, is

$${}_{50}P_2 = \frac{50!}{48!} = (50)(49) = 2450.$$

(b) Since A will serve only if he is president, we have two situations here: (i) A is selected as the president, which yields 49 possible outcomes for the treasurer's position, or (ii) officers are selected from the remaining 49 people without A, which has the number of choices ${}_{49}P_2 = (49)(48) = 2352$. Therefore, the total number of choices is $49 + 2352 = 2401$.

(c) The number of selections when B and C serve together is 2. The number of selections when both B and C are not chosen is ${}_{48}P_2 = 2256$. Therefore, the total number of choices in this situation is $2 + 2256 = 2258$.

(d) The number of selections when D serves as an officer but not E is $(2)(48) = 96$, where 2 is the number of positions D can take and 48 is the number of selections of the other officer from the remaining people in the club except E. The number of selections when E serves as an officer but not D is also $(2)(48) = 96$. The number of selections when both D and E are not chosen is ${}_{48}P_2 = 2256$. Therefore, the total number of choices is $(2)(96) + 2256 = 2448$. This problem also has another short solution: Since D and E can only serve together in 2 ways, the answer is $2450 - 2 = 2448$.

2.3 Counting Sample Points

The number of distinct permutations of n things of which n_1 are of one kind, n_2 of a second kind, $\ldots$, n_k of a kth kind is

$$\frac{n!}{n_1!n_2!\cdots n_k!}.$$

ample 2.20: In a college football training session, the defensive coordinator needs to have 10 players standing in a row. Among these 10 players, there are 1 freshman, 2 sophomores, 4 juniors, and 3 seniors. How many different ways can they be arranged in a row if only their class level will be distinguished?

***Solution*:** Directly using Theorem 2.4, we find that the total number of arrangements is

$$\frac{10!}{1!\;2!\;4!\;3!} = 12,600.$$

2.3 Counting Sample Points

The number of ways of partitioning a set of n objects into r cells with n_1 elements in the first cell, n_2 elements in the second, and so forth, is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!},$$

where $n_1 + n_2 + \cdots + n_r = n$.

mple 2.21: In how many ways can 7 graduate students be assigned to 1 triple and 2 double hotel rooms during a conference?

***Solution*:** The total number of possible partitions would be

$$\binom{7}{3,2,2} = \frac{7!}{3!\ 2!\ 2!} = 210.$$

2.3 Counting Sample Points

- **Combinations**

In many problems, we are interested in the number of ways of selecting r objects from n without regard to order. These selections are called **combinations**. A combination is actually a partition with two cells, the one cell containing the r objects selected and the other cell containing the $(n-r)$ objects that are left. The number of such combinations, denoted by

$$\binom{n}{r, n-r}, \text{ is usually shortened to } \binom{n}{r},$$

since the number of elements in the second cell must be $n-r$.

The number of combinations of n distinct objects taken r at a time is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

2.3 Counting Sample Points

- **Combinations**

xample 2.23: How many different letter arrangements can be made from the letters in the word *STATISTICS*?

***Solution*:** Using the same argument as in the discussion for Theorem 2.6, in this example we can actually apply Theorem 2.5 to obtain

$$\binom{10}{3,3,2,1,1} = \frac{10!}{3!\,3!\,2!\,1!\,1!} = 50,400.$$

Here we have 10 total letters, with 2 letters (S, T) appearing 3 times each, letter I appearing twice, and letters A and C appearing once each. On the other hand, this result can be directly obtained by using Theorem 2.4.

xample 2.22: A young boy asks his mother to get 5 Game-Boy™ cartridges from his collection of 10 arcade and 5 sports games. How many ways are there that his mother can get 3 arcade and 2 sports games?

?

2.3 Counting Sample Points

- **Exercises**

2.22 In a medical study, patients are classified in 8 ways according to whether they have blood type AB^+, AB^-, A^+, A^-, B^+, B^-, O^+, or O^-, and also according to whether their blood pressure is low, normal, or high. Find the number of ways in which a patient can be classified.

2.26 A California study concluded that following 7 simple health rules can extend a man's life by 11 years on the average and a woman's life by 7 years. These 7 rules are as follows: no smoking, get regular exercise, use alcohol only in moderation, get 7 to 8 hours of sleep, maintain proper weight, eat breakfast, and do not eat between meals. In how many ways can a person adopt 5 of these rules to follow

(a) if the person presently violates all 7 rules?

(b) if the person never drinks and always eats breakfast?

2.3 Counting Sample Points

- **Exercises**

2.28 A drug for the relief of asthma can be purchased from 5 different manufacturers in liquid, tablet, or capsule form, all of which come in regular and extra strength. How many different ways can a doctor prescribe the drug for a patient suffering from asthma?

2.33 If a multiple-choice test consists of 5 questions, each with 4 possible answers of which only 1 is correct,

(a) in how many different ways can a student check off one answer to each question?

(b) in how many ways can a student check off one answer to each question and get all the answers wrong?

2.3 Counting Sample Points

- **Exercises**

2.37 In how many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?

2.41 Find the number of ways that 6 teachers can be assigned to 4 sections of an introductory psychology course if no teacher is assigned to more than one section.

2.47 How many ways are there to select 3 candidates from 8 equally qualified recent graduates for openings in an accounting firm?

2.4 Probability of an Event

- The likelihood of the occurrence of an event resulting from such a statistical experiment is evaluated by means of a set of real numbers, called **weights** or **probabilities**, ranging from 0 to 1.
- If we have reason to believe that a certain sample point is quite likely to occur when the experiment is conducted, the probability assigned should be close to 1.
- On the other hand, a probability closer to 0 is assigned to a sample point that is not likely to occur.
- For points outside the sample space, that is, for simple events that cannot possibly occur, we assign a probability of 0.

2.4 Probability of an Event

The **probability** of an event A is the sum of the weights of all sample points in A. Therefore,

$$0 \leq P(A) \leq 1, \quad P(\phi) = 0, \quad \text{and} \quad P(S) = 1.$$

Furthermore, if $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots.$$

nple 2.24: A coin is tossed twice. What is the probability that at least 1 head occurs?

***Solution*:** The sample space for this experiment is

$$S = \{HH, HT, TH, TT\}.$$

If the coin is balanced, each of these outcomes is equally likely to occur. Therefore, we assign a probability of ω to each sample point. Then $4\omega = 1$, or $\omega = 1/4$. If A represents the event of at least 1 head occurring, then

$$A = \{HH, HT, TH\} \text{ and } P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

Example 2.25: A die is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number less than 4 occurs on a single toss of the die, find $P(E)$.

?

2.4 Probability of an Event

mple 2.26: In Example 2.25, let A be the event that an even number turns up and let B be the event that a number divisible by 3 occurs. Find $P(A \cup B)$ and $P(A \cap B)$.

***Solution*:** For the events $A = \{2, 4, 6\}$ and $B = \{3, 6\}$, we have

$$A \cup B = \{2, 3, 4, 6\} \text{ and } A \cap B = \{6\}.$$

By assigning a probability of $1/9$ to each odd number and $2/9$ to each even number, we have

$$P(A \cup B) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9} \quad \text{and} \quad P(A \cap B) = \frac{2}{9}.$$

If the sample space for an experiment contains N elements, all of which are equally likely to occur, we assign a probability equal to $1/N$ to each of the N points. The probability of any event A containing n of these N sample points is then the ratio of the number of elements in A to the number of elements in S.

Rule 2.3: If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the probability of event A is

$$P(A) = \frac{n}{N}.$$

2.4 Probability of an Event

mple 2.28: In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

?

2.4 Probability of an Event

mple 2.28: In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

Solution: The number of ways of being dealt 2 aces from 4 cards is

$$\binom{4}{2} = \frac{4!}{2!\,2!} = 6,$$

and the number of ways of being dealt 3 jacks from 4 cards is

$$\binom{4}{3} = \frac{4!}{3!\,1!} = 4.$$

By the multiplication rule (Rule 2.1), there are $n = (6)(4) = 24$ hands with 2 aces and 3 jacks. The total number of 5-card poker hands, all of which are equally likely, is

$$N = \binom{52}{5} = \frac{52!}{5!\,47!} = 2{,}598{,}960.$$

Therefore, the probability of getting 2 aces and 3 jacks in a 5-card poker hand is

$$P(C) = \frac{24}{2,598,960} = 0.9 \times 10^{-5}.$$

2.4 Probability of an Event

– If the outcomes of an experiment are not equally likely to occur, the probabilities must be assigned on the **basis of prior knowledge** or experimental evidence.

- For example, if a coin is not balanced, we could estimate the probabilities of heads and tails by tossing the coin a large number of times and **recording the outcomes**.
- According to the **relative frequency** definition of probability, the true probabilities would be the fractions of heads and tails that occur in the long run.

- To find a numerical value that represents adequately the probability of **winning at tennis**, we must depend on our **past performance** at the game as well as that of the opponent and, to some extent, our **belief in our ability to win.**

2.5 Additive Rules

If A and B are two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

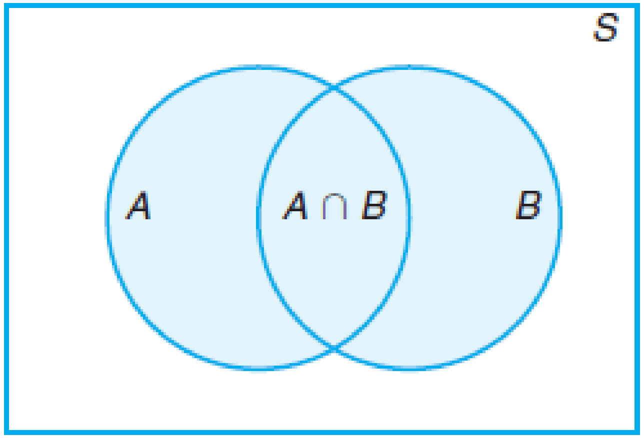


Figure 2.7: Additive rule of probability.

2.5 Additive Rules

If A and B are mutually exclusive, then

$$P(A \cup B) = P(A) + P(B).$$

If $A_1,\ A_2, \ldots, A_n$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

If $A_1,\ A_2, \ldots, A_n$ is a partition of sample space S, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) = P(S) = 1.$$

For three events A, B, and C,

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) \\ & - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

2.5 Additive Rules

xample 2.29: John is going to graduate from an industrial engineering department in a university by the end of the semester. After being interviewed at two companies he likes, he assesses that his probability of getting an offer from company A is 0.8, and his probability of getting an offer from company B is 0.6. If he believes that the probability that he will get offers from both companies is 0.5, what is the probability that he will get at least one offer from these two companies?

xample 2.30: What is the probability of getting a total of 7 or 11 when a pair of fair dice is tossed?

2.5 Additive Rules

xample 2.29: John is going to graduate from an industrial engineering department in a university by the end of the semester. After being interviewed at two companies he likes, he assesses that his probability of getting an offer from company A is 0.8, and his probability of getting an offer from company B is 0.6. If he believes that the probability that he will get offers from both companies is 0.5, what is the probability that he will get at least one offer from these two companies?

***Solution*:** Using the additive rule, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.6 - 0.5 = 0.9.$$

xample 2.30: What is the probability of getting a total of 7 or 11 when a pair of fair dice is tossed?

***Solution*:** Let A be the event that 7 occurs and B the event that 11 comes up. Now, a total of 7 occurs for 6 of the 36 sample points, and a total of 11 occurs for only 2 of the sample points. Since all sample points are equally likely, we have $P(A) = 1/6$ and $P(B) = 1/18$. The events A and B are mutually exclusive, since a total of 7 and 11 cannot both occur on the same toss. Therefore,

$$P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}.$$

This result could also have been obtained by counting the total number of points for the event $A \cup B$, namely 8, and writing

$$P(A \cup B) = \frac{n}{N} = \frac{8}{36} = \frac{2}{9}.$$

2.5 Additive Rules

If A and A' are complementary events, then

$$P(A) + P(A') = 1.$$

Example 2.32: If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are, respectively, 0.12, 0.19, 0.28, 0.24, 0.10, and 0.07, what is the probability that he will service at least 5 cars on his next day at work?

***Solution*:** Let E be the event that at least 5 cars are serviced. Now, $P(E) = 1 - P(E')$, where E' is the event that fewer than 5 cars are serviced. Since

$$P(E') = 0.12 + 0.19 = 0.31,$$

it follows from Theorem 2.9 that

$$P(E) = 1 - 0.31 = 0.69.$$

2.5 Additive Rules

- **Exercises**

2.52 Suppose that in a senior college class of 500 students it is found that 210 smoke, 258 drink alcoholic beverages, 216 eat between meals, 122 smoke and drink alcoholic beverages, 83 eat between meals and drink alcoholic beverages, 97 smoke and eat between meals, and 52 engage in all three of these bad health practices. If a member of this senior class is selected at random, find the probability that the student

(a) smokes but does not drink alcoholic beverages;

(b) eats between meals and drinks alcoholic beverages but does not smoke;

(c) neither smokes nor eats between meals.

2.53 The probability that an American industry will locate in Shanghai, China, is 0.7, the probability that it will locate in Beijing, China, is 0.4, and the probability that it will locate in either Shanghai or Beijing or both is 0.8. What is the probability that the industry will locate

(a) in both cities?

(b) in neither city?

2.5 Additive Rules

- **Exercises**

2.58 A pair of fair dice is tossed. Find the probability of getting

(a) a total of 8;

(b) at most a total of 5.

2.59 In a poker hand consisting of 5 cards, find the probability of holding

(a) 3 aces;

(b) 4 hearts and 1 club.

2.5 Additive Rules

- **Exercises**

2.63 According to *Consumer Digest* (July/August 1996), the probable location of personal computers (PC) in the home is as follows:

Adult bedroom:	0.03
Child bedroom:	0.15
Other bedroom:	0.14
Office or den:	0.40
Other rooms:	0.28

(a) What is the probability that a PC is in a bedroom?

(b) What is the probability that it is not in a bedroom?

(c) Suppose a household is selected at random from households with a PC; in what room would you expect to find a PC?

2.6 Conditional Probability, Independence, and the Product Rule

- **Conditional Probability**

The probability of an event B occurring when it is known that some event A has occurred is called a **conditional probability** and is denoted by $P(B|A)$. The symbol $P(B|A)$ is usually read "the probability that B occurs given that A occurs" or simply "the probability of B, given A."

2.6 Conditional Probability, Independence, and the Product Rule

Consider the event B of getting a perfect square when a die is tossed. The die is constructed so that the even numbers are twice as likely to occur as the odd numbers. Based on the sample space $S = \{1, 2, 3, 4, 5, 6\}$, with probabilities of 1/9 and 2/9 assigned, respectively, to the odd and even numbers, the probability of B occurring is 1/3. Now suppose that it is known that the toss of the die resulted in a number greater than 3. We are now dealing with a reduced sample space $A = \{4, 5, 6\}$, which is a subset of S. To find the probability that B occurs, relative to the space A, we must first assign new probabilities to the elements of A proportional to their original probabilities such that their sum is 1. Assigning a probability of w to the odd number in A and a probability of $2w$ to the two even numbers, we have $5w = 1$, or $w = 1/5$. Relative to the space A, we find that B contains the single element 4. Denoting this event by the symbol $B|A$, we write $B|A = \{4\}$, and hence

$$P(B|A) = \frac{2}{5}.$$

This example illustrates that events may have different probabilities when considered relative to different sample spaces.

We can also write

$$P(B|A) = \frac{2}{5} = \frac{2/9}{5/9} = \frac{P(A \cap B)}{P(A)},$$

where $P(A \cap B)$ and $P(A)$ are found from the original sample space S. In other words, a conditional probability relative to a subspace A of S may be calculated directly from the probabilities assigned to the elements of the original sample space S.

2.6 Conditional Probability, Independence, and the Product Rule

The conditional probability of B, given A, denoted by $P(B|A)$, is defined by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad \text{provided} \quad P(A) > 0.$$

le 2.34: The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$. Find the probability that a plane

(a) arrives on time, given that it departed on time, and (b) departed on time, given that it has arrived on time.

??

2.6 Conditional Probability, Independence, and the Product Rule

The conditional probability of B, given A, denoted by $P(B|A)$, is defined by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad \text{provided} \quad P(A) > 0.$$

le 2.34: The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$. Find the probability that a plane

(a) arrives on time, given that it departed on time, and (b) departed on time, given that it has arrived on time.

***Solution*:** Using Definition 2.10, we have the following.

(a) The probability that a plane arrives on time, given that it departed on time is

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) The probability that a plane departed on time, given that it has arrived or time, is

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

2.6 Conditional Probability, Independence, and the Product Rule

- **Independent Events**

4. Now consider an experiment in which 2 cards are drawn in succession from an ordinary deck, with replacement. The events are defined as

A: the first card is an ace,

B: the second card is a spade.

Since the first card is replaced, our sample space for both the first and the second draw consists of 52 cards, containing 4 aces and 13 spades. Hence,

$$P(B|A) = \frac{13}{52} = \frac{1}{4} \quad \text{and} \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

That is, $P(B|A) = P(B)$. When this is true, the events A and B are said to be **independent**.

Two events A and B are **independent** if and only if

$$P(B|A) = P(B) \quad \text{or} \quad P(A|B) = P(A),$$

assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**.

2.6 Conditional Probability, Independence, and the Product Rule

- **The Product Rule**

If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0.$$

$$P(A \cap B) = P(B \cap A) = P(B)P(A|B).$$

ample 2.36: Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

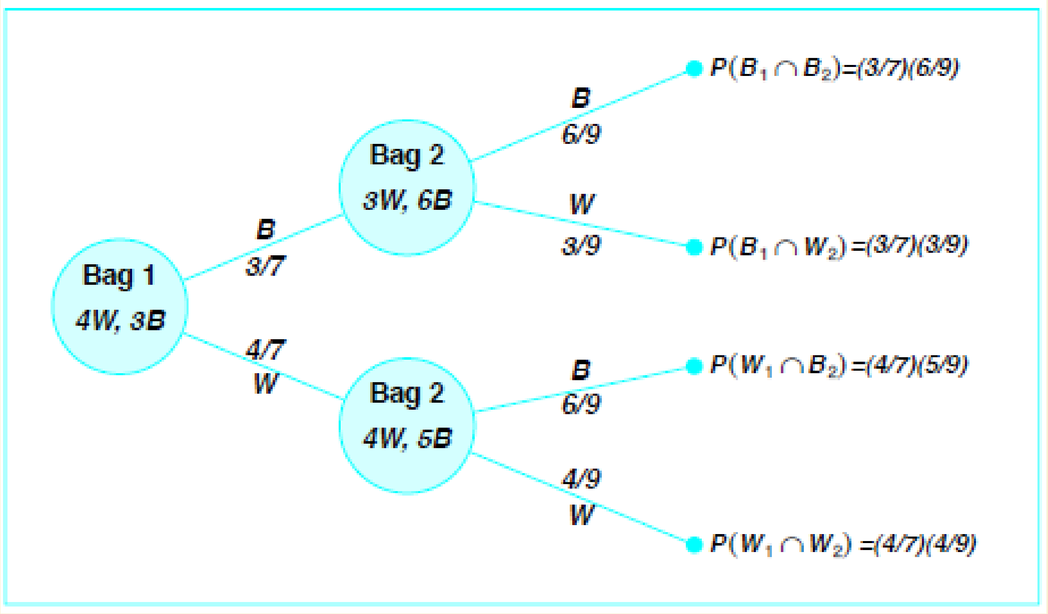
***Solution*:** We shall let A be the event that the first fuse is defective and B the event that the second fuse is defective; then we interpret $A \cap B$ as the event that A occurs and then B occurs after A has occurred. The probability of first removing a defective fuse is $1/4$; then the probability of removing a second defective fuse from the remaining 4 is $4/19$. Hence,

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{4}{19}\right) = \frac{1}{19}.$$

nple 2.37: One bag contains 4 white balls and 3 black balls, and a second bag contains 3 white balls and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. What is the probability that a ball now drawn from the second bag is black?

***Solution*:** Let B_1, B_2, and W_1 represent, respectively, the drawing of a black ball from bag 1, a black ball from bag 2, and a white ball from bag 1. We are interested in the union of the mutually exclusive events $B_1 \cap B_2$ and $W_1 \cap B_2$. The various possibilities and their probabilities are illustrated in Figure 2.8. Now

$$\begin{aligned} P[(B_1 \cap B_2) \text{ or } (W_1 \cap B_2)] &= P(B_1 \cap B_2) + P(W_1 \cap B_2) \\ &= P(B_1)P(B_2|B_1) + P(W_1)P(B_2|W_1) \\ &= \left(\frac{3}{7}\right)\left(\frac{6}{9}\right) + \left(\frac{4}{7}\right)\left(\frac{5}{9}\right) = \frac{38}{63}. \end{aligned}$$



2.6 Conditional Probability, Independence, and the Product Rule

- **The Product Rule**

Two events A and B are independent if and only if

$$P(A \cap B) = P(A)P(B).$$

Therefore, to obtain the probability that two independent events will both occur, we simply find the product of their individual probabilities.

mple 2.38: A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

***Solution*:** Let A and B represent the respective events that the fire engine and the ambulance are available. Then

$$P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016.$$

2.6 Conditional Probability, Independence, and the Product Rule

- **The Product Rule**

If, in an experiment, the events $A_1, A_2, \ldots, A_k$ can occur, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}).$$

If the events $A_1, A_2, \ldots, A_k$ are independent, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \cdots P(A_k).$$

2.6 Conditional Probability, Independence, and the Product Rule

- **The Product Rule**

ımple 2.40: Three cards are drawn in succession, without replacement, from an ordinary deck of playing cards. Find the probability that the event $A_1 \cap A_2 \cap A_3$ occurs, where A_1 is the event that the first card is a red ace, A_2 is the event that the second card is a 10 or a jack, and A_3 is the event that the third card is greater than 3 but less than 7.

***Solution*:** First we define the events

A_1: the first card is a red ace,

A_2: the second card is a 10 or a jack,

A_3: the third card is greater than 3 but less than 7.

Now

$$P(A_1) = \frac{2}{52}, \quad P(A_2|A_1) = \frac{8}{51}, \quad P(A_3|A_1 \cap A_2) = \frac{12}{50},$$

and hence, by Theorem 2.12,

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\ &= \left(\frac{2}{52}\right)\left(\frac{8}{51}\right)\left(\frac{12}{50}\right) = \frac{8}{5525}. \end{aligned}$$

2.7 Bayes' Rule

- **Theorem of total probability** or the **rule of elimination**

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A of S,

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i).$$

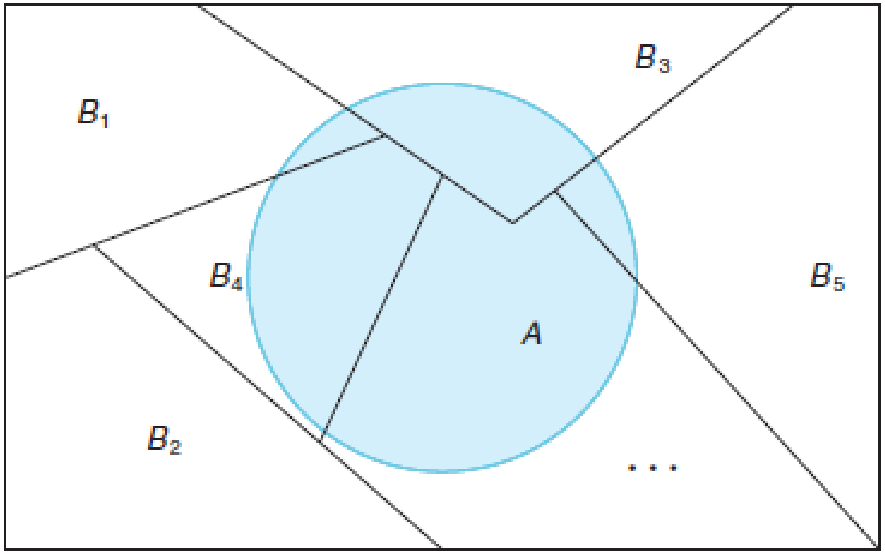


Figure 2.14: Partitioning the sample space S.

ple 2.41: In a certain assembly plant, three machines, B_1, B_2, and B_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

***Solution*:** Consider the following events:

A: the product is defective,

B_1: the product is made by machine B_1,

B_2: the product is made by machine B_2,

B_3: the product is made by machine B_3.

Applying the rule of elimination, we can write

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

Referring to the tree diagram of Figure 2.15, we find that the three branches give the probabilities

$$P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006,$$
$$P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135,$$
$$P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005,$$

and hence

$$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245.$$

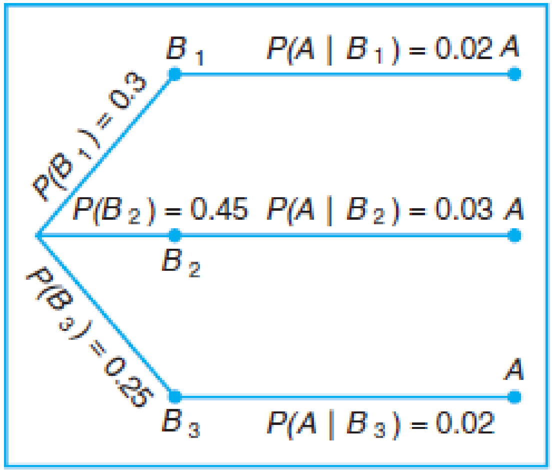


Figure 2.15: Tree diagram for Example 2.41.

2.7 Bayes' Rule

- **Bayes' Rule**

Instead of asking for $P(A)$ in Example 2.41, by the rule of elimination, suppose that we now consider the problem of finding the conditional probability $P(B_i|A)$. In other words, suppose that a product was randomly selected and it is defective. What is the probability that this product was made by machine B_i? Questions of this type can be answered by using the following theorem, called **Bayes' rule**:

(Bayes' Rule) If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \text{for } r = 1, 2, \ldots, k.$$

2.7 Bayes' Rule

- **Bayes' Rule**

mple 2.42: With reference to Example 2.41, if a product was chosen randomly and found to be defective, what is the probability that it was made by machine B_3?

***Solution*:** Using Bayes' rule to write

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)},$$

and then substituting the probabilities calculated in Example 2.41, we have

$$P(B_3|A) = \frac{0.005}{0.006 + 0.0135 + 0.005} = \frac{0.005}{0.0245} = \frac{10}{49}.$$

In view of the fact that a defective product was selected, this result suggests that it probably was not made by machine B_3.

ple 2.43: A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \qquad P(D|P_2) = 0.03, \qquad P(D|P_3) = 0.02,$$

where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

?

ple 2.43: A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \qquad P(D|P_2) = 0.03, \qquad P(D|P_3) = 0.02,$$

where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

***olution*:** From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

we must find $P(P_j|D)$ for $j = 1, 2, 3$. Bayes' rule (Theorem 2.14) shows

$$P(P_1|D) = \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)}$$
$$= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158.$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \text{ and } P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3.

2.7 Bayes' Rule

- **Exercises**

2.74 A class in advanced physics is composed of 10 juniors, 30 seniors, and 10 graduate students. The final grades show that 3 of the juniors, 10 of the seniors, and 5 of the graduate students received an A for the course. If a student is chosen at random from this class and is found to have earned an A, what is the probability that he or she is a senior?

2.81 The probability that a married man watches a certain television show is 0.4, and the probability that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does, is 0.7. Find the probability that

(a) a married couple watches the show;

(b) a wife watches the show, given that her husband does;

(c) at least one member of a married couple will watch the show.

2.7 Bayes' Rule

- **Exercises**

2.76 In an experiment to study the relationship of hypertension and smoking habits, the following data are collected for 180 individuals:

	Nonsmokers	Moderate Smokers	Heavy Smokers
H	21	36	30
NH	48	26	19

where H and NH in the table stand for *Hypertension* and *Nonhypertension*, respectively. If one of these individuals is selected at random, find the probability that the person is

(a) experiencing hypertension, given that the person is a heavy smoker;

(b) a nonsmoker, given that the person is experiencing no hypertension.

2.7 Bayes' Rule

- **Exercises**

2.78 A manufacturer of a flu vaccine is concerned about the quality of its flu serum. Batches of serum are processed by three different departments having rejection rates of 0.10, 0.08, and 0.12, respectively. The inspections by the three departments are sequential and independent.

(a) What is the probability that a batch of serum survives the first departmental inspection but is rejected by the second department?

(b) What is the probability that a batch of serum is rejected by the third department?

2.7 Bayes' Rule

- **Exercises**

2.96 Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations L_1, L_2, L_3, and L_4 will be operated 40%, 30%, 20%, and 30% of the time. If a person who is speeding on her way to work has probabilities of 0.2, 0.1, 0.5, and 0.2, respectively, of passing through these locations, what is the probability that she will receive a speeding ticket?

2.98 If the person in Exercise 2.96 received a speeding ticket on her way to work, what is the probability that she passed through the radar trap located at L_2?

Chapter 3
Random Variable &
Probability Distributions

Chapter 3: Random Variable & Probability Distributions

- The concept of a random variable
- Discrete, continuous & empirical distributions

- Discrete Distributions:
  - Random Variables
  - Binomial Distributions
  - Geometric Distributions

- Continuous Distributions:
  - Density Curves
  - The Normal Distribution
  - Standard Normal Calculations

3.1 Concept of a Random Variable

A **random variable** is a function that associates a real number with each element in the sample space.

Example 3.1: Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. The possible outcomes and the values y of the random variable Y, where Y is the number of red balls, are

Sample Space	y
RR	2
RB	1
BR	1
BB	0

3.4: Statisticians use **sampling plans** to either accept or reject batches or lots of material. Suppose one of these sampling plans involves sampling independently 10 items from a lot of 100 items in which 12 are defective.

Let X be the random variable defined as the number of items found defective in the sample of 10. In this case, the random variable takes on the values $0, 1, 2, \ldots, 9, 10$.

Discrete random variable

3.1 Concept of a Random Variable

A **random variable** is a function that associates a real number with each element in the sample space.

3.6: Interest centers around the proportion of people who respond to a certain mail order solicitation. Let X be that proportion. X is a random variable that takes on all values x for which $0 \le x \le 1$.

3.7: Let X be the random variable defined by the waiting time, in hours, between successive speeders spotted by a radar unit. The random variable X takes on all values x for which $x \ge 0$.

Continuous random variable

3.1 Concept of a Random Variable

A **random variable** is a function that associates a real number with each element in the sample space.

If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers, it is called a **discrete sample space**.

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a **continuous sample space**.

3.2 Discrete Probability Distributions

The set of ordered pairs $(x, f(x))$ is a **probability function, probability mass function**, or **probability distribution** of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

nple 3.8: A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Solution: Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can only take the numbers 0, 1, and 2. Now

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95}, \quad f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190},$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}.$$

Thus, the probability distribution of X is

x	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

3.2 Discrete Probability Distributions

mple 3.9: If a car agency sells 50% of its inventory of a certain foreign car equipped with side airbags, find a formula for the probability distribution of the number of cars with side airbags among the next 4 cars sold by the agency.

Solution: Since the probability of selling an automobile with side airbags is 0.5, the $2^4 = 16$ points in the sample space are equally likely to occur. Therefore, the denominator for all probabilities, and also for our function, is 16. To obtain the number of ways of selling 3 cars with side airbags, we need to consider the number of ways of partitioning 4 outcomes into two cells, with 3 cars with side airbags assigned to one cell and the model without side airbags assigned to the other. This can be done in $\binom{4}{3} = 4$ ways. In general, the event of selling x models with side airbags and $4 - x$ models without side airbags can occur in $\binom{4}{x}$ ways, where x can be 0, 1, 2, 3, or 4. Thus, the probability distribution $f(x) = P(X = x)$ is

$$f(x) = \frac{1}{16}\binom{4}{x}, \quad \text{for } x = 0, 1, 2, 3, 4.$$

3.2 Discrete Probability Distributions

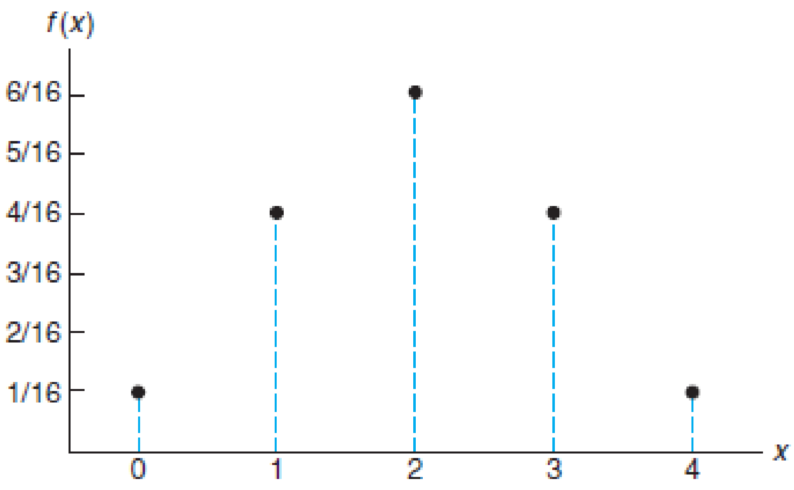


Figure 3.1: Probability mass function plot.

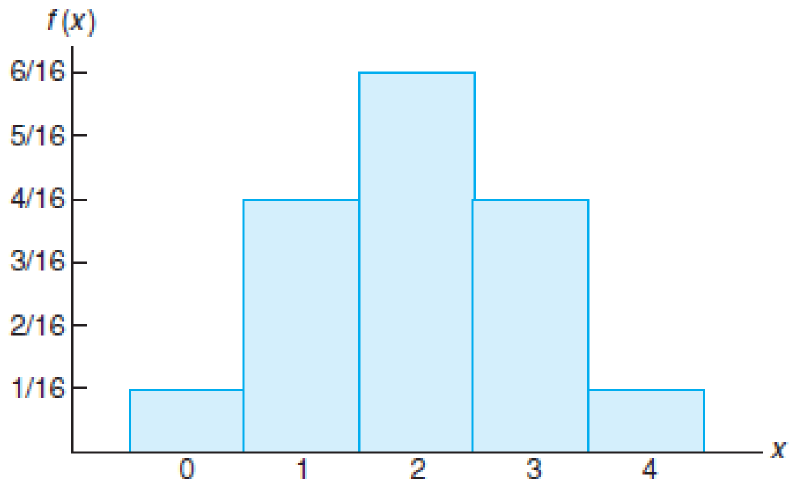


Figure 3.2: Probability histogram.

3.2 Discrete Probability Distributions

The **cumulative distribution function** $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty.$$

ıple 3.10: Find the cumulative distribution function of the random variable X in Example 3.9. Using $F(x)$, verify that $f(2) = 3/8$.

Solution: Direct calculations of the probability distribution of Example 3.9 give $f(0)= 1/16$, $f(1) = 1/4$, $f(2)= 3/8$, $f(3)= 1/4$, and $f(4)= 1/16$. Therefore,

$$F(0) = f(0) = \frac{1}{16},$$
$$F(1) = f(0) + f(1) = \frac{5}{16},$$
$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16},$$
$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16},$$
$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1.$$

Hence,

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ \frac{1}{16}, & \text{for } 0 \le x < 1, \\ \frac{5}{16}, & \text{for } 1 \le x < 2, \\ \frac{11}{16}, & \text{for } 2 \le x < 3, \\ \frac{15}{16}, & \text{for } 3 \le x < 4, \\ 1 & \text{for } x \ge 4. \end{cases}$$

Now

$$f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}.$$

3.3 Continuous Probability Distributions

The function $f(x)$ is a **probability density function** (pdf) for the continuous random variable X, defined over the set of real numbers, if

1. $f(x) \geq 0$, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\ dx = 1$.
3. $P(a < X < b) = \int_a^b f(x)\ dx$.

ple 3.11: Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

.

(a) Verify that $f(x)$ is a density function.

(b) Find $P(0 < X \leq 1)$.

***Solution*:** We use Definition 3.6.

(a) Obviously, $f(x) \geq 0$. To verify condition 2 in Definition 3.6, we have

$$\int_{-\infty}^{\infty} f(x)\ dx = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9}\Big|_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1.$$

(b) Using formula 3 in Definition 3.6, we obtain

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \left.\frac{x^3}{9}\right|_0^1 = \frac{1}{9}.$$

3.3 Continuous Probability Distributions

$$P(a < X < b) = \int_a^b f(x)\, dx.$$

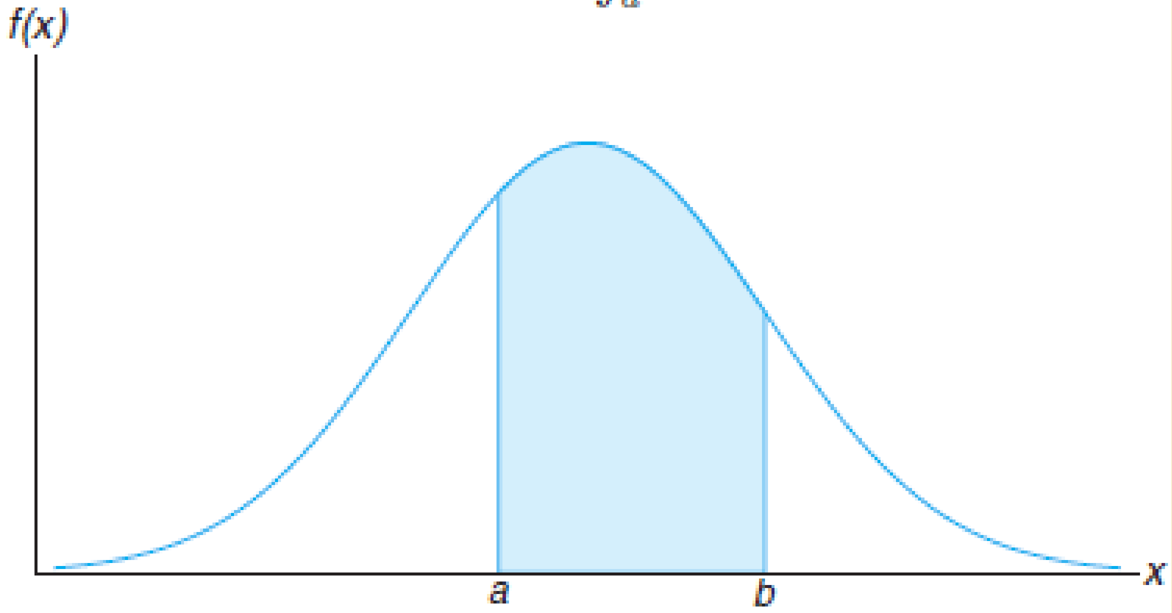


Figure 3.5: $P(a < X < b)$.

3.3 Continuous Probability Distributions

The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt, \quad \text{for } -\infty < x < \infty.$$

As an immediate consequence of Definition 3.7, one can write the two results

$$P(a < X < b) = F(b) - F(a) \text{ and } f(x) = \frac{dF(x)}{dx},$$

if the derivative exists.

3.3 Continuous Probability Distributions

ıple 3.12: For the density function of Example 3.11, find $F(x)$, and use it to evaluate $P(0 < X \leq 1)$.

***Solution*:** For $-1 < x < 2$,

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-1}^{x} \frac{t^2}{3}dt = \left.\frac{t^3}{9}\right|_{-1}^{x} = \frac{x^3+1}{9}.$$

Therefore,

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{x^3+1}{9}, & -1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

The cumulative distribution function $F(x)$ is expressed in Figure 3.6. Now

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9},$$

which agrees with the result obtained by using the density function in Example 3.11.

3.3 Continuous Probability Distributions

- **Exercises**

3.1 Classify the following random variables as discrete or continuous:

X: the number of automobile accidents per year in Virginia.

Y: the length of time to play 18 holes of golf.

M: the amount of milk produced yearly by a particular cow.

N: the number of eggs laid each month by a hen.

P: the number of building permits issued each month in a certain city.

Q: the weight of grain produced per acre.

3.3 Let W be a random variable giving the number of heads minus the number of tails in three tosses of a coin. List the elements of the sample space S for the three tosses of the coin and to each sample point assign a value w of W.

3.3 Continuous Probability Distributions

- **Exercises**

3.5 Determine the value c so that each of the following functions can serve as a probability distribution of the discrete random variable X:

(a) $f(x) = c(x^2 + 4)$, for $x = 0, 1, 2, 3$;

(b) $f(x) = c\binom{2}{x}\binom{3}{3-x}$, for $x = 0, 1, 2$.

3.6 The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20{,}000}{(x+100)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that a bottle of this medicine will have a shell life of

(a) at least 200 days;

(b) anywhere from 80 to 120 days.

3.3 Continuous Probability Distributions

- **Exercises**

3.13 The probability distribution of X, the number of imperfections per 10 meters of a synthetic fabric in continuous rolls of uniform width, is given by

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

Construct the cumulative distribution function of X.

3.9 The proportion of people who respond to a certain mail-order solicitation is a continuous random variable X that has the density function

$$f(x) = \begin{cases} \frac{2(x+2)}{5}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Show that $P(0 < X < 1) = 1$.

(b) Find the probability that more than 1/4 but fewer than 1/2 of the people contacted will respond to this type of solicitation.

3.3 Continuous Probability Distributions

- **Exercises**

3.18 A continuous random variable X that can assume values between $x = 2$ and $x = 5$ has a density function given by $f(x) = 2(1 + x)/27$. Find

(a) $P(X < 4)$;

(b) $P(3 \le X < 4)$.

3.22 Three cards are drawn in succession from a deck without replacement. Find the probability distribution for the number of spades.

3.3 Continuous Probability Distributions

- **Exercises**

3.29 An important factor in solid missile fuel is the particle size distribution. Significant problems occur if the particle sizes are too large. From production data in the past, it has been determined that the particle size (in micrometers) distribution is characterized by

$$f(x) = \begin{cases} 3x^{-4}, & x > 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify that this is a valid density function.

(b) Evaluate $F(x)$.

(c) What is the probability that a random particle from the manufactured fuel exceeds 4 micrometers?

3.4 Joint Probability Distributions

- **Several random variables**
  - $f(x, y) = P(X = x, Y = y)$; that is, the values $f(x, y)$ give the probability that outcomes x and y occur at the same time.
  - For example, if an 18-wheeler is to have its tires serviced and X represents the number of miles these tires have been driven and Y represents the number of tires that need to be replaced, then $f(30000, 5)$ is the probability that the tires are used over 30,000 miles and the truck needs 5 new tires.

3.4 Joint Probability Distributions

The function $f(x, y)$ is a **joint probability distribution** or **probability mass function** of the discrete random variables X and Y if

1. $f(x, y) \geq 0$ for all (x, y),
2. $\sum_{x} \sum_{y} f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$.

For any region A in the xy plane, $P[(X, Y) \in A] = \sum \sum_{A} f(x, y)$.

Discrete

ple 3.14: Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

(a) the joint probability function $f(x,y)$,

(b) $P[(X,Y) \in A]$, where A is the region $\{(x,y)|x+y \leq 1\}$.

***Solution*:** The possible pairs of values (x,y) are $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, $(0,2)$, and $(2,0)$.

(a) Now, $f(0,1)$, for example, represents the probability that a red and a green pens are selected. The total number of equally likely ways of selecting any 2 pens from the 8 is $\binom{8}{2} = 28$. The number of ways of selecting 1 red from 2 red pens and 1 green from 3 green pens is $\binom{2}{1}\binom{3}{1} = 6$. Hence, $f(0,1) = 6/28 = 3/14$. Similar calculations yield the probabilities for the other cases, which are presented in Table 3.1. Note that the probabilities sum to 1. In Chapter 5, it will become clear that the joint probability distribution of Table 3.1 can be represented by the formula

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}},$$

for $x = 0, 1, 2$; $y = 0, 1, 2$; and $0 \leq x + y \leq 2$.

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

(b) The probability that (X,Y) fall in the region A is

$$P[(X,Y) \in A] = P(X+Y \leq 1) = f(0,0) + f(0,1) + f(1,0)$$
$$= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}.$$

Discrete

3.4 Joint Probability Distributions

The function $f(x,y)$ is a **joint density function** of the continuous random variables X and Y if

1. $f(x,y) \geq 0$, for all (x,y),
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$,
3. $P[(X,Y) \in A] = \int\int_A f(x,y)\,dx\,dy$, for any region A in the xy plane.

ple 3.15: A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify condition 2 of Definition 3.9.

(b) Find $P[(X,Y) \in A]$, where $A = \{(x,y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$.

Continuous

?

3.4 Joint Probability Distributions

Solution: (a) The integration of $f(x, y)$ over the whole region is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\,dx\,dy$$
$$= \int_0^1 \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Big|_{x=0}^{x=1} dy$$
$$= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5}\right) dy = \left(\frac{2y}{5} + \frac{3y^2}{5}\right)\Big|_0^1 = \frac{2}{5} + \frac{3}{5} = 1.$$

(b) To calculate the probability, we use

$$P[(X,Y) \in A] = P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right)$$
$$= \int_{1/4}^{1/2}\int_0^{1/2} \frac{2}{5}(2x+3y)\,dx\,dy$$
$$= \int_{1/4}^{1/2} \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Big|_{x=0}^{x=1/2} dy = \int_{1/4}^{1/2} \left(\frac{1}{10} + \frac{3y}{5}\right) dy$$
$$= \left(\frac{y}{10} + \frac{3y^2}{10}\right)\Big|_{1/4}^{1/2}$$
$$= \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}.$$

Continuous

3.4 Joint Probability Distributions

- **Marginal distributions**

The **marginal distributions** of X alone and of Y alone are

$$g(x) = \sum_y f(x,y) \quad \text{and} \quad h(y) = \sum_x f(x,y)$$

for the discrete case, and

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

for the continuous case.

ple 3.16: Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone.

***Solution*:** For the random variable X, we see that

$$g(0) = f(0,0) + f(0,1) + f(0,2) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$
$$g(1) = f(1,0) + f(1,1) + f(1,2) = \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28},$$

and

$$g(2) = f(2,0) + f(2,1) + f(2,2) = \frac{3}{28} + 0 + 0 = \frac{3}{28},$$

which are just the column totals of Table 3.1. In a similar manner we could show that the values of $h(y)$ are given by the row totals. In tabular form, these marginal distributions may be written as follows:

x	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

ple 3.17: Find $g(x)$ and $h(y)$ for the joint density function of Example 3.15.

***Solution*:** By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{2}{5}(2x+3y)\,dy = \left(\frac{4xy}{5} + \frac{6y^2}{10}\right)\Big|_{y=0}^{y=1} = \frac{4x+3}{5},$$

for $0 \le x \le 1$, and $g(x) = 0$ elsewhere. Similarly,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{2}{5}(2x+3y)\,dx = \frac{2(1+3y)}{5},$$

for $0 \le y \le 1$, and $h(y) = 0$ elsewhere.

3.4 Joint Probability Distributions

- **Conditional probability distribution**
  - In the beginning of this chapter, we stated that the value x of the random variable X represents an event that is a subset of the sample space. If we use the definition of conditional probability as stated in Chapter 2,

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) > 0,$$

where A and B are now the events defined by X = x and Y = y, respectively, then

$$P(Y = y \mid X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x, y)}{g(x)}, \text{ provided } g(x) > 0,$$

where X and Y are discrete random variables.

3.4 Joint Probability Distributions

Let X and Y be two random variables, discrete or continuous. The **conditional distribution** of the random variable Y given that $X = x$ is

$$f(y|x) = \frac{f(x,y)}{g(x)}, \text{ provided } g(x) > 0.$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x|y) = \frac{f(x,y)}{h(y)}, \text{ provided } h(y) > 0.$$

If we wish to find the probability that the discrete random variable X falls between a and b when it is known that the discrete variable $Y = y$, we evaluate

$$P(a < X < b \mid Y = y) = \sum_{a<x<b} f(x|y),$$

where the summation extends over all values of X between a and b. When X and Y are continuous, we evaluate

$$P(a < X < b \mid Y = y) = \int_a^b f(x|y)\, dx.$$

ple 3.18:| Referring to Example 3.14, find the conditional distribution of X, given that $Y = 1$, and use it to determine $P(X = 0 \mid Y = 1)$.

***Solution*:** We need to find $f(x|y)$, where $y = 1$. First, we find that

$$h(1) = \sum_{x=0}^{2} f(x, 1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Now

$$f(x|1) = \frac{f(x,1)}{h(1)} = \left(\frac{7}{3}\right) f(x, 1), \quad x = 0, 1, 2.$$

Therefore,

$$f(0|1) = \left(\frac{7}{3}\right) f(0, 1) = \left(\frac{7}{3}\right)\left(\frac{3}{14}\right) = \frac{1}{2}, \; f(1|1) = \left(\frac{7}{3}\right) f(1, 1) = \left(\frac{7}{3}\right)\left(\frac{3}{14}\right) = \frac{1}{2},$$

$$f(2|1) = \left(\frac{7}{3}\right) f(2, 1) = \left(\frac{7}{3}\right)(0) = 0,$$

and the conditional distribution of X, given that $Y = 1$, is

x	0	1	2
$f(x\|1)$	$\frac{1}{2}$	$\frac{1}{2}$	0

Finally,

$$P(X = 0 \mid Y = 1) = f(0|1) = \frac{1}{2}.$$

Therefore, if it is known that 1 of the 2 pen refills selected is red, we have a probability equal to 1/2 that the other refill is not blue.

ple 3.19: The joint density for the random variables (X, Y), where X is the unit temperature change and Y is the proportion of spectrum shift that a certain atomic particle produces, is

$$f(x,y) = \begin{cases} 10xy^2, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal densities $g(x)$, $h(y)$, and the conditional density $f(y|x)$.

(b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.

***Solution*:** (a) By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_x^1 10xy^2\,dy$$

$$= \frac{10}{3}xy^3\Big|_{y=x}^{y=1} = \frac{10}{3}x(1-x^3),\ 0 < x < 1,$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^y 10xy^2\,dx = 5x^2y^2\Big|_{x=0}^{x=y} = 5y^4,\ 0 < y < 1.$$

Now

$$f(y|x) = \frac{f(x,y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1-x^3)} = \frac{3y^2}{1-x^3},\ 0 < x < y < 1.$$

(b) Therefore,

$$P\left(Y > \frac{1}{2}\ \Big|\ X = 0.25\right) = \int_{1/2}^{1} f(y \mid x = 0.25)\,dy = \int_{1/2}^{1} \frac{3y^2}{1-0.25^3}\,dy = \frac{8}{9}.$$

iple 3.20: Given the joint density function

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1, \\ 0, & \text{elsewhere,} \end{cases}$$

find $g(x)$, $h(y)$, $f(x|y)$, and evaluate $P(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3})$.

***Solution*:** By definition of the marginal density. for $0 < x < 2$,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\ dy = \int_0^1 \frac{x(1+3y^2)}{4} dy$$
$$= \left(\frac{xy}{4} + \frac{xy^3}{4}\right)\Big|_{y=0}^{y=1} = \frac{x}{2},$$

and for $0 < y < 1$,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\ dx = \int_0^2 \frac{x(1+3y^2)}{4} dx$$
$$= \left(\frac{x^2}{8} + \frac{3x^2y^2}{8}\right)\Big|_{x=0}^{x=2} = \frac{1+3y^2}{2}.$$

Therefore, using the conditional density definition, for $0 < x < 2$,

$$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{x(1+3y^2)/4}{(1+3y^2)/2} = \frac{x}{2},$$

and

$$P\left(\frac{1}{4} < X < \frac{1}{2} \,\middle|\, Y = \frac{1}{3}\right) = \int_{1/4}^{1/2} \frac{x}{2} dx = \frac{3}{64}.$$

3.4 Joint Probability Distributions

Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables X and Y are said to be **statistically independent** if and only if

$$f(x, y) = g(x)h(y)$$

for all (x, y) within their range.

3.4 Joint Probability Distributions

- **Exercises**

3.37 Determine the values of c so that the following functions represent joint probability distributions of the random variables X and Y:

(a) $f(x, y) = cxy$, for $x = 1, 2, 3$; $y = 1, 2, 3$;

(b) $f(x, y) = c|x - y|$, for $x = -2, 0, 2$; $y = -2, 3$.

3.39 From a sack of fruit containing 3 oranges, 2 apples, and 3 bananas, a random sample of 4 pieces of fruit is selected. If X is the number of oranges and Y is the number of apples in the sample, find

(a) the joint probability distribution of X and Y;

(b) $P[(X, Y) \in A]$, where A is the region that is given by $\{(x, y) \mid x + y \leq 2\}$.

3.48 Referring to Exercise 3.39, find

(a) $f(y|2)$ for all values of y;

(b) $P(Y = 0 \mid X = 2)$.

3.4 Joint Probability Distributions

- **Exercises**

3.43 Let X denote the reaction time, in seconds, to a certain stimulus and Y denote the temperature (°F) at which a certain reaction starts to take place. Suppose that two random variables X and Y have the joint density

$$f(x,y) = \begin{cases} 4xy, & 0 < x < 1,\ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find

(a) $P(0 \le X \le \frac{1}{2}$ and $\frac{1}{4} \le Y \le \frac{1}{2})$;

(b) $P(X < Y)$.

3.4 Joint Probability Distributions

- **Exercises**

3.40 A fast-food restaurant operates both a drive-through facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-through and walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y), & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal density of X.

(b) Find the marginal density of Y.

(c) Find the probability that the drive-through facility is busy less than one-half of the time.

Chapter 4
Mathematical Expectation

Chapter 4: Mathematical Expectation

- Mean of a random variable
- Variance and co-variance
- Chebyshev's theorem

- Show how to compute the mean and variance of a random variable as well as the covariance/correlation among two variables

4.1 Mean of a Random Variable

Let X be a random variable with probability distribution $f(x)$. The **mean**, or **expected value**, of X is

$$\mu = E(X) = \sum_x x f(x)$$

if X is discrete, and

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

if X is continuous.

Note that the way to calculate the expected value, or mean, shown here is different from the way to calculate the sample mean described in Chapter 1, where the sample mean is obtained by using data.

In mathematical expectation, the expected value is calculated by using the probability distribution.

However, the mean is usually understood as a "center" value of the underlying distribution if we use the expected value, as in the previous definition.

4.1 Mean of a Random Variable

mple 4.1: A lot containing 7 components is sampled by a quality inspector; the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

***Solution*:** Let X represent the number of good components in the sample. The probability distribution of X is

$$f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, \qquad x = 0, 1, 2, 3.$$

Simple calculations yield $f(0) = 1/35$, $f(1) = 12/35$, $f(2) = 18/35$, and $f(3) = 4/35$. Therefore,

$$\mu = E(X) = (0)\left(\frac{1}{35}\right) + (1)\left(\frac{12}{35}\right) + (2)\left(\frac{18}{35}\right) + (3)\left(\frac{4}{35}\right) = \frac{12}{7} = 1.7.$$

Thus, if a sample of size 3 is selected at random over and over again from a lot of 4 good components and 3 defective components, it will contain, on average, 1.7 good components.

Discrete random variable

4.1 Mean of a Random Variable

mple 4.3: Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \frac{20{,}000}{x^3}, & x > 100, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected life of this type of device.

***Solution*:** Using Definition 4.1, we have

$$\mu = E(X) = \int_{100}^{\infty} x \frac{20{,}000}{x^3}\, dx = \int_{100}^{\infty} \frac{20{,}000}{x^2}\, dx = 200.$$

Therefore, we can expect this type of device to last, *on average*, 200 hours.

Continuous random variable

4.1 Mean of a Random Variable

mple 4.2: A salesperson for a medical device company has two appointments on a given day. At the first appointment, he believes that he has a 70% chance to make the deal, from which he can earn \$1000 commission if successful. On the other hand, he thinks he only has a 40% chance to make the deal at the second appointment, from which, if successful, he can make \$1500. What is his expected commission based on his own probability belief? Assume that the appointment results are independent of each other.

?

4.1 Mean of a Random Variable

mple 4.2: A salesperson for a medical device company has two appointments on a given day. At the first appointment, he believes that he has a 70% chance to make the deal, from which he can earn \$1000 commission if successful. On the other hand, he thinks he only has a 40% chance to make the deal at the second appointment, from which, if successful, he can make \$1500. What is his expected commission based on his own probability belief? Assume that the appointment results are independent of each other.

Solution: First, we know that the salesperson, for the two appointments, can have 4 possible commission totals: \$0, \$1000, \$1500, and \$2500. We then need to calculate their associated probabilities. By independence, we obtain

$$f(\$0) = (1 - 0.7)(1 - 0.4) = 0.18, \quad f(\$2500) = (0.7)(0.4) = 0.28,$$
$$f(\$1000) = (0.7)(1 - 0.4) = 0.42, \text{ and } f(\$1500) = (1 - 0.7)(0.4) = 0.12.$$

Therefore, the expected commission for the salesperson is

$$E(X) = (\$0)(0.18) + (\$1000)(0.42) + (\$1500)(0.12) + (\$2500)(0.28)$$
$$= \$1300.$$

4.1 Mean of a Random Variable

- Random variable *g*(*X*)

Now let us consider a new random variable $g(X)$, which depends on X; that is, each value of $g(X)$ is determined by the value of X. For instance, $g(X)$ might be X^2 or $3X-1$, and whenever X assumes the value 2, $g(X)$ assumes the value $g(2)$. In particular, if X is a discrete random variable with probability distribution $f(x)$, for $x=-1,0,1,2$, and $g(X)=X^2$, then

$$\begin{aligned} P[g(X)=0] &= P(X=0)=f(0),\\ P[g(X)=1] &= P(X=-1)+P(X=1)=f(-1)+f(1),\\ P[g(X)=4] &= P(X=2)=f(2), \end{aligned}$$

and so the probability distribution of $g(X)$ may be written

$g(x)$	0	1	4
$P[g(X)=g(x)]$	$f(0)$	$f(-1)+f(1)$	$f(2)$

By the definition of the expected value of a random variable, we obtain

$$\begin{aligned} \mu_{g(X)} &= E[g(x)] = 0f(0)+1[f(-1)+f(1)]+4f(2)\\ &= (-1)^2 f(-1)+(0)^2 f(0)+(1)^2 f(1)+(2)^2 f(2)=\sum_x g(x)f(x). \end{aligned}$$

This result is generalized in the next theorem

4.1 Mean of a Random Variable

- Random variable *g*(*X*)

Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$$

if X is discrete, and

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

if X is continuous.

4.1 Mean of a Random Variable

nple 4.4: Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $g(X) = 2X - 1$ represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

***Solution*:** By Theorem 4.1, the attendant can expect to receive

$$E[g(X)] = E(2X-1) = \sum_{x=4}^{9}(2x-1)f(x)$$
$$= (7)\left(\frac{1}{12}\right) + (9)\left(\frac{1}{12}\right) + (11)\left(\frac{1}{4}\right) + (13)\left(\frac{1}{4}\right) + (15)\left(\frac{1}{6}\right) + (17)\left(\frac{1}{6}\right) = \$12.67.$$

Discrete random variable

4.1 Mean of a Random Variable

mple 4.5: Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of $g(X) = 4X + 3$.

***Solution*:** By Theorem 4.1, we have

$$E(4X+3) = \int_{-1}^{2} \frac{(4x+3)x^2}{3}\, dx = \frac{1}{3}\int_{-1}^{2} (4x^3 + 3x^2)\, dx = 8.$$

Continuous random variable

4.1 Mean of a Random Variable

- We shall now extend our concept of mathematical expectation to the case of two random variables X and Y with joint probability distribution f (x, y).

Let X and Y be random variables with joint probability distribution $f(x,y)$. The mean, or expected value, of the random variable $g(X,Y)$ is

$$\mu_{g(X,Y)} = E[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy$$

if X and Y are continuous.

nple 4.6: Let X and Y be the random variables with joint probability distribution indicated in Table 3.1 on page 96. Find the expected value of $g(X,Y) = XY$. The table is reprinted here for convenience.

	$f(x,y)$	$x=0$	$x=1$	$x=2$	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

***Solution*:** By Definition 4.2, we write

$$\begin{aligned}
E(XY) &= \sum_{x=0}^{2}\sum_{y=0}^{2} xyf(x,y) \\
&= (0)(0)f(0,0) + (0)(1)f(0,1) \\
&\quad + (1)(0)f(1,0) + (1)(1)f(1,1) + (2)(0)f(2,0) \\
&= f(1,1) = \frac{3}{14}.
\end{aligned}$$

nple 4.7: Find $E(Y/X)$ for the density function

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

***Solution*:** We have

$$E\left(\frac{Y}{X}\right) = \int_0^1 \int_0^2 \frac{y(1+3y^2)}{4}\, dx dy = \int_0^1 \frac{y+3y^3}{2}\, dy = \frac{5}{8}.$$

4.2 Variance and Covariance of Random Variables

– Figure 4.1, we have the histograms of two discrete probability distributions that have the same mean, μ = 2, but differ considerably in variability, or the dispersion of their observations about the mean.

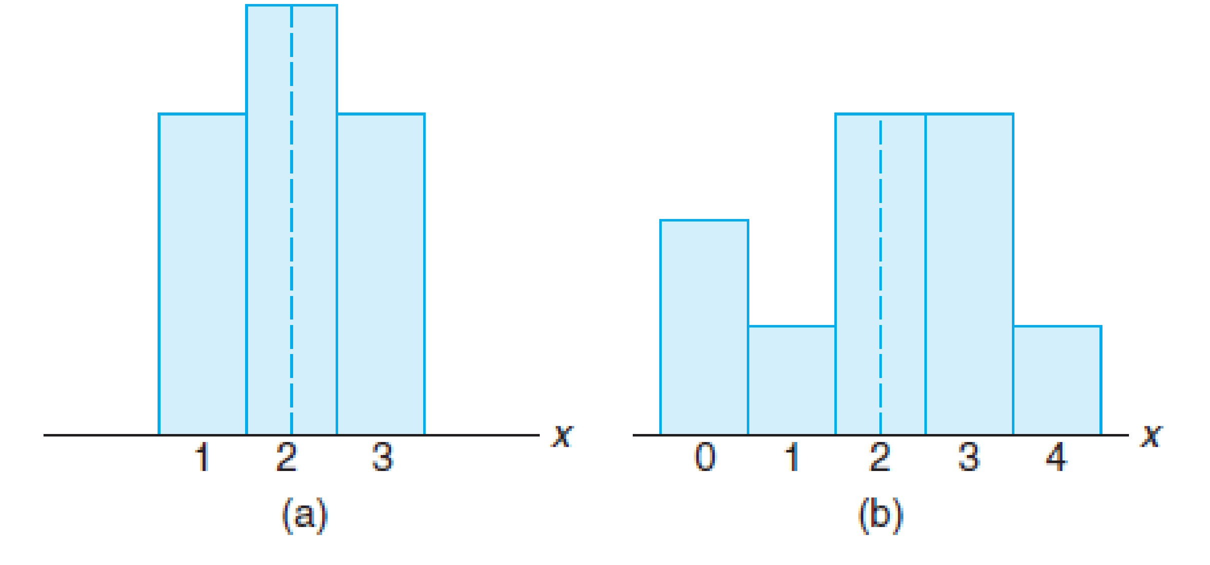


Figure 4.1: Distributions with equal means and unequal dispersions.

4.2 Variance and Covariance of Random Variables

- **Variance of the random variable X** or the **variance of the probability distribution of X** is denoted by **Var(X)** or the symbol **σ^2**

Let X be a random variable with probability distribution $f(x)$ and mean μ. The variance of X is

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x), \qquad \text{if } X \text{ is discrete, and}$$

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx, \qquad \text{if } X \text{ is continuous.}$$

The positive square root of the variance, σ, is called the **standard deviation** of X.

nple 4.8: Let the random variable X represent the number of automobiles that are used for official business purposes on any given workday. The probability distribution for company A [Figure 4.1(a)] is

x	1	2	3
$f(x)$	0.3	0.4	0.3

and that for company B [Figure 4.1(b)] is

x	0	1	2	3	4
$f(x)$	0.2	0.1	0.3	0.3	0.1

Show that the variance of the probability distribution for company B is greater than that for company A.

***Solution*:** For company A, we find that

$$\mu_A = E(X) = (1)(0.3) + (2)(0.4) + (3)(0.3) = 2.0,$$

and then

$$\sigma_A^2 = \sum_{x=1}^{3}(x-2)^2 = (1-2)^2(0.3) + (2-2)^2(0.4) + (3-2)^2(0.3) = 0.6.$$

For company B, we have

$$\mu_B = E(X) = (0)(0.2) + (1)(0.1) + (2)(0.3) + (3)(0.3) + (4)(0.1) = 2.0,$$

and then

$$\begin{aligned}\sigma_B^2 &= \sum_{x=0}^{4}(x-2)^2 f(x)\\ &= (0-2)^2(0.2) + (1-2)^2(0.1) + (2-2)^2(0.3)\\ &\quad + (3-2)^2(0.3) + (4-2)^2(0.1) = 1.6.\end{aligned}$$

Clearly, the variance of the number of automobiles that are used for official business purposes is greater for company B than for company A.

4.2 Variance and Covariance of Random Variables

– An alternative and preferred formula for finding σ^2, which often simplifies the calculations, is stated in the following theorem.

The variance of a random variable X is

$$\sigma^2 = E(X^2) - \mu^2.$$

mple 4.9: Let the random variable X represent the number of defective parts for a machine when 3 parts are sampled from a production line and tested. The following is the probability distribution of X.

x	0	1	2	3
$f(x)$	0.51	0.38	0.10	0.01

Using Theorem 4.2, calculate σ^2.

***Solution*:** First, we compute

$$\mu = (0)(0.51) + (1)(0.38) + (2)(0.10) + (3)(0.01) = 0.61.$$

Now,

$$E(X^2) = (0)(0.51) + (1)(0.38) + (4)(0.10) + (9)(0.01) = 0.87.$$

Therefore,

$$\sigma^2 = 0.87 - (0.61)^2 = 0.4979.$$

4.2 Variance and Covariance of Random Variables

– We shall now extend our concept of the variance of a random variable X to include random variables related to X.

For the random variable g(X), the variance is denoted by $\sigma^2_{g(X)}$ and is calculated by means of the following theorem.

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x)$$

if X is discrete, and

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x)\, dx$$

if X is continuous.

nple 4.11: Calculate the variance of $g(X) = 2X + 3$, where X is a random variable with probability distribution

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

***Solution*:** First, we find the mean of the random variable $2X+3$. According to Theorem 4.1,

$$\mu_{2X+3} = E(2X+3) = \sum_{x=0}^{3}(2x+3)f(x) = 6.$$

Now, using Theorem 4.3, we have

$$\sigma^2_{2X+3} = E\{[(2X+3) - \mu_{2x+3}]^2\} = E[(2X+3-6)^2]$$
$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4.$$

nple 4.12: Let X be a random variable having the density function given in Example 4.5 on page 115. Find the variance of the random variable $g(X) = 4X + 3$.

***Solution*:** In Example 4.5, we found that $\mu_{4X+3} = 8$. Now, using Theorem 4.3,

$$\sigma^2_{4X+3} = E\{[(4X+3) - 8]^2\} = E[(4X-5)^2]$$
$$= \int_{-1}^{2}(4x-5)^2\frac{x^2}{3}\,dx = \frac{1}{3}\int_{-1}^{2}(16x^4 - 40x^3 + 25x^2)\,dx = \frac{51}{5}.$$

4.2 Variance and Covariance of Random Variables

- **Covariance of *X* and *Y***

Let X and Y be random variables with joint probability distribution $f(x, y)$. The covariance of X and Y is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_y) f(x, y)$$

if X and Y are discrete, and

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_y) f(x, y)\, dx\, dy$$

if X and Y are continuous.

The alternative and preferred formula for σ_{XY} is stated by following theorem.

The covariance of two random variables X and Y with means μ_X and μ_Y, respectively, is given by

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

nple 4.13: Example 3.14 on page 95 describes a situation involving the number of blue refills X and the number of red refills Y. Two refills for a ballpoint pen are selected at random from a certain box, and the following is the joint probability distribution:

	$f(x,y)$	$x=0$	$x=1$	$x=2$	$h(y)$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Find the covariance of X and Y.

***Solution*:** From Example 4.6, we see that $E(XY) = 3/14$. Now

$$\mu_X = \sum_{x=0}^{2} xg(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{3}{4},$$

and

$$\mu_Y = \sum_{y=0}^{2} yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) = \frac{1}{2}.$$

Therefore,

Discrete

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = -\frac{9}{56}.$$

ıple 4.14: The fraction X of male runners and the fraction Y of female runners who compete in marathon races are described by the joint density function

$$f(x,y) = \begin{cases} 8xy, & 0 \le y \le x \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the covariance of X and Y.

***Solution*:** We first compute the marginal density functions. They are

$$g(x) = \begin{cases} 4x^3, & 0 \le x \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$h(y) = \begin{cases} 4y(1-y^2), & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

From these marginal density functions, we compute

$$\mu_X = E(X) = \int_0^1 4x^4 \, dx = \frac{4}{5} \text{ and } \mu_Y = \int_0^1 4y^2(1-y^2) \, dy = \frac{8}{15}.$$

From the joint density function given above, we have

$$E(XY) = \int_0^1 \int_y^1 8x^2y^2 \, dx \, dy = \frac{4}{9}.$$

Then

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225}.$$

Continuous

4.2 Variance and Covariance of Random Variables

– Although the covariance between two random variables does provide information regarding the nature of the relationship, the magnitude of σ_{XY} does not indicate anything regarding the strength of the relationship, since σ_{XY} is not scale-free.
  - Its magnitude will depend on the units used to measure both X and Y.

– There is a scale-free version of the covariance called the **correlation coefficient** that is used widely in statistics.

Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

It should be clear to us that ρ_{XY} is free of the units of X and Y.

ple 4.15: Find the correlation coefficient between X and Y in Example 4.13.
***Solution*:** Since

$$E(X^2) = (0^2)\left(\frac{5}{14}\right) + (1^2)\left(\frac{15}{28}\right) + (2^2)\left(\frac{3}{28}\right) = \frac{27}{28}$$

and

$$E(Y^2) = (0^2)\left(\frac{15}{28}\right) + (1^2)\left(\frac{3}{7}\right) + (2^2)\left(\frac{1}{28}\right) = \frac{4}{7},$$

we obtain

$$\sigma_X^2 = \frac{27}{28} - \left(\frac{3}{4}\right)^2 = \frac{45}{112} \text{ and } \sigma_Y^2 = \frac{4}{7} - \left(\frac{1}{2}\right)^2 = \frac{9}{28}.$$

Therefore, the correlation coefficient between X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-9/56}{\sqrt{(45/112)(9/28)}} = -\frac{1}{\sqrt{5}}.$$

ple 4.16: Find the correlation coefficient of X and Y in Example 4.14.
***Solution*:** Because

$$E(X^2) = \int_0^1 4x^5\,dx = \frac{2}{3} \text{ and } E(Y^2) = \int_0^1 4y^3(1-y^2)\,dy = 1 - \frac{2}{3} = \frac{1}{3},$$

we conclude that

$$\sigma_X^2 = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75} \text{ and } \sigma_Y^2 = \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225}.$$

Hence,

$$\rho_{XY} = \frac{4/225}{\sqrt{(2/75)(11/225)}} = \frac{4}{\sqrt{66}}.$$

4.4 Chebyshev's Theorem

- The Russian mathematician P. L. Chebyshev (1821–1894) discovered that the fraction of the area between any two values symmetric about the mean is related to the standard deviation. Since the area under a probability distribution curve or in a probability histogram adds to 1, the area between any two numbers is the probability of the random variable assuming a value between these numbers.
- The following theorem, due to Chebyshev, gives a conservative estimate of the probability that a random variable assumes a value within k standard deviations of its mean for any real number k.

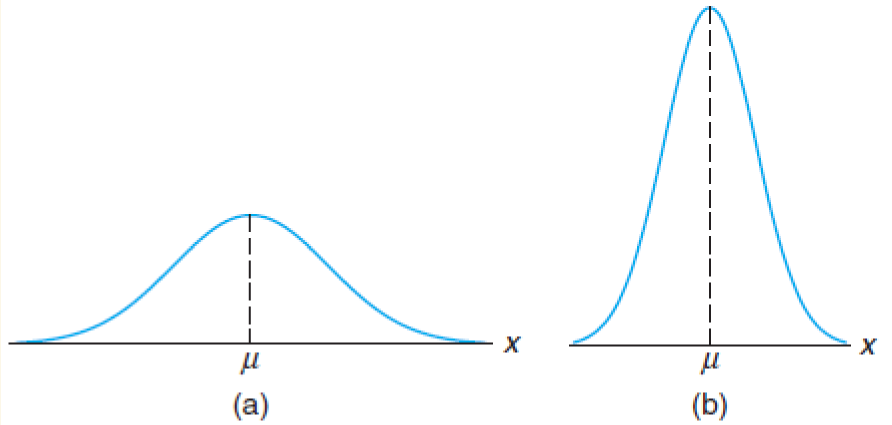


Figure 4.2: Variability of continuous observations about the mean.

4.4 Chebyshev's Theorem

(Chebyshev's Theorem) The probability that any random variable X will assume a value within k standard deviations of the mean is at least $1 - 1/k^2$. That is,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

ple 4.27: A random variable X has a mean $\mu = 8$, a variance $\sigma^2 = 9$, and an unknown probability distribution. Find

(a) $P(-4 < X < 20)$,

(b) $P(|X - 8| \geq 6)$.

Solution: (a) $P(-4 < X < 20) = P[8 - (4)(3) < X < 8 + (4)(3)] \geq \frac{15}{16}$.

(b) $P(|X - 8| \geq 6) = 1 - P(|X - 8| < 6) = 1 - P(-6 < X - 8 < 6)$

$$= 1 - P[8 - (2)(3) < X < 8 + (2)(3)] \leq \frac{1}{4}.$$

The use of Chebyshev's theorem is relegated to situations where the form of the distribution is unknown.

4.4 Chebyshev's Theorem

- **Exercises**

4.2 The probability distribution of the discrete random variable X is

$$f(x) = \binom{3}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3.$$

Find the mean of X.

4.4 A coin is biased such that a head is three times as likely to occur as a tail. Find the expected number of tails when this coin is tossed twice.

4.10 Two tire-quality experts examine stacks of tires and assign a quality rating to each tire on a 3-point scale. Let X denote the rating given by expert A and Y denote the rating given by B. The following table gives the joint distribution for X and Y.

$f(x,y)$		$y=1$	$y=2$	$y=3$
x	1	0.10	0.05	0.02
	2	0.10	0.35	0.05
	3	0.03	0.10	0.20

Find μ_X and μ_Y.

4.4 Chebyshev's Theorem

- **Exercises**

4.7 By investing in a particular stock, a person can make a profit in one year of $4000 with probability 0.3 or take a loss of $1000 with probability 0.7. What is this person's expected gain?

4.14 Find the proportion X of individuals who can be expected to respond to a certain mail-order solicitation if X has the density function

$$f(x) = \begin{cases} \frac{2(x+2)}{5}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

4.22 The hospitalization period, in days, for patients following treatment for a certain type of kidney disorder is a random variable $Y = X + 4$, where X has the density function

$$f(x) = \begin{cases} \frac{32}{(x+4)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the average number of days that a person is hospitalized following treatment for this disorder.

4.4 Chebyshev's Theorem

- **Exercises**

4.17 Let X be a random variable with the following probability distribution:

x	-3	6	9
$f(x)$	$1/6$	$1/2$	$1/3$

Find $\mu_{g(X)}$, where $g(X) = (2X+1)^2$.

4.20 A continuous random variable X has the density function

$$f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of $g(X) = e^{2X/3}$.

4.26 Let X and Y be random variables with joint density function

$$f(x, y) = \begin{cases} 4xy, & 0 < x,\ y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of $Z = \sqrt{X^2 + Y^2}$.

4.4 Chebyshev's Theorem

- **Exercises**

4.29 Exercise 3.29 on page 93 dealt with an important particle size distribution characterized by

$$f(x) = \begin{cases} 3x^{-4}, & x > 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Plot the density function.
(b) Give the mean particle size.

4.32 In Exercise 3.13 on page 92, the distribution of the number of imperfections per 10 meters of synthetic fabric is given by

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

(a) Plot the probability function.
(b) Find the expected number of imperfections, $E(X) = \mu$.
(c) Find $E(X^2)$.

Chapter 5
Some Discrete Probability Distributions

Chapter 5: Some Discrete Probability Distributions

– Hypergeometric distributions and their Discrete

– Uniform, Binomial, Poisson & probability

- We will learn when and how to model random variables
  - Using the discrete uniform, Binomial, Poisson or Hypergeometric Distribution
- How to compute the chances of certain events from such distributions

5.2 Binomial and Multinomial Distributions

- An experiment often consists of repeated trials, each with two possible outcomes that may be labeled success or failure.
- The most obvious application deals with the testing of items as they come off an assembly line, where each trial may indicate a defective or a non defective item. We may choose to define either outcome as a success.
- The process is referred to as a Bernoulli process. Each trial is called a Bernoulli trial.
  - For example, if one were drawing cards from a deck, the probabilities for repeated trials change if the cards are not replaced. That is, the probability of selecting a heart on the first draw is 1/4, but on the second draw it is a conditional probability having a value of 13/51 or 12/51, depending on whether a heart appeared on the first draw: this, then, would no longer be considered a set of Bernoulli trials.

5.2 Binomial and Multinomial Distributions

– The Bernoulli Process

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Example

Outcome	NNN	NDN	NND	DNN	NDD	DND	DDN	DDD
x	0	1	1	1	2	2	2	3

Since the items are selected independently and we assume that the process produces 25% defectives, we have

$$P(NDN) = P(N)P(D)P(N) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{9}{64}.$$

Similar calculations yield the probabilities for the other possible outcomes. The probability distribution of X is therefore

x	0	1	2	3
$f(x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

5.2 Binomial and Multinomial Distributions

– Binomial Distribution

- The number X of successes in **n Bernoulli trials** is called a binomial random variable. The probability distribution of this discrete random variable is called the binomial distribution, and its values will be denoted by **b(x; n, p)**

$$P(X=2)=f(2)=b\left(2;3,\frac{1}{4}\right)=\frac{9}{64}.$$

Binomial Distribution A Bernoulli trial can result in a success with probability p and a failure with probability $q=1-p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$b(x;n,p)=\binom{n}{x}p^x q^{n-x}, \quad x=0,1,2,\dots,n.$$

Note that when $n=3$ and $p=1/4$, the probability distribution of X, the number of defectives, may be written as

$$b\left(x;3,\frac{1}{4}\right)=\binom{3}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x}, \quad x=0,1,2,3,$$

5.2 Binomial and Multinomial Distributions

– **Binomial Distribution**

Example 5.1: The probability that a certain kind of component will survive a shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

***Solution*:** Assuming that the tests are independent and $p = 3/4$ for each of the 4 tests, we obtain

$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!\,2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

5.2 Binomial and Multinomial Distributions

Example 5.1: The probability that a certain kind of component will survive a shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

Solution: Assuming that the tests are independent and $p = 3/4$ for each of the 4 tests, we obtain

$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!\,2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

– Frequently, we are interested in problems where it is necessary to find P(X <r) or P(a ≤ X ≤ b).

Binomial sums $B(r; n, p) = \sum_{x=0}^{r} b(x; n, p)$

5.2 Binomial and Multinomial Distributions

– Binomial Sums Distribution

Example 5.2: The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?

Solution: Let X be the number of people who survive.

$$\text{(a)} \qquad P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662 = 0.0338$$

$$\text{(b)} \quad P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4) = \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4) = 0.9050 - 0.0271 = 0.8779$$

$$\text{(c)} \qquad P(X = 5) = b(5; 15, 0.4) = \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4) = 0.4032 - 0.2173 = 0.1859$$

5.2 Binomial and Multinomial Distributions

Example 5.3: A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%.

(a) The inspector randomly picks 20 items from a shipment. What is the probability that there will be at least one defective item among these 20?

(b) Suppose that the retailer receives 10 shipments in a month and the inspector randomly tests 20 devices per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective device among the 20 that are selected and tested from the shipment?

Solution: (a) Denote by X the number of defective devices among the 20. Then X follows a $b(x; 20, 0.03)$ distribution. Hence,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - b(0; 20, 0.03)$$
$$= 1 - (0.03)^0(1 - 0.03)^{20-0} = 0.4562.$$

(b) In this case, each shipment can either contain at least one defective item or not. Hence, testing of each shipment can be viewed as a Bernoulli trial with $p = 0.4562$ from part (a). Assuming independence from shipment to shipment

and denoting by Y the number of shipments containing at least one defective item, Y follows another binomial distribution $b(y; 10, 0.4562)$. Therefore,

$$P(Y = 3) = \binom{10}{3} 0.4562^3 (1 - 0.4562)^7 = 0.1602.$$

5.2 Binomial and Multinomial Distributions

– Mean and Variance

The mean and variance of the binomial distribution $b(x; n, p)$ are

$$\mu = np \text{ and } \sigma^2 = npq.$$

Example 5.4: It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight into the true extent of the problem, it is determined that some testing is necessary. It is too expensive to test all of the wells in the area, so 10 are randomly selected for testing.

(a) Using the binomial distribution, what is the probability that exactly 3 wells have the impurity, assuming that the conjecture is correct?

(b) What is the probability that more than 3 wells are impure?

Solution: (a) We require

$$b(3; 10, 0.3) = \sum_{x=0}^{3} b(x; 10, 0.3) - \sum_{x=0}^{2} b(x; 10, 0.3) = 0.6496 - 0.3828 = 0.2668.$$

(b) In this case, $P(X > 3) = 1 - 0.6496 = 0.3504$.

5.2 Binomial and Multinomial Distributions

– Mean and Variance

Example 5.5: Find the mean and variance of the binomial random variable of Example 5.2, and then use Chebyshev's theorem (on page 137) to interpret the interval $\mu \pm 2\sigma$.

Solution: Since Example 5.2 was a binomial experiment with $n = 15$ and $p = 0.4$, by Theorem 5.1, we have

$$\mu = (15)(0.4) = 6 \text{ and } \sigma^2 = (15)(0.4)(0.6) = 3.6.$$

Taking the square root of 3.6, we find that $\sigma = 1.897$. Hence, the required interval is $6 \pm (2)(1.897)$, or from 2.206 to 9.794. Chebyshev's theorem states that the number of recoveries among 15 patients who contracted the disease has a probability of at least 3/4 of falling between 2.206 and 9.794 or, because the data are discrete, between 2 and 10 inclusive.

5.2 Binomial and Multinomial Distributions

– Multinomial Experiments and the Multinomial Distribution

Multinomial Distribution If a given trial can result in the k outcomes $E_1, E_2, \ldots, E_k$ with probabilities $p_1, p_2, \ldots, p_k$, then the probability distribution of the random variables $X_1, X_2, \ldots, X_k$, representing the number of occurrences for $E_1, E_2, \ldots, E_k$ in n independent trials, is

$$f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

with

$$\sum_{i=1}^{k} x_i = n \text{ and } \sum_{i=1}^{k} p_i = 1.$$

The multinomial distribution derives its name from the fact that the terms of the multinomial expansion of $(p_1 + p_2 + \cdots + p_k)^n$ correspond to all the possible values of $f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n)$.

5.2 Binomial and Multinomial Distributions

– Multinomial Experiments and the Multinomial Distribution

Example 5.7: The complexity of arrivals and departures of planes at an airport is such that computer simulation is often used to model the "ideal" conditions. For a certain airport with three runways, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet:

Runway 1: $p_1 = 2/9$,
Runway 2: $p_2 = 1/6$,
Runway 3: $p_3 = 11/18$.

What is the probability that 6 randomly arriving airplanes are distributed in the following fashion?

Runway 1: 2 airplanes,
Runway 2: 1 airplane,
Runway 3: 3 airplanes

***Solution*:** Using the multinomial distribution, we have

$$f\left(2,1,3;\frac{2}{9},\frac{1}{6},\frac{11}{18},6\right) = \binom{6}{2,1,3}\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3$$
$$= \frac{6!}{2!\,1!\,3!}\cdot\frac{2^2}{9^2}\cdot\frac{1}{6}\cdot\frac{11^3}{18^3} = 0.1127.$$

5.3 Hypergeometric Distribution

– The hypergeometric distribution does not require independence and is based on sampling done **without replacement**

In general, we are interested in the probability of selecting x successes from the k items labeled successes and $n - x$ failures from the $N - k$ items labeled failures when a random sample of size n is selected from N items. This is known as a **hypergeometric experiment**, that is, one that possesses the following two properties:

1. A random sample of size n is selected without replacement from N items.
2. Of the N items, k may be classified as successes and $N - k$ are classified as failures.

The number X of successes of a hypergeometric experiment is called a **hypergeometric random variable**. Accordingly, the probability distribution of the hypergeometric variable is called the **hypergeometric distribution**, and its values are denoted by $h(x; N, n, k)$, since they depend on the number of successes k in the set N from which we select n items.

5.3 Hypergeometric Distribution

– The hypergeometric distribution does not require independence and is based on sampling done **without replacement**

Hypergeometric Distribution The probability distribution of the hypergeometric random variable X, the number of successes in a random sample of size n selected from N items of which k are labeled **success** and $N-k$ labeled **failure**, is

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad \max\{0, n-(N-k)\} \le x \le \min\{n, k\}.$$

The range of x can be determined by the three binomial coefficients in the definition, where x and $n-x$ are no more than k and $N-k$, respectively, and both of them cannot be less than 0. Usually, when both k (the number of successes) and $N-k$ (the number of failures) are larger than the sample size n, the range of a hypergeometric random variable will be $x = 0, 1, \ldots, n$.

5.3 Hypergeometric Distribution

Example 5.9: Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

***Solution*:** Using the hypergeometric distribution with $n = 5$, $N = 40$, $k = 3$, and $x = 1$, we find the probability of obtaining 1 defective to be

$$h(1; 40, 5, 3) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011.$$

Once again, this plan is not desirable since it detects a bad lot (3 defectives) only about 30% of the time.

The mean and variance of the hypergeometric distribution $h(x; N, n, k)$ are

$$\mu = \frac{nk}{N} \text{ and } \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right).$$

5.5 Poisson Distribution and the Poisson Process

– Experiments yielding numerical values of a random variable X, the number of outcomes occurring **during a given time interval** or in a **specified region**, are called **Poisson experiments.**

- **The given time interval** may be of any length, such as a minute, a day, a week, a month, or even a year.
  – For example, a Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, the number of days school is closed due to snow during the winter, or the number of games postponed due to rain during a baseball season.
- **The specified region** could be a line segment, an area, a volume, or perhaps a piece of material. In such instances, X might represent the number of field mice *per acre*, the number of bacteria in a given culture, or the number of typing errors *per page*.

– A Poisson experiment is derived from the Poisson process and possesses the following properties. |*next slide*

5.5 Poisson Distribution and the Poisson Process

– **Properties of the Poisson Process**

1. The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region. In this sense we say that the Poisson process has no memory.
2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.
3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution.**

5.5 Poisson Distribution and the Poisson Process

Poisson Distribution The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

$$p(x; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

where λ is the average number of outcomes per unit time, distance, area, or volume and $e = 2.71828\ldots$.

The mean number of outcomes is computed from $\mu = \lambda t$, where t is the specific "time," "distance," "area," or "volume" of interest.

ıple 5.17: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution: Using the Poisson distribution with $x = 6$ and $\lambda t = 4$ and referring to Table A.2, we have

$$p(6; 4) = \frac{e^{-4}4^6}{6!} = \sum_{x=0}^{6} p(x; 4) - \sum_{x=0}^{5} p(x; 4) = 0.8893 - 0.7851 = 0.1042.$$

5.5 Poisson Distribution and the Poisson Process

ıple 5.18: Ten is the average number of oil tankers arriving each day at a certain port. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

***Solution*:** Let X be the number of tankers arriving each day. Then, using Table A.2, we have

$$P(X > 15) = 1 - P(X \leq 15) = 1 - \sum_{x=0}^{15} p(x; 10) = 1 - 0.9513 = 0.0487.$$

> Both the mean and the variance of the Poisson distribution $p(x; \lambda t)$ are λt.

Chapter 6
Some Continuous Probability Distributions

Chapter 6: Some Continuous Probability Distributions

- Probability density function
- Normal distribution & probability

- We will be familiarize with distributions needed to make inferences and to model certain random variables.

6.1 Continuous Uniform Distribution

– One of the simplest continuous distributions in all of statistics is the **continuous uniform distribution.** This distribution is characterized by a density function that is "flat," and thus the probability is uniform in a closed interval, say [A, B].

Uniform Distribution The density function of the continuous uniform random variable X on the interval $[A, B]$ is

$$f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A \le x \le B, \\ 0, & \text{elsewhere.} \end{cases}$$

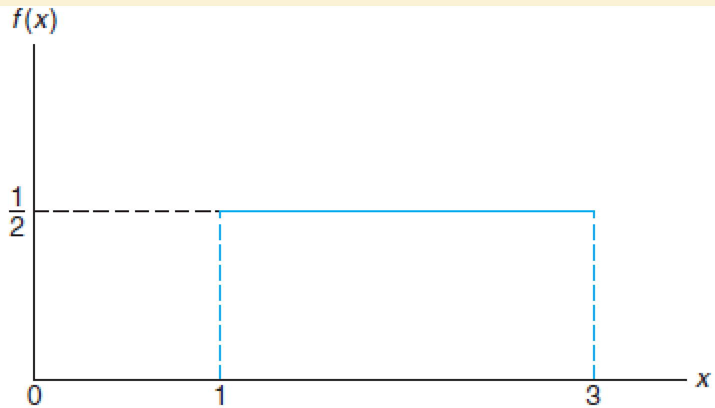


Figure 6.1: The density function for a random variable on the interval [1, 3].

6.1 Continuous Uniform Distribution

nple 6.1: Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $[0, 4]$.

(a) What is the probability density function?

(b) What is the probability that any given conference lasts at least 3 hours?

Solution: (a) The appropriate density function for the uniformly distributed random variable X in this situation is

$$f(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 4, \\ 0, & \text{elsewhere.} \end{cases}$$

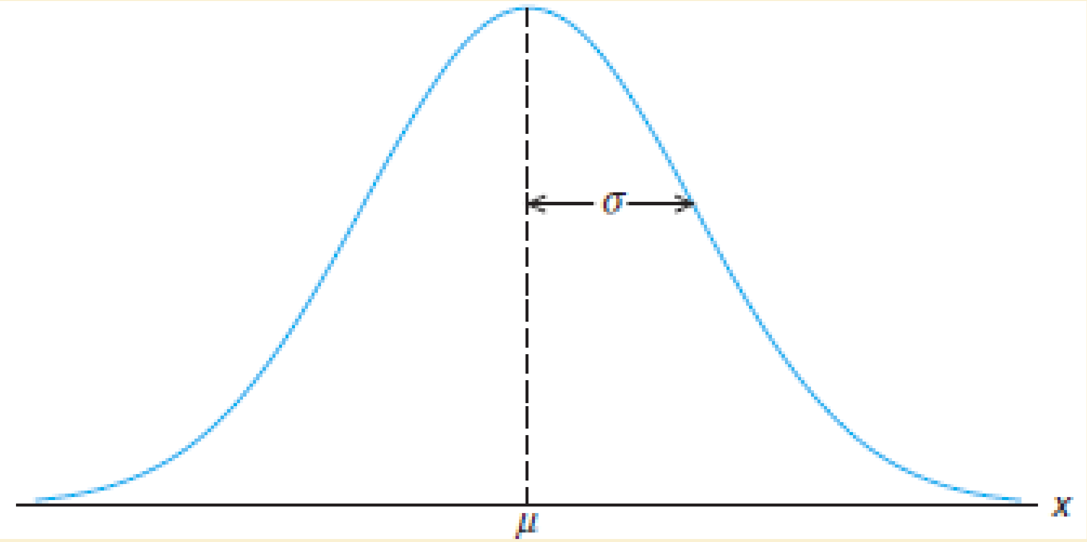
(b) $P[X \ge 3] = \int_3^4 \frac{1}{4}\, dx = \frac{1}{4}$.

The mean and variance of the uniform distribution are

$$\mu = \frac{A+B}{2} \text{ and } \sigma^2 = \frac{(B-A)^2}{12}.$$

6.2 Normal Distribution

- The **most important** continuous probability **distribution** in the entire field of statistics is the normal distribution. Its graph, called the **normal curve**, is the bell-shaped curve of following Figure, which approximately **describes many phenomena** that occur in nature, industry, and research.
  - For example, **physical measurements** are often more than adequately explained with a normal distribution. In addition, errors in scientific measurements are extremely well approximated by a normal distribution.
  - The normal distribution is often referred to as the **Gaussian distribution**, in honor of Karl Friedrich Gauss (1777–1855)



6.2 Normal Distribution

- The mathematical equation for the probability distribution of the normal variable depends on the two parameters μ and σ, its mean and standard deviation, respectively. Hence, we denote the values of the density of X by n(x; μ, σ).

Normal Distribution The density of the normal random variable X, with mean μ and variance σ^2, is

$$n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty,$$

where $\pi = 3.14159\ldots$ and $e = 2.71828\ldots$.

6.2 Normal Distribution

Once μ and σ are specified, the normal curve is completely determined. For example, if $\mu = 50$ and $\sigma = 5$, then the ordinates $n(x; 50, 5)$ can be computed for various values of x and the curve drawn. In Figure 6.3, we have sketched two normal curves having the same standard deviation but different means. The two curves are identical in form but are centered at different positions along the horizontal axis.

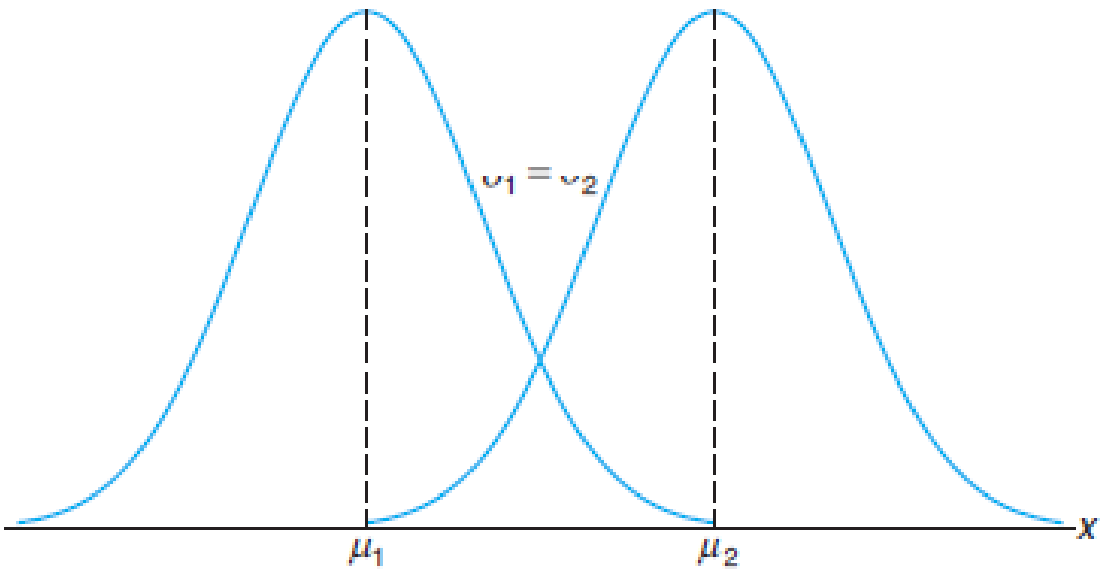


Figure 6.3: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 = \sigma_2$.

6.2 Normal Distribution

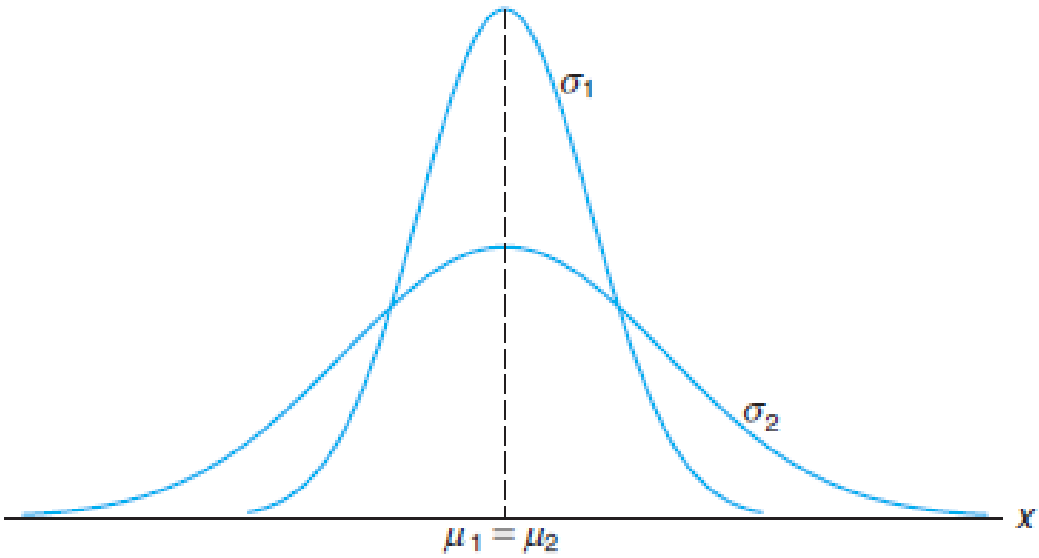


Figure 6.4: Normal curves with $\mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$.

In Figure 6.4, we have sketched two normal curves with the same mean but different standard deviations. This time we see that the two curves are centered at exactly the same position on the horizontal axis, but the curve with the larger standard deviation is lower and spreads out farther. Remember that the area under a probability curve must be equal to 1, and therefore the more variable the set of observations, the lower and wider the corresponding curve will be.

Figure 6.5 shows two normal curves having different means and different standard deviations. Clearly, they are centered at different positions on the horizontal axis and their shapes reflect the two different values of σ.

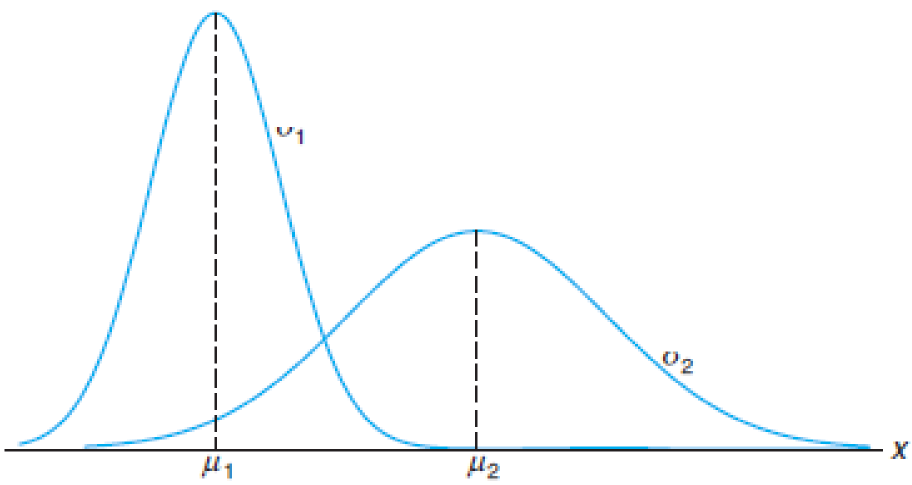


Figure 6.5: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 < \sigma_2$.

Based on inspection of Figures 6.2 through 6.5 and examination of the first and second derivatives of $n(x;\mu,\sigma)$, we list the following properties of the normal curve:

1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
2. The curve is symmetric about a vertical axis through the mean μ.
3. The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < X < \mu + \sigma$ and is concave upward otherwise.
4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.
5. The total area under the curve and above the horizontal axis is equal to 1.

6.3 Areas under the Normal Curve

The curve of any continuous probability distribution or density function is constructed so that the area under the curve bounded by the two ordinates $x = x_1$ and $x = x_2$ equals the probability that the random variable X assumes a value between $x = x_1$ and $x = x_2$. Thus, for the normal curve in Figure 6.6,

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x;\mu,\sigma)\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\,dx$$

is represented by the area of the shaded region.

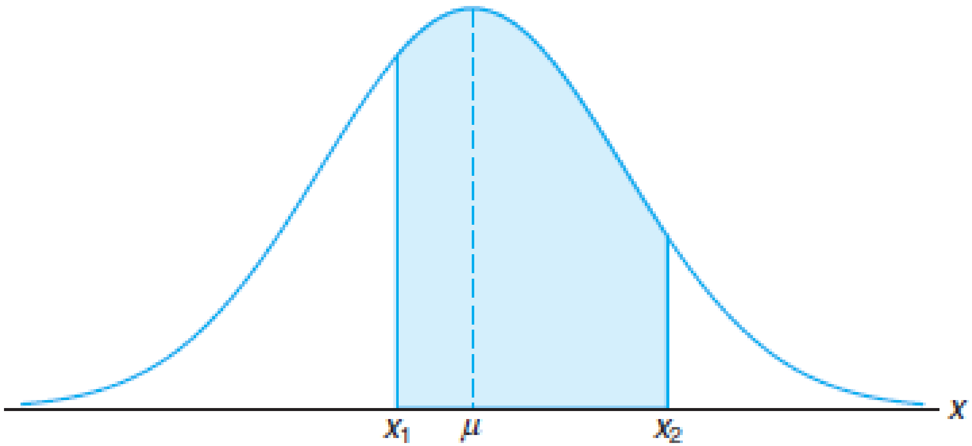


Figure 6.6: $P(x_1 < X < x_2) =$ area of the shaded region.

6.3 Areas under the Normal Curve

- There are many types of statistical software that can be used in calculating areas under the normal curve (for **many distributions)**. The **difficulty** encountered in **solving integrals** of normal density functions necessitates the tabulation of normal curve areas for quick reference. However, it would be a hopeless task to attempt to set up separate tables for every conceivable value of μ and σ. Fortunately, we are able to **transform all the observations of any normal random variable X** into a new set of observations of a **normal random variable Z with mean 0 and variance 1.** This can be done by means of the transformation :

$$Z = \frac{X - \mu}{\sigma}.$$

Whenever X assumes a value x, the corresponding value of Z is given by $z = (x - \mu)/\sigma$. Therefore, if X falls between the values $x = x_1$ and $x = x_2$, the random variable Z will fall between the corresponding values $z_1 = (x_1 - \mu)/\sigma$ and $z_2 = (x_2 - \mu)/\sigma$. Consequently, we may write

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{x1}^{x2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx = \frac{1}{\sqrt{2\pi}} \int_{z1}^{z2} e^{-\frac{1}{2}z^2}\, dz$$
$$= \int_{z1}^{z2} n(z;0,1)\, dz = P(z_1 < Z < z_2),$$

where Z is seen to be a normal random variable with mean 0 and variance 1.

6.3 Areas under the Normal Curve

mple 6.4: Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$, find the probability that X assumes a value between 45 and 62.

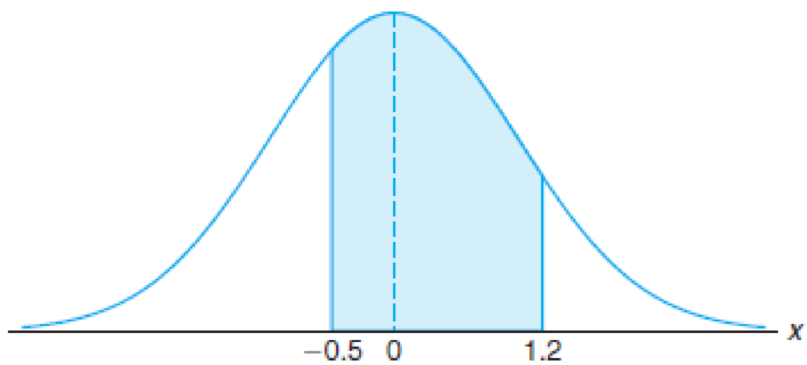


Figure 6.11: Area for Example 6.4.

***Solution*:** The z values corresponding to $x_1 = 45$ and $x_2 = 62$ are

$$z_1 = \frac{45 - 50}{10} = -0.5 \text{ and } z_2 = \frac{62 - 50}{10} = 1.2.$$

Therefore,

$$P(45 < X < 62) = P(-0.5 < Z < 1.2).$$

The distribution of a normal random variable with mean 0 and variance 1 is called a **standard normal distribution**.

Chapter 8
Sampling Distributions

Chapter 8: Sampling Distributions

- Show you how statistics computed vary from one sample to another
- What are the features of such variations in terms of mean, variance, and distribution

8.1 Random Sampling

- In this chapter, we focus on sampling from distributions or populations and study such important quantities as the sample mean and sample variance
  - Example;

    If the students of a certain school are given blood tests and the type of blood is of interest, then a descriptive representation might be more useful.

    A person’s blood can be classified in 8 ways: AB, A, B, or O, each with a plus or minus sign, depending on the presence or absence of the Rh antigen.

8.1 Random Sampling

- **Populations and Samples**
  - The totality of observations with which we are concerned, whether their number be finite or infinite, constitutes what we call a **population**
    - The number of observations in the population is defined to be the size of the population. If there are 600 students in the school whom we classified according to blood type, we say that we have a population of size 600.
  - A **sample** is a subset of a population.

Let $X_1, X_2, \ldots, X_n$ be n independent random variables, each having the same probability distribution $f(x)$. Define $X_1, X_2, \ldots, X_n$ to be a **random sample** of size n from the population $f(x)$ and write its joint probability distribution as

$$f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\cdots f(x_n).$$

8.2 Some Important Statistics

– **We shall, define some important statistics that describe corresponding measures of a random sample**

(a) Sample mean:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Note that the statistic $\bar{X}$ assumes the value $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$ when X_1 assumes the value x_1, X_2 assumes the value x_2, and so forth. The term *sample mean* is applied to both the statistic $\bar{X}$ and its computed value $\bar{x}$.

(b) Sample median:

$$\tilde{x} = \begin{cases} x_{(n+1)/2}, & \text{if } n \text{ is odd,} \\ \frac{1}{2}(x_{n/2} + x_{n/2+1}), & \text{if } n \text{ is even.} \end{cases}$$

The sample median is also a location measure that shows the middle value of the sample. Examples for both the sample mean and the sample median can be found in Section 1.3. The sample mode is defined as follows.

(c) The sample mode is the value of the sample that occurs most often.

8.2 Some Important Statistics

– We shall, define some important statistics that describe corresponding measures of a random sample

(a) Sample variance:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \quad (8.2.1)$$

mple 8.2: A comparison of coffee prices at 4 randomly selected grocery stores in San Diego showed increases from the previous month of 12, 15, 17, and 20 cents for a 1-pound bag. Find the variance of this random sample of price increases.

Solution: Calculating the sample mean, we get

$$\bar{x} = \frac{12 + 15 + 17 + 20}{4} = 16 \text{ cents}.$$

Therefore,

$$s^2 = \frac{1}{3}\sum_{i=1}^{4}(x_i - 16)^2 = \frac{(12-16)^2 + (15-16)^2 + (17-16)^2 + (20-16)^2}{3}$$

$$= \frac{(-4)^2 + (-1)^2 + (1)^2 + (4)^2}{3} = \frac{34}{3}.$$

8.2 Some Important Statistics

– We shall, define some important statistics that describe corresponding measures of a random sample

As in Chapter 1, the **sample standard deviation** and the **sample range** are defined below.

(b) Sample standard deviation:

$$S = \sqrt{S^2},$$

where S^2 is the sample variance.

Let $X_{\max}$ denote the largest of the X_i values and $X_{\min}$ the smallest.

(c) Sample range:

$$R = X_{\max} - X_{\min}.$$

8.3 Sampling Distributions

– The probability distribution of a statistic is called a **sampling distribution.**

What Is the Sampling Distribution of $\bar{X}$?

We should view the sampling distributions of $\bar{X}$ and S^2 as the mechanisms from which we will be able to make inferences on the parameters μ and σ^2. The sampling distribution of $\bar{X}$ with sample size n is the distribution that results when an **experiment is conducted over and over** (always with sample size n) **and the many values of $\bar{X}$ result**. This sampling distribution, then, describes the variability of sample averages around the population mean μ. In the case of the soft-drink machine, knowledge of the sampling distribution of $\bar{X}$ arms the analyst with the knowledge of a "typical" discrepancy between an observed $\bar{x}$ value and true μ. The same principle applies in the case of the distribution of S^2. The sampling distribution produces information about the variability of s^2 values around σ^2 in repeated experiments.

8.4 Sampling Distribution of Means and the Central Limit Theorem

Central Limit Theorem: If $\bar{X}$ is the mean of a random sample of size n taken from a population with mean μ and finite variance σ^2, then the limiting form of the distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}},$$

as $n \to \infty$, is the standard normal distribution $n(z; 0, 1)$.

8.4 Sampling Distribution of Means and the

mple 8.4: An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

***Solution*:** The sampling distribution of $\bar{X}$ will be approximately normal, with $\mu_{\bar{X}} = 800$ and $\sigma_{\bar{X}} = 40/\sqrt{16} = 10$. The desired probability is given by the area of the shaded region in Figure 8.2.

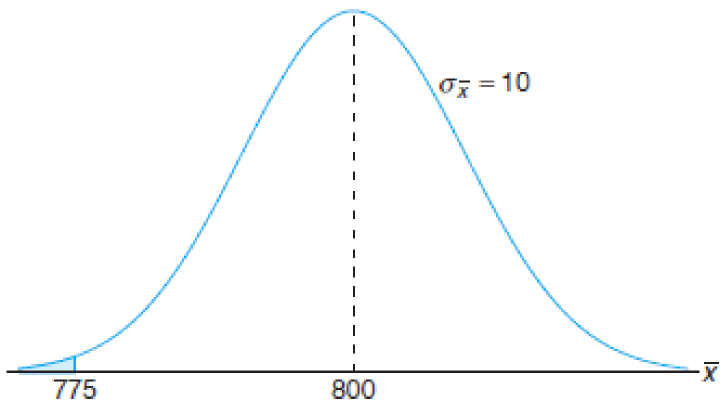


Figure 8.2: Area for Example 8.4.

Corresponding to $\bar{x} = 775$, we find that

$$z = \frac{775 - 800}{10} = -2.5,$$

and therefore

$$P(\bar{X} < 775) = P(Z < -2.5) = 0.0062.$$

Chapter 11
Simple Linear Regression and Correlation

Chapter 11: Simple Linear Regression and Correlation

- Confidence Intervals
- Test for Slope of Regression Lines
- Scatter Plots
- Correlation
- The Least Squares Regression Line
- Residuals
- Non-Linear Models
- Relations in Categorical Data

11.2 The Simple Linear Regression (SLR) Model

Simple Linear Regression Model

$$Y = \beta_0 + \beta_1 x + \epsilon.$$

In the above, β_0 and β_1 are unknown intercept and slope parameters, respectively, and ϵ is a random variable that is assumed to be distributed with $E(\epsilon) = 0$ and $\mathrm{Var}(\epsilon) = \sigma^2$. The quantity σ^2 is often called the error variance or residual variance.

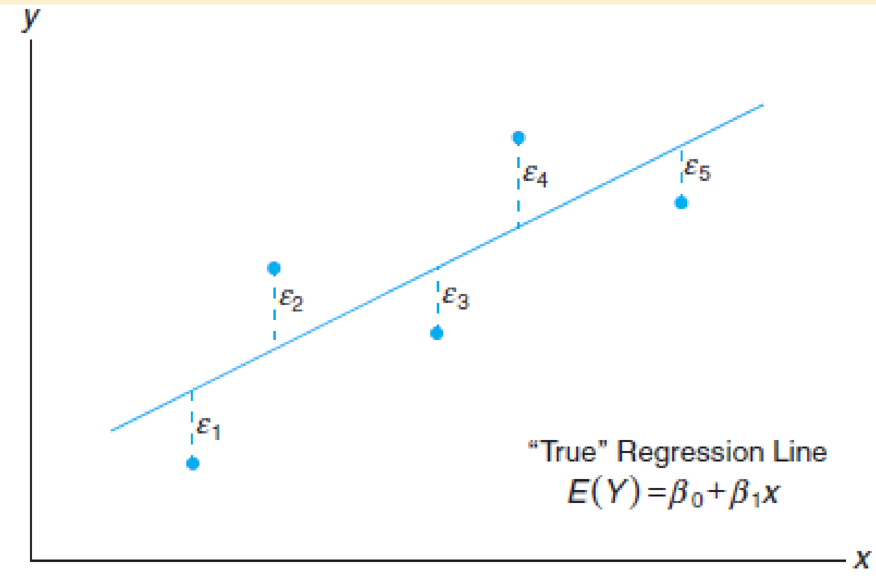


Figure 11.2: Hypothetical (x, y) data scattered around the true regression line for $n = 5$.

11.3 Least Squares and the Fitted Model

– In this section, we discuss the method of **fitting an estimated regression line** to the data.

Residual: Error in Fit Given a set of regression data $\{(x_i, y_i); i = 1, 2, \ldots, n\}$ and a fitted model, $\hat{y}_i = b_0 + b_1 x_i$, the ith residual e_i is given by

$$e_i = y_i - \hat{y}_i, \quad i = 1, 2, \ldots, n.$$

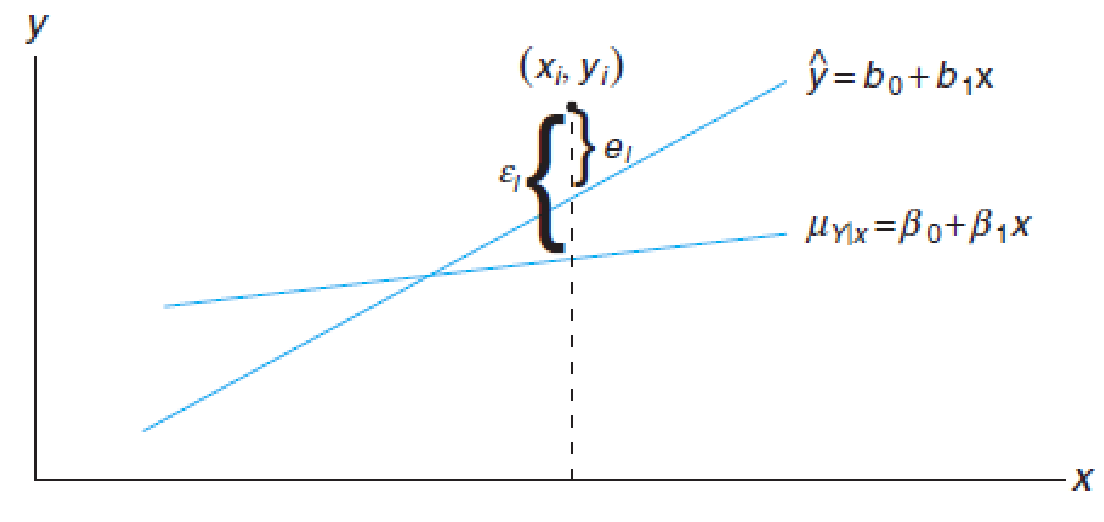


Figure 11.5: Comparing ϵ_i with the residual, e_i.

11.3 Least Squares and the Fitted Model

- **The Method of Least Squares**
- We shall find b0 and b1, the estimates of β0 and β1, so that the sum of the squares of the residuals is a minimum.

Estimating the Regression Coefficients

Given the sample $\{(x_i, y_i);\ i = 1, 2, \ldots, n\}$, the least squares estimates b_0 and b_1 of the regression coefficients β_0 and β_1 are computed from the formulas

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \text{ and}$$

$$b_0 = \frac{\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i}{n} = \bar{y} - b_1\bar{x}.$$

ple 11.1: Estimate the regression line for the pollution data of Table 11.1.

Table 11.1: Measures of Reduction in Solids and Oxygen Demand

Solids Reduction, x (%)	Oxygen Demand Reduction, y (%)	Solids Reduction, x (%)	Oxygen Demand Reduction, y (%)
3	5	36	34
7	11	37	36
11	21	38	38
15	16	39	37
18	16	39	36
27	28	39	45
29	27	40	39
30	25	41	41
30	35	42	40
31	30	42	44
31	40	43	37
32	32	44	44
33	34	45	46
33	32	46	46
34	34	47	49
36	37	50	51
36	38		

11.3 Least Squares and the Fitted Model

ple 11.1: Estimate the regression line for the pollution data of Table 11.1.

***Solution*:**

$$\sum_{i=1}^{33} x_i = 1104, \quad \sum_{i=1}^{33} y_i = 1124, \quad \sum_{i=1}^{33} x_i y_i = 41{,}355, \quad \sum_{i=1}^{33} x_i^2 = 41{,}086$$

Therefore,

$$b_1 = \frac{(33)(41{,}355) - (1104)(1124)}{(33)(41{,}086) - (1104)^2} = 0.903643 \text{ and}$$

$$b_0 = \frac{1124 - (0.903643)(1104)}{33} = 3.829633.$$

Thus, the estimated regression line is given by

$$\hat{y} = 3.8296 + 0.9036x.$$