University of Anbar
College of Engineering
Department of Electrical Engineering

Lecture Notes for the Class

Engineering Probability and Statistics

Lecturer: **Dr. Mohammed Ali AlMahamdy**
mohammed.almahamdy@uoanbar.edu.iq

بلا احصائيات انت مجرد شخص اخر لديه رأي.
"Without data, you're just another person with an opinion."
W. Edwards Deming

This print includes a collection of selected Materials from the books:

- Statistics for Engineers and Scientists, 4th edition, by Navidi
- Probability & Statistics for Engineers & Scientists, 9th edition, by Walpole et.al.
- Schaum's Outline Probability and Statistics 4th edition, by Spiegel et.al.

2021-2022

Contents

Introduction

Statistics in Engineering:

An engineer is someone who solves problems of interest to society by the efficient application of scientific principles. Engineers accomplish this by either refining an existing product or process or by designing a new product or process that meets customers' needs.

The field of **statistics** deals with the collection, presentation, analysis, and use of data to make decisions, solve problems, and design products and processes.

Variability: successive observations of a system or phenomenon do not produce the same result.

[illegible]

Random variable: let X be a measurement, μ is constant and ε is random disturbance. $X = \mu + \varepsilon$.

Sample versus **population**:

Probability versus Statistics: Probability deals with predicting the likelihood of future events, while statistics involves the analysis of the frequency of past events. Probability is primarily a theoretical branch of mathematics, which studies the consequences of mathematical definitions. Statistics is primarily an applied branch of mathematics, which tries to make sense of observations in the real world.

Possibility vs Probability: Technically possibility of any event is always 1 or 0 (i.e. 'yes' or 'no'). If an event is possible, how likely will it occur under the certain given situation is probability.

Skiena, Steven S., "Calculated Bets: Computers, Gambling, and Mathematical Modeling to Win," Cambridge University Press, 2001.

Accuracy vs Precision

Word	Meaning	Examples
Precision	• the quality of being exact. • Precision is how close the measured values are to **each other**.	• Extra precision is required to align the very best ... • Precision bombing was used to destroy enemy airbases.
Accuracy	• the fact of being exact or correct • the ability to do something without making mistakes • Accuracy is how close a measured value is to **the actual (true) value**.	• We can predict changes with a surprising degree of accuracy.

High Accuracy
Low Precision

Low Accuracy
High Precision

High Accuracy
High Precision

Mistakes in Data Analyses

- **Cherry picking** is the act of pointing to individual cases or data that seem to confirm a particular position while ignoring a significant portion of related cases or data that may contradict that position. It is a kind of fallacy of selective attention, the most common example of which is the confirmation bias. Cherry picking may be committed intentionally or unintentionally. This fallacy is a major problem in public debate.
  The term is based on the perceived process of harvesting fruit, such as cherries. The picker would be expected to only select the ripest and healthiest fruits. An observer who only sees the selected fruit may thus wrongly conclude that most, or even all, of the tree's fruit is in a likewise good condition. This can also give a false impression of the quality of the fruit (since it is only a sample and is not a representative sample).
  In other words: Cherry-picking occurs when someone researches a claim about a subject and then selects only the articles that support the claim or ignores any resolution of the data that could refute the claim. In other words, it's about selecting your facts to prove your claim.

- **Confirmation bias** is the tendency to search for, interpret, favor, and recall information in a way that confirms one's preexisting beliefs or hypotheses. It is a type of cognitive bias and a systematic error of inductive reasoning. People display this bias when they gather or remember information selectively, or when they interpret it in a biased way. The effect is stronger for emotionally charged issues and for deeply entrenched beliefs. Confirmation bias is a variation of the more general tendency of apophenia.
  People also tend to interpret ambiguous evidence as supporting their existing position. Biased search, interpretation and memory have been invoked to explain attitude polarization (when a disagreement becomes more extreme even though the different parties are exposed to the same evidence), belief perseverance (when beliefs persist after the evidence for them is shown to be false), the irrational primacy effect (a greater reliance on information encountered early in a series) and illusory correlation (when people falsely perceive an association between two events or situations).

- A **hasty generalization** is a fallacy in which a conclusion is not logically justified by sufficient or unbiased evidence. It's also called an *insufficient sample*, *a faulty generalization*, *a biased generalization*, and *jumping to a conclusion*.
  Drawing a conclusion based on a small sample size, rather than looking at statistics that are much more in line with the typical average of humans.
  - Example #1: My father smoked four packs of cigarettes a day since age fourteen and lived until age sixty-nine. Therefore, smoking really can't be that bad for you.
    Explanation: It is extremely unreasonable (and dangerous) to draw a universal conclusion about the health risks of smoking by the case study of one man.
  - Example #2: Four out of five dentists recommend Happy Glossy Smiles toothpaste brand. Therefore, it must be great.
    Explanation: It turns out that only five dentists were actually asked. When a random sampling of 1000 dentists was polled, only 20% actually recommended the brand. The

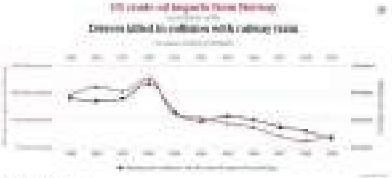
[illegible] out of five [illegible] or necessarily unusual sample [illegible]. It just happened to be [illegible] among small samples.

- **Spurious Relationships**: the old adage says "correlation does not imply causation." Just because two [illegible] fluctuate in tandem, this only proves that [illegible] doesn't prove that they are meaningfully related to one another.

  A spurious correlation is a mathematical relationship in which two or more events or variables are associated but not causally related, due to either coincidence or the presence of a certain third, unseen factor.

  The website (www.tylervigen.com) lists many strange correlations in public data sets.

  For example, the number of people who drowned by falling into a swimming pool correlates with the number of films Nicolas Cage appeared in. Another example: the per capita consumption of cheese correlates with the number of people who died by becoming tangled in their bedsheets. Also



The Wisdom of Crowds

In 1907, Sir Francis Galton asked ~800 villagers to guess the weight of an ox. None of them got the right answer, but when Galton averaged their guesses, he arrived at a near perfect estimate. This is a classic demonstration of the "wisdom of the crowds", where groups of people pool their abilities to show collective intelligence.

[illegible] the collective wisdom of a group will surpass that of even its smartest individual members.

Watch: [illegible]

Data Representation

Example 1: The following data is gathered from a survey:

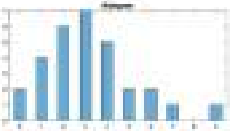
[illegible]

We can plot these data as:

The above plot is not that helpful to understand the results of the survey. One way is to better present the data is to use the **Histogram**.

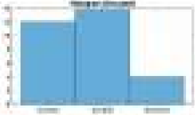
The data ranges from 0 to 9. We just count the number of occurrences of each possible value:

x	Frequency of x
0	2
1	4
2	6
3	7
4	5
5	2
6	2
7	1
8	0
9	1



We can group the data into different categories to yield a more meaningful histogram:

category	Frequency of x
$0 \le x < 3$	12
$3 \le x < 6$	14
$6 \le x \le 9$	4



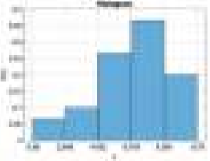
Using categories for histograms is practical in many applications.

Example2: plot the histogram for the following data using 5 categories:

[illegible]

Solution: here the number of points is $N = 30$, and the number of categories is $c = 5$, hence the range of each category is $\Delta = \frac{[illegible]}{5} = \frac{[illegible]}{5} = 1.7[illegible]$. So, we construct the table as:

Category label i	Category	Frequency of i n_i	Relative frequency $f(i) = \frac{n_i}{N}$
1	[illegible]	2	[illegible]
2	[illegible]	3	[illegible]
3	[illegible]	8	[illegible]
4	[illegible]	10	[illegible]
5	[illegible]	6	[illegible]



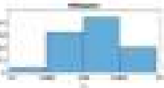
[illegible]

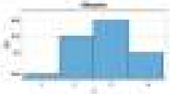
Example: plot the histogram for the following data using 4 categories:

[illegible]

Solution: here the number of points is $N = 25$, and the number of categories is $c = 4$, hence the range of each category is $\Delta = \frac{13 - 3.7}{4} = 2.325$. So we construct the table as:

Category label	Category description	Frequency of x_i n_i	Relative frequency $f(x) = \frac{n_i}{N}$
1	$3.7 \le x < 6.025$	1	0.04
2	$6.025 \le x < 8.35$	8	0.32
3	$8.35 \le x < 10.675$	11	0.44
4	$10.675 \le x \le 13$	5	0.20

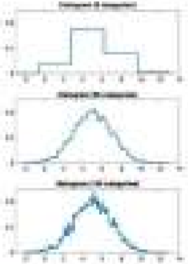


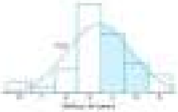


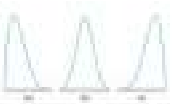
HW1: plot the histogram for the following data using 5 categories: [illegible]

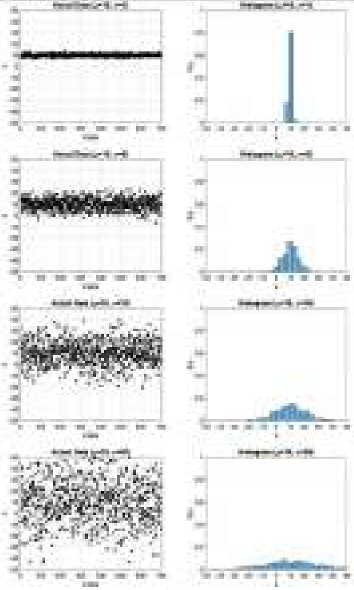
HW2: plot the histogram for the following data using 5 categories: [illegible]

As N increases, we can choose larger r. Hence, the shape of $f(x)$ approaches to the "normal distribution".









Probability

Experiment

[illegible]

Sample Space

[illegible]

Tree Diagram

[illegible]

Statement

[illegible]

- [illegible]
- [illegible]
- [illegible]

Event:

It is a subset of a sample space.

Example: Consider the experiment of rolling a six-sided die. The sample space for the chosen numbers is $S = \{1, 2, 3, 4, 5, 6\}$. Let the event A be the collection of the odd numbers only. Then $A = \{x \in S \mid x \text{ is odd}\} = \{1, 3, 5\}$.

Complement

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: Consider the sample space S = {book, cell phone, mp3, paper, laptop}. Let the event A be A = {cell phone, laptop}. Then A' = {book, mp3, paper}.

Combining Events

We often construct events by combining simpler events. Because events are subsets of sample spaces, it is traditional to use the notation of sets to describe events constructed in this way. We review the necessary notation here.

- The **union** of two events A and B, denoted $A \cup B$, is the set of outcomes that belong either to A, to B, or to both. In words, $A \cup B$ means "A or B." Thus the event $A \cup B$ occurs whenever either A or B (or both) occurs.
- The **intersection** of two events A and B, denoted $A \cap B$, is the set of outcomes that belong both to A and to B. In words, $A \cap B$ means "A and B." Thus the event $A \cap B$ occurs whenever both A and B occur.
- The **complement** of an event A, denoted A^c, is the set of outcomes that do not belong to A. In words, A^c means "not A." Thus the event A^c occurs whenever A does not occur.

Events can be graphically illustrated with Venn diagrams. Figure 2.1 illustrates the events $A \cup B$, $A \cap B$, and $B \cap A^c$.

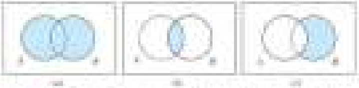


FIGURE 2.1 Venn diagrams illustrating various events: (a) $A \cup B$, (b) $A \cap B$, (c) $B \cap A^c$.

Mutually Exclusive Events

There are some events that can never occur together. For example, it is impossible that a coin can come up both heads and tails, and it is impossible that a steel pin can be both too long and too short. Events like this are said to be **mutually exclusive**.

Definition

- The events A and B are said to be **mutually exclusive** if they have no outcomes in common.
- More generally, a collection of events $A_1, A_2, \ldots, A_n$ is said to be mutually exclusive if no two of them have any outcomes in common.

The Venn diagram in Figure 2.2 illustrates mutually exclusive events.

FIGURE 2.2 The events A and B are mutually exclusive.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Counting Sample Points

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.

The Fundamental Principle of Counting
Suppose that k operations are to be performed. If there are n_1 ways to perform the first operation, and if for each of these ways there are n_2 ways to perform the second operation, and if for each choice of ways to perform the first two operations there are n_3 ways to perform the third operation, and so on, then the total number of ways to perform the sequence of k operations is $n_1 n_2 \cdots n_k$.

Example 2.18

When ordering a certain type of computer, there are 3 choices of hard drive, 4 choices for the amount of memory, 2 choices of video card, and 3 choices of monitor. In how many ways can a computer be ordered?

Solution:
The total number of ways to order a computer is $(3)(4)(2)(3) = 72$.

Example: 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.

Example: A developer of a new subdivision offers prospective home buyers a choice of Tudor, rustic, colonial, and traditional exterior styling in ranch, two-story, and split-level floor plans. In how many different ways can a buyer order one of these homes?

Solution: Since $n_1 = 4$ and $n_2 = 3$, a buyer must choose from

$$n_1 n_2 = (4)(3) = 12 \text{ possible homes.}$$

Example: If a 22-member club needs to elect a chair and a treasurer, how many different ways can these two to be elected?

Solution: For the chair position, there are 22 total possibilities. For each of those 22 possibilities, there are 21 possibilities to elect the treasurer. Using the multiplication rule, we obtain $n_1 \times n_2 = 22 \times 21 = 462$ different ways. ■

Permutations

A permutation is an ordering of a collection of objects. For example, there are six permutations of the letters A, B, C: ABC, ACB, BAC, BCA, CAB and CBA.

For any positive integer n, $n! = n(n-1)(n-2)\cdots(3)(2)(1)$.

Also, we define $0! = 1$.

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

Example 2.12

Five people stand in line at a movie theater. Into how many different orders can they be arranged?

Solution

The number of permutations of a collection of five people is $5! = (5)(4)(3)(2)(1) = 120$.

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

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

Example: 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

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

Example: 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.

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 $49 + 2352 = 2401$.

Example [illegible]

Five lifeguards are available for duty one Saturday afternoon. There are three lifeguard stations. In how many ways can three lifeguards be chosen and ordered among the stations?

Solution:

We use the fundamental principle of counting. There are 5 ways to choose a lifeguard to occupy the first station, then 4 ways to choose a lifeguard to occupy the second station, and finally 3 ways to choose a lifeguard to occupy the third station. The total number of permutations of three lifeguards chosen from 5 is therefore (5)(4)(3) = 60.

Combinations

In some cases, when choosing a set of objects from a larger set, we don't care about the ordering of the chosen objects; we care only which objects are chosen.

> The number of combinations of k objects chosen from a group of n objects is
>
> $$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad (2.32)$$

Example 2.14

In a certain event, 30 people attend, and 5 will be chosen at random to receive door prizes. The prizes are all the same, so the order in which the people are chosen does not matter. How many different groups of five people can be chosen?

Solution

Since the order of the five chosen people does not matter, we need to compute the number of combinations of 5 chosen from 30. This is

$$\binom{30}{5} = \frac{30!}{5!25!} = \frac{(30)(29)(28)(27)(26)}{(5)(4)(3)(2)(1)} = 142{,}506$$

Example: 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?

Solution: The number of ways of selecting 3 cartridges from 10 is

$$\binom{10}{3} = \frac{10!}{3!(10-3)!} = 120$$

The number of ways of selecting 2 cartridges from 5 is

$$\binom{5}{2} = \frac{5!}{2!3!} = 10$$

Using the multiplication rule (Rule 2.1) with $n_1 = 120$ and $n_2 = 10$, we have (120)(10) = 1200 ways.

Example: If ${}^nP_r = 3024$ and ${}^nC_r = 126$ then find n and r.

Solution:

$${}^nP_r = \frac{n!}{(n-r)!}$$

$${}^nC_r = \frac{n!}{(n-r)!\,r!}$$

$$\frac{{}^nP_r}{{}^nC_r} = \frac{3024}{126} = 24$$

$$\left[\frac{n!}{(n-r)!}\right] \div \left[\frac{n!}{(n-r)!\,r!}\right] = 24$$

$$\therefore r! = 24 \rightarrow \text{By comparison we find } r = 4$$

$${}^nP_4 = \frac{n!}{(n-4)!} = n(n-1)(n-2)(n-3) = 3024$$

By comparison we find n = 9.

H.W: If ${}^nP_r$ = [illegible] and ${}^nC_r$ = [illegible] then find n and r.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Probability of an Event

Definition

Given any experiment and any event A:

- The expression $P(A)$ denotes the probability that the event A occurs.
- $P(A)$ is the proportion of times that event A would occur in the long run, if the experiment were to be repeated over and over again.

If S is a sample space containing N equally likely outcomes, and if A is an event containing k outcomes, then

$$P(A) = \frac{k}{N} \quad (2.1)$$

Axioms of Probability

The Axioms of Probability

1. Let S be a sample space. Then $P(S) = 1$.
2. For any event A, $0 \le P(A) \le 1$.
3. If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$. More generally, if $A_1, A_2, \ldots$ are mutually exclusive events, then $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$

For any event A,

$$P(A^c) = 1 - P(A) \quad (2.2)$$

Let $\emptyset$ denote the empty set. Then

$$P(\emptyset) = 0 \quad (2.3)$$

Example: 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: 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 $P(E)$.

Solution: The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We assign a probability of w to each odd number and a probability of $2w$ to each even number. Since the sum of the probabilities must be 1, we have $9w = 1$ or $w = 1/9$. Hence, probabilities of 1/9 and 2/9 are assigned to each odd and even number, respectively. Therefore,

$$E = \{1, 2, 3\} \text{ and } P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

Example: In the example above, 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

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

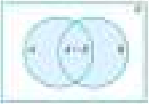
By assigning a probability of 1/9 to each odd number and 2/9 to each even number, we have

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

The Additive Rules

Let A and B be any events. Then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (2.5)$$



If A and B are mutually exclusive, then

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

For three events A, B and C,

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

Example: 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

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

Example: 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

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

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

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

If A and A' are complementary events, then

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

Example: 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

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

Exercise #[illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

[illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

Conditional Probability

A probability that is based on a part of a sample space is called a conditional probability. While the unconditional probability is based on the entire sample space.

Let A and B be events with $P(B) \neq 0$.

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

As an additional illustration, suppose that our sample space S is the population of adults in a small town who have completed the requirements for a college degree. We shall categorize them according to gender and employment status. The data are given in Table 2.1.

Table 2.1: Categorization of the Adults in a Small Town

	Employed	Unemployed	Total
Male	460	40	500
Female	140	260	400
Total	600	300	900

One of these individuals is to be selected at random for a tour throughout the country to publicize the advantages of establishing new industries in the town. We shall be concerned with the following events:

M: a man is chosen,

E: the one chosen is employed.

Using the reduced sample space E, we find that

$$P(M|E) = \frac{460}{600} = \frac{23}{30}.$$

Let $n(A)$ denote the number of elements in any set A. Using this notation, since each adult has an equal chance of being selected, we can write

$$P(M|E) = \frac{n(E \cap M)}{n(E)} = \frac{n(E \cap M)/n(S)}{n(E)/n(S)} = \frac{P(E \cap M)}{P(E)},$$

where $P(E \cap M)$ and $P(E)$ are found from the original sample space S. To verify this result, note that

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{and} \quad P(E \cap M) = \frac{460}{900} = \frac{23}{45}.$$

Hence,

$$P(M|E) = \frac{23/45}{2/3} = \frac{23}{30},$$

as before.

Example: 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 on time is

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

The Multiplication Rule

If A and B are two events with $P(A) \neq 0$, then

$$P(A \cap B) = P(A)P(B|A) \qquad (2.11)$$

Example: 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

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

Independent Events

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**.

If A and B are independent events, then

$$P(A \cap B) = P(A)P(B) \qquad (2.12)$$

Example [illegible]

A system contains two components, A and B. Both components must function for the system to work. The probability that component A fails is 0.08, and the probability that component B fails is 0.05. Assume the two components function independently. What is the probability that the system functions?

Solution

The probability that the system functions is the probability that both components function. Therefore

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

Since the components function independently,

$$\begin{aligned} P(\text{A functions and B functions}) &= P(\text{A functions})P(\text{B functions}) \\ &= [1 - P(\text{A fails})][1 - P(\text{B fails})] \\ &= (1 - 0.08)(1 - 0.05) \\ &= 0.874 \end{aligned}$$

Example [illegible]

Of the microprocessors manufactured by a certain process, 20% are defective. Five microprocessors are chosen at random. Assume they function independently. What is the probability that they all work?

Solution

For $i = 1, \ldots, 5$, let A_i denote the event that the ith microprocessor works. Then

$$\begin{aligned} P(\text{all 5 work}) &= P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\ &= P(A_1)P(A_2)P(A_3)P(A_4)P(A_5) \\ &= (1 - 0.20)^5 \\ &= 0.328 \end{aligned}$$

Example [illegible]

In Example [illegible], what is the probability that at least one of the microprocessors works?

Solution

The easiest way to solve this problem is to notice that

$$P(\text{at least one works}) = 1 - P(\text{all are defective})$$

Now, letting D_i denote the event that the ith microprocessor is defective,

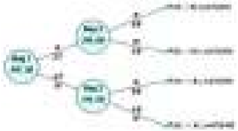
$$\begin{aligned} P(\text{all are defective}) &= P(D_1 \cap D_2 \cap D_3 \cap D_4 \cap D_5) \\ &= P(D_1)P(D_2)P(D_3)P(D_4)P(D_5) \\ &= (0.20)^5 \\ &= 0.00032 \end{aligned}$$

Therefore $P(\text{at least one works}) = 1 - 0.00032 = 0.99968$.

Example: 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

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



Reliability Analysis

Example 2.36

A certain system has two components, A and B, connected in series as shown in the following diagram.

The system will function only if both components function. The probability that A functions is given by $P(A) = 0.98$ and the probability that B functions is given by $P(B) = 0.95$. Assume that A and B function independently. Find the probability that the system functions.

Solution

Since the system will function only if both components function, it follows that

$$
\begin{aligned}
P(\text{system functions}) &= P(A \cap B) \\
&= P(A)P(B) \text{ by the assumption of independence} \\
&= (0.98)(0.95) \\
&= 0.931
\end{aligned}
$$

Example 2.34

A certain system has two components, C and D, connected in parallel as shown in the following diagram.

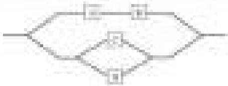
The system will function if either C or D functions. The probability that C functions is 0.95 and the probability that D functions is 0.85. Assuming C and D function independently, find the probability that the system functions.

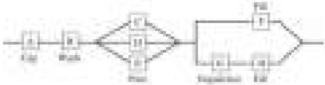
Solution

Since the system will function as long as either of the two components functions, it follows that

$$P(\text{system functions}) = P(C \cup D)$$
$$= P(C) + P(D) - P(C \cap D)$$
$$= P(C) + P(D) - P(C)P(D),$$
by the assumption of independence,
$$= 0.95 + 0.85 - (0.95)(0.85)$$
$$= 0.9925.$$

The reliability of more complex systems can often be determined by decomposing the system into a series of subsystems each of which contains components connected either in series or in parallel.





[illegible]

2.[illegible] Suppose the diagram of an electrical system is as given in Figure [illegible]. What is the probability that the system works? Assume the components fail independently.

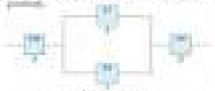


Figure [illegible]: Diagram for Exercise [illegible]

2.[illegible] A circuit system is given in Figure [illegible]. Assume the components fail independently.

(a) What is the probability that the entire system works?

(b) Given that the system works, what is the probability that the component A is not working?

Figure [illegible]: Diagram for Exercise [illegible]

The Law of Total Probability

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,

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

Example: 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

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

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

and hence

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

Bayes' Rule

If A and B are two events, we have seen that in most cases $P(A|B) \neq P(B|A)$. Bayes' rule provides a formula that allows us to calculate one of the conditional probabilities if we know the other one. To see how it works, assume that we know $P(B|A)$ and we wish to calculate $P(A|B)$. Start with the definition of conditional probability (Equation 2.13):

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

Now use Equation (2.13) to substitute $P(B|A)P(A)$ for $P(A \cap B)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad (2.25)$$

Equation (2.25) is essentially Bayes' rule. When Bayes' rule is written, the expression $P(B)$ in the denominator is usually replaced with a more complicated expression derived from the law of total probability. Specifically, since the events A and A^c are mutually exclusive and exhaustive, the law of total probability shows that

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) \qquad (2.26)$$

[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 $P(A) \neq 0$,

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

Review Examples

1.[illegible] A ball is drawn at random from a box containing 6 red balls, 4 white balls, and 5 blue balls. Determine the probability that it is (a) red, (b) white, (c) blue, (d) not red, (e) red or white.

(a) **Method 1**

Let R, W, and B denote the events of drawing a red ball, white ball and blue ball respectively. Then

$$P(R) = \frac{\text{ways of choosing a red ball}}{\text{total ways of choosing a ball}} = \frac{6}{6+4+5} = \frac{6}{15} = \frac{2}{5}$$

Method 2

Our sample space consists of $6 + 4 + 5 = 15$ sample points. Then if we assign equal probabilities $1/15$ to each sample point, we see that $P(R) = 6/15 = 2/5$, since there are 6 sample points corresponding to "red ball".

(b) $P(W) = \frac{4}{6+4+5} = \frac{4}{15}$

(c) $P(B) = \frac{5}{6+4+5} = \frac{5}{15} = \frac{1}{3}$

(d) $P(\text{not red}) = P(R^c) = 1 - P(R) = 1 - \frac{2}{5} = \frac{3}{5}$ by part (a).

(e) **Method 1**

$$P(\text{red or white}) = P(R \cup W) = \frac{\text{ways of choosing a red or white ball}}{\text{total ways of choosing a ball}}$$
$$= \frac{6+4}{6+4+5} = \frac{10}{15} = \frac{2}{3}$$

This can also be worked using the sample space as in part (a).

Method 2

$$P(R \cup W) = P(B^c) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3} \text{ by part (c).}$$

Method 3

Since events R and W are mutually exclusive, it follows from [illegible] that

$$P(R \cup W) = P(R) + P(W) = \frac{2}{5} + \frac{4}{15} = \frac{2}{3}$$

1.8. A fair die is tossed twice. Find the probability of getting a 4, 5, or 6 on the first toss and a 1, 2, 3, or 4 on the second toss.

Let A_1 be the event "4, 5, or 6 on first toss" and A_2 be the event "1, 2, 3, or 4 on second toss." Then we are looking for $P(A_1 \cap A_2)$.

Method 1

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = P(A_1)P(A_2) = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = \frac{1}{3}$$

We have used here the fact that the result of the second toss is independent of the first so that $P(A_2 \mid A_1) = P(A_2)$. Also we have used $P(A_1) = 3/6$ (since 4, 5, or 6 are 3 out of 6 equally likely possibilities) and $P(A_2) = 4/6$ (since 1, 2, 3, or 4 are 4 out of 6 equally likely possibilities).

Method 2

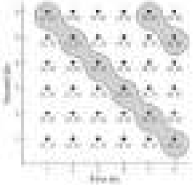
Each of the 6 ways in which a die can fall on the first toss can be associated with each of the 6 ways in which it can fall on the second toss, a total of $6 \cdot 6 = 36$ ways, all equally likely.

Each of the 3 ways in which A_1 can occur can be associated with each of the 4 ways in which A_2 can occur to give $3 \cdot 4 = 12$ ways in which both A_1 and A_2 can occur. Then

$$P(A_1 \cap A_2) = \frac{12}{36} = \frac{1}{3}$$

This shows directly that A_1 and A_2 are independent since

$$P(A_1 \cap A_2) = \frac{1}{3} = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = P(A_1)P(A_2)$$



1.11. Three balls are drawn successively from the box of Problem 1.7. Find the probability that they are drawn in the order red, white, and blue if each ball is (a) replaced, (b) not replaced.

Let R_1 = event "red on first draw," W_2 = event "white on second draw," B_3 = event "blue on third draw." We require $P(R_1 \cap W_2 \cap B_3)$.

(a) If each ball is replaced, then the events are independent and

$$P(R_1 \cap W_2 \cap B_3) = P(R_1)P(W_2 \mid R_1)P(B_3 \mid R_1 \cap W_2)$$
$$= P(R_1)P(W_2)P(B_3)$$
$$= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{6+4+5}\right)\left(\frac{5}{6+4+5}\right) = \frac{8}{225}$$

(b) If each ball is not replaced, then the events are dependent and

$$P(R_1 \cap W_2 \cap B_3) = P(R_1)P(W_2 \mid R_1)P(B_3 \mid R_1 \cap W_2)$$
$$= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{5+4+5}\right)\left(\frac{5}{5+3+5}\right) = \frac{4}{91}$$

1.12. Find the probability of a 4 turning up at least once in two tosses of a fair die.

Let A_1 = event "4 on first toss" and A_2 = event "4 on second toss." Then

$A_1 \cup A_2$ = event "4 on first toss or 4 on second toss or both"
= event "at least one 4 turns up,"

and we require $P(A_1 \cup A_2)$.

Method 1

Events A_1 and A_2 are not mutually exclusive, but they are independent [illegible]. Hence,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$
$$= P(A_1) + P(A_2) - P(A_1)P(A_2)$$
$$= \frac{1}{6} + \frac{1}{6} - \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{11}{36}$$

Method 2

$$P(\text{at least one 4 comes up}) + P(\text{no 4 comes up}) = 1$$

Then

$$P(\text{at least one 4 comes up}) = 1 - P(\text{no 4 comes up})$$
$$= 1 - P(\text{no 4 on 1st toss and no 4 on 2nd toss})$$
$$= 1 - P(A_1' \cap A_2') = 1 - P(A_1')P(A_2')$$
$$= 1 - \left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{11}{36}$$

Method 3

Total number of equally likely ways in which both dice can fall $= 6 \cdot 6 = 36$.

Also
Number of ways in which A_1 occurs but not A_2 = 5
Number of ways in which A_2 occurs but not A_1 = 5
Number of ways in which both A_1 and A_2 occur = 1

Then the number of ways in which at least one of the events A_1 or A_2 occurs $= 5 + 5 + 1 = 11$. Therefore, $P(A_1 \cup A_2) = 11/36$.

1.17. One bag contains 4 white balls and 2 black balls; another contains 3 white balls and 5 black balls. If one ball is drawn from each bag, find the probability that (a) both are white, (b) both are black, (c) one is white and one is black.

Let W_1 = event "white ball from first bag", W_2 = event "white ball from second bag".

(a) $P(W_1 \cap W_2) = P(W_1)P(W_2 \mid W_1) = P(W_1)P(W_2) = \left(\frac{4}{4+2}\right)\left(\frac{3}{3+5}\right) = \frac{1}{4}$

(b) $P(W_1' \cap W_2') = P(W_1')P(W_2' \mid W_1') = P(W_1')P(W_2') = \left(\frac{2}{4+2}\right)\left(\frac{5}{3+5}\right) = \frac{5}{24}$

(c) The required probability is

$$1 - P(W_1 \cap W_2) - P(W_1' \cap W_2') = 1 - \frac{1}{4} - \frac{5}{24} = \frac{13}{24}$$

1.18. Box I contains 3 red and 2 blue marbles while Box II contains 2 red and 8 blue marbles. A fair coin is tossed. If the coin turns up heads, a marble is chosen from Box I; if it turns up tails, a marble is chosen from Box II. Find the probability that a red marble is chosen.

Let R denote the event "a red marble is chosen" while I and II denote the events that Box I and Box II are chosen, respectively. Since a red marble can result by choosing either Box I or II, we can use the results of Problem 1.14 with $n = 2$, $A_1 = I$, $A_2 = II$. Therefore, the probability of choosing a red marble is

$$P(R) = P(I)P(R \mid I) + P(II)P(R \mid II) = \left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right) = \frac{2}{5}$$

1.19. In how many ways can 5 differently colored marbles be arranged in a row?

We must arrange the 5 marbles in 5 positions thus: – – – – –. The first position can be occupied by any one of 5 marbles, i.e., there are 5 ways of filling the first position. When this has been done, there are 4 ways of filling the second position. Then there are 3 ways of filling the third position, 2 ways of filling the fourth position, and finally only 1 way of filling the last position. Therefore:

$$\text{Number of arrangements of 5 marbles in a row} = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$$

In general,

$$\text{Number of arrangements of } n \text{ different objects in a row} = n(n-1)(n-2)\cdots 1 = n!$$

This is also called the number of permutations of n different objects taken n at a time and is denoted by ${}_nP_n$.

1.38 In how many ways can 10 people be seated on a bench if only 4 seats are available?

The first seat can be filled in any one of 10 ways, and when this has been done there are 9 ways of filling the second seat, 8 ways of filling the third seat, and 7 ways of filling the fourth seat. Therefore,

$$\text{Number of arrangements of 10 people taken 4 at a time} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

In general,

$$\text{Number of arrangements of } n \text{ different objects taken } r \text{ at a time} = n(n-1)\cdots(n-r+1)$$

This is also called the number of permutations of n different objects taken r at a time and is denoted by ${}_nP_r$. Note that when $r = n$, ${}_nP_n = n!$ as in Problem 1.37.

1.39 Evaluate (a) ${}_8P_3$, (b) ${}_6P_4$, (c) ${}_{15}P_1$, (d) ${}_3P_3$.

(a) ${}_8P_3 = 8 \cdot 7 \cdot 6 = 336$ (b) ${}_6P_4 = 6 \cdot 5 \cdot 4 \cdot 3 = 360$ (c) ${}_{15}P_1 = 15$ (d) ${}_3P_3 = 3 \cdot 2 \cdot 1 = 6$

1.40 It is required to seat 5 men and 4 women in a row so that the women occupy the even places. How many such arrangements are possible?

The men may be seated in ${}_5P_5$ ways, and the women in ${}_4P_4$ ways. Each arrangement of the men may be associated with each arrangement of the women. Hence,

$$\text{Number of arrangements} = {}_5P_5 \cdot {}_4P_4 = 5!\,4! = (120)(24) = 2880$$

1.41 How many 4-digit numbers can be formed with the 10 digits 0, 1, 2, …, 9 if (a) repetitions are allowed, (b) repetitions are not allowed, (c) the last digit must be zero and repetitions are not allowed?

(a) The first digit can be any one of 9 (since 0 is not allowed). The second, third, and fourth digits can be any one of 10. Then $9 \cdot 10 \cdot 10 \cdot 10 = 9000$ numbers can be formed.

(b) The first digit can be any one of 9 (any one but 0).

The second digit can be any one of 9 (any but that used for the first digit).

The third digit can be any one of 8 (any but those used for the first two digits).

The fourth digit can be any one of 7 (any but those used for the first three digits).

Then $9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

Another method

The first digit can be any one of 9, and the remaining three can be chosen in ${}_9P_3$ ways. Then $9 \cdot {}_9P_3 = 9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

(c) The first digit can be chosen in 9 ways, the second in 8 ways, and the third in 7 ways. Then $9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

Another method

The first digit can be chosen in 9 ways and the next two digits in ${}_8P_2$ ways. Then $9 \cdot {}_8P_2 = 9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

1.27. [illegible]

[illegible]

[illegible]

$$[illegible]$$

1.28. [illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

$$[illegible]$$

1.29. [illegible]

$$[illegible]$$

1.30. [illegible]

(a) **Method 1**

[illegible]

$$[illegible]$$

Method 2

$$[illegible]$$

(b) [illegible]

$$[illegible]$$

[illegible]

(c) [illegible]

$$[illegible]$$

[illegible]

Another method

[illegible]

[illegible] A box contains 5 red and 4 white marbles. Two marbles are drawn successively from the box without replacement and it is noted that the second one is white. What is the probability that the first is also white?

Method 1

[illegible]

$$[illegible]$$

Method 2

Since the second is known to be white, there are only 3 ways out of the remaining 8 in which the first can be white, so that the probability is 3/8.

[illegible] The probabilities that a husband and wife will be alive 20 years from now are given by 0.8 and 0.9, respectively. Find the probability that in 20 years (a) both, (b) neither, (c) at least one, will be alive.

Let H, W be the events that the husband and wife, respectively, will be alive in 20 years. Then $P(H) = 0.8$, $P(W) = 0.9$. We suppose that H and W are independent events, which may or may not be reasonable.

(a) $P(\text{both will be alive}) = P(H \cap W) = P(H)P(W) = (0.8)(0.9) = 0.72$.

(b) $P(\text{neither will be alive}) = P(H' \cap W') = P(H')P(W') = (0.2)(0.1) = 0.02$.

(c) $P(\text{at least one will be alive}) = 1 - P(\text{neither will be alive}) = 1 - 0.02 = 0.98$.

Random Variables

Concept of a Random Variable

It is often important to allocate a numerical description to the outcome. For example, the sample space giving a detailed description of each possible outcome when three electronic components are tested may be written $S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$, where N denotes nondefective and D denotes defective. One is naturally concerned with the number of defectives that occur. Thus each point in the sample space will be assigned a numerical value of 0, 1, 2, or 3. These values are, of course, random quantities determined by the outcome of the experiment. They may be viewed as values assumed by the random variable X, the number of defective items when three electronic components are tested. We shall use a capital letter, say X, to denote a random variable and its corresponding small letter, x in this case, for one of its values.

> A **random variable** assigns a numerical value to each outcome in a sample space.

Example: 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

Discrete Random Variables

> A random variable is **discrete** if its possible values form a discrete set. This means that if the possible values are arranged in order, there is a gap between each value and the next one. The set of possible values may be infinite; for example, the set of all integers and the set of all positive integers are both discrete sets.

Discrete Probability Distributions

A discrete random variable assumes each of its values with a certain probability. In the case of tossing a coin three times, the variable X, representing the number of heads, assumes the value 2 with probability 3/8, since 3 of the 8 equally likely sample points result in two heads and one tail.

x	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

The **probability mass function** of a discrete random variable X is the function $p(x) = P(X = x)$. The probability mass function is sometimes called the **probability distribution**.

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. $P(X = x) = f(x)$.

The Cumulative Distribution Function of a Discrete Random Variable

The probability mass function specifies the probability that a random variable is equal to a given value. A function called the **cumulative distribution function** specifies the probability that a random variable is less than or equal to a given value. The cumulative distribution function of the random variable X is the function $F(x) = P(X \leq x)$.

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

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

Example: Find the cumulative distribution function of the random variable X described in the following table. Using $F(x)$, verify that $f(2) = \frac{3}{8}$.

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Solution:

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

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,

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

Figure 3.1: Probability mass function plot

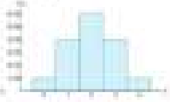
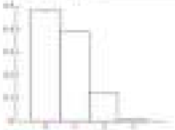


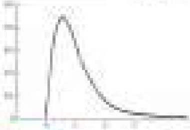
Figure 3.2: Probability histogram

Figure 3.3: Discrete cumulative distribution function

Continuous Random Variables

A random variable is continuous if its probabilities are given by areas under a curve. The curve is called a probability density function (PDF) for the random variable. The probability density function is sometimes called the probability distribution.

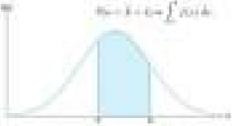




Computing Probabilities with the PDF

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

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



Let X be a continuous random variable with probability density function $f(x)$. Let a and b be any two numbers, with $a < b$. Then

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

In addition,

$$P(X \le b) = P(X < b) = \int_{-\infty}^{b} f(x)\,dx \qquad (2.32)$$

$$P(X \ge a) = P(X > a) = \int_{a}^{\infty} f(x)\,dx \qquad (2.33)$$

Example: 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 $P(0 < X \le 1)$.

Solution: We use Definition 10.

(a) Obviously, $f(x) \ge 0$. To verify condition 2 in Definition 10, 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 10, we obtain

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

■

The Cumulative Distribution Function of a Continuous Random Variable

> The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is
>
> $$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt.$$

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.

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

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

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,

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

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

Example: The Department of Energy (DOE) puts projects out on bid and generally estimates what a reasonable bid should be. Call the estimate b. The DOE has determined that the density function of the winning (low) bid is

$$f(y) = \begin{cases} [illegible] & [illegible] \\ 0, & \text{elsewhere.} \end{cases}$$

Find $F(y)$ and use it to determine the probability that the winning bid is less than the DOE's preliminary estimate b.

Solution: For [illegible],

$$F(y) = \int [illegible] \, dt = [illegible]$$

Figure 3.6: Continuous cumulative distribution function.

Thus,

$$F(y) = \begin{cases} 0, & [illegible] \\ [illegible], & [illegible] \\ 1, & [illegible] \end{cases}$$

To determine the probability that the winning bid is less than the preliminary bid estimate b, we have

$$P(Y \le b) = F(b) = [illegible]$$

Example [illegible]

Refer to Example [illegible]. Find the cumulative distribution function $F(x)$ and plot it.

Solution

The probability density function of X is given by $f(x) = 0$ if $x \le 0$, $f(x) = 1.5(1 - x^2)$ if $0 < x < 1$, and $f(x) = 0$ if $x \ge 1$. The cumulative distribution function is given by $F(x) = \int_{-\infty}^{x} f(t)\,dt$. Since $f(x)$ is defined separately on three different intervals, the computation of the cumulative distribution function involves three separate cases.

If $x \le 0$:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt \\ &= \int_{-\infty}^{x} 0\,dt \\ &= 0 \end{aligned}$$

If $0 < x < 1$:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt \\ &= \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{x} f(t)\,dt \\ &= \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} 1.5(1 - t^2)\,dt \\ &= 0 + 1.5\left(t - \frac{t^3}{3}\right)\Big|_{0}^{x} \\ &= 1.5\left(x - \frac{x^3}{3}\right) \end{aligned}$$

If $x \ge 1$:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt \\ &= \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{1} f(t)\,dt + \int_{1}^{x} f(t)\,dt \\ &= \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} 1.5(1 - t^2)\,dt + \int_{1}^{x} 0\,dt \\ &= 0 + 1.5\left[t - \frac{t^3}{3}\right]_{0}^{1} + 0 \\ &= 0 + 1 + 0 \\ &= 1 \end{aligned}$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

[illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

Joint Probability Distributions

When two or more random variables are associated with each item in a population, the random variables are said to be jointly distributed.

Jointly Discrete Random Variables

If X and Y are two discrete random variables, the probability distribution for their simultaneous occurrence can be represented by a function with values $f(x, y)$ for any pair of values (x, y) within the range of the random variables X and Y. It is customary to refer to this function as the **joint probability distribution** of X and Y.

Hence, in the discrete case,

$$f(x, y) = P(X = x, Y = y);$$

that is, the values $f(x, y)$ give the probability that outcomes x and y occur at the same time.

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, $P[(X, Y) \in A] = \sum \sum_{A} f(x, y)$.

Example [illegible]

Find the probability that $X = $ [illegible] mm.

x	y	$P(X = x$ and $Y = y)$
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]

Solution

$P(X = $ [illegible]$) = P(X = $ [illegible] and $Y = $ [illegible]$) + P(X = $ [illegible] and $Y = $ [illegible]$)$

$=$ [illegible] $+$ [illegible]

$=$ [illegible]

Example [illegible]

Find the probability that [illegible]

Solution

$P(\ldots) = $ [illegible]

[illegible]

[illegible]

[illegible]

It is [illegible] the joint probability mass function $f(x, y)$ as follows:

[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]

Example: A joint discrete probability function $f(x, y)$ is expressed below. Find $P(X + Y \le 1)$.

$$f(x,y) = \begin{cases} [\text{illegible}] & [\text{illegible}] \\ 0 & \text{elsewhere} \end{cases}$$

Solution: First we construct the table for the joint discrete probability distribution:

	$f(x, y)$	0	1	2
	0	[illegible]	[illegible]	[illegible]
[illegible]	1	[illegible]	[illegible]	[illegible]
	2	[illegible]	[illegible]	[illegible]

$$P(X + Y \le 1) = f(0,0) + f(0,1) + f(1,0)$$
$$= [\text{illegible}]$$

Jointly Continuous Random Variables

If X and Y are jointly continuous random variables, with joint probability density function $f(x, y)$, and $a \le b$, $c \le d$, then

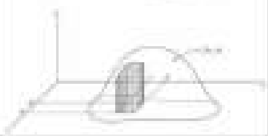
$$P(a \le X \le b \text{ and } c \le Y \le d) = \int_a^b \int_c^d f(x, y)\,dy\,dx$$

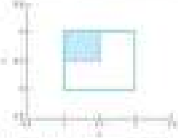
The joint probability density function has the following properties:

$$f(x, y) \ge 0 \text{ for all } x \text{ and } y$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$$

$$P(a < X < b, c < Y < d) = \int_c^d \int_a^b f(x, y)\,dx\,dy$$





Example: A joint continuous probability function $f(x, y)$ is expressed below. (a) Verify it is a valid probability distribution function for the given boundaries. Find $P[(X, Y) \in A]$, where $A = \{(x, y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$.

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

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

$$P[(X, Y) \in A] = P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right)$$

$$= \int_{1/4}^{1/2}\int_0^{1/2} \frac{2}{5}(2x + 3y)\, dx\, dy$$

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

$$= \left(\frac{y}{10} + \frac{3y^2}{10}\right)\Big|_{1/4}^{1/2}$$

$$= \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}$$

Marginal Distributions

> The **marginal distributions** of X alone and of Y alone are
>
> $$f_X(x) = g(x) = \sum_y f(x,y) \quad \text{and} \quad f_Y(y) = h(y) = \sum_x f(x,y)$$
>
> for the discrete case, and
>
> $$f_X(x) = g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{and} \quad f_Y(y) = h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$
>
> for the continuous case.

The term marginal is used here because, in the discrete case, the values of $g(x)$ and $h(y)$ are just the marginal totals of the respective columns and rows when the values of $f(x,y)$ are displayed in a rectangular table.

Example: Compute the marginal distributions for the following distribution table.

	$f(x,y)$	$x=0$	$x=1$	$x=2$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0
	2	$\frac{1}{28}$	0	0

	$f(x,y)$	$x=0$	$x=1$	$x=2$	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: 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. 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
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Example: Find $g(x)$ and $h(y)$ for the joint density function

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

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.

Example [illegible]

[illegible]

Solution

[illegible]

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy = [illegible]$$

and

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx = [illegible]$$

Example [illegible]

Refer to Example [illegible]. Find the marginal densities of X and of Y.

Solution

To compute $f_X(x)$, the marginal density of X, we fix x and integrate the joint density along the vertical line through x, as shown in Figure [illegible]. The integration is with respect to y, and the limits of integration are $y = 0$ to $y = x$.

$$f_X(x) = \int_0^x 8xy \, dy$$

$$= 4xy^2 \Big|_{y=0}^{y=x}$$

$$= 4x^3 \qquad \text{for } 0 < x < 1$$

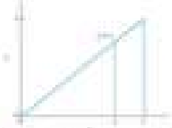


FIGURE [illegible] The marginal density $f_X(x)$ is computed by integrating the joint density along the vertical line through x.

To compute $f_Y(y)$, the marginal density of Y, we fix y and integrate the joint density along the horizontal line through y, as shown in Figure [illegible]. The integration is with respect to x, and the limits of integration are $x = y$ to $x = 1$.

$$f_Y(y) = \int_y^1 8xy \, dx$$

$$= 4x^2 y \Big|_{x=y}^{x=1}$$

$$= 4y - 4y^3 \qquad \text{for } 0 < y < 1$$

Note that

$$\int_{-\infty}^{\infty} g(x)\, dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dy\, dx = 1,$$

and

$$P(a < X < b) = P(a < X < b, -\infty < Y < \infty)$$
$$= \int_a^b \int_{-\infty}^{\infty} f(x, y)\, dy\, dx = \int_a^b g(x)\, dx.$$

Conditional Probability Distribution

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

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

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

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.

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

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

Example: Find the conditional distribution of X, given that $Y = 1$, and use it to determine $P(X = 0 \mid Y = 1)$.

$f(x,y)$		$x=0$	$x=1$	$x=2$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0
	2	$\frac{1}{28}$	0	0

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,

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

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

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,

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

Example: 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,

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

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

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

Then

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

(b) Therefore,

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

Example: Given the joint density function

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

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

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

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

and for $0<y<1$,

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

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

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

PS: $f(x|y)$ here does not depend on y.

Example [illegible]

Table [illegible] presents the joint probability mass function of the length X and width Y of [illegible]. Compute the conditional probability mass function $p_{Y|X}(y \mid 130)$.

Solution

The possible values for Y are [illegible]. From Table [illegible], [illegible]. Therefore,

$$
\begin{aligned}
p_{Y|X}(15 \mid 130) &= P(Y = 15 \mid X = 130) \\
&= \frac{P(Y = 15 \text{ and } X = 130)}{P(X = 130)} \\
&= \frac{[illegible]}{[illegible]} \\
&= [illegible]
\end{aligned}
$$

[illegible]

TABLE [illegible] [illegible]

x	[illegible]	[illegible]	$p_X(x)$
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
$p_Y(y)$	[illegible]	[illegible]	

Example

[illegible]

Solution

[illegible]

Independent Random Variables

If $f(x|y)$ does not depend on y, then $f(x|y) = g(x)$ and $f(x, y) = g(x)h(y)$.

> 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.

Example: Show that the random variables X and Y are not statistically independent.

	$f(x,y)$	$x=0$	$x=1$	$x=2$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0
	2	$\frac{1}{28}$	0	0

Proof: Let us consider the point $(0,1)$. From Table we find the three probabilities $f(0,1)$, $g(0)$, and $h(1)$ to be

$$f(0,1)=\frac{3}{14},$$

$$g(0)=\sum_{y=0}^{2} f(0,y)=\frac{3}{28}+\frac{3}{14}+\frac{1}{28}=\frac{5}{14},$$

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

Clearly,

$$f(0,1)\neq g(0)h(1),$$

and therefore X and Y are not statistically independent. ∎

Example [illegible]

[illegible]

Solution

[illegible]

Example [illegible]

[illegible]

Solution

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] [illegible] random variable X [illegible]

$$[illegible]$$

(a) [illegible] X [illegible]

(b) [illegible]

(c) [illegible] X.

[illegible] [illegible] random variables X and Y [illegible]

$$[illegible]$$

(a) Find the [illegible] of X and Y.

(b) Find [illegible]

[illegible] [illegible] random variables X and Y.

$$[illegible]$$

(a) Find the [illegible] of X and Y.

(b) Are X and Y independent?

(c) Find [illegible]

[illegible] [illegible]

$$[illegible]$$

(a) [illegible]

(b) [illegible]

Review Examples

Discrete random variables and probability distributions

2.1. Suppose that a pair of fair dice are to be tossed, and let the random variable X denote the sum of the points. Obtain the probability distribution for X.

Answer

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2.2. (a) Find the distribution function $F(x)$ for the random variable X of Problem 2.1, and (b) graph this distribution function.

(a) We have $F(x) = P(X \le x) = \sum_{u \le x} f(u)$. Then from the results of Problem 2.1, we find

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \le x < 3 \\ 3/36 & 3 \le x < 4 \\ 6/36 & 4 \le x < 5 \\ \vdots & \vdots \\ 35/36 & 11 \le x < 12 \\ 1 & 12 \le x < \infty \end{cases}$$

(b) See Fig. 2-6.

Fig. 2-6

Continuous random variables and probability distributions

2.3. A random variable X has the density function $f(x) = c/(x^2 + 1)$, where $-\infty < x < \infty$. (a) Find the value of the constant c. (b) Find the probability that X^2 lies between 1/3 and 1.

(a) We must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$, i.e.,

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2+1} = c \tan^{-1} x \Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

so that $c = 1/\pi$.

2.18 Find the distribution function corresponding to the density function of Problem 2.17.

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{du}{u^2+1} = \frac{1}{\pi}\left[\tan^{-1} u \Big|_{-\infty}^{x}\right]$$

$$= \frac{1}{\pi}\left[\tan^{-1} x - \tan^{-1}(-\infty)\right] = \frac{1}{\pi}\left[\tan^{-1} x + \frac{\pi}{2}\right]$$

$$= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$$

2.19 The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

(a)
$$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

(b)
$$P(X > 2) = \int_{2}^{\infty} 2e^{-2u}\,du = -e^{-2u}\Big|_{2}^{\infty} = e^{-4}$$

Another method

By definition, $P(X \leq 2) = F(2) = 1 - e^{-4}$. Hence,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

(c)
$$P(-3 < X \leq 4) = \int_{-3}^{4} f(u)\,du = \int_{-3}^{0} 0\,du + \int_{0}^{4} 2e^{-2u}\,du$$
$$= -e^{-2u}\Big|_{0}^{4} = 1 - e^{-8}$$

Another method

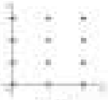
$$P(-3 < X \leq 4) = P(X \leq 4) - P(X \leq -3)$$
$$= F(4) - F(-3)$$
$$= (1 - e^{-8}) - (0) = 1 - e^{-8}$$

Joint distributions and independent variables

2.20 The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c. (b) Find $P(X = 2, Y = 1)$.
(c) Find $P(X \geq 1, Y \leq 2)$.

(a) The sample points (x, y) for which probabilities are different from zero are indicated in Fig. 2-6. The probabilities associated with these points, given by $c(2x + y)$, are shown in Table 2-6. Since the grand total, $42c$, must equal 1, we have $c = 1/42$.



2.40. Show that the random variables X and Y of Problem 2.8 are dependent.

If the random variables X and Y are independent, then we must have, for all x and y,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

But, as seen from Problems 2.8(b) and 2.9,

$$P(X = 2, Y = 1) = \frac{5}{42} \qquad P(X = 2) = \frac{14}{42} \qquad P(Y = 1) = \frac{11}{42}$$

so that

$$P(X = 2, Y = 1) \neq P(X = 2)P(Y = 1)$$

The result also follows from the fact that the joint probability function $(2x + y)/42$ cannot be expressed as a function of x alone times a function of y alone.

2.41. The joint density function of two continuous random variables X and Y is

$$f(x, y) = \begin{cases} cxy & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of the constant c. (c) Find $P(X \geq 3, Y \leq 2)$.
(b) Find $P(1 < X < 2, 2 < Y < 3)$.

(a) We must have the total probability equal to 1, i.e.,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$$

Using the definition of $f(x, y)$, the integral has the value

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\,dx\,dy = c\int_{x=0}^{4}\left[\int_{y=1}^{5} xy\,dy\right]dx$$

$$= c\int_{x=0}^{4} \frac{xy^2}{2}\Bigg|_{y=1}^{5} dx = c\int_{x=0}^{4}\left(\frac{25x}{2} - \frac{x}{2}\right)dx$$

$$= c\int_{x=0}^{4} 12x\,dx = c(6x^2)\Big|_{x=0}^{4} = 96c$$

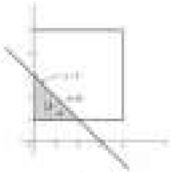
Then $96c = 1$ and $c = 1/96$.

(b) Using the value of c found in (a), we have

$$P(1 < X < 2, 2 < Y < 3) = \int_{x=1}^{2}\int_{y=2}^{3} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,dy\right]dx = \frac{1}{96}\int_{x=1}^{2} \frac{xy^2}{2}\Bigg|_{y=2}^{3} dx$$

$$= \frac{1}{96}\int_{x=1}^{2} \frac{5x}{2}\,dx = \frac{5}{192}\left(\frac{x^2}{2}\right)\Bigg|_{1}^{2} = \frac{5}{128}$$



[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

Miscellaneous problems

[illegible]

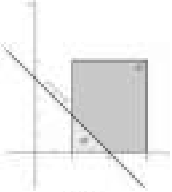
$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]



Mathematical Expectation

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.

Example 2.26

[illegible]

[illegible]

Find the mean of X.

Solution

[illegible]

[illegible]

[illegible]

Example: Find $E(X)$ of X whose pdf is expressed as:

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

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

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

Example: 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.

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.

Example: Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 9:00 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:

$$E[g(X)] = E(2X-1) = \sum_{x=4}^{9}(2x-1)f(x)$$

$$= (7)\left(\frac{1}{12}\right) + (9)\left(\frac{1}{12}\right) + (11)\left(\frac{1}{4}\right) + (13)\left(\frac{1}{4}\right)$$

$$+ (15)\left(\frac{1}{6}\right) + (17)\left(\frac{1}{6}\right) = \$12.67$$

Example: 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

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

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.

Example: Let X and Y be the random variables with joint probability distribution indicated in Table. Find the expected value of $g(X,Y) = XY$.

$f(x,y)$		$x=0$	$x=1$	$x=2$	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:

$$E(XY) = \sum_{x=0}^{2}\sum_{y=0}^{2} xy f(x,y)$$
$$= (0)(0)f(0,0) + (0)(1)f(0,1)$$
$$+ (1)(0)f(1,0) + (1)(1)f(1,1) + (2)(0)f(2,0)$$
$$= f(1,1) = \frac{3}{14}$$

Example: 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

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

Example [illegible]

The displacement of a piston in an internal combustion engine is defined to be the volume that the top of the piston moves through from the top to the bottom of its stroke. Let X represent the diameter of the cylinder bore, in millimeters, and let Y represent the length of the piston stroke in millimeters. The displacement is given by $D = \pi X^2 Y/4$. Assume X and Y are jointly distributed with joint probability mass function

$$f(x,y) = \begin{cases} 100 & 129.9 < x < 130.1 \text{ and } 99.8 < y < 100.0 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean of D.

Solution

$$\mu_D = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\pi x^2 y}{4} f(x,y)\,dy\,dx$$
$$= \int_{129.9}^{130.1}\int_{99.8}^{100.0} \frac{\pi x^2 y}{4}(100)\,dy\,dx$$
$$= 1{,}325{,}501$$

The mean displacement is 1,325,501 mm³, or approximately 1325 mL.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]

(a) [illegible]
(b) [illegible]

[illegible]
(a) [illegible]
(b) [illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

4.28 Consider the information in Exercise 3.28 on page 93. The problem deals with the weight in ounces of the product in a cereal box, with

$$f(x) = \begin{cases} \frac{2}{5}, & 23.75 \le x \le 26.25, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Plot the density function.

(b) Compute the expected value, or mean weight, in ounces.

(c) Are you surprised at your answer in (b)? Explain why or why not.

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.

4.30 In Exercise 3.31 on page 94, the distribution of times before a major repair of a washing machine was given as

$$f(y) = \begin{cases} \frac{1}{4}e^{-y/4}, & y \ge 0, \\ 0, & \text{elsewhere.} \end{cases}$$

What is the population mean of the times to repair?

4.31 Consider Exercise 3.32 on page 94.

(a) What is the mean proportion of the budget allocated to environmental and pollution control?

(b) What is the probability that a company selected at random will have allocated to environmental and pollution control a proportion that exceeds the population mean given in (a)?

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 $E(X^2)$.

Variance of Random Variables

The mean, or expected value, of a random variable X is of special importance in statistics because it describes where the probability distribution is centered. By itself, however, the mean does not give an adequate description of the shape of the distribution. We also need to characterize the variability in the distribution. In Figure 4.1, we have the histograms of two discrete probability distributions that have the same mean, $\mu = 2$, but differ considerably in variability, or the dispersion of their observations about the mean.

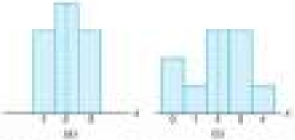


Figure 4.1: Distributions with equal means and unequal dispersions.

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.

The quantity $x - \mu$ in the definition above is called the deviation of an observation from its mean. Since the deviations are squared and then averaged, σ^2 will be much smaller for a set of x values that are close to μ than it will be for a set of values that vary considerably from μ.

Example: 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

f(x)

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

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

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

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

> The variance of a random variable X is
> $$\sigma^2 = E(X^2) - \mu^2.$$

Example: 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,

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

Example: The weekly demand for a drinking-water product, in thousands of liters, from a local chain of efficiency stores is a continuous random variable X having the probability density

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

Find the mean and variance of X.

Solution: Calculating $E(X)$ and $E(X^2)$, we have

$$\mu = E(X) = 2\int_1^2 x(x-1)\,dx = \frac{5}{3}$$

and

$$E(X^2) = 2\int_1^2 x^2(x-1)\,dx = \frac{17}{6}.$$

Therefore,

$$\sigma^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}.$$

Example [illegible]

[illegible]

Solution

[illegible]

$$\sigma_X^2 = [illegible] = 1.97$$

The standard deviation is $\sigma_X = \sqrt{1.97} = [illegible]$

Example 2.28

Use the alternate formula (Equation (2.31)) to compute the variance of X, the number of times a process is recalibrated.

Solution

In Example [illegible] the mean was computed to be $\mu_X = 1.30$. The variance is therefore

$$\sigma_X^2 = 0^2 P(X=0) + 1^2 P(X=1) + 2^2 P(X=2) + 3^2 P(X=3)$$
$$+ 4^2 P(X=4) - (1.30)^2$$
$$= 0^2(0.35) + 1^2(0.25) + 2^2(0.20) + 3^2(0.15) + 4^2(0.05) - 1.30^2$$
$$= 1.51$$

Example 2.29

A resistor in a certain circuit is specified to have a resistance in the range 99 Ω–101 Ω. An engineer obtains two resistors. The probability that both of them meet the specification is 0.36, the probability that exactly one of them meets the specification is 0.48, and the probability that neither of them meets the specification is 0.16. Let X represent the number of resistors that meet the specification. Find the probability mass function, and the mean, variance, and standard deviation of X.

Solution

The probability mass function is $P(X=0) = 0.16$, $P(X=1) = 0.48$, $P(X=2) = 0.36$, and $P(X=x) = 0$ for $x \neq 0, 1$, or 2. The mean is

$$\mu_X = 0(0.16) + 1(0.48) + 2(0.36)$$
$$= 1.20$$

The variance is

$$\sigma_X^2 = (0-1.20)^2(0.16) + (1-1.20)^2(0.48) + (2-1.20)^2(0.36)$$
$$= 0.4800$$

The standard deviation is $\sigma_X = \sqrt{0.4800} = 0.693$.

Example [illegible]

[illegible]

Solution

[illegible]

$$\mu_X = \int_{-\infty}^{\infty} x f(x)\,dx$$

[illegible]

[illegible]

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.

Example: 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 $2X+3$.

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

$$\sigma^2_{2X+3} = E\{[(2X+3) - \mu_{2X+3}]^2\} = E[(2X+3-6)^2]$$

$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4.$$

Example: 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 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,

$$\sigma^2_{4X+3} = E\{[(4X+3) - 8]^2\} = E[(4X-5)^2]$$

$$= \int_{-1}^{2}(4x-5)^2\frac{x^2}{3}\,dx = \frac{1}{3}\int_{-1}^{2}(16x^4 - 40x^3 + 25x^2)\,dx = \frac{51}{5}.$$

Covariance of Random Variables

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

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = E(XY) - \mu_X\mu_Y.$$

$$\mathrm{Cov}(X,Y) = \begin{cases} \text{Positive} \\ 0 \\ \text{Negative} \end{cases}$$

Example: The following is the joint probability distribution. Find the covariance of X and Y.

	$f(x,y)$	$x=0$	$x=1$	$x=2$	$h(y)$
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

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

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

and

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

Therefore,

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

Example [illegible]

Continuing Example [illegible] [illegible] the line $x = 1$ and the line $y = x$ [illegible] the joint density of X and Y is given by

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

Find $\mathrm{Cov}(X,Y)$.

We now compute [illegible] as well as [illegible]:

$$
\begin{aligned}
\mu_X &= \int_{-\infty}^{\infty} x f_X(x)\,dx \\
&= \int_0^1 4x^4\,dx \\
&= \frac{4}{5}
\end{aligned}
$$

$$
\begin{aligned}
\mu_Y &= \int_{-\infty}^{\infty} y f_Y(y)\,dy \\
&= \int_0^1 (4y^2 - 4y^4)\,dy \\
&= \frac{8}{15}
\end{aligned}
$$

$$
\mathrm{Cov}(X,Y) = \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225} \approx 0.01778
$$

Example [illegible]

Quality control checks on [illegible] panel. [illegible] X [illegible] Y [illegible] Find the covariance of X and Y.

	y			
x	0	1	2	$p_X(x)$
0	[illegible]	[illegible]	[illegible]	[illegible]
1	[illegible]	[illegible]	[illegible]	[illegible]
2	[illegible]	[illegible]	[illegible]	[illegible]
$p_Y(y)$	[illegible]	[illegible]	[illegible]	

Solution

We will use the formula $\mathrm{Cov}(X, Y) = \mu_{XY} - \mu_X \mu_Y$ [illegible]. First we compute μ_{XY}:

$$\mu_{XY} = \sum_{x=0}^{2} \sum_{y=0}^{2} xy\, p(x, y)$$

[illegible]

We use the marginals to compute μ_X and μ_Y:

$\mu_X = $ [illegible]

$\mu_Y = $ [illegible]

It follows that $\mathrm{Cov}(X, Y) = $ [illegible]

Correlation Coefficient

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 jointly distributed random variables with standard deviations σ_X and σ_Y. The correlation between X and Y is denoted $\rho_{X,Y}$ and is given by

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} \qquad \text{[illegible]}$$

For any two random variables X and Y:

$$-1 \le \rho_{XY} \le 1$$

Note that:

- ρ_{XY} is free of the units of X and Y.
- Value meaning: $\begin{cases} \rho_{XY} = 1 & \text{when exact linear dependency} \\ 0 < \rho_{XY} < 1 & \\ \rho_{XY} = 0 & \text{when } \sigma_{XY} = 0 \\ 0 > \rho_{XY} > -1 & \\ \rho_{XY} = -1 & \text{when exact negative linear dependency} \end{cases}$

Example: Find the correlation coefficient between X and Y.

	$f(x,y)$	$x=0$	$x=1$	$x=2$	$h(y)$
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

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

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

Example 2.69

Refer to Example 2.67. Find $\rho_{X,Y}$.

Solution

In Example 2.67, we computed [illegible] [illegible] [illegible] and [illegible]. We must now compute σ_X and σ_Y. To do this, we use the marginal densities of X and of Y, which were computed in Example 2.[illegible]. [illegible] We obtain

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - \mu_X^2$$

$$= \int [illegible]\,dx - \left([illegible]\right)^2$$

$$= [illegible]$$

$$\sigma_Y^2 = \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy - \mu_Y^2$$

$$= \int [illegible]\,dy - \left([illegible]\right)^2$$

$$= [illegible]$$

It follows that $\rho_{X,Y} = \dfrac{[illegible]}{\sqrt{[illegible]}} = [illegible]$

Example 2.70

Refer to Example 2.68. Find $\rho_{X,Y}$.

Solution

In Example 2.68, we computed $\operatorname{Cov}(X, Y) = [illegible]$, $\mu_X = [illegible]$, and $\mu_Y = [illegible]$. We must now compute σ_X and σ_Y. To do this, we use the marginal densities of X and of Y, which were presented in the table in Example 2.[illegible]. We obtain

$$\sigma_X^2 = \sum_{x} x^2 p_X(x) - \mu_X^2$$

$$= [illegible]$$

$$= [illegible]$$

$$\sigma_Y^2 = \sum_{y} y^2 p_Y(y) - \mu_Y^2$$

$$= [illegible]$$

$$= [illegible]$$

It follows that

$$\rho_{X,Y} = \frac{[illegible]}{\sqrt{[illegible]}} = [illegible]$$

For any random variable X, $\operatorname{Cov}(X, X) = \sigma_X^2$ and $\rho_{X,X} = 1$.

- If $Cov(X,Y) = \sigma_{XY} = 0$, then X and Y are said to be uncorrelated.
- If X and Y are independent, then X and Y are uncorrelated.
- It is mathematically possible for X and Y to be uncorrelated without being independent. This rarely occurs in practice.

[illegible]

[illegible]

Linear Combinations of Random Variables

One Random Variable

If a and b are constants, then

$$E(aX + b) = aE(X) + b$$

Proof: By the definition of expected value,

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b) f(x)\, dx = a\int_{-\infty}^{\infty} x f(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx$$

The first integral on the right is $E(X)$ and the second integral equals 1. Therefore, we have

$$E(aX + b) = aE(X) + b$$

Setting $a = 0$, we see that $E(b) = b$.

Setting $b = 0$, we see that $E(aX) = aE(X)$.

If X is a random variable, a and b are constants, for $Y = aX + b$:

$$E(Y) = aE(X) + b$$

$$\sigma_Y^2 = a^2\sigma_X^2 \rightarrow \sigma_Y = |a|\sigma_X$$

Example: Find $E(2X - 1)$

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

Solution:

Since $E(2X - 1) = 2E(X) - 1$

$$\mu = E(X) = \sum_{x=4}^{9} x f(x)$$

$$= (4)\left(\frac{1}{12}\right) + (5)\left(\frac{1}{12}\right) + (6)\left(\frac{1}{4}\right) + (7)\left(\frac{1}{4}\right) + (8)\left(\frac{1}{6}\right) + (9)\left(\frac{1}{6}\right) = \frac{41}{6}$$

Therefore,

$$E(2X - 1) = (2)\left(\frac{41}{6}\right) - 1 = [illegible]$$

Example: 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:

$$E(4X + 3) = 4E(X) + 3.$$

Now

$$E(X) = \int_{-1}^{2} x\left(\frac{x^2}{3}\right) dx = \int_{-1}^{2} \frac{x^3}{3} dx = \frac{5}{4}.$$

Therefore,

$$E(4X + 3) = (4)\left(\frac{5}{4}\right) + 3 = 8.$$

Example 2.10

[illegible] Let X represent the [illegible] diameter of a cylinder in millimeters. Assume that the probability density function of X is:

$$f(x) = \begin{cases} [illegible] & [illegible] < x < [illegible] \\ 0 & \text{otherwise} \end{cases}$$

[illegible]

Solution:

[illegible]

[illegible]

The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)].$$

Proof: By definition,

$$\begin{aligned} E[g(X) \pm h(X)] &= \int_{-\infty}^{\infty} [g(x) \pm h(x)] f(x)\, dx \\ &= \int_{-\infty}^{\infty} g(x) f(x)\, dx \pm \int_{-\infty}^{\infty} h(x) f(x)\, dx \\ &= E[g(X)] \pm E[h(X)]. \end{aligned}$$

Example: Let X be a random variable with probability distribution as follows:

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

Find the expected value of $Y = (X - 1)^2$.

Solution: We can write

$$E[(X-1)^2] = E(X^2 - 2X + 1) = E(X^2) - 2E(X) + E(1).$$

$E(1) = 1$, and by direct computation,

$$E(X) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (2)(0) + (3)\left(\frac{1}{6}\right) = 1 \text{ and}$$

$$E(X^2) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (4)(0) + (9)\left(\frac{1}{6}\right) = 2.$$

Hence,

$$E[(X-1)^2] = 2 - (2)(1) + 1 = 1.$$

Example: The weekly demand for a certain drink, in thousands of liters, at a chain of convenience stores is a continuous random variable $g(X) = X^2 + X - 2$, where X has the density function

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

Find the expected value of the weekly demand for the drink.

Solution: By Theorem 4.6, we write

$$E(X^2 + X - 2) = E(X^2) + E(X) - E(2).$$

$E(2) = 2$, and by direct integration,

$$E(X) = \int_1^2 2x(x-1)\,dx = \frac{5}{3} \text{ and } E(X^2) = \int_1^2 2x^2(x-1)\,dx = \frac{17}{6}.$$

Now

$$E(X^2 + X - 2) = \frac{17}{6} + \frac{5}{3} - 2 = \frac{5}{2}.$$

Two Random Variables

For the random variables X and Y, let $Z = aX + bY + c$ where a, b, and c are constants.

$$E(Z) = aE(X) + bE(Y) + c$$

$$\sigma_Z^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\mathrm{Cov}(X, Y)$$

If X and Y are independent, $\sigma_Z^2 = a^2\sigma_X^2 + b^2\sigma_Y^2$.

> The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions. That is,
>
> $$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)].$$

> Let X and Y be two independent random variables. Then
>
> $$E(XY) = E(X)E(Y).$$

Proof: By Definition [illegible],

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x, y)\, dx\, dy.$$

Since X and Y are independent, we may write

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

where $g(x)$ and $h(y)$ are the marginal distributions of X and Y, respectively. Hence,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy g(x)h(y)\, dx\, dy = \int_{-\infty}^{\infty} x g(x)\, dx \int_{-\infty}^{\infty} y h(y)\, dy$$
$$= E(X)E(Y).$$

Example: X and Y are independent random variables with the joint density function

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

Show that $E(XY) = E(X)E(Y)$.

Solution: By definition,

$$E(XY) = \int_0^1 \int_0^2 \frac{x^2 y(1+3y^2)}{4}\,dx\,dy = \frac{5}{6},\quad E(X) = \frac{4}{3},\ \text{and } E(Y) = \frac{5}{8}.$$

Hence,

$$E(X)E(Y) = \left(\frac{4}{3}\right)\left(\frac{5}{8}\right) = \frac{5}{6} = E(XY).$$

Example 2.71

Continuing Example [illegible], assume that the mobile computer moves from a random position (X, Y) vertically to the point $(X, 0)$, and then along the x axis to the origin. Find the mean and variance of the distance traveled.

Solution

The distance traveled is the sum $X + Y$. The means of X and Y were computed in Example [illegible]. They are $\mu_X = E(X) = 0.6000$ and $\mu_Y = E(Y) = 0.533$. We compute

$$\begin{aligned} \mu_{X+Y} &= \mu_X + \mu_Y \\ &= 0.6000 + 0.5333 \\ &= 1.1333 \end{aligned}$$

To compute σ_{X+Y} we use Equation [illegible]. In Example [illegible] we computed $\mathrm{Cov}(X, Y) = 0.01778$. In Example [illegible] we computed $\sigma_X^2 = 0.06667$ and $\sigma_Y^2 = 0.08889$. Therefore

$$\begin{aligned} \sigma_{X+Y}^2 &= \sigma_X^2 + \sigma_Y^2 + 2\,\mathrm{Cov}(X, Y) \\ &= 0.06667 + 0.08889 + 2(0.01778) \\ &= 0.1911 \end{aligned}$$

Example

[illegible]

[illegible]

Solution

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible] [illegible]

[illegible] [illegible]

[illegible] [illegible]

$$[illegible]$$

[illegible] [illegible]

$$[illegible]$$

[illegible]

[illegible] [illegible]

$$[illegible]$$

[illegible]

[illegible] [illegible]

$$[illegible]$$

[illegible]

[illegible] [illegible]

[illegible]

Find
(a) [illegible]
(b) [illegible]

[illegible] [illegible]

[illegible] [illegible]

[illegible] [illegible]

$$[illegible]$$

[illegible]

[illegible] [illegible]

$$[illegible]$$

(a) [illegible]
(b) [illegible]
(c) [illegible]
(d) [illegible]

[illegible] [illegible]

$$[illegible]$$

(a) [illegible]
(b) Find [illegible] and Var([illegible]).

[illegible] [illegible]

$$[illegible]$$

(a) [illegible]
(b) [illegible]

[illegible] [illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

(a) [illegible]
(b) [illegible]
(c) [illegible]

[illegible]

$$[illegible]$$

(a) Find the mean life of the [illegible].
(b) Find $E(X^2)$.
(c) Find the variance and standard deviation of the random variable X.

[illegible] Consider the joint density function

$$[illegible]$$

Compute the correlation coefficient [illegible].

Review Examples

[illegible]

3.5. The joint density function of two random variables X and Y is given by

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X)$, (b) $E(Y)$, (c) $E(XY)$, (d) $E(2X + 3Y)$.

(a) $$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} x\left(\frac{xy}{96}\right)dx\,dy = \frac{8}{3}$$

(b) $$E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} y\left(\frac{xy}{96}\right)dx\,dy = \frac{31}{9}$$

(c) $$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (xy)\left(\frac{xy}{96}\right)dx\,dy = \frac{248}{27}$$

(d) $$E(2X + 3Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (2x + 3y)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (2x + 3y)\left(\frac{xy}{96}\right)dx\,dy = \frac{47}{3}$$

Another method

(c) Since X and Y are independent, we have, using parts (a) and (b),

$$E(XY) = E(X)E(Y) = \left(\frac{8}{3}\right)\left(\frac{31}{9}\right) = \frac{248}{27}$$

(d) By Theorems 3-1 and 3-2, pages 76–77, together with (a) and (b),

$$E(2X + 3Y) = 2E(X) + 3E(Y) = 2\left(\frac{8}{3}\right) + 3\left(\frac{31}{9}\right) = \frac{47}{3}$$

Variance and standard deviation

3.6. Find (a) the variance, (b) the standard deviation of the sum obtained in tossing a pair of fair dice.

(a) Referring to Problem 3.2, we have $E(X) = E(Y) = 7/2$. Moreover,

$$E(X^2) = E(Y^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + \cdots + 6^2\left(\frac{1}{6}\right) = \frac{91}{6}$$

Then, by Theorem 3-4,

$$\operatorname{Var}(X) = \operatorname{Var}(Y) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

and, since X and Y are independent, Theorem 3-7 gives

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) = \frac{35}{6}$$

(b) $$\sigma_{X+Y} = \sqrt{\operatorname{Var}(X + Y)} = \sqrt{\frac{35}{6}}$$

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

Another method
[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible] Find the first four moments about the origin for a random variable X having density function [illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

Conditional expectation, variance, and moments

[illegible]

The Percentiles of a Population

Let $0 < p < 100$. The p^{th} percentile of a population is the value x_p such that $p\%$ of the population values are less than or equal to x_p. Thus if X is a continuous random variable with probability density function $f(x)$, the p^{th} percentile of X is the point x_p that solves the equation:

$$P(X \le x_p) = \int_{-\infty}^{x_p} f(x)\,dx = \frac{p}{100}$$

The Median of a Population

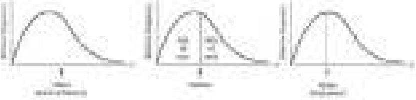
The median is a special case of a percentile; it is the 50th percentile.

Let X be a continuous random variable with probability mass function $f(x)$ and cumulative distribution function $F(x)$.

- The median of X is the point x_m that solves the equation $F(x_m) = P(X \le x_m) = \int_{-\infty}^{x_m} f(x)\,dx = 0.5$.

The Mode of a Population

The mode of a discrete random variable is that value which occurs most often or, in other words, has the greatest probability of occurring. The mode of a continuous random variable X is the value (or values) of X where the probability density function has a relative maximum.



Example [illegible]

[illegible]

$$[illegible]$$

[illegible]

Solution

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

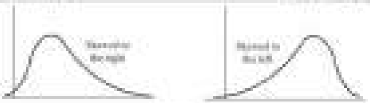
[illegible]

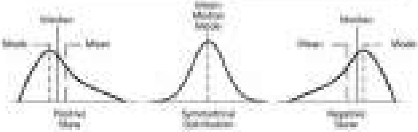
Skewness

[illegible]

$$[illegible]$$

[illegible]





Kurtosis

In some cases a distribution may have its values concentrated near the mean so that the distribution has a large peak as indicated by the solid curve of the figure below. In other cases, the distribution may be relatively flat as in the dashed curve.

Measures of the degree of peak around the mean of a distribution are called coefficients of kurtosis, or briefly kurtosis.

$$\alpha_4 = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4}$$

This is usually compared with the normal curve, which has a coefficient of kurtosis equal to 3.

Example: The density function of a continuous random variable X is

$$f(x) = \begin{cases} 4x(9 - x^2)/81 & 0 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the mode. (b) Find the median. (c) Compare mode, median, and mean.

Solution:

(a) The mode is obtained by finding where the density $f(x)$ has a relative maximum. The relative maxima of $f(x)$ occur where the derivative is zero, i.e.,

$$\frac{d}{dx}\left[\frac{4x(9 - x^2)}{81}\right] = \frac{36 - 12x^2}{81} = 0$$

Then $x = \sqrt{3} = 1.73$ approx., which is the required mode. Note that this does give the maximum since the second derivative, $-24x/81$, is negative for $x = \sqrt{3}$.

(b) The median is that value a for which $P(X \le a) = 1/2$. Now, for $0 < a < 3$,

$$P(X \le a) = \frac{4}{81}\int_0^a x(9 - x^2)\,dx = \frac{4}{81}\left(\frac{9a^2}{2} - \frac{a^4}{4}\right)$$

Setting this equal to 1/2, we find that

$$2a^4 - 36a^2 + 81 = 0$$

from which

$$a^2 = \frac{36 \pm \sqrt{(36)^2 - 4(2)(81)}}{2(2)} = \frac{36 \pm \sqrt{648}}{4} = 9 \pm \frac{9}{2}\sqrt{2}$$

Therefore, the required median, which must lie between 0 and 3, is given by

$$a^2 = 9 - \frac{9}{2}\sqrt{2}$$

from which $a = 1.62$ approx.

(c) $$E(X) = \frac{4}{81}\int_0^3 x^2(9 - x^2)\,dx = \frac{4}{81}\left(3x^3 - \frac{x^5}{5}\right)\Bigg|_0^3 = 1.60$$

which is practically equal to the median. The mode, median, and mean are shown in Fig. 3-6.

Fig. 3-6

Example:

Determine the (a) 10th, (b) 25th, (c) 75th percentile values for the distribution of the example above.

Solution:

From that example, we have

$$F(x) = [illegible]$$

(a) The 10th percentile is the value of x for which [illegible] $= 0.10$ [illegible]

(b) The 25th percentile is the value of x such that [illegible] $= 0.25$ [illegible]

(c) The 75th percentile is the value of x such that [illegible] $= 0.75$ [illegible]

Homework

For each of the density functions of a continuous random variable X shown below, find (a) the mode, (b) the median and (c) [illegible]

(1) $f(x) = \begin{cases} \frac{625}{114}[illegible] & [illegible] \\ 0 & \text{otherwise} \end{cases}$

(2) $f(x) = \begin{cases} [illegible] & [illegible] \\ 0 & \text{otherwise} \end{cases}$

(3) $f(x) = \begin{cases} [illegible] & [illegible] \\ 0 & \text{otherwise} \end{cases}$

Commonly Used Distributions

The Uniform Distribution

This distribution is characterized by a density function that is "flat" and thus the probability is uniform in a closed bounded interval say [a, b].

Discrete uniform distribution

A finite number of values are equally likely to be observed; every one of n values has equal probability 1/n.

$$P(X = k) = f(k) = \begin{cases} \frac{1}{n} & k = k_1, k_2, k_3, \ldots, k_n \\ 0 & \text{elsewhere} \end{cases}$$

Example: A simple example of the discrete uniform distribution is throwing a fair die. The possible values are 1, 2, 3, 4, 5, 6; and each time the die is thrown the probability of a given score is 1/6.

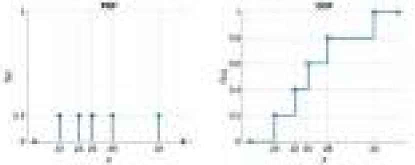
The probability of getting 2 is $P(X = 2) = f(2) = \frac{1}{6}$

The probability of getting a number less than 4 is:

$P(X < 4) = P(X \le 3) = F(3) = P(1) + P(2) + P(3) = \frac{3}{6} = 0.5$

Note that if two dice are thrown and their values added, the resulting distribution is no longer uniform because not all sums have equal probability.

Example: the discrete random variable X are uniformly distributed on [illegible]. Plot PDF and CDF.



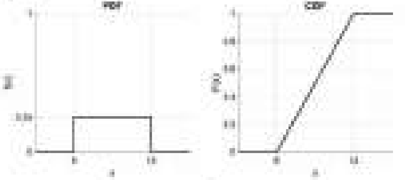
Continuous Uniform Distribution

It is also called the rectangular distribution. It has constant probability on the interval [a, b].

$$P(X=x)=f(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & \text{otherwise}\end{cases}$$

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\le x\le b\\ 1 & x\ge b\end{cases}$$

Example: the continuous random variable X are uniformly distributed on the range [8, 12]. Plot PDF and CDF.



$$E(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$E\left[X^2\right] = \int_a^b x^2 f(x)\,dx = \frac{b^2+ab+a^2}{3}$$

$$Var(X) = E\left[X^2\right] - (E(X))^2 = \frac{(b-a)^2}{12}$$

Example: 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 Bernoulli Discrete Distribution

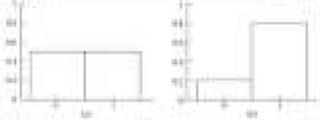
Imagine an experiment that can result in one of two outcomes. One outcome is labeled "success," and the other outcome is labeled "failure." The probability of success is denoted by p. The probability of failure q is therefore $1 - p$. Such a trial is called a Bernoulli trial with success probability p. The simplest Bernoulli trial is the toss of a coin. The two outcomes are heads and tails. If we define heads to be the success outcome then p is the probability that the coin comes up heads. For a fair coin, $p = \frac{1}{2}$.

For any Bernoulli trial, we define a random variable X as follows: If the experiment results in success, then $X = 1$. Otherwise $X = 0$. It follows that X is a discrete random variable, with probability mass function $p(x)$ defined by:

$p(0) = P(X = 0) = 1 - p$

$p(1) = P(X = 1) = p$

$p(x) = 0$ for any value of x other than 0 or 1.



[illegible]

Example

[illegible]

Solution

[illegible]

Example

[illegible]

Solution

[illegible]

Example

[illegible]

Solution

[illegible]

Mean and Variance of a Bernoulli Random Variable

[illegible]

$$\mu_X = p$$

$$\sigma_X^2 = p(1-p)$$

Example 4.x

Refer to Example 4.3. Find μ_X and σ_X^2.

Solution

Here, $X \sim$ Bernoulli(0.1), the success probability p is equal to 0.1. Using Equations (4.3) and (4.4), $\mu_X = 0.1$ and $\sigma_X^2 = 0.1(1 - 0.1) = 0.09$.

The Binomial Discrete Distribution

If a total of n Bernoulli trials are conducted, and

- The trials are independent
- Each trial has the same success probability p
- X is the number of successes in the n trials

then X has the binomial distribution with parameters n and p, denoted $X \sim \text{Bin}(n, p)$.

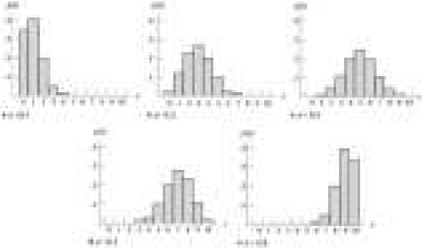
Probability Mass Function of a Binomial Random Variable

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

$$p(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

$$F(x) = \sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}$$

FIGURE 4.x (a) The Bin(10, 0.4) probability histogram. (b) The Bin(20, 0.1) probability histogram.



Example [illegible]

[illegible]

Solution:

[illegible]

$$
\begin{aligned}
P(\ldots) &= [illegible]\\
&= [illegible]\\
&= [illegible]\\
&\quad + [illegible]\\
&= [illegible]\\
&= [illegible]
\end{aligned}
$$

Where Does the Name *Binomial* Come From?

The binomial distribution derives its name from the fact that the $n + 1$ terms in the binomial expansion of $(q + p)^n$ correspond to the various values of $b(x; n, p)$ for $x = 0, 1, 2, \ldots, n$. That is,

$$
\begin{aligned}
(q + p)^n &= \binom{n}{0} q^n + \binom{n}{1} p q^{n-1} + \binom{n}{2} p^2 q^{n-2} + \cdots + \binom{n}{n} p^n \\
&= b(0; n, p) + b(1; n, p) + b(2; n, p) + \cdots + b(n; n, p).
\end{aligned}
$$

Since $p + q = 1$, we see that

$$\sum_{x=0}^{n} b(x; n, p) = 1,$$

a condition that must hold for any probability distribution.

Frequently, we are interested in problems where it is necessary to find $P(X < r)$ or $P(a \le X \le b)$. Binomial sums

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

Example: 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.

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

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

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

Example: 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,

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

Example: 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, $P(X > 3) = 1 - 0.6496 = 0.3504$ ∎

The Mean and Variance of a Binomial Random Variable

If $X \sim \text{Bin}(n, p)$, then the mean and variance of X are given by

$$\mu_X = np \qquad (4.5)$$
$$\sigma_X^2 = np(1-p) \qquad (4.6)$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Review Examples

Method 1

Alternatively

Method 2 (using formula):

Alternatively

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

6.5. Find the probability that in a family of 4 children there will be (a) at least 1 boy, (b) at least 1 boy and at least 1 girl. Assume that the probability of a male birth is 1/2.

(a) $P(1\text{ boy}) = \binom{4}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3 = \frac{1}{4}$ $\qquad P(2\text{ boys}) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{3}{8}$

$P(3\text{ boys}) = \binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1 = \frac{1}{4}$ $\qquad P(4\text{ boys}) = \binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{1}{16}$

Then

$$P(\text{at least 1 boy}) = P(1\text{ boy}) + P(2\text{ boys}) + P(3\text{ boys}) + P(4\text{ boys})$$
$$= \frac{1}{4} + \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = \frac{15}{16}$$

Another method

$$P(\text{at least 1 boy}) = 1 - P(\text{no boy}) = 1 - \left(\frac{1}{2}\right)^4 = 1 - \frac{1}{16} = \frac{15}{16}$$

(b) $P(\text{at least 1 boy and at least 1 girl}) = 1 - P(\text{no boy}) - P(\text{no girl})$
$$= 1 - \frac{1}{16} - \frac{1}{16} = \frac{7}{8}$$

We could also have solved this problem by letting X be a random variable denoting the number of boys in families with 4 children. Then, for example, (a) becomes

$$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4) = \frac{15}{16}$$

6.6. Out of 2000 families with 4 children each, how many would you expect to have (a) at least 1 boy, (b) 2 boys, (c) 1 or 2 girls, (d) no girls?

Referring to Problem 6.5, we see that

(a) Expected number of families with at least 1 boy $= 2000\left(\frac{15}{16}\right) = 1875$

(b) Expected number of families with 2 boys $= 2000 \cdot P(2\text{ boys}) = 2000\left(\frac{3}{8}\right) = 750$

(c) $P(1\text{ or }2\text{ girls}) = P(1\text{ girl}) + P(2\text{ girls})$
$$= P(3\text{ boys}) + P(2\text{ boys}) = \frac{1}{4} + \frac{3}{8} = \frac{5}{8}$$

Expected number of families with 1 or 2 girls $= (2000)\left(\frac{5}{8}\right) = 1250$

(d) Expected number of families with no girls $= (2000)\left(\frac{1}{16}\right) = 125$

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

[illegible] Find the probability of getting a total of 7 at least once in three tosses of a pair of fair dice.

[illegible]

$$[illegible]$$

and

$$[illegible]$$

[illegible]

[illegible]

$$[illegible]$$

The Normal Continuous Distribution

The normal distribution (also called the Gaussian distribution) is by far the most commonly used distribution in statistics. This distribution provides a good model for many, although not all, continuous populations.

The normal distribution is continuous rather than discrete. The mean of a normal random variable may have any value, and the variance may have any positive value. The probability density function of a normal random variable with mean μ and variance σ^2 is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \qquad -\infty < x < \infty$$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right) dt$$

If $X \sim N(\mu, \sigma^2)$, then the mean and variance of X are given by:

$$\mu_X = \mu$$
$$\sigma_X^2 = \sigma^2$$

Figure 4.4 presents a plot of the normal probability density function with mean μ and standard deviation σ. The normal probability density function is sometimes called the normal curve. Note that the normal curve is symmetric around μ, so that μ is the median as well as the mean. It is also the case that for any normal population:

- About 68% of the population is in the interval $\mu \pm \sigma$.
- About 95% of the population is in the interval $\mu \pm 2\sigma$.
- About 99.7% of the population is in the interval $\mu \pm 3\sigma$.

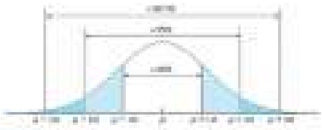
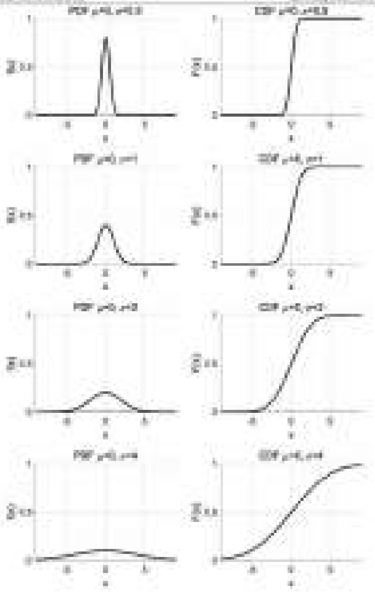


FIGURE 4.4 Probability density function of a normal random variable with mean μ and variance σ^2.



The following are some important properties of the normal curve:

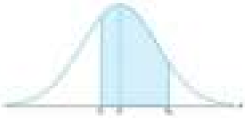
1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
2. The curve is symmetric about a vertical axis through the mean μ.
3. The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < X < \mu + \sigma$ 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.

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,

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

is represented by the area of the shaded region.



The area under the curve between any two ordinates must then also depend on the values μ and σ. This is evident in the following figure, where we have shaded regions corresponding to $P(x_1 < X < x_2)$ for two curves with different means and variances. Obviously, the two shaded regions are different in size; therefore, the probability associated with each distribution will be different for the two given values of X.

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 transform x to the standard score z as:

$$z = \frac{x - \mu}{\sigma}$$

Thus, if x is an observation sampled from a normal population with mean μ and variance σ^2, the standard score equivalent of x is the number z. The number z is an observation sampled from a normal population with mean 0 and standard deviation 1. This normal population is called the standard normal population.

Example [illegible]

Aluminum sheets used to make beverage cans have thicknesses (in thousandths of an inch) that are normally distributed with mean 10 and standard deviation 1.3. A particular sheet is 10.8 thousandths of an inch thick. Find the z-score.

Solution:

The quantity 10.8 is an observation from a normal population with mean $\mu = 10$ and standard deviation $\sigma = 1.3$. Therefore

$$z = \frac{10.8 - 10}{1.3}$$

[illegible]

Example 4.40

Refer to Example 4.39. The thickness of a certain sheet has a z-score of [illegible]. Find the thickness of the sheet in the original units of thousandths of inches.

Solution

We use Equation [illegible], substituting [illegible] for z and solving for x. We obtain

$$[illegible]$$

Solving for x yields x = [illegible]. The sheet is [illegible] thousandths of an inch thick.

Areas under the standard normal curve (mean 0, variance 1) have been extensively tabulated. A typical such table, called a standard normal table, or z table, is given as Table A.2 (in Appendix A). To find areas under a normal curve with a different mean and variance, we convert to standard units and use the z table. Table A.2 provides areas in the left-hand tail of the curve for values of z.

Example 4.41

Find the area under the normal curve to the left of $z = 0.47$.

Solution

From the z table, the area is 0.6808. See Figure 4.5.

FIGURE 4.5 Solution to Example 4.41.

Example 4.42

Find the area under the normal curve to the right of $z = 1.38$.

Solution

From the z table, the area to the left of $z = 1.38$ is 0.9162. Therefore the area to the right is $1 - 0.9162 = 0.0838$. See Figure 4.6.

FIGURE 4.6 Solution to Example 4.42.

Example 4.43

Find the area under the normal curve between $z = 0.71$ and $z = 1.28$.

Solution

From the z table, the area to the left of $z = 1.28$ is 0.8997. The area to the left of $z = 0.71$ is 0.7611. The area between $z = 0.71$ and $z = 1.28$ is therefore $0.8997 - 0.7611 = 0.1386$. See Figure 4.7.

FIGURE 4.7 Solution to Example 4.43.

Example 4.44

What z-score corresponds to the 75th percentile of a normal curve? The 25th percentile? The median?

Solution

To answer this question, we use the z table in reverse. We need to find the z-score for which 75% of the area of the curve is to the left. From the body of the table, the closest area to 75% is 0.7486, corresponding to a z-score of 0.67. Therefore the 75th percentile is approximately 0.67. By the symmetry of the curve, the 25th percentile is $z = -0.67$ (this can also be looked up in the table directly). See Figure 4.8. The median is $z = 0$.

FIGURE 4.8 Solution to Example 4.44.

Example 4.49

Refer to Examples 4.47 and 4.48. Assume that the process has been recalibrated so that the mean diameter is now 2.5 cm. To what value must the standard deviation be lowered so that 95% of the diameters will meet the specification?

Solution

The specification interval is 2.49–2.51 cm. We must find a value for σ so that this interval spans the middle 95% of the population of ball bearing diameters. See Figure 4.15. The z-score that has 2.5% of the area to its left is $z = -1.96$. The z-score that has 2.5% of the area to its right is $z = 1.96$ (this follows from the symmetry of the curve). It follows that the lower specification limit, 2.49, has a z-score of -1.96, while the upper limit of 2.51 has a z-score of 1.96. Either of these facts may be used to find σ. From Equation (4.25),

$$1.96 = \frac{2.51 - 2.50}{\sigma}$$

Solving for σ yields $\sigma = 0.0051$ cm.

FIGURE 4.15 Solution to Example 4.49. If $\sigma = 0.0051$, then approximately 95% of the population will fall between 2.49 and 2.51.

Example: Given a standard normal distribution, find the area under the curve that lies

(a) to the right of $z = 1.84$ and

(b) between $z = -1.97$ and $z = 0.86$.

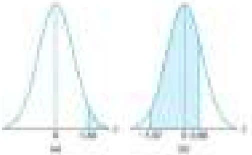


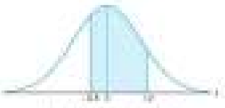
Figure 6.9: Areas for Example 6.2.

Solution: See Figure 6.9 for the specific areas.

(a) The area in Figure 6.9(a) to the right of $z = 1.84$ is equal to 1 minus the area in Table A.3 to the left of $z = 1.84$, namely, $1 - 0.9671 = 0.0329$.

(b) The area in Figure 6.9(b) between $z = -1.97$ and $z = 0.86$ is equal to the area to the left of $z = 0.86$ minus the area to the left of $z = -1.97$. From Table A.3 we find the desired area to be $0.8051 - 0.0244 = 0.7807$.

[illegible]



Therefore,

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

$P(-0.5 < Z < 1.2)$ is shown by the area of the shaded region in Figure 6.12. This area may be found by subtracting the area to the left of the ordinate $z = -0.5$ from the entire area to the left of $z = 1.2$. Using Table A.3, we have

$$\begin{aligned} P(45 < X < 62) &= P(-0.5 < Z < 1.2) = P(Z < 1.2) - P(Z < -0.5) \\ &= 0.8849 - 0.3085 = 0.5764. \end{aligned}$$

Example 6.5: Given that X has a normal distribution with $\mu = 300$ and $\sigma = 50$, find the probability that X assumes a value greater than 362.

Solution: The normal probability distribution with the desired area shaded is shown in Figure 6.13. To find $P(X > 362)$, we need to evaluate the area under the normal curve to the right of $x = 362$. This can be done by transforming $x = 362$ to the corresponding z value, obtaining the area to the left of z from Table A.3, and then subtracting this area from 1. We find that

$$z = \frac{362 - 300}{50} = 1.24.$$

Hence,

$$P(X > 362) = P(Z > 1.24) = 1 - P(Z < 1.24) = 1 - 0.8925 = 0.1075.$$

Figure 6.13: Area for Example 6.5.

According to Chebyshev's theorem on page 137, the probability that a random variable assumes a value within 2 standard deviations of the mean is at least 3/4. If the random variable has a normal distribution, the z values corresponding to $x_1 = \mu - 2\sigma$ and $x_2 = \mu + 2\sigma$ are easily computed to be

$$z_1 = \frac{(\mu - 2\sigma) - \mu}{\sigma} = -2 \text{ and } z_2 = \frac{(\mu + 2\sigma) - \mu}{\sigma} = 2.$$

Hence,

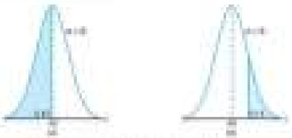
$$\begin{aligned} P(\mu - 2\sigma < X < \mu + 2\sigma) &= P(-2 < Z < 2) = P(Z < 2) - P(Z < -2) \\ &= 0.9772 - 0.0228 = 0.9544, \end{aligned}$$

which is a much stronger statement than that given by Chebyshev's theorem.

Using the Normal Curve in Reverse

[illegible]

[illegible]



[illegible]

[illegible]

6.4 Applications of the Normal Distribution

Some of the many problems for which the normal distribution is applicable are treated in the following examples. The use of the normal curve to approximate binomial probabilities is considered in Section 6.5.

Example 6.7: A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Solution: First construct a diagram such as Figure 6.14, showing the given distribution of battery lives and the desired area. To find $P(X < 2.3)$, we need to evaluate the area under the normal curve to the left of 2.3. This is accomplished by finding the area to the left of the corresponding z value. Hence, we find that

$$z = \frac{2.3 - 3}{0.5} = -1.4,$$

and then, using Table A.3, we have

$$P(X < 2.3) = P(Z < -1.4) = 0.0808.$$

Figure 6.14: Area for Example 6.7.

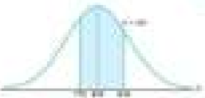


Figure 6.15: Area for Example 6.8.

Example 6.8: An electrical firm manufactures light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a bulb burns between 778 and 834 hours.

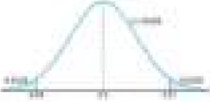
Solution: The distribution of light bulb life is illustrated in Figure 6.15. The z values corresponding to $x_1 = 778$ and $x_2 = 834$ are

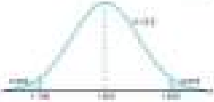
$$z_1 = \frac{778 - 800}{40} = -0.55 \text{ and } z_2 = \frac{834 - 800}{40} = 0.85.$$

Hence,

$$P(778 < X < 834) = P(-0.55 < Z < 0.85) = P(Z < 0.85) - P(Z < -0.55)$$
$$= 0.8023 - 0.2912 = 0.5111.$$

Example 6.9: In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01 cm. The





Example 8.11: A certain machine makes electrical resistors having a mean resistance of 40 ohms and a standard deviation of 2 ohms. Assuming that the resistance follows a normal distribution and can be measured to any degree of accuracy, what percentage of resistors will have a resistance exceeding 43 ohms?

Solution: A percentage is found by multiplying the relative frequency by 100%. Since the relative frequency for an interval is equal to the probability of a value falling in the interval, we must find the area to the right of $x = 43$ in Figure 8.19. This can be done by transforming $x = 43$ to the corresponding z value, obtaining the area to the left of z from Table A.3, and then subtracting this area from 1. We find

$$z = \frac{43 - 40}{2} = 1.5.$$

Therefore,

$$P(X > 43) = P(Z > 1.5) = 1 - P(Z < 1.5) = 1 - 0.9332 = 0.0668.$$

Hence, 6.68% of the resistors will have a resistance exceeding 43 ohms. ∎

Figure 8.19: Area for Example 8.11.

Figure 8.20: Area for Example 8.12.

Example 8.12: Find the percentage of resistances exceeding 43 ohms for Example 8.11 if resistance is measured to the nearest ohm.

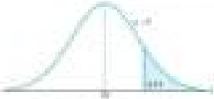
Solution: This problem differs from that in Example 8.11 in that we now assign a measurement of 43 ohms to all resistors whose resistances are greater than 42.5 and less than 43.5. We are actually approximating a discrete distribution by means of a continuous normal distribution. The required area is the region shaded to the right of 43.5 in Figure 8.20. We now find that

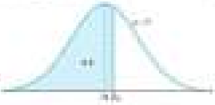
$$z = \frac{43.5 - 40}{2} = 1.75.$$

Hence,

$$P(X > 43.5) = P(Z > 1.75) = 1 - P(Z < 1.75) = 1 - 0.9599 = 0.0401.$$

Therefore, 4.01% of the resistances exceed 43 ohms when measured to the nearest ohm. The difference $6.68\% - 4.01\% = 2.67\%$ between this answer and that of Example 8.11 represents all those resistance values greater than 43 and less than 43.5 that are now being recorded as 43 ohms. ∎





[illegible]

(c) [illegible]

(d) [illegible]

(e) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

[illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

(d) [illegible]

[illegible]

(c) [illegible]

(d) [illegible]

[illegible]

(a) [illegible]

(b) [illegible]

(c) [illegible]

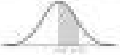
[illegible]

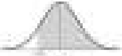
(a) [illegible]

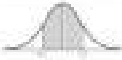
(b) [illegible]

[illegible]

Review Examples







Data Sampling

Populations and Samples

A **population** consists of the totality of the observations with which we are concerned. In other words, it is all the possible outcomes from an experiment or a system. Populations can be finite or infinite.

In the field of **statistical inference**, statisticians are interested in arriving at conclusions concerning a population when it is impossible or impractical to observe the entire set of observations that make up the population. For example, in attempting to determine the average length of life of a certain brand of light bulb, it would be impossible to test all such bulbs if we are to have any left to sell. Therefore, we must depend on a subset of observations from the population to help us make inferences concerning that same population. This brings us to consider the notion of sampling.

A **sample** is a subset of a population. We must obtain samples that are representative of the population if our inferences from the sample to the population are to be valid. Hence we are tempted to choose a sample by selecting the most convenient members of the population. Such a procedure may lead to erroneous inferences concerning the population. Any sampling procedure that produces inferences that consistently overestimate or consistently underestimate some characteristic of the population is said to be **biased**. To eliminate any possibility of bias in the sampling procedure, it is desirable to choose a random sample in the sense that the observations are made independently and at random.

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

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

Any function of the random variables constituting a random sample is called a **statistic**.

The Sample Mean, Median, and Mode

(a) Sample mean:

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

Note that the statistic $\bar{X}$ assumes the value $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ 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.

The Sample Variance and Standard Deviation

Sample variance:

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

If S^2 is the variance of a random sample of size n, we may write

$$S^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right].$$

The sample standard deviation is $S = \sqrt{S^2}$.

Example: 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,

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

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

Example: Find the variance of the data 3, 4, 5, 6, 6, and 7, representing the number of trout caught by a random sample of 6 fishermen on June 19, 1996, at Lake Muskoka.

Solution: We find that $\sum_{i=1}^{6} x_i^2 = 171$, $\sum_{i=1}^{6} x_i = 31$, and $n = 6$. Hence,

$$s^2 = \frac{1}{(6)(5)}[(6)(171) - (31)^2] = \frac{13}{6}.$$

Thus, the sample standard deviation $s = \sqrt{13/6} = 1.47$ and the sample range is $7 - 3 = 4$.

Exercises #12

[illegible] [illegible]
(a) the mean;
(b) the median;
(c) the mode.

[illegible] [illegible]
(a) the mean;
(b) the median.

[illegible] [illegible]
(a) the mean;
(b) the median;
(c) the mode.

[illegible] [illegible]
(a) the mean;
(b) the median.

[illegible] [illegible]

Laundry Detergent	Phosphates per Load (grams)
[illegible]	[illegible]

For the given phosphate data, find
(a) the mean;
(b) the median;
(c) the mode.

[illegible] [illegible]
(a) the range;
(b) the standard deviation.

[illegible] [illegible]
(a) the range;
(b) the variance [illegible]

[illegible] [illegible]

[illegible] [illegible]
(a) the mean;
(b) the variance.

[illegible] [illegible]

Calculate the standard deviation.

[illegible] [illegible]

[illegible] [illegible]

Find
(a) the mean [illegible];
(b) the median [illegible].

The Central Limit Theorem

The first important sampling distribution to be considered is that of the mean $\bar{X}$. Suppose that a random sample of n observations is taken from a normal population with mean μ and variance σ^2. Each observation X_i, $i = 1, 2, \ldots, n$, of the random sample will then have the same normal distribution as the population being sampled. Hence, by the reproductive property of the normal distribution established in Theorem 7.11, we conclude that

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

has a normal distribution with mean

$$\mu_{\bar{X}} = \frac{1}{n}(\underbrace{\mu + \mu + \cdots + \mu}_{n \text{ terms}}) = \mu \text{ and variance } \sigma^2_{\bar{X}} = \frac{1}{n^2}(\underbrace{\sigma^2 + \sigma^2 + \cdots + \sigma^2}_{n \text{ terms}}) = \frac{\sigma^2}{n}.$$

If we are sampling from a population with unknown distribution, either finite or infinite, the sampling distribution of $\bar{X}$ will still be approximately normal with mean μ and variance σ^2/n, provided that the sample size is large. This amazing result is an immediate consequence of the following theorem, called 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 $n(z; 0, 1)$.

The normal approximation for $\bar{X}$ will generally be good if $n \geq 30$, provided the population distribution is not terribly skewed. If $n < 30$, the approximation is good only if the population is not too different from a normal distribution and, as stated above, if the population is known to be normal, the sampling distribution of $\bar{X}$ will follow a normal distribution exactly, no matter how small the size of the samples.

The sample size $n = 30$ is a guideline to use for the Central Limit Theorem. However, as the statement of the theorem implies, the presumption of normality on the distribution of $\bar{X}$ becomes more accurate as n grows larger. In fact, Figure 8.1 illustrates how the theorem works. It shows how the distribution of $\bar{X}$ becomes closer to normal as n grows larger, beginning with the clearly nonsymmetric distribution of an individual observation ($n = 1$). It also illustrates that the mean of $\bar{X}$ remains μ for any sample size and the variance of $\bar{X}$ gets smaller as n increases.

Figure 8.1: Illustration of the Central Limit Theorem (distribution of $\bar{X}$ for $n = 1$, moderate n, and large n).

Example: 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

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

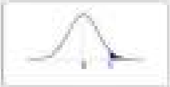
APPENDIX A

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

APPENDIX B

Areas under the Standard Normal Curve from z to ∞

$P(Z > z)$



APPENDIX C

Areas under the Standard Normal Curve from 0 to z

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]