Chapter 5
Some Discrete Probability Distributions

1. Introduction

Often, the observations generated by different statistical experiment have the same general type of behavior.

1. Introduction

Often, the observations generated by different statistical experiment have the same general type of behavior.

Consequently, discrete random variables associated with these experiments can be described by essentially the same probability distribution.

1. Introduction

Often, the observations generated by different statistical experiment have the same general type of behavior.

Consequently, discrete random variables associated with these experiments can be described by essentially the same probability distribution.

In fact, one needs only a handful of important probability distributions to describe many of the discrete random variables encountered in practice.

2. Bernoulli Random Variable

Example When a student calls a university help desk for technical support, he/she will either immediately be able to speak to someone (S, for success) or will be placed on hold (F, for failure). With $\mathcal{S} = \{S, F\}$, define an r.v. X by

$$X(S) = 1, X(F) = 0$$

Bernoulli Random Variable

Any random variable whose only possible values are 0 and 1 is called a **Bernoulli random variable**.

Bernoulli Random Variable

Example Consider whether the next person buying a computer at a certain electronic store buys a laptop or a desktop model. Let

$$X = \begin{cases} 1, & \text{if the customer purchases a desktop computer} \\ 0, & \text{if the customer purchases a laptop computer.} \end{cases}$$

Bernoulli Random Variable

Example Consider whether the next person buying a computer at a certain electronic store buys a laptop or a desktop model. Let

$$X = \begin{cases} 1, & \text{if the customer purchases a desktop computer} \\ 0, & \text{if the customer purchases a laptop computer.} \end{cases}$$

If 20% of all purchasers during that week select a desktop, the pmf for X is

$$p(x) = \begin{cases} 0.8, & \text{if } x = 0 \\ 0.2, & \text{if } x = 1. \end{cases}$$

Bernoulli Random Variable

the general case

The pmf of any Bernoulli r.v. can be expressed in the form $p(1) = \alpha$ and $p(0) = 1 - \alpha$, where $0 < \alpha < 1$.

Bernoulli Random Variable

the general case

The pmf of any Bernoulli r.v. can be expressed in the form $p(1) = \alpha$ and $p(0) = 1 - \alpha$, where $0 < \alpha < 1$.

We call α the **parameter** of the distribution.

3. Binomial and Multinomial Distributions

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

3. Binomial and Multinomial Distributions

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

Example

The testing of items as they come off an assembly line, where each test or trial may indicate a defective or a nondefective item.

3. Binomial and Multinomial Distributions

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

Example

The testing of items as they come off an assembly line, where each test or trial may indicate a defective or a nondefective item. We may choose to define either outcomes as a **success**.

3. Binomial and Multinomial Distributions

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

Example

The testing of items as they come off an assembly line, where each test or trial may indicate a defective or a nondefective item. We may choose to define either outcomes as a **success**.

The process is referred to as a **Bernoulli process**.

3. Binomial and Multinomial Distributions

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

Example

The testing of items as they come off an assembly line, where each test or trial may indicate a defective or a nondefective item. We may choose to define either outcomes as a **success**.

The process is referred to as a **Bernoulli process**. Each trial is called a **Bernoulli trial**.

The Bernoulli Process

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

The Bernoulli Process

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

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

The Bernoulli Process

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

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

X: **the number of successes in the process**

Binomial Distribution

Example

Three items are selected at random from a manufacturing process, inspected, and classified defective or nondefective.

Binomial Distribution

Example

Three items are selected at random from a manufacturing process, inspected, and classified defective or nondefective.

X: the number of defective

Binomial Distribution

Example

Three items are selected at random from a manufacturing process, inspected, and classified defective or nondefective.

X: the number of defective

What is distribution of X?

Binomial Distribution

Example

Three items are selected at random from a manufacturing process, inspected, and classified defective or nondefective.

X: the number of defective

What is distribution of X? (the possible values and corresponding probability)

Binomial Distribution

X: the number of defective

$$S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$$

$$0, \quad 1, \quad 1, \quad 1, \quad 2, \quad 2, \quad 2, \quad 3$$

Binomial Distribution

The items are selected independently from a process that we shall assume produces 25% defectives.

Binomial Distribution

The items are selected independently from a process that we shall assume produces 25% defectives.

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

Binomial Distribution

The items are selected independently from a process that we shall assume produces 25% defectives.

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

Similarly calculation yield the probabilities for the other possible outcomes.

$$S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$$

$$0, \quad 1, \quad 1, \quad 1, \quad 2, \quad 2, \quad 2, \quad 3$$

Binomial Distribution

The probability distribution of X is

0	1	2	3
$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

Binomial Distribution

In general, the number X of successes in n Bernoulli trials is called a **binomial random variable.**

Binomial Distribution

In general, the number X of successes in n Bernoulli trials is called a **binomial random variable.** The probability distribution of this discrete random variable is called the **binomial distribution**,

Binomial Distribution

In general, the number X of successes in n Bernoulli trials is called a **binomial random variable.** The probability distribution of this discrete random variable is called the **binomial distribution**,

The pmf of a binomial r.v. X depends on the two parameters n and p, we denote the **pmf** by $b(x; n, p)$,

Binomial Distribution

In general, the number X of successes in n Bernoulli trials is called a **binomial random variable.** The probability distribution of this discrete random variable is called the **binomial distribution**,

The pmf of a binomial r.v. X depends on the two parameters n and p, we denote the **pmf** by $b(x; n, p)$,n is the number of trials, p is the probability of a success on a given trial.

Binomial Distribution

Binomial Distribution

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

$$b(x; n, p) = C_n^x p^x q^{n-x}, \qquad x = 0, 1, 2, \cdots, n.$$

Binomial Distribution

Binomial Distribution

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

$$b(x; n, p) = C_n^x p^x q^{n-x}, \qquad x = 0, 1, 2, \cdots, n.$$

Why?

Binomial Distribution

Previous example:

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

$$b(x; 3, \frac{1}{4}) = C_3^x (\frac{1}{4})^x (\frac{3}{4})^{3-x}, \quad x = 0, 1, 2, 3.$$

Binomial Distribution

Example 5.4 The probability that a certain kind of component will survive a given shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

Binomial Distribution

Example 5.4 The probability that a certain kind of component will survive a given shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

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

$$b(2; 4, \frac{3}{4}) = C_4^2(\frac{3}{4})^2(\frac{1}{4})^2$$

Binomial Distribution

Where does the name Binomial come from?

Binomial Distribution

Where does the name Binomial come from?

The binomial expansion of $(p+q)^n$ is

Binomial Distribution

Where does the name Binomial come from?

The binomial expansion of $(p+q)^n$ is

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

Binomial Distribution

Where does the name Binomial come from?

The binomial expansion of $(p+q)^n$ is

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

Since $p+q=1$ we see that

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

Binomial Distribution

Where does the name Binomial come from?

The binomial expansion of $(p+q)^n$ is

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

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.

Binomial Distribution

For $X \sim Bin(n, p)$ the cdf will be denoted by

$$B(x; n, p) = P(X \leq x) = \sum_{y=0}^{x} b(y; n, p), \quad x = 0, 1, \ldots, n$$

Appendix Table A.1 tabulates the cdf $B(x; n, p)$ for $n = 5, 10, 15, 20, 25$ in combination with selected values of p.

Binomial Distribution

For $X \sim Bin(n,p)$ the cdf will be denoted by

$$B(x;n,p) = P(X \leq x) = \sum_{y=0}^{x} b(y;n,p), \quad x = 0,1,\ldots,n$$

Appendix Table A.1 tabulates the cdf $B(x;n,p)$ for $n = 5, 10, 15, 20, 25$ in combination with selected values of p.

Binomial Distribution

Example 5.5 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,

(c) exactly 5 survive?

Binomial Distribution

Theorem 5.2 The mean and variance of the binomial distribution $b(x; n, p)$ are

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

Binomial Distribution

Theorem 5.2 The mean and variance of the binomial distribution $b(x; n, p)$ are

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

Proof

Let the outcome on the jth trial be represented by a **Bernoulli random variable** I_j,

Binomial Distribution

Theorem 5.2 The mean and variance of the binomial distribution $b(x; n, p)$ are

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

Proof

Let the outcome on the jth trial be represented by a **Bernoulli random variable** I_j, which assumes the value 0(failure) and 1(success) with probabilities q and p, respectively.

Binomial Distribution

Theorem 5.2 The mean and variance of the binomial distribution $b(x; n, p)$ are

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

Proof

Let the outcome on the jth trial be represented by a **Bernoulli random variable** I_j, which assumes the value 0(failure) and 1(success) with probabilities q and p, respectively.

Therefore, in a binomial experiment the number of successes $X = I_1 + I_2 + \cdots + I_n$.

Binomial Distribution

The mean of any I_j is $E(I_j) = p$. Therefore, the mean of the binomial distribution is

$$E(X) = E(I_1) + E(I_2) + \cdots + E(I_n) = np.$$

Binomial Distribution

The mean of any I_j is $E(I_j) = p$. Therefore, the mean of the binomial distribution is

$$E(X) = E(I_1) + E(I_2) + \cdots + E(I_n) = np.$$

The variance of any I_j is $\sigma^2_{I_j} = pq$.

Binomial Distribution

The mean of any I_j is $E(I_j) = p$. Therefore, the mean of the binomial distribution is

$$E(X) = E(I_1) + E(I_2) + \cdots + E(I_n) = np.$$

The variance of any I_j is $\sigma^2_{I_j} = pq$. The variance of the binomial distribution is

$$\sigma^2_X = \sigma^2_{I_1} + \sigma^2_{I_2} + \cdots + \sigma^2_{I_n} = npq.$$

Binomial Distribution

Example 5.7 Find the mean and variance of the binomial random variable of Example 5.5, 15 people are known to have contracted this disease, the probability that a patient recovers from a rare blood disease is 0.4.

Binomial Distribution

Example 5.7 Find the mean and variance of the binomial random variable of Example 5.5, 15 people are known to have contracted this disease, the probability that a patient recovers from a rare blood disease is 0.4.

Solution

Since Example 5.5 was a binomial experiment with $n = 15$ and $p = 0.4$, by Theorem 5.2, we have

$$\mu = 15 \times 0.4 \quad \text{and} \quad \sigma^2 = 15 \times 0.4 \times 0.6 = 3.6$$

Binomial Distribution

Example 5.8 It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight on this problem, it is determined that some tests should be made. It is too expensive to test all of the many wells in the area so 10 were randomly selected for testing.

Binomial Distribution

Example 5.8 It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight on this problem, it is determined that some tests should be made. It is too expensive to test all of the many wells in the area so 10 were randomly selected for testing.

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

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

Binomial Distribution

Example 5.9 Consider the situation of Example 5.8. The '30% are impure' is merely a conjecture put forth by the area water board. Suppose 10 wells are randomly selected and 6 are found to contain the impurity. What does this imply about the conjecture? Use a probability statement.

Binomial Distribution

Example 5.9 Consider the situation of Example 5.8. The '30% are impure' is merely a conjecture put forth by the area water board. Suppose 10 wells are randomly selected and 6 are found to contain the impurity. What does this imply about the conjecture? Use a probability statement.

Solution We must first ask 'if the conjecture is correct, is it likely that we could have found 6 impure wells'?

Binomial Distribution

Example 5.9 Consider the situation of Example 5.8. The '30% are impure' is merely a conjecture put forth by the area water board. Suppose 10 wells are randomly selected and 6 are found to contain the impurity. What does this imply about the conjecture? Use a probability statement.

Solution We must first ask 'if the conjecture is correct, is it likely that we could have found 6 impure wells'?

$$P(X = 6) = 0.0367$$

Binomial Distribution

Example 5.9 Consider the situation of Example 5.8. The '30% are impure' is merely a conjecture put forth by the area water board. Suppose 10 wells are randomly selected and 6 are found to contain the impurity. What does this imply about the conjecture? Use a probability statement.

Solution We must first ask 'if the conjecture is correct, is it likely that we could have found 6 impure wells'?

$$P(X = 6) = 0.0367$$

As a result it is very unlikely that 6 wells would be found impure if only 30% of all are impure. The impurity problem is much more severe.

Multinomial Experiments

In many applications, there are more than two possible outcomes. For example, the color of guinea pigs produced as offspring may be red, black, or white.

Multinomial Experiments

In many applications, there are more than two possible outcomes. For example, the color of guinea pigs produced as offspring may be red, black, or white.

The binomial experiment becomes a **multinomial experiment** if we let each trial have more than 2 possible outcomes.

Multinomial Distribution

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

$$p(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

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

Multinomial Distribution

Example 5.10 For a certain airport containing three runways it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet
Runway 1: $p_1 = 2/9$,
Runway 2: $p_2 = 1/6$,
Runway 3: $p_3 = 11/18$.

Multinomial Distribution

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

Runway 1: 2 airplanes (2/9)

Runway 2: 1 airplane (1/6)

Runway 3: 3 airplanes (11/18)

Multinomial Distribution

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

Runway 1: 2 airplanes (2/9)

Runway 2: 1 airplane (1/6)

Runway 3: 3 airplanes (11/18)

Solution

$$p(2,1,3;\frac{2}{9},\frac{1}{6},\frac{11}{18},6) = \frac{6!}{2!1!3!}\cdot\frac{2^2}{9^2}\cdot\frac{1}{6}\cdot\frac{11^3}{18^3} = 0.1127$$

4. Poisson Distribution and Poisson Process

X: the number of outcomes occurring during **a given time interval** or **in a specified region**

4. Poisson Distribution and Poisson Process

X: the number of outcomes occurring during **a given time interval** or **in a specified region**

The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

4. Poisson Distribution and Poisson Process

X: the number of outcomes occurring during **a given time interval** or **in a specified region**

The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

Examples

the number of telephone calls per hour received by an office,

4. Poisson Distribution and Poisson Process

X: the number of outcomes occurring during **a given time interval** or **in a specified region**

The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

Examples

the number of telephone calls per hour received by an office,
the number of days school is closed due to snow during the winter,

4. Poisson Distribution and Poisson Process

X: the number of outcomes occurring during **a given time interval** or **in a specified region**

The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

Examples

the number of telephone calls per hour received by an office,
the number of days school is closed due to snow during the winter,
the number of postponed games due to rain during a baseball season.

Poisson experiments

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Poisson experiments

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Examples

the number of field mice per acre,

Poisson experiments

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Examples

the number of field mice per acre,

the number of bacteria in a given culture,

Poisson experiments

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Examples

the number of field mice per acre,

the number of bacteria in a given culture,

the number of typing errors per page.

Poisson experiments

The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Examples

the number of field mice per acre,

the number of bacteria in a given culture,

the number of typing errors per page.

The number X is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**.

Poisson Distribution

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

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

where μ is the parameter $(\mu > 0)$, and $e = 2.71828\ldots$.

Poisson Distribution

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

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

where μ is the parameter $(\mu > 0)$, and $e = 2.71828\ldots$.

Table A.2 contains Poisson probability sum

Poisson Distribution

Theorem 5.5 The mean and variance of the Poisson distribution $p(x; \mu)$ both have the value μ.

Poisson Distribution

Theorem 5.5 The mean and variance of the Poisson distribution $p(x;\mu)$ both have the value μ.

Proof......

Poisson Distribution

Example 5.19 During a laboratory experiment the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Poisson Distribution

Example 5.19 During a laboratory experiment the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution

$$p(6;4) = \frac{e^{-4}4^6}{6!} = 0.1042.$$

Poisson Distribution

Example 5.20 Ten is the average number of oil tankers arriving each day at a certain port city. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

Poisson Distribution

Example 5.20 Ten is the average number of oil tankers arriving each day at a certain port city. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

Solution

Let X be the number of tankers arriving each day.

Poisson Distribution

Example 5.20 Ten is the average number of oil tankers arriving each day at a certain port city. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

Solution

Let X be the number of tankers arriving each day.Then,

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

Poisson Distribution

The Poisson Distribution As a Limiting Form of the Binomial

Theorem 5.6 Let X be a binomial random variable with probability distribution $b(x; n, p)$ When $n \to \infty$, $p \to 0$, and $\mu = np$ remains constant,

$$b(x; n, p) \to p(x; \mu).$$

Poisson Distribution

The Poisson Distribution As a Limiting Form of the Binomial

Theorem 5.6 Let X be a binomial random variable with probability distribution $b(x; n, p)$ When $n \to \infty$, $p \to 0$, and $\mu = np$ remains constant,

$$b(x; n, p) \to p(x; \mu).$$

In any binomial experiment in which n is large and p is small, $b(x; n, p) \approx p(x; \mu)$, where $\mu = np$.

Poisson Distribution

The Poisson Distribution As a Limiting Form of the Binomial

Theorem 5.6 Let X be a binomial random variable with probability distribution $b(x; n, p)$ When $n \to \infty$, $p \to 0$, and $\mu = np$ remains constant,

$$b(x; n, p) \to p(x; \mu).$$

In any binomial experiment in which n is large and p is small, $b(x; n, p) \approx p(x; \mu)$, where $\mu = np$.($n > 50$ and $np < 5$)

Poisson Distribution

Example 5.21 In a certain industrial facility accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

(a) What is the probability that in any given period of 400 days there will be an accident on one day?

(b) What is the probability that there are at most three days with an accident?