Chapter 5

Some Discrete Probability Distributions

Prof. N. G. J. Dias

Senior Professor and Head

Department of Computer Systems Engineering

Faculty of Computing and Technology

University of Kelaniya

5.1 Introduction

- Often, the observations generated by different statistical experiments have the same general type of behavior.
- Consequently, discrete random variables associated with these experiments can be described by essentially the same probability distribution and therefore can be represented by a single formula.
- In fact, one needs only a handful of important probability distributions to describe many of the discrete random variables encountered in practice.
- In this chapter we will summarize following discrete probability distributions that you have studied in the first year:
    - Discrete uniform distribution
    - Bernoulli distribution
    - Binomial distribution
    - Poisson distribution

- Hypergeometric distribution
- Negative binomial distribution
- Geometric distribution

- We also compute mean, variance and MGF for each of these distributions with various examples.

5.2 Discrete Uniform Distribution

5.2.1 Defining Discrete Uniform Distribution

- Definition 1

  Let X be a discrete random variable with k possible values $1,2,\cdots,k$.

  X has a *discrete uniform probability distribution* if all of these values have the same probability, i.e., if the probability mass function of X has the form

$$f(x) = P(X = x) = \begin{cases} \frac{1}{k}, & \text{for } x = 1,2,\cdots,k \\ 0, & \text{otherwise.} \end{cases}$$

- A random variable having a probability mass function as above is called a *discrete uniform random variable*.
- A simple example of the discrete uniform distribution is the distribution of the score of a fair die, with $k = 6$.

5.2.2 Mean, Variance and MGF of a Discrete Uniform Distribution

- Mean $= \mu = E(X) = \sum_{x=1}^{k} xf(x) = \sum_{x=1}^{k} x \cdot \frac{1}{k} = \frac{1+2+\cdots+k}{k} = \frac{\frac{1}{2}k(k+1)}{k} = \frac{k+1}{2}$.
- Variance σ^2:

$$E(X^2) = \sum_{x=1}^{k} x^2 f(x) = \sum_{x=1}^{k} x^2 \cdot \frac{1}{k} = \frac{1^2+2^2+\cdots+k^2}{k} = \frac{\frac{1}{6}k(k+1)(2k+1)}{k}$$
$$= \frac{1}{6}(k+1)(2k+1).$$

Therefore, $\sigma^2 = Var(X) = E(X^2) - [E(X)]^2$

$$= \frac{1}{6}(k+1)(2k+1) - \left(\frac{k+1}{2}\right)^2$$
$$= \frac{k^2-1}{12}.$$

- MGF of X:

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{k} e^{tx} f(x) = \sum_{x=1}^{k} e^{tx} \frac{1}{k} = \frac{1}{k} \sum_{x=1}^{k} (e^{t})^{x} = \frac{e^{t}(1-e^{tk})}{k(1-e^{t})}.$$

5.3 Bernoulli Distribution

5.3.1 Bernoulli Process

- An experiment often consists of repeated trials, each with two possible outcomes that may be labeled *success* or *failure*.
- Such a process is referred to as a *Bernoulli process* and each trial is called a *Bernoulli trial*.

  In other words, a Bernoulli trial is an experiment with only two possible outcomes (*success* or *failure*)

  For example, the experiment of tossing a coin repeatedly.
- A Bernoulli process must possess the following properties:
  1. The experiment consists of repeated trials.
  2. Each trial results in an outcome that may be classified as a success or a failure.

3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

- Let p be the probability that an event will happen in any single Bernoulli trial (called the *probability of success*).

  Then $q = 1 - p$ is the probability that the event will fail to happen in any single trial (called the *probability of failure*).

5.3.2 Defining Bernoulli Distribution

- The *Bernoulli distribution* is the distribution of the outcome of a *single* Bernoulli trial.
- The two outcomes of the experiment are denoted by 1 and 0, and they represent the success and failure respectively.

- Definition 2 - Bernoulli Distribution

  The distribution of a random variable X with the following probability mass function:

$$f(x) = P(X = x) = \begin{cases} p^x(1-p)^{1-x}, & \text{for } x = 0,1 \\ 0, & \text{otherwise.'} \end{cases}$$

  is called a *Bernoulli Distribution*.

- Note that, $P(X = 1) = p$ and $P(X = 0) = 1 - p$, and no other values are possible.

- A Bernoulli distribution of a random variable X with parameter p is often written as:

$$X \sim Bernoulli(p).$$

5.3.3 Mean, Variance and MGF of a Bernoulli Distribution

- Let $X \sim Bernoulli(p)$.

  Then $\mu = E(X) = \sum_{x=0}^{1} xf(x) = 0 \cdot p^0(1-p)^{1-0} + 1 \cdot p^1(1-p)^{1-1} = p$, and

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{1} x^2 f(x) \\
&= \sum_{x=0}^{1} x^2 \cdot p^x(1-p)^{1-x} \\
&= 0^2 \cdot p^0(1-p)^{1-0} + 1^1 \cdot p^1(1-p)^{1-1} \\
&= p.
\end{aligned}
$$

  Thus, $\sigma^2 = Var(X) = E(X^2) - [E(X)]^2 = p - p^2$.

$$
\begin{aligned}
M_X(t) = E(e^{tX}) = \sum_{x=0}^{1} e^{tx} f(x) &= \sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x} \\
&= \sum_{x=0}^{1} (pe^t)^x (1-p)^{1-x} \\
&= (pe^t)^0(1-p)^{1-0} + (pe^t)^1(1-p)^{1-1} \\
&= (1-p) + pe^t.
\end{aligned}
$$

5.4 Binomial Distribution

5.4.1 Defining Binomial Distribution

- The number X of successes in n Bernoulli trials is called a *binomial random variable*.
- The probability distribution of this discrete random variable X is called the *binomial distribution*, and its probabilities depend on the number of trials n and the probability of a success p on a given trial.

- Definition 3 - Binomial Distribution

  Let a Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$.

  The probability distribution of the binomial random variable X, the number of successes x in n independent trials, is given by the probability mass function

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

- Any random variable X whose probability mass function is given by the above equation is said to be a *binomial random variable* with parameters n and p.
- A Binomial distribution of a random variable X with parameters n and p is often written as:

$$X \sim Binomial(n, p).$$

- Note that in a binomial experiment, if we have x successes in n trials (each with probability p), then there are $n - x$ failures (each with probability $q = 1 - p$).

Example 1

Consider the set of Bernoulli trials where three items are selected at random from a manufacturing process, inspected, and classified as defective (D) or nondefective (N).

A defective item is designated a success.

The number of successes is a random variable X assuming integral values from 0 through 3.

The eight possible outcomes and the corresponding values of X are

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

Since the 3 items are selected independently and we assume that the process produces 25% defectives, we have $n = 3$ and $p = \frac{1}{4}$.

Thus with $n = 3$ and $p = \frac{1}{4}$, the probability distribution of X, the number of defectives, may be written as

$$f(x) = P(X = x) = \binom{3}{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{3-x}, \text{ for } x = 0,1,2,3,$$

rather than in the tabular form.

Example 2

The probability that a certain kind of component will survive a shock test is $\frac{3}{4}$. Find the probability that exactly 2 of the next 4 components tested survive.

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

$$f(x) = P(X = x) = \binom{4}{2}\left(\frac{3}{4}\right)^{2}\left(\frac{1}{4}\right)^{4-2} = \left(\frac{4!}{2!\cdot 2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

- Note: The binomial distribution derives its name from the fact that the $n+1$ terms in the binomial expansion of $(p+q)^n$ correspond to the various values of probabilities $P(X=x)$ for $x=0,1,2,\cdots,n$.

  That is,

$$\begin{aligned}(p+q)^n &= \binom{n}{0}p^0q^n + \binom{n}{1}p^1q^{n-1} + \binom{n}{2}p^2q^{n-2} + \cdots + \binom{n}{n}p^nq^0 \\ &= P(X=0) + P(X=1) + P(X=2) + \cdots + P(X=n) \\ &= \sum_{x=0}^{n} P(X=x)\end{aligned}$$

  Since $p+q=1$, we see that $\sum_{x=0}^{n} P(X=x)=1$, a condition that must hold for any probability distribution.

Example 3

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. We have $p = 0.4$ and $n = 15$.

(a) $P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} P(X = x) = 1 - 0.9662 = 0.0338.$

(b) $P(3 \leq X \leq 8) = \sum_{x=3}^{8} P(X = x) = \sum_{x=0}^{8} P(X = x) - \sum_{x=0}^{2} P(X = x)$

$= 0.9050 - 0.0271 = 0.8779.$

(c) $P(X = 5) = 0.1859.$

Example 4

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 binomial distribution with $n = 20$ and $p = 0.03$.

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

(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 with $\mathrm{n} = 10$ and $p = 0.4562$.

Therefore,

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

5.4.2 Computing Probabilities of a Binomial Distribution

- Suppose that the binomial random variable X has a distribution with parameters n and p. The key to computing its probability mass function

  $$f(x) = \mathrm{P}(\mathrm{X} \leq x) = \sum_{i=0}^{x} \binom{n}{x} p^i (1-p)^{n-i}, x = 0,1,2,\cdots,n$$

  is to utilize the following relationship between $P(X = i + 1)$ and $P(X = i)$:

  $$P(X = i + 1) = \left(\frac{p}{1-p}\right)\left(\frac{n-i}{i+1}\right) P(X = i).$$

  The proof of this equation is left as an exercise.

- For example, let X be a binomial random variable with parameters $n = 6, p = 0.4$.

  Then, starting with $P(X = 0) = (0.6)^6 = 0.467$ and recursively employing the above equation, we obtain

  $$P(X = 1) = \left(\frac{0.4}{1-0.4}\right)\left(\frac{6-0}{0+1}\right) P(X = 0) = \left(\frac{4}{6}\right) \cdot \left(\frac{6}{1}\right) \cdot (0.467) = 0.1866,$$

$$P(X=2)=\left(\frac{0.4}{1-0.4}\right)\left(\frac{6-1}{1+1}\right)P(X=1)=\left(\frac{4}{6}\right)\cdot\left(\frac{5}{2}\right)\cdot(0.1866)=0.3110,$$

$$P(X=3)=\left(\frac{0.4}{1-0.4}\right)\left(\frac{6-2}{2+1}\right)P(X=2)=\left(\frac{4}{6}\right)\cdot\left(\frac{4}{3}\right)\cdot(0.3110)=0.2765,$$

$$P(X=4)=\left(\frac{0.4}{1-0.4}\right)\left(\frac{6-3}{3+1}\right)P(X=3)=\left(\frac{4}{6}\right)\cdot\left(\frac{3}{4}\right)\cdot(0.2765)=0.1382,$$

$$P(X=5)=\left(\frac{0.4}{1-0.4}\right)\left(\frac{6-4}{4+1}\right)P(X=4)=\left(\frac{4}{6}\right)\cdot\left(\frac{2}{5}\right)\cdot(0.1382)=0.0369, \text{ and}$$

$$P(X=6)=\left(\frac{0.4}{1-0.4}\right)\left(\frac{6-5}{5+1}\right)P(X=5)=\left(\frac{4}{6}\right)\cdot\left(\frac{1}{6}\right)\cdot(0.0369)=0.0041.$$

Exercises

1. The probability that a randomly chosen member of a reading group is left-handed is 0.15. A random sample of 20 members of the group is taken.

    (a) Suggest a suitable model for the random variable X, the number of members in the sample who are left-handed. Justify your choice.

    (b) Use your model to calculate the probability that

        (i) exactly 7 of the members in the sample are left-handed

        (ii) fewer than two of the members in the sample are left-handed.

2. It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee that at most 1 of the 10 disks is defective. What proportion of packages is returned? If someone buys three packages, what is the probability that exactly one of them will be returned? (R142)

5.4.3 Mean, Variance and MGF of a Binomial Distribution

- The MGF of a binomial distribution can be easily found as follows.

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx} f(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{1-x}$$

$$= \sum_{x=0}^{1} \binom{n}{x} (pe^t)^x (1-p)^{1-x}$$

$$= [(1-p) + pe^t]^n,$$

for all real values of t.

- The mean μ and the variance σ^2 of X may be computed from $M_X(t)$.

Since $M'_X(t) = n[(1-p) + pe^t]^{n-1}(pe^t)$, and

$$M''_X(t) = n[(1-p) + pe^t]^{n-1}(pe^t)$$

$$+n(n-1)[(1-p) + pe^t]^{n-2}(pe^t)^2,$$

it follows that, $\mu = M'_X(0) = np$ and

$$\sigma^2 = M''_X(0) - \mu^2 = np + n(n-1)p^2 - (np)^2 = np(1-p) = npq.$$

Example 5

Let X be the number of heads (successes) in $n = 7$ independent tosses of an unbiased coin.

The probability mass function of X is

$$f(x) = \begin{cases} \binom{7}{x}\left(\frac{1}{2}\right)^x\left(1-\frac{1}{2}\right)^{7-x}, & x = 0,1,2,3,4,5,6,7, \\ 0, & elsewhere. \end{cases}$$

Then X has MGF, $M_X(t) = \left(\frac{1}{2}+\frac{1}{2}e^t\right)^7$, mean $\mu = np = \frac{7}{2}$, and

variance $\sigma^2 = np(1-p) = \frac{7}{4}$.

Furthermore, we have $P(0 \leq X \leq 1) = \sum_{x=0}^{1} f(x) = \frac{1}{128} + \frac{7}{128} = \frac{8}{128}$, and

$$P(X = 5) = f(5) = \binom{7}{5}\left(\frac{1}{2}\right)^5\left(1-\frac{1}{2}\right)^{7-5} = \binom{7}{5}\left(\frac{1}{2}\right)^7 = \frac{21}{128}.$$

Example 6

If the MGF of a random variable X is $M_X(t) = \left(\frac{2}{3} + \frac{1}{3}e^t\right)^5$, then X has a binomial distribution with $n = 5$ and $p = \frac{1}{3}$; that is, the probability mass function of X is

$$f(x) = \begin{cases} \binom{5}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{5-x}, & x = 0,1,2,3,4,5, \\ 0, & elsewhere. \end{cases}$$

Here $\mu = np = \frac{5}{3}$ and variance $\sigma^2 = np(1-p) = \frac{10}{9}$.

5.5 Poisson Distribution

5.5.1 Introduction

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

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

  For example, a Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, the number of days school is closed due to snow during the winter, or the number of games postponed due to rain during a baseball season.

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

  In such instances, X might represent the number of field mice per acre, the number of bacteria in a given culture, or the number of typing errors per page.

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

5.5.2 Defining Poisson Distribution

- Definition 4 - Poisson Distribution

  Let X be a Poisson random variable which can take on the values $0,1,2,\cdots$ such that the probability distribution of X is given by the probability mass function

$$f(x) = P(X = x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!}, & x = 0,1,2,\cdots, \\ 0, & elsewhere. \end{cases}$$

  where λ is a given positive constant.

- Any random variable X whose probability mass function is given by the above equation is said to be a *Poisson random variable with parameter* λ and the distribution is called the *Poisson distribution* (after S. D. Poisson, who discovered it in the early part of the nineteenth century).

- The Poisson distribution of a random variable X with parameter λ is often written as:

$$X \sim Poisson(\lambda).$$

- Note that by the definition, since $\lambda > 0$ we have $f(x) \geq 0$ and

$$\sum_x f(x) = \sum_x P(X = x) = \sum_x \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_x \frac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1,$$

i.e., $f(x)$ satisfies the conditions of being a probability mass function of a discrete random variable X.

5.5.3 Mean, Variance and MGF of a Poisson Distribution

- The MGF of a Poisson distribution is given by

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^t}$$

$$= e^{\lambda(e^t - 1)},$$

for all vales of t.

- Since $M'_X(t) = e^{\lambda(e^t-1)}(\lambda e^t)$, and

$$M''_X(t) = e^{\lambda(e^t-1)}(\lambda e^t) + e^{\lambda(e^t-1)}(\lambda e^t)^2,$$

we have $\mu = M'_X(0) = \lambda$ and $\sigma^2 = M''_X(0) - \mu^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$.

That is, a Poisson distribution has $\mu = \sigma^2 = \lambda > 0$.

On this account, a Poisson probability mass function is frequently written

$$f(x) = P(X = x) = \begin{cases} \frac{\mu^x e^{-\mu}}{x!}, & x = 0,1,2,\cdots \\ 0, & elsewhere. \end{cases}.$$

Thus the parameter λ in a Poisson probability mass function is the mean μ.

Example 7

Suppose that X has a Poisson distribution with $\mu = 2$. Then the probability mass function of X is

$$f(x) = P(X = x) = \begin{cases} \frac{2^x e^{-2}}{x!}, & x = 0,1,2,\cdots, \\ 0, & elsewhere. \end{cases}$$

The variance of this distribution $\sigma^2 = \mu = 2$.

If we wish to compute $P(1 \leq X)$, we have

$$P(1 \leq X) = 1 - P(X = 0) = 1 - e^{-2} \approx 0.865.$$

Example 8

If the MGF of a random variable X is $M_X(t) = e^{4(e^t - 1)}$, then X has a Poisson distribution with $\mu = 4$.

Accordingly, $P(X = 3) = \frac{4^3 e^{-4}}{3!} = 0.195$.

Example 9

An internet service provider has a large number of users regularly connecting to the internet. On average, 4 users every hour fail to connect to the internet on their first attempt.

(a) Give two reasons why a Poisson distribution might be a suitable model for the number of failed connections every hour.

(b) Find the probability that in a randomly chosen hour:

  (i) 2 users fail to connect on their first attempt.

  (ii) more than 6 users fail to connect on their first attempt.

(c) Find the probability that in a randomly chosen 90-minute period:

  (i) 5 users fail to connect on their first attempt.

  (ii) fewer than 7 users fail to connect on their first attempt.

Solution:

(a) Failed connections occur singly and at a constant rate of 4 users per hour.

(b) Let X represents the number of failed connections in hour.

Then $X \sim Poisson(4)$.

(i) $P(X = 2) = 0.1465$

(ii) $P(X > 6) = 1 - P(X \leq 6) = 1 - 0.88932 = 0.1107$ (4 d.p.)

(c) Let Y represents the number of failed connections in 90 minutes.

Then $X \sim Poisson(6)$.

(i) $P(X = 5) = 0.1606$

(ii) $P(Y < 7) = P(Y \leq 6) = 0.6063$ (4 d.p.)

Example 10

The number of patients visiting a clinic that treats insect bites can be modelled as a Poisson distribution with a rate of 3 patients per day.

(a) Find the probability that there are more than 4 patients on a given day.

(b) An extra doctor is required at the clinic if, within 5 day period, at least 4 of the days have more than 4 patients.

Find the probability that an extra doctor is required.

Solution: Let X represent the number of patients visiting the clinic per day.

Then $X \sim Poisson(3)$.

(a) Need to find, $P(X > 4) = 1 - P(X \leq 4) = 1 - 0.8153 = 0.1847$.

(b) Now we have to consider the number of days in which there are more than 4 patients visiting the clinic.

Let Y represent the number of days when there are more than 4 patients visiting the clinic in 5 day period.

Then $Y \sim Binomial(5, 0.1847)$.

We need to find,

$$
\begin{aligned}
P(Y \geq 4) &= P(Y = 4) + P(Y = 5) \\
&= \binom{5}{4}(0.1847)^4(0.8153)^1 + \binom{5}{5}(0.1847)^5(0.8153)^0 \\
&= 0.004744113 + 0.000214949 \\
&= 0.00496.
\end{aligned}
$$

5.5.4 Computing Probabilities of a Poisson Distribution

- If X is Poisson with mean λ, then the probability mass function is given by

$$f(x) = P(X = x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!}, & x = 0,1,2,\cdots, \\ 0, & elsewhere. \end{cases}$$

then $\frac{P(X=i+1)}{P(X=i)} = \frac{\frac{\lambda^{i+1}e^{-\lambda}}{(i+1)!}}{\frac{\lambda^i e^{-\lambda}}{i!}} = \frac{\lambda}{i+1} \Rightarrow P(X = i+1) = \left(\frac{\lambda}{i+1}\right) P(X = i).$

- Starting with $P(X = 0) = e^{-\lambda}$, we can use equation above to successively compute

$$P(X = 1) = \lambda P(X = 0),$$

$$P(X = 2) = \frac{\lambda}{2} P(X = 1),$$

$$\vdots$$

$$P(X = i+1) = \left(\frac{\lambda}{i+1}\right) P(X = i).$$

Exercises

1. If a random variable X has a Poisson distribution such that $P(X=1) = P(X=2)$, find $P(X=4)$.
2. The MGF of a random variable X is $e^{4(e^t-1)}$.

   Show that $P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.931$.
3. In a lengthy manuscript, it is discovered that only 13.5 percent of the pages contains no typing errors. If we assume that the number of errors per page is a random variable with a Poisson distribution, find the percentage of pages that have exactly one error.

5.5.5 Poisson Approximation to Binomial Distribution

- Suppose that $X \sim Binomial(n, p)$, n is large and p is small.

  Under such circumstances, the distribution of X is well-approximated by a Poison distribution with $\lambda = np$.

  The connection is exact at the limit, i.e.,

  $$Binomial(n, p) \rightarrow Poisson(\lambda)$$

  if $n \rightarrow \infty$ and $p \rightarrow 0$ in such a way that $np = \lambda$ remains constant.

Example 11

Ten percent of the tools produced in a certain manufacturing process turn out to be defective.

Find the probability that in a sample of 10 tools chosen at random, exactly 2 will be defective, by using

(a) the binomial distribution,

(b) the Poisson approximation to the binomial distribution.

Solution:

(a) The probability of a defective tool is $p = 0.1$.

Let X denote the number of defective tools out of 10 chosen.

Then, according to the binomial distribution,

$$P(X = 2) = \binom{10}{2}(0.1)^2(1 - 0.1)^{10-2} = \binom{10}{2}(0.1)^2(0.9)^8 = 0.1937 \text{ or } 0.19.$$

(b) We have $np = 10 \cdot (0.1) = 1$.

Then, according to the Poisson distribution, $P(X = 2) = \frac{(1)^2 e^{-1}}{2!} = 0.1839$ or 0.18.

- In general, the approximation is good if $p \leq 0.1$ and $\lambda = np \leq 5$.

Example 12

If the probability that an individual will suffer a bad reaction from injection of a given serum is 0.001, determine the probability that out of 2000 individuals,

(a) exactly 3,

(b) more than 2,

individuals will suffer a bad reaction.

Solution: Let X denote the number of individuals suffering a bad reaction.

X is Bernoulli distributed, but since bad reactions are assumed to be rare events, we can suppose that X is Poisson distributed, i.e.,

$$f(x) = P(X = x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!}, & x = 0,1,2,\cdots, \\ 0, & elsewhere. \end{cases}$$

where $\lambda = np = 2000 \cdot (0.001) = 2.$

(a) $P(X = 3) = \frac{2^3 e^{-2}}{3!} = 0.180.$

(b) $P(X > 2) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$

$$= 1 - \left[\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!}\right]$$

$$= 1 - 5e^{-2}$$

$$= 0.323.$$

- Note that, an exact evaluation of the probabilities using the binomial distribution may require much more work.

Example 13

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?

Solution: Let X be a binomial random variable with $n = 400$ and $p = 0.005$.

Thus, $\lambda = np = 400 \cdot (0.005) = 2$.

Using the Poisson approximation,

(a) $P(X = 1) = \frac{2^1 e^{-2}}{1!} = 0.271$.

(b) $P(X \leq 3) = \sum_{x=0}^{3} \frac{2^x e^{-2}}{x!} = 0.857$.

Example 14

In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles.

What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

Solution: This is essentially a binomial experiment with $n = 8000$ and $p = 0.001$.

Since p is very close to 0 and n is quite large, we shall approximate with the Poisson distribution using $\lambda = np = 8000 \cdot (0.001) = 8$.

Hence, if X represents the number of bubbles, we have

$$P(X < 7) = \sum_{x=0}^{6} \binom{8000}{x} (0.001)^x (0.999)^{8000-x} \approx \sum_{x=0}^{6} \frac{8^x e^{-8}}{x!} = 0.3134.$$

Exercises

1. An airline is selling tickets to a flight with 198 seats. It knows that, on average, about 1% of customers who have bought tickets fail to arrive for the flight. Because of this, the airline over books the flight by selling 200 tickets. What is the probability that every one who arrives for the flight will get a seat?

2. The chance that a lottery ticket has a winning number is 0.0000001.

   (a) If 10,000,000 people buy tickets which are independently numbered, what is the probability there is no winner?

   (b) What is the probability that there is exactly 1 winner?

   (c) What is the probability that there are exactly 2 winners?

5.6 Hypergeometric Distribution

5.6.1 Introduction

- The types of applications for the hypergeometric are very similar to those for the binomial distribution.
- We are interested in computing probabilities for the number of observations that fall into a particular category.
- But in the case of the binomial distribution, independence among trials is required.

  As a result, if that distribution is applied to, say, sampling from a lot of items (deck of cards, batch of production items), the sampling must be done *with replacement* of each item after it is observed.
- The hypergeometric distribution does not require independence and is based on sampling done *without replacement*.

- Applications for the hypergeometric distribution are found in many areas, with heavy use in acceptance sampling, electronic testing, and quality assurance.
- Obviously, in many of these fields, testing is done at the expense of the item being tested.

  That is, the item is destroyed and hence cannot be replaced in the sample.

  Thus, sampling without replacement is necessary.

- Let us consider the following problem:

  To find the probability of observing 3 red cards in 5 draws from an ordinary deck of 52 playing cards, the binomial distribution does not apply unless each card is replaced and the deck reshuffled before the next draw is made.

  To solve the problem of sampling without replacement, let us restate the problem as follows.

If 5 cards are drawn at random, we are interested in the probability of selecting 3 red cards from the 26 available in the deck and 2 black cards from the 26 available in the deck.

There are $\binom{26}{3}$ ways of selecting 3 red cards, and for each of these ways we can choose 2 black cards in $\binom{26}{2}$ ways.

Therefore, the total number of ways to select 3 red and 2 black cards in 5 draws is the product $\binom{26}{3}\binom{26}{2}$.

The total number of ways to select any 5 cards from the 52 that are available is

$\binom{52}{5}$.

Hence, the probability of selecting 5 cards without replacement of which 3 are red and 2 are black is given by

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = 0.3251.$$

- In general, we are interested in the probability of selecting x successes from the k items labeled successes and $n - x$ failures from the $N - k$ items labeled failures when a random sample of size n is selected from N items.

  This is known as a *hypergeometric experiment*, that is, one that possesses the following two properties:

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

- The number X of successes of a hypergeometric experiment is called a *hypergeometric random variable*.
- Accordingly, the probability distribution of the hypergeometric variable is called the *hypergeometric distribution*.

5.6.2 Defining Hypergeometric Distribution

- Let us now generalize the concept in order to find a formula for the probability mass function of a hypergeometric random variable.

  Let X be the random variable representing the number of successes of a hypergeometric experiment.

  The total number of samples of size n chosen from N items is $\binom{N}{n}$.

  These samples are assumed to be equally likely.

  There are $\binom{k}{x}$ ways of selecting x successes from the k that are available, and for each of these ways we can choose the $n - x$ failures in $\binom{N-k}{n-x}$ ways.

  Thus, the total number of favorable samples among the $\binom{N}{n}$ possible samples is given by $\binom{k}{x}\binom{N-k}{n-x}$.

  Now, we can define the hypergeometric distribution as follows.

- Definition 5 – Hypergeometric Distribution

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

$$f(x) = P(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \max\{0, n-(N-k)\} \leq x \leq \min\{n, k\}.$$

- Any random variable X whose probability mass function is given by the above equation is said to be a *hypergeometric random variable* with parameters N, k and n.

- A hypergeometric distribution of a random variable X with parameters N, k and n is often written as:

$$X \sim H(N, k, n).$$

- The range of x can be determined by the three binomial coefficients in the definition, where x and $n-x$ are no more than k and $N-k$, respectively, and both of them cannot be less than 0.

  Usually, when both k (the number of successes) and $N-k$ (the number of failures) are larger than the sample size n, the range of a hypergeometric random variable will be $x = 0,1,2,\cdots,n$.

Example 15

Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found.

What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

Solution: Let X be the number of defectives found in the sample of size 5.

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

$$P(X = x) = \frac{\binom{3}{1}\binom{40-3}{5-1}}{\binom{40}{5}} = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011.$$

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

Example 16

A box contains 6 blue marbles and 4 red marbles. An experiment is performed in which a marble is chosen at random and its color observed, but the marble is not replaced. Find the probability that after 5 trials of the experiment, 3 blue marbles will have been chosen.

Solution: Method 1

The number of different ways of selecting 3 blue marbles out of 6 blue marbles is $\binom{6}{3}$.

The number of different ways of selecting the remaining 2 marbles out of the 4 red marbles is $\binom{4}{2}$.

Therefore, the number of different samples containing 3 blue marbles and 2 red marbles is $\binom{6}{3}\binom{4}{2}$.

Now the total number of different ways of selecting 5 marbles out of the 10 marbles (6+4) in the box is $\binom{10}{5}$.

Therefore, the required probability is given by $\frac{\binom{6}{3}\binom{4}{2}}{\binom{10}{5}} = \frac{10}{21}$.

Method 2 (Using formula)

Let X be the random variable representing the number of blue marbles chosen.

We have $N = 6 + 4 = 10, n = 5, k = 6$ and $x = 3$.

Then the required probability is

$$P(X = 3) = \frac{\binom{6}{3}\binom{10-6}{5-3}}{\binom{10}{5}} = \frac{\binom{6}{3}\binom{4}{2}}{\binom{10}{5}} = \frac{10}{21}.$$

5.6.3 Mean and Variance of a Hypergeometric Distribution

- The mean and variance of the hypergeometric distribution are given by

$\mu = \frac{nk}{N}$, and

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

Example 17

Find the mean and variance of the random variable of Example 16 and then use Chebyshev's theorem to interpret the interval $\mu \pm 2\sigma$.

Solution: Since Example 16 was a hypergeometric experiment with $N = 40, n = 5$, and $k = 3$, we have

$$\mu = \frac{5 \cdot 3}{40} = \frac{3}{8} = 0.375, \text{ and}$$

$$\sigma^2 = \left(\frac{40-5}{40-1}\right) \cdot 5 \cdot \left(\frac{3}{40}\right) \cdot \left(1 - \frac{3}{40}\right) = 0.3113.$$

Taking the square root of 0.3113, we find that $\sigma = 0.558$.

Hence, the required interval is $0.375 \pm 2 \cdot (0.558)$, or from -0.741 to 1.491.

Chebyshev's theorem states that the number of defectives obtained when 5 components are selected at random from a lot of 40 components of which 3 are defective has a probability of at least $\frac{3}{4}$ of falling between -0.741 and 1.491.

Exercises

1. The components of a 6-component system are to be randomly chosen from a bin of 20 used components. The resulting system will be functional if at least 4 of its 6 components are in working condition. If 15 of the 20 components in the bin are in working condition, what is the probability that the resulting system will be functional? (Ans: ≈ 0.8687) R156

5.6.4 Relationship to the Binomial Distribution

- There is an interesting relationship between the hypergeometric and the binomial distribution.
- As one might expect, if n is small compared to N, the nature of the N items changes very little in each draw.
- So a binomial distribution can be used to approximate the hypergeometric distribution when n is small compared to N.
- In fact, as a rule of thumb, the approximation is good when $\frac{n}{N} \leq 0.05$.
- Thus, the quantity $\frac{k}{N}$ plays the role of the binomial parameter p.
- As a result, the binomial distribution may be viewed as a large-population version of the hypergeometric distribution.
- The mean and variance then come from the formulas

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

- Comparing these formulas with those of Section 5.6.3, we see that the mean is the same but the variance differs by a correction factor of $\left(\frac{N-n}{N-1}\right)$, which is negligible when n is small relative to N.

Example 18

A manufacturer of automobile tires reports that among a shipment of 5000 sent to a local distributor, 1000 are slightly blemished. If one purchases 10 of these tires at random from the distributor, what is the probability that exactly 3 are blemished?

Solution: Since $N = 5000$ is large relative to the sample size $n = 10$, we shall approximate the desired probability by using the binomial distribution.

The probability of obtaining a blemished tire is 0.2.

Therefore, the probability of obtaining exactly 3 blemished tires is

$$P(X = 3) = \frac{\binom{1000}{3}\binom{5000-1000}{10-3}}{\binom{5000}{10}} = 0.2015.$$

Using the binomialapproximation, we get

$$P(X = 3) \approx \binom{10}{3}(0.2)^3(1 - 0.2)^{10-3} = 0.2013.$$

5.7 Negative Binomial Distribution

5.7.1 Introduction

- Let us consider an experiment where the properties are the same as those listed for a binomial experiment, with the exception that the trials will be repeated until a *fixed* number of successes occur.
- Therefore, instead of the probability of x successes in n trials, where n is fixed, we are now interested in the probability that the kth success occurs on the xth trial.
- Experiments of this kind are called *negative binomial experiments*.
- To illustrate this type of experiment consider the following problem:

  Consider the use of a drug that is known to be effective in 60% of the cases where it is used.

  The drug will be considered a success if it is effective in bringing some degree of relief to the patient.

We are interested in finding the probability that the fifth patient to experience relief is the seventh patient to receive the drug during a given week.

Designating a success by S and a failure by F, a possible order of achieving the desired result is *SFSSSFS*, which occurs with probability

$$(0.6)(0.4)(0.6)(0.6)(0.6)(0.4)(0.6) = (0.6)^5(0.4)^2.$$

We could list all possible orders by rearranging the F's and S's except for the last outcome, which must be the fifth success.

The total number of possible orders is equal to the number of partitions of the first six trials into two groups with 2 failures assigned to the one group and 4 successes assigned to the other group.

This can be done in $\binom{6}{4}$ mutually exclusive ways.

Hence, if X represents the outcome on which the fifth success occurs, then

$$P(X = 6) = \binom{6}{4}(0.6)^5(0.4)^2 = 0.1866.$$

5.7.2 Defining Negative Binomial Distribution

- The number X of trials required to produce k successes in a negative binomial experiment is called a *negative binomial random variable*, and its probability distribution is called the *negative binomial distribution*.
- Probabilities of a negative binomial distribution depend on the number of successes desired and the probability of a success on a given trial.
- To obtain the general formula for probability distribution of the random variable X, consider the probability of a success on the xth trial preceded by $k-1$ successes and $x-k$ failures in some specified order.

  Since the trials are independent, we can multiply all the probabilities corresponding to each desired outcome.

  Each success occurs with probability p and each failure with probability $q = 1 - p$.

  Therefore, the probability for the specified order ending in success is

$$p^{k-1} q^{x-k} p = p^k q^{x-k}.$$

The total number of sample points in the experiment ending in a success, after the occurrence of $k-1$ successes and $x-k$ failures in any order, is equal to the number of partitions of $x-1$ trials into two groups with $k-1$ successes corresponding to one group and $x-k$ failures corresponding to the other group.

This number is specified by the term $\binom{x-1}{k-1}$, each mutually exclusive and occurring with equal probability $p^k q^{x-k}$.

We obtain the general formula by multiplying $p^k q^{x-k}$ by $\binom{x-1}{k-1}$.

Now, we can define the negative binomial distribution as follows.

- ## Definition 6 - Negative Binomial Distribution

  If repeated independent trials can result in a success with probability p and a failure with probability $q = 1 - p$, then the probability distribution of the negative binomial random variable X, the number of the trial on which the kth success occurs, is given by the probability mass function

  $$f(x) = P(X = x) = \binom{x-1}{k-1} p^k q^{x-k}, x = k, k+1, k+2, \cdots.$$

- Any random variable X whose probability mass function is given by the above equation is said to be a *negative binomial random variable* with parameters k and p.

- A Negative Binomial distribution of a random variable X with parameters k and p is often written as:

  $$X \sim B^*(k, p).$$

Example 19

In an NBA (National Basketball Association) championship series, the team that wins four games out of seven is the winner.

Suppose that teams A and B face each other in the championship games and that team A has probability 0.55 of winning a game over team B.

(a) What is the probability that team A will win the series in 6 games?

(b) What is the probability that team A will win the series?

(c) If teams A and B were facing each other in a regional playoff series, which is decided by winning three out of five games, what is the probability that team A would win the series?

Solution: We have $x = 6, k = 4, p = 0.55$ and $q = 1 - 0.55 = 0.45$.

(a) $P(X = 6) = \binom{6-1}{4-1}(0.55)^4(0.45)^{6-4} = \binom{5}{3}(0.55)^4(0.45)^2 = 0.1853.$

(b) $P(team\ A\ wins\ the\ championship\ series)$

$$= P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)$$
$$= 0.0915 + 0.1647 + 0.1853 + 0.1668$$
$$= 0.6083.$$

(c) For this part, $k = 3, p = 0.55$ and $q = 1 - 0.55 = 0.45.$

$$P(team\ A\ wins\ the\ playoff) = P(X = 3) + P(X = 4) + P(X = 5)$$
$$= 0.1664 + 0.2246 + 0.2021$$
$$= 0.5931.$$

5.7.3 Mean, Variance and MGF of a Negative Binomial Distribution

- The mean and variance of the negative binomial distribution are given by

$$\mu = \frac{r}{p}, \text{ and } \sigma^2 = \frac{rq}{p^2}.$$

- The MGF of a Negative Binomial Distribution is given by

$$M_X(t) = \frac{pe^t}{(1-qe^t)^r}.$$

5.8 Geometric Distribution

- If we consider the special case of the negative binomial distribution where $k = 1$, we have a probability distribution for the number of trials required for a single success.

  For example, consider the experiment of tossing of a coin until a head occurs. We might be interested in the probability that the first head occurs on the fourth toss.

- When $k = 1$, the negative binomial distribution reduces to the form

$$P(X = x) = pq^{x-1}, x = 1,2,3,\cdots.$$

  Since the successive terms constitute a geometric progression, it is customary to refer to this special case of the negative binomial distribution as the *geometric distribution*.

  Now we can define the geometric distribution formally as follows.

- Definition 7 – Geometric Distribution

  If repeated independent trials can result in a success with probability p and a failure with probability $q = 1 - p$, then the probability distribution of the geometric random variable X, the number of the trial on which the first success occurs, is given by the probability mass function

  $$f(x) = P(X = x) = pq^{x-1}, x = 1,2,3, \cdots.$$

  The above distribution of the random variable X is called the *Geometric Distribution* of X.

- A geometric distribution of a random variable X with parameter p is often written as:

  $$X \sim G(p).$$

Example 20

For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective.

What is the probability that the fifth item inspected is the first defective item found?

Solution: Using the geometric distribution with $x = 5$ and $p = 0.01$, we have

$P(X = 5) = (0.01)(0.99)^4 = 0.0096.$

Example 21

At a "busy time," a telephone exchange is very near capacity, so callers have difficulty placing their calls. It may be of interest to know the number of attempts necessary in order to make a connection.

Suppose that we let $p = 0.05$ be the probability of a connection during a busy time. We are interested in knowing the probability that 5 attempts are necessary for a successful call.

Solution: Using the geometric distribution with $x = 5$ and $p = 0.05$ yields

$$P(X = 5) = (0.05)(0.95)^4 = 0.041.$$

- Quite often, in applications dealing with the geometric distribution, the mean and variance are important.

  The mean and variance of the geometric distribution are given by

  $$\mu = \frac{1}{p}, \text{ and } \sigma^2 = \frac{q}{p^2}.$$

- The MGF of a geometric distribution is given by $M_X(t) = \frac{pe^t}{1-qe^t}$.

  We can obtain these results by letting $r = 1$ in mean, variance and MGF of the negative binomial distribution.