CHAPTER II
DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

Second Year of Engineering Program
Department of Foundation Year

LIN Mongkolsery
sery@itc.edu.kh

2021-2022

Contents

Random Variables

Definition 1

Let $\mathcal{S}$ be a sample space of random experiments. A function $X : \mathcal{S} \to \mathbb{R}$ is called a **random variable (rv)**.

- $D = \{x = X(s) : s \in \mathcal{S}\}$ is called **range** of random variable X. It is the set of all possible value of X.
- If D is countable set then X is called **discrete random variable**.
- If D is uncountable set then X is called **continuous random variable**.

NOTE

In this whole chapter, X will be denoted a discrete random variable (drv). Otherwise, it needs to be specified.

Random Variables

Example 1

A rat is selected at random from a cage and its sex is determined. The set of possible outcomes is female and male. Thus, the outcome space is $S = \{\mathrm{Female}, \mathrm{Male}\} = \{F, M\}$. We defind $X : S \to \mathbb{R}$ by $X(F) = 0$ and $X(M) = 1$. Then we have X is a random variable, moreover, it is a discrete random variable since $D = \{0, 1\}$.

Example 2

Suppose a pair of fair dices are rolled. Then the sample space is $S = \{(i, j) : 1 \leq i, j \leq 6\}$. We define $X : S \to \mathbb{R}$ by $X(i, j) = i + j$. Then X is a discrete random variable. The set of all possible value of X is $\mathcal{D} = \{2, \ldots, 12\}$.

Probability Distributions for Discrete Random Variables

Definition 2

The **probability distribution** or **probability mass function (pmf)** of a discrete random variable is denoted $p_X(x)$ and defined by by

$$p_X(x) = P(X = x) = P(s \in S : X(s) = x).$$

Properties

The pmf $p_X(x)$ must satisfy the following properties:

1. $p_X(x) \geq 0, \forall x \in \mathcal{D}$,
2. $\sum_{x \in \mathcal{D}} p_X(x) = 1$.

Probability Distributions for Discrete Random Variables

Example 3

Let X be the sum of the up-faces on a roll of a pair of fair 6-sided dice, each with the numbers 1 through 6 on it. The sample space is $S = \{(i,j) : 1 \leq i,j \leq 6\}$. Because the dice are fair, $P[\{(i,j)\}] = 1/36$. The random variable X is $X(i,j) = i + j$. The set of all possible value of X is $\mathcal{D} = \{2, \ldots, 12\}$. By enumeration, the pmf of X is given by

X	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Suppose $A_1 = \{x : x = 7, 11\}$ and $A_2 = \{x : x = 2, 3, 12\}$, then $P(A_1) = 8/36$ and $P(A_2) = 4/36$.

The Cumulative Distribution Function

Definition 3

For a dev X, the **cumulative distribution function (cdf)**, denoted by F_X, is defined by

$$F_X(x) = P(X \leq x) = \sum_{y \leq x} p_X(y).$$

The Cumulative Distribution Function

Example 4

From Example 3, we have

X	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

We then able to compute the values of the cdf as follow.

$$F_X(1.7) = P(X \leq 1.7) = P(\emptyset) = 0$$

$$F_X(2) = P(X \leq 2) = P(X = 2) = \frac{1}{36}$$

$$F_X(3) = P(X \leq 3) = P(X = 2) + P(X = 3) = \frac{1}{12}$$

The Cumulative Distribution Function

The cdf $F_X(x)$ is given by

$$F_X(x) = \begin{cases} 0, & x < 2 \\ 1/36, & 2 \le x < 3 \\ 1/12, & 3 \le x < 4 \\ 1/6, & 4 \le x < 5 \\ 5/18, & 5 \le x < 6 \\ 5/12, & 6 \le x < 7 \\ 21/36, & 7 \le x < 8 \\ 7/12, & 8 \le x < 9 \\ 5/6, & 9 \le x < 10 \\ 11/12, & 10 \le x < 11 \\ 35/36, & 11 \le x < 12 \\ 1, & 12 \le x \end{cases}$$

The Cumulative Distribution Function

Theorem 1

Let $F_X(x)$ be a cdf of drv X. Then,

1. If $a < b$, then $F_X(a) \leq F_X(b)$, (F is a nondecreasing function).
2. $\lim_{x \to -\infty} F_X(x) = 0$, (the lower limit of F_X is 0).
3. $\lim_{x \to +\infty} F_X(x) = 1$, (the upper limit of F_X is 1).
4. $\lim_{x \to x_0^+} F_X(x) = F_X(x_0)$, ($F_X$ is right continuous).

Theorem 2

Let F_X be a cdf of drv X. Then,

1. $P(a < X \leq b) = F_X(b) - F_X(a)$.
2. $P(X = x) = F_X(x) - F_X(x^-)$, where $F_X(x^-) = \lim_{z \to x^-} F_X(z)$

Probability distribution for discrete random variables

Example 5

Given the cdf F_X as follow.

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x/2, & 0 \leq x < 1 \\ 1, & 1 \leq x. \end{cases}$$

Then,

$$P(-1 < X \leq 1/2) = F_X(1/2) - F_X(-1) = 1/4 - 0 = 1/4$$

and

$$P(X = 1) = F_X(1) - F_X(1^-) = 1 - 1/2 = 1/2.$$

The value $1/2$ equals the value of the step of F_X at $x = 1$.

Probability distribution for discrete random variables

Example 6

An automobile service facility specializing in engine tune-ups knows that 45% of all tune-ups are done on four-cylinder automobiles, 40% on six-cylinder automobiles, and 15% on eight-cylinder automobiles. Let X be number of cylinders on the next car to be tuned.

(a) What is the pmf of X?

(b) Draw both a line graph and a probability histogram for the pmf of part (a).

(c) What is the probability that the next car tuned has at least six cylinders? More than six cylinders?

Expected Values

Definition 4

The **expected value** or **mean value** of X, denoted by $E(X)$ or μ_X or just μ, is

$$E(X) = \sum_{x \in \mathcal{D}} x.p_X(x)$$

Theorem 3

The expected value of any function $h(X)$, denoted by $E[h(X)]$, is given by

$$E[h(X)] = \sum_{x \in \mathcal{D}} h(x).p_X(x).$$

In particular, if a and b are constants

1. $E(b) = b$
2. $E(aX + b) = aE(X) + b.$

The Variance of X

Definition 5

The **variance** of drv X, denoted by $V(X)$ or σ_X^2, or just σ^2, is defined by

$$V(X) = E[(X-\mu)^2] = \sum_{x \in \mathcal{D}} (x-\mu)^2 . p_X(x).$$

The **standard deviation** (SD) of X is

$$\sigma_X = \sqrt{V(X)}.$$

Theorem 4

1. $V(X) = E(X^2) - [E(X)]^2$
2. $V(aX+b) = a^2 V(X)$
3. $\sigma_{aX+b} = |a| \sigma_X$.

The Variance of X

Example 7

The pmf of the amount of memory $X(GB)$ in a purchased flash drive was given in as

x	1	2	4	8	18
$p(x)$	.05	.10	.35	.40	.10

Compute the following:

(a) $E(X)$.

(b) $V(X)$ directly from the definition.

(c) The standard deviation of X.

(d) $V(X)$ using the shortcut formula.

The Moment-Generating Function of X

Definition 6

The **moment-generating function (mgf)** of X, denoted by $M(t)$, is defined by

$$M(t) = E\left(e^{tX}\right)$$

if it exists.

Theorem 5

1. $M^{(n)}(t) = E[X^n e^{tX}]$
2. $E[X] = M'(0)$
3. $V(X) = M''(0) - [M'(0)]^2$

The Binomial Probability Distribution

Definition 7

A **binomial experiment** is an experiment satisfying the following properties:

1. It consists of a sequence of n smaller experiments called **trials**, where n is a (non-random) constant.
2. Each trial can result in one of the same two possible outcomes (dichotomous trials), which we generically denote by success and failure.
3. The trials are independent
4. The probability of success is constant from trial to trial, we denote this probability by p.

Note that when $n = 1$, it is called **Bernoulli experiment**.

The Binomial Random Variable and Distribution

Definition 8

The **binomial random variable** X associated with a binomial experiment consisting of n trials is defined by

$$X = \text{the number of success among the } n \text{ trials.}$$

We write $X \sim \mathrm{Bin}(n, p)$ to indicate that X is a binomial rv based on n trials with success probability p.

Theorem 6

If $X \sim \mathrm{Bin}(n, p)$, then the pmf of the binomial rv X is given by

$$p_X(x) = \begin{cases} C_n^x . p^x (1-p)^{n-x}, & x = 0, 1, 2, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

Using Binomial Tables

Theorem 7

If $X \sim \mathrm{Bin}(n, p)$, then

1. $E(X) = np$
2. $V(X) = npq$
3. $\sigma_X = \sqrt{npq}$
4. $M(t) = (q + pe^t)^n$

where $q = 1 - p$.

Binomial Distribution

Example 8

Many utility companies promote energy conservation by offering discount rates to consumers who keep their energy usage below certain established subsidy standards. A recent report notes that 70% of the people live in Phnom Penh have reduced their electricity usage sufficiently to qualify for discounted rates. If five residential subscribers are randomly selected from Phnom Penh, find the probability of each of the following events:

1. All five qualify for the favorable rates.
2. At least four qualify for the favorable rates.
3. At least two do not qualify the favorable rates.
4. How many subscribers would you expect to qualify the favorable rates?

The Hypergeometric Distribution

Definition 9

The **hypergeometric experiment** is an experiment satisfying the following properties:

1. The population or set to be sampled consists of N individuals, objects, or elements (a finite population).
2. Each individual can be characterized as a success (S) or a failure (F), and there are M successes in the population.
3. A sample of n individuals is selected without replacement in such a way that each subset of size n is equally likely to be chosen.

The Hypergeometric Distribution

Definition 10

The **hypergeometric random variable** X associated with hypergeometric experiment is defined by

$$X = \text{the number of success in the sample.}$$

$X \sim \mathrm{Hp}(n, M, N)$ denotes the hypergeometric rv of sample size n drawn from a population of size N consisting of M success.

Theorem 8

If $X \sim \mathrm{Hp}(n, M, N)$, then the pmf of the hypergeometrix rv X is given by

$$p_X(x) = \frac{C_M^x . C_{N-M}^{n-x}}{C_N^n}$$

for all integers x satisfying $\max(0, n - N + M) \leq x \leq \min(n, M)$.

The Hypergeometric Distribution

Theorem 9

If $X \sim \mathrm{Hp}(n, M, N)$, then

1. $E(X) = n.\dfrac{M}{N}$
2. $V(X) = \left(\dfrac{N-n}{N-1}\right) n\dfrac{M}{N}\left(1 - \dfrac{M}{N}\right)$

Example 9

Five individuals from an animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into the population. After they have had an opportunity to mix, a random sample of 10 of these animals is selected. Let $X =$ the number of tagged animals in the second sample. If there are actually 25 animals of this type in the region, what is the probability that $(a) X = 2$? $(b) X \leq 2$? and then find $E(X)$ and $V(X)$.

The relationship between Hypergeometric and Binomial Distributions

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 $\frac{n}{N} \leq 0.05$.

Example 10

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?

Answer: 0.2013

The Negative Binomial Distribution

Definition 11

The **negative binomial experiment** is an experiment satisfying the following conditions:

1. The experiment consists of a sequence of independent trials.
2. Each trial can result in either a success (S) or a failure (F).
3. The probability of success is constant from trial to trial, so $P(\text{S on trial } i) = p$ for $i = 1, 2, 3, \ldots$.
4. The experiment continues (trials are performed) until a total of r successes have been observed, where r is a specified positive integer.

The Negative Binomial Distribution

Definition 12

The **negative binomial rv** X associated with the negative binomial experiment is defined by

$$X = \text{the number of failures that precede the } r\text{th success.}$$

$X \sim \mathrm{Nb}(r, p)$ denotes the negative binomial rv X with parameters $r =$ the number of success and $P(\text{success}) = p$.

The Negative Binomial Distribution

Theorem 10

If $X \sim \mathrm{Nb}(r, p)$, then the pmf of the negative binomial rv X is given by

$$p_X(x) = C_{x+r-1}^{r-1} p^r (1-p)^x$$

where $x = 0, 1, 2, \ldots$.

Remark:

- if $r = 1$ then X is a **geometric random variable** with pmf

$$p_X(x) = p(1-p)^x, \qquad x = 0, 1, 2, \ldots.$$

The Negative Binomial Distribution

Theorem 11

If $X \sim \mathrm{Nb}(r, p)$, then

1. $E(X) = \dfrac{r(1-p)}{p}$
2. $V(X) = \dfrac{r(1-p)}{p^2}$
3. $M(t) = \left(\dfrac{p}{1 - e^t + pe^t}\right)^r$.

The Negative Binomial Distribution

Example 11

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?

Example 12

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 5 games?

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

(c) What is the probability that team B will win the series in 6 games?

Poisson Probability Distribution

Definition 13

A drv X is called **Poisson distribution** or **Poisson rv** with parameter $\lambda (\lambda > 0)$ if the pmf of X is

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

We write $X \sim \mathrm{Po}(\lambda)$.

Poisson Distribution as a Limit

Theorem 12

Suppose $X \sim \mathrm{Bin}(n, p)$. When $n \to \infty,\ p \to 0$, and $np \to \lambda$ remains constant, then

$$X \sim \mathrm{Po}(\lambda).$$

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?

The Mean and Variance of X

Theorem 13

If $X \sim \mathrm{Po}(\lambda)$, then

1. $E(X) = \lambda$
2. $V(X) = \lambda$
3. $M(t) = e^{\lambda(e^t - 1)}$.

Example 14

If a publisher of nontechnical books takes great pains to ensure that its books are free of typographical errors, so that the probability of any given page containing at least one such error is .005 and errors are independent from page to page, what is the probability that one of its 400-page novels will contain exactly one page with errors? At most three pages with errors?