Chapter S3

Discrete Random Variables

In this unit, students will

- understand concepts of discrete random variables, probability distributions, expectations and variances.

Introduction

Consider the following question:

What is the probability of obtaining three 2's when six dice are tossed?

We can solve this question using the technique learnt in the Chapter S2. However, we will like to generalise the method to solving similar type of questions. As such, we will introduce the idea of a discrete random variable. As a quick introduction, when 6 dice are tossed, the number of two's obtained could take on the following values: 0, 1, 2, 3, 4, 5, 6. Each possible value has an associated probability. Thus if we let the variable X be the number of two's obtained, then this variable is random since the possible values has a certain probability of occurring.

The convention is to use capital letter, say X, as the name of the random variable and the corresponding lower case letter, x, to represent one of the values it can take. For the example above, X will denote the number of two's obtained, and we are interested in $\mathrm{P}(X = x)$ when $x = 2$, or simply written as $\mathrm{P}(X = 2)$. In this chapter, we will formally introduce this and its associated properties.

3.1 Discrete Random Variables

Definition:

Let X be a variable which has the following properties:

a) It is a discrete variable and can take only values $x_1,\ x_2,\ \ldots,\ x_n$.

b) The probability that X assumes $x_1,\ x_2,\ \ldots,\ x_n$ are $p_1,\ p_2,\ \ldots,\ p_n$ respectively.

Then X is a **discrete random variable** if $\sum_{i=1}^{n} p_i = 1$.

A table showing all the possible occurrences and probabilities is called a **probability distribution table**. The function that allocates probabilities $\mathrm{P}(X = x)$ is known as the **probability density function (p.d.f) of X**.

Example 1: Obtain a formula for the probability distribution of the random variable X defined as the result of rolling a fair six-sided die.

Solution:

Since the die is fair, $\mathrm{P}(X = x) = \frac{1}{6}$ for $x = 1, 2, \ldots, 6$. This can be presented by the following table:

x	1	2	3	4	5	6
$\mathrm{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Remark: This is an example of a *uniform distribution.*

Example 2: Two tetrahedral dice, each with faces labelled 1, 2, 3 and 4 are thrown and the score, X, is noted, where X is the sum of the two numbers on which the dice land. Compute the probability distribution table of X.

Solution:

The outcome table for X is as follows:

	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

From the outcome table, the possible values of X are 2, 3, 4, 5, 6, 7, and 8.
To find the probability, one can refer to the outcome table and compute them.
The required probability distribution table is as follows:

x	2	3	4	5	6	7	8
$\mathrm{P}(X = x)$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

Remark: This distribution is *symmetric.*

Note: One way to check your answer is to add up all the probabilities and see if it equates to 1, i.e. $\sum_{\text{all } x} \mathrm{P}(X = x) = 1$.

Example 3:
Three coins are tossed.
(i) Construct the probability distribution table for the number of heads obtained.
(ii) Write down a formula for the probability distribution function.
(iii) Find the probability that there are

- **(a)** at least 1 head,
- **(b)** not more than 2 heads,
- **(c)** at least 1 head but not more than 2 heads,
- **(d)** not more than 2 heads given that there is at least 1 head.

Solution:

(i) Let X be the number of heads obtained when three coins are tossed.

x	0	1	2	3
$\mathrm{P}(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

(ii) $\mathrm{P}(X=x)=\binom{3}{x}\left(\frac{1}{2}\right)^3$, for $x=0,\ 1,\ 2,\ 3$.

(iii) (a) $\mathrm{P}(X\geq 1)=1-\mathrm{P}(X=0)=\frac{7}{8}$

(b) $\mathrm{P}(X\leq 2)=\frac{7}{8}$

(c) $\mathrm{P}(1\leq X\leq 2)=\mathrm{P}(X=1)+\mathrm{P}(X=2)=\frac{3}{4}$

(d) $\mathrm{P}(X\leq 2\mid X\geq 1)=\frac{\mathrm{P}(1\leq X\leq 2)}{\mathrm{P}(X\geq 1)}=\frac{3/4}{7/8}=\frac{6}{7}$

Thinkzone:
It is necessary to define the discrete random variable if the question did not do so.

For (ii), suppose r heads are obtained, we must consider all the possible arrangement of the appearance of the heads in the 3 tosses, that is, $\binom{3}{r}$.

Remark: $\mathrm{P}(X\leq 2)$ in (iii)(b) is an example of a value of the cumulative distribution function of X, $\mathrm{P}(X\leq x)$.

Example 4: A box contains three red marbles and five green marbles which are indistinguishable except for colour. Five marbles are taken at random **with replacement** and the colour of the marble is noted for each draw. Let X be the number of times a green marble is drawn. Find the probability distribution function of X.

Solution:

The possible values of X are $x=0,1,\ldots,5$.

Thus
$$\mathrm{P}(X=x)=\binom{5}{x}\left(\frac{5}{8}\right)^x\left(\frac{3}{8}\right)^{5-x}\quad\text{for } x=0,1,\ldots,5.$$

Thinkzone:
Probability of obtaining a green marble is $\frac{5}{8}$ and obtaining a red marble is $\frac{3}{8}$.
Suppose a green marble is drawn r times. We must consider all the possible arrangement of the appearance of the green marble in the five **draws**, that is, $\binom{5}{r}$.

Remark: The probability distributions in Example 3 and 4 are examples of the binomial distribution, which will be discussed in detail in the next chapter.

Example 5:

Emma plays a game in which she throws two dice. If she get two sixes, she wins 20 cents, if she gets one six she wins 10 cents, otherwise she wins nothing. She has to pay 5 cents to enter the game.

(i) Write out the probability distribution of X, the amount Emma gains in one game.

(ii) What is the mode of the distribution?

Solution:

The possible values of X are $x = -5,\ 5$ and 15.

Thus $\mathrm{P}(X=-5)=\left(\frac{5}{6}\right)^2=\frac{25}{36}$

$\mathrm{P}(X=5)=2\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)=\frac{5}{18}$

$\mathrm{P}(X=15)=\left(\frac{1}{6}\right)^2=\frac{1}{36}$

Thus the probability distribution of X is

x	-5	5	15
$\mathrm{P}(X=x)$	$\frac{25}{36}$	$\frac{5}{18}$	$\frac{1}{36}$

where x is in cents.

The mode is -5.

Thinkzone:

The gain is obtained by subtracting from her wins the cost of entering the game.

Check if the probabilities add up to 1.

The **mode** is the value with the **highest probability**.

Example 6:

The probability density function of the discrete random variable M is given by

$$\mathrm{P}(M=m)=\begin{cases}\frac{1}{10}|m| & \text{for } m=-1,\ 0,\ \frac{1}{2}\\ km^2 & \text{for } m=1,\ 2\end{cases}$$

Find k.

Solution:

Since $\sum_{\text{all } m}\mathrm{P}(M=m)=1$,

$$\frac{1}{10}|-1|+0+\frac{1}{10}\left(\frac{1}{2}\right)+k(1)^2+k(2)^2=1$$

$$5k=\frac{17}{20}\Rightarrow k=\frac{17}{100}$$

Note: As seen from Example 5 and 6, a discrete random variable can take on negative and/or non-integer values. Examples include shoe sizes which can have values like $8\frac{1}{2}$ and gains in a game.

Self-Review 1: Packets of Golden Cornflakes are sold for \$3.50 each. One in twenty of the packets contains a \$1 coin. Mr. Wong buys two packets and \$$X$ is the net cost of the two packets. Find the probability distribution for X.

3.2 Expectation

Definition:

The **expectation** of a discrete random variable X, denoted by E(X), is given by

$$\mathrm{E}(X)=\sum_{\text{all }x} x\mathrm{P}(X=x).$$

We often use the symbol μ for the expectation, $\mu=\mathrm{E}(X)$.

The expectation is another word for the average of the values of X. This definition is motivated by the formula $\bar{x}=\dfrac{\Sigma fx}{\Sigma f}=\sum x\left(\dfrac{f}{\Sigma f}\right)$ where f represent the frequency of the value associated with x.

Remark: E(X) is a **constant** and not a function of X. E(X) represents the theoretical average (or mean) that is expected if a sufficiently large number of the same experiment were repeated.

Properties of Expectation:

a) If a and b are constants, then
- $\mathrm{E}(a)=a$,
- $\mathrm{E}(aX)=a\mathrm{E}(X)$ and
- $\mathrm{E}(aX\pm b)=a\mathrm{E}(X)\pm b$.

b) If g(X) is a function of the discrete random variable of X, then

$$\mathrm{E}\big(\mathrm{g}(X)\big)=\sum_{\text{all }x}\mathrm{g}(x)\mathrm{P}(X=x).$$

c) If X and Y are discrete random variables, then

$$\mathrm{E}(aX\pm bY)=a\mathrm{E}(X)\pm b\mathrm{E}(Y) \text{ for constants } a \text{ and } b.$$

Example 7:

Find the expectation of the random variable X with the following probability distribution function.

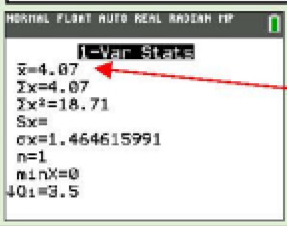
x	0	1	2	3	4	5	6
P($X=x$)	0.03	0.04	0.06	0.12	0.4	0.15	0.2

Solution:

$$\mathrm{E}(X)=\sum_{\text{all }x} x\mathrm{P}(X=x)=0(0.03)+1(0.04)+2(0.06)+3(0.12)$$
$$+4(0.4)+5(0.15)+6(0.2)$$
$$=4.07$$

Thinkzone:

We can also use graphing calculator to compute the mean. Refer to **Annex 1** for details to get the following.



Note that $\bar{x}$ refers to the mean. In the event when probabilities are entered in L_2, Σx and Σx^2 give the value of $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$ respectively.

Example 8: A discrete random variable X is defined as the result of rolling a fair four-sided die that has sides labelled 1, 2, 3 and 4. Find E(X), E(X^3) and E($3X - 1$).

Solution:

We have

x	1	2	3	4
x^3	1	8	27	64
$3x-1$	2	5	8	11
$\mathrm{P}(X=x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

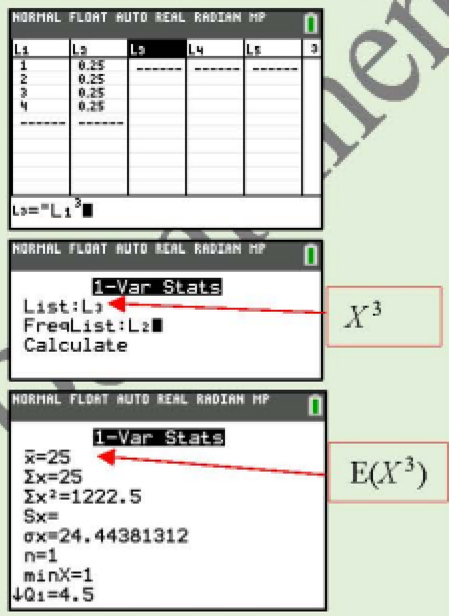
$$\mathrm{E}(X)=1\left(\frac{1}{4}\right)+2\left(\frac{1}{4}\right)+3\left(\frac{1}{4}\right)+4\left(\frac{1}{4}\right)=\frac{5}{2}.$$

$$\mathrm{E}\left(X^3\right)=1\left(\frac{1}{4}\right)+8\left(\frac{1}{4}\right)+27\left(\frac{1}{4}\right)+64\left(\frac{1}{4}\right)=25,$$

$$\mathrm{E}\left(3X-1\right)=3\mathrm{E}\left(X\right)-1=3\left(\frac{5}{2}\right)-1=\frac{13}{2}$$

$$\left[\text{Alternatively, } \mathrm{E}\left(3X-1\right)=2\left(\frac{1}{4}\right)+5\left(\frac{1}{4}\right)+8\left(\frac{1}{4}\right)+11\left(\frac{1}{4}\right)=\frac{13}{2}\right]$$

Thinkzone:
Check your results using GC.
How can you use the GC in this case to obtain the value of E(X^3) and E($3X - 1$)? Refer to **Annex 2**.



Self-Review 2:

The random variable X has distribution given by $\mathrm{P}(X=-1)=\frac{1}{5}$, $\mathrm{P}(X=3)=\frac{2}{5}$, $\mathrm{P}(X=8)=\frac{2}{5}$.

Find (a) $\mathrm{E}(X)$, (b) $\mathrm{E}(X^2)$, (c) $\mathrm{E}(|X|)$ and (d) $\mathrm{E}\left(\frac{1}{X}\right)$. [4.2, 29.4, 4.6, –0.0167]

Example 9: In a gambling game, Pascal is paid \$$a$ if he gets all heads or tails when 3 fair coins are tossed. He pays \$3 if either 1 head or 2 heads are shown. Calculate the expected gain when $a=5$. Find the value of a such that the game is fair. (Note: A game is fair if the expected gain is 0.)

Solution:

Let X be the gain from one game. Thus

x	5	-3
$\mathrm{P}(X=x)$	$\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^3=\frac{1}{4}$	$1-\left[\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^3\right]=\frac{3}{4}$

Expected gain = $\mathrm{E}(X)=5\left(\frac{1}{4}\right)-3\left(\frac{3}{4}\right)=-1$

To make the game fair, $\mathrm{E}(X)=a\left(\frac{1}{4}\right)-3\left(\frac{3}{4}\right)=0 \quad \Rightarrow \quad a=9$

Example 10: A random variable X has probability density function as follows:

x	1	2	3	4	5
$P(X=x)$	0.1	0.2	$2p$	0.2	0.1

Find the value of
(i) p,
(ii) the expectation, $E(X)$ and
(iii) $E(X^2)$.
Hence, evaluate the value of $E(X^2)-[E(X)]^2$.

Solution:

(i) Since $\sum_{\text{all } x} P(X=x)=1$, thus $0.1+0.2+2p+0.2+0.1=1 \Rightarrow p=0.2$

(ii) Method 1: Since the distribution is symmetrical about $x=3$, $E(X)=3$.

Method 2: $E(X)=1\cdot(0.1)+2(0.2)+3(0.4)+4(0.2)+5(0.1)=3$

(iii) $E(X^2)=1\cdot(0.1)+2^2(0.2)+3^2(0.4)+4^2(0.2)+5^2(0.1)=10.2$

$E(X^2)-[E(X)]^2=10.2-3^2=1.2$

Important note: In Example 10, it can be seen from the table that the distribution is symmetrical about x = 3 and we computed and found that $E(X)$ = 3. In fact, **for any distribution that is symmetrical about x_0, then $E(X) = x_0$.**

The value computed for part (c) in Example 10 for $E(X^2)-[E(X)]^2$ is called the **variance** of X.

3.3 Variance

There are various ways to describe or summarise a certain set of data values. The **mean** is an example of a measure of central tendency. There are other measures such as the **median** and **mode**. These measures attempt to measure a typical value about which the distribution clusters.
In order to describe the data set more accurately, we must look at other factors such as the extent to which the data is scattered or spread out. Measures commonly used to describe the spread of a data are the **range**, the **variance** and the **standard deviation**. These are called measures of variation.

Definition:
The **variance** of X, written **Var(X)**, is given by

$$\text{Var}(X)=E\left[(X-\mu)^2\right]=E(X^2)-[E(X)]^2$$

We use the symbol, σ^2, for Var(X).

The **standard deviation** of X is $\sigma=\sqrt{\text{Var}(X)}$, where $\sigma > 0$.

To establish the second equality in the definition, we have

$$\begin{aligned}\mathrm{E}\left[(X-\mu)^2\right] &= \mathrm{E}\left(X^2-2\mu X+\mu^2\right)\\ &= \mathrm{E}(X^2)-2\mu\mathrm{E}(X)+\mu^2\\ &= \mathrm{E}(X^2)-2\mu^2+\mu^2\\ &= \mathrm{E}(X^2)-\mu^2\\ &= \mathrm{E}(X^2)-\left[\mathrm{E}(X)\right]^2\end{aligned}$$

Properties of Variance:

(a) If a and b are constants, then

- $\mathrm{Var}(a)=0$,
- $\mathrm{Var}(aX)=a^2\mathrm{Var}(X)$ and
- $\mathrm{Var}(aX\pm b)=a^2\mathrm{Var}(X)$

(b) If X and Y are **independent** discrete random variables, then

$$\mathrm{Var}(aX\pm bY)=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y).$$

Example 11:

Find the variance of the random variable X with the following probability distribution function.

x	0	1	2	3	4	5	6
$\mathrm{P}(X=x)$	0.03	0.04	0.06	0.12	0.4	0.15	0.2

Solution:

From Example 7, $\mathrm{E}(X)=4.07$

$$\mathrm{E}(X^2)=\sum_{\text{all } x}x^2\mathrm{P}(X=x)$$

$$=0(0.03)+1^2(0.04)+2^2(0.06)+3^2(0.12)+4^2(0.4)+5^2(0.15)+6^2(0.2)$$

$$=18.71$$

$$\mathrm{Var}(X)=\mathrm{E}(X^2)-\left[\mathrm{E}(X)\right]^2$$

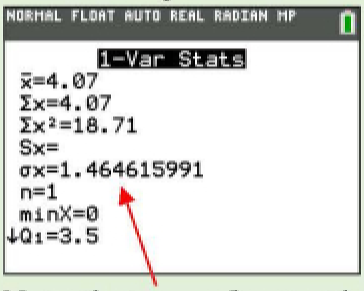
$$=18.71-(4.07)^2$$

$$=2.1451$$

$$=2.15 \quad \text{(to 3 s.f.)}$$

Thinkzone:

Using the same method as describe earlier in Example 7, we can use 1-Var Stats to compute the variance.



Note that σ_x refers to the standard deviation. Thus squaring this quantity will give us the variance.

Example 12: The random variable X has probability distribution given by:

x	2	3	4
$\mathrm{P}(X=x)$	p	p	$1-2p$

Find the expectation of X and show that X has variance equal to $p(5-9p)$.

Determine the range of values of p such that the variance of X is more than $2-6p$.

Solution:

$\mathrm{E}(X) = 2p + 3p + 4(1-2p) = 4-3p$

$\mathrm{E}(X^2) = 4p + 9p + 16(1-2p) = 16-19p$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \\ &= 16-19p-(4-3p)^2 \\ &= 5p-9p^2 \\ &= p(5-9p)\end{aligned}$$

$$\begin{aligned}& p(5-9p) > 2-6p \\ \Rightarrow\ & 9p^2 - 11p + 2 < 0 \\ \Rightarrow\ & (9p-2)(p-1) < 0 \\ \Rightarrow\ & \frac{2}{9} < p < 1\end{aligned}$$

Since $1-2p>0$, that is $p<\frac{1}{2}$. Thus $\frac{2}{9} < p < \frac{1}{2}$.

Thinkzone:

Example 13: (MB 1976/I/12)

Ten playing cards, two of which are Aces, are lying face down on a table. A player turns the cards over, one by one in random order, and the N^{th} card turned over is the first Ace found $(1 \le N \le 9)$. Calculate the probabilities P(N = 2) and P(N = 9).

A player pays a stake of \$10 to try his luck at finding an Ace, if the first card he turns over is an Ace he gets his stake \$10 back and in addition he gets a prize of \$10. If he fails to find an Ace with his first card, and if he turns a second card, and if this is an Ace his \$10 stake is returned, but otherwise he gets nothing at all and his stake is lost. Calculate the player's expectation, stating whether it is a gain or a loss.

Solution:

$\mathrm{P}(N=2) = \frac{8}{10}\left(\frac{2}{9}\right) = \frac{8}{45}$

$\mathrm{P}(N=9) = \left(\frac{8}{10}\right)\left(\frac{7}{9}\right)\left(\frac{6}{8}\right)\left(\frac{5}{7}\right)\left(\frac{4}{6}\right)\left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)\left(\frac{2}{2}\right) = \frac{1}{45}$

Let Y be the player's profit.

y	10	0	−10
$\mathrm{P}(Y=y)$	$\mathrm{P}(N=1) = \frac{2}{10}$	$\mathrm{P}(N=2) = \frac{8}{45}$	$\mathrm{P}(N \ge 3) = 1 - \frac{2}{10} - \frac{8}{45} = \frac{28}{45}$

$$\begin{aligned}\mathrm{E}(Y) &= 10\left(\frac{2}{10}\right) - 10\left(\frac{28}{45}\right) \\ &= -\frac{38}{9}\end{aligned}$$

Since $\mathrm{E}(Y) < 0$, thus the player's expectation is a loss.

Thinkzone:

$\mathrm{P}(N=9) = \frac{8!2!}{10!} = \frac{1}{45}$

There are 10! ways to arrange the cards. If N = 9, the first 8 cards are not Aces (8! ways to arrange) while the 9^{th} and 10^{th} cards are both aces (2! ways to arrange).

Example 14: Suppose that the number of cars, X, that pass through a car wash between 4 p.m. and 5 p.m. on any sunny Friday has the following distribution

x	4	5	6	7	8	9
$P(X=x)$	0.1	0.2	0.3	p	0.1	0.2

(i) Find p.

(ii) Let $g(X) = 3X + 1$ represent the amount of money in dollars, paid to the attendant by the manager. Find the expected earnings of the attendant for this particular time period. Find, also, the standard deviation of the attendant's earning for this particular time period.

(iii) Given that for this particular time period, the company makes a net profit of $X^2 - 2X$. Find the expected net profit and the standard deviation of the net profit of the company in this particular time period.

Solution:

(i) $0.1+0.2+0.3+p+0.1+0.2=1$

$\Rightarrow p=0.1$

(ii) $E(X)=4(0.1)+5(0.2)+6(0.3)+7(0.1)+8(0.1)+9(0.2)$

$=6.5$

$E(g(X))=E(3X+1)$

$=3E(X)+1=20.5$

Thus the attendant's expected earning is \$20.50.

$E(X^2)=4^2(0.1)+5^2(0.2)+6^2(0.3)+7^2(0.1)+8^2(0.1)+9^2(0.2)$

$=44.9$

$Var(X)=E(X^2)-[E(X)]^2=44.9-(6.5)^2=2.65$

$Var(g(X))=Var(3X+1)$

$=3^2 Var(X)=23.85$

Thus the standard deviation of the attendant's earning is $\sqrt{23.85}=\$4.88$ (to 3 s.f).

(iii) Let M be the company's net profit for this time period. Hence $M=X^2-2X$.

$E(M)=E(X^2-2X)=E(X^2)-2E(X)$

$=44.9-2(6.5)=31.9$

$Var(M)=E(M^2)-[E(M)]^2$

$E(M^2)=E\left[\left(X^2-2X\right)^2\right]$

$=E(X^4)-4E(X^3)+4E(X^2)$

Since $E(X^3)=4^3(0.1)+5^3(0.2)+6^3(0.3)+7^3(0.1)+8^3(0.1)+9^3(0.2)=327.5$

$E(X^4)=4^4(0.1)+5^4(0.2)+6^4(0.3)+7^4(0.1)+8^4(0.1)+9^4(0.2)=2501.3$

Thus $E(M^2)=2501.3-4(327.5)+4(44.9)=1370.9$

$Var(M)=E(M^2)-[E(M)]^2=1370.9-(31.9)^2=353.29$

Thus the company expected earning and the standard deviation of the net profit in this particular period are \$31.90 and $\sqrt{353.29} = \$18.80$ respectively.

Remark: We can also make use of the graphing calculator to compute the expectation and variance of functions of discrete random variable. This is useful for us to check our answers especially when the function of the discrete random variable may be a complicated one. Refer to **ANNEX 2** for the keystrokes.

Independent Observations of the Same Discrete Random Variable

Consider an experiment that is conducted n times under identical conditions (e.g. when a fair 6-sided die is repeatedly tossed). Let the random variable X denote the result when the experiment is conducted once. Then $X_1, X_2, ..., X_n$, are said to be independent random variables representing the result of the respective experiments (e.g. X_1, X_2, … is result of the first, second, … toss of the die respectively) and have the same probability distribution as X (i.e. identically distributed).

Example 15:
Rods of length 2 m or 3 m are selected at random with probabilities 0.4 and 0.6 respectively.

(i) Find the expectation and variance of the length of a rod.
(ii) Find the expectation and variance of twice the length of a rod selected at random.
(iii) Two rods are now selected at random. Find the expectation and variance of the sum of the lengths of the two rods.
(iv) Find the probability that the sum of the two lengths is equal to 5 m.

Solution:

(i) Let X be the length of a randomly chosen rod.
The probability distribution of X is as follows.

X	2	3
$P(X=x)$	0.4	0.6

Using G.C., $E(X) = 2.6$,
$Var(X) = 0.24$.

(ii) $E(2X) = 2(2.6) = 5.2$
$Var(2X) = 2^2 Var(X) = 4(0.24) = 0.96$

(iii) Let X_1 and X_2 be the length of the two chosen rods.
$E(X_1 + X_2) = E(X_1) + E(X_2) = 2E(X) = 5.2$
$Var(X_1 + X_2) = Var(X_1) + Var(X_2) = 2Var(X) = 0.48$

Important Note: $2X \neq X_1 + X_2$. $2X$ refers to twice the length of a single rod. $X_1 + X_2$ refers to sum of the lengths of two rods. The variances are different.

Thinkzone:

X_1 and X_2 are two **independent** observations of X.

(iv)	The required probability is $\text{P}(X_1 + X_2 = 5)$ $= \text{P}(X_1 = 2 \text{ and } X_2 = 3) + \text{P}(X_1 = 3 \text{ and } X_2 = 2)$ $= \text{P}(X_1 = 2)\text{P}(X_2 = 3) + \text{P}(X_1 = 3)\text{P}(X_2 = 2)$ $= (0.4)(0.6) + (0.6)(0.4)$ $= 0.48$	$\text{P}(X_1 = 2 \text{ and } X_2 = 3)$ $= \text{P}(X_1 = 2)\text{P}(X_2 = 3)$ as X_1 and X_2 are independent.

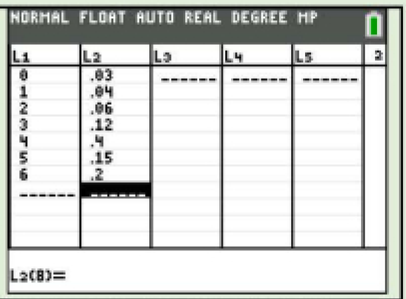
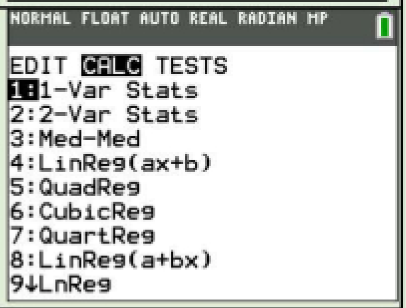
In general, if $X_1, X_2, ..., X_n$ are *n* **independent** observations of the random variable *X*, then

- $\text{E}(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n) = a_1\text{E}(X_1) \pm a_2\text{E}(X_2) \pm ... \pm a_n\text{E}(X_n)$
- $\text{Var}(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n) = a_1^2\text{Var}(X_1) + a_2^2\text{Var}(X_2) + ... + a_n^2\text{Var}(X_n)$

Self Review 3:
Austin removes the labels from 3 tins of tomato soup and 4 tins of peaches. He sends the labels off to the manufacturers in order to win himself a huggable teddy bear. Delighted with the prospect of the forthcoming bear, he forgot to mark the tins, which, devoid of their labels, then appear identical. The next week, Austin was hungry and wanted to have peaches. He chooses tins at random, opening each in turn until a tin of peach has been located. Let p_x be the probability that the x^{th} tin is the first tin containing peaches. Determine the values of p_1, p_2, p_3 and p_4. Determine the expectation and variance of the number of tins that are opened. [1.6, 0.64]

ANNEX 1: Use of Graphing Calculator to Compute Mean and Variance of Discrete Random Variables

Using Example 7 and Example 11 to illustrate how the graphing calculator can be used.

1.	Press [stat] and select **1:Edit.** Key in the values of *x* in list L_1 and the corresponding probabilities P(*X* = *x*) in list L_2.	
2.	Press [stat]. Under the **CALC** tab, select **1: 1-Var Stats.**	

3. At the List option, key in L_1.
At the FreqList option, key in L_2.

[Note: You can input L_1 by pressing [2nd] [1], L_2 by pressing [2nd] [2], …]

Select Calculate and press [enter].

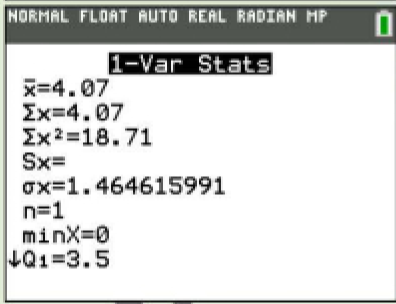
$\overline{x}$ refers to the mean i.e $\text{E}(X)$.

σ_x refers to the standard deviation i.e. $\sqrt{\text{Var}(X)}$.

In the event when probabilities are entered in L_2, Σx and Σx^2 give the value of $\text{E}(X)$ and $\text{E}(X^2)$ respectively. [Why? Recall that $\Sigma x = \Sigma x_i f_i$ and $\Sigma x^2 = \Sigma x_i^2 f_i$ where f_i represent the frequency of x_i. If L_2 are probabilities instead of frequencies, then $\Sigma x = \Sigma x_i \text{P}(X = x_i) = \text{E}(X)$ and $\Sigma x^2 = \Sigma x_i^2 \text{P}(X = x_i) = \text{E}(X^2)$.]

Hence $\text{E}(X) = 4.07$, $\text{E}(X^2) = 18.71$, $\text{Var}(X) = (1.4646)^2$

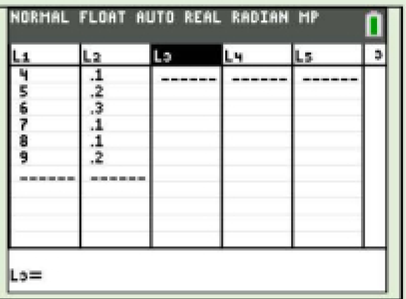
[Note: You can also use the stored value of σ_x to compute $\text{Var}(X) = \sigma_x^2$ by pressing [vars] and select **5:Statistics** followed by **4: σx** [x^2]]



ANNEX 2: Use of Graphing Calculator to Compute Mean and Variance of Functions of Discrete Random Variables

We will use Example 14(iii) to illustrate how the graphing calculator can be used to find $\text{E}\left(X^3\right)$ and $\text{E}\left(X^4\right)$.

1. Press [stat] and select **1:Edit.**
Key in the values of x in list L_1 and the corresponding probabilities $\text{P}(X = x)$ in list L_2.



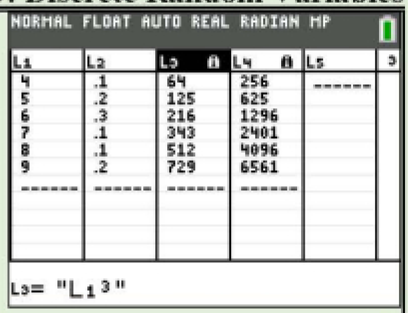
2. Navigate to the list L_3 header.

Press [alpha] [+] to key in " and then press [2nd] [1] [^] [3] to key in $L_1^{\,3}$. Press [enter] and the values for X^3 will be calculated at once under list L_3.

Navigate to the list L_4 header.

Press [alpha] [+] to key in " and then press [2nd] [1] [^] [4] to key in $L_1^{\,4}$. Press [enter] and the values for X^4 will be calculated at once under list L_4.

[Note: Keying in " before $L_1^{\,3}$ and $L_1^{\,4}$ retains their formulae. This is useful so that when you edit the values in L_1, the corresponding values of $L_1^{\,3}$ and $L_1^{\,4}$ will be changed immediately.]



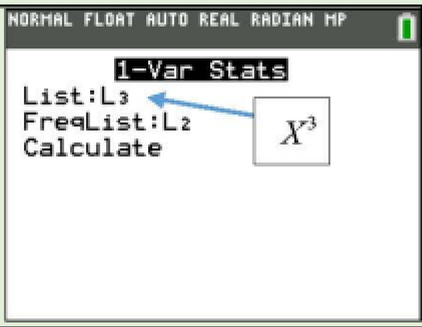
3. Press [stat]. Under the **CALC** tab, select **1: 1-Var Stats.**

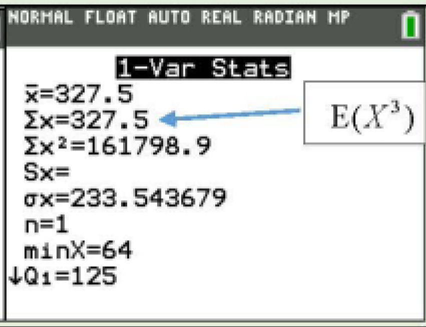
4. To compute $\mathrm{E}(X^3)$, at the List option, key in L_3. At the FreqList option, key in L_2.

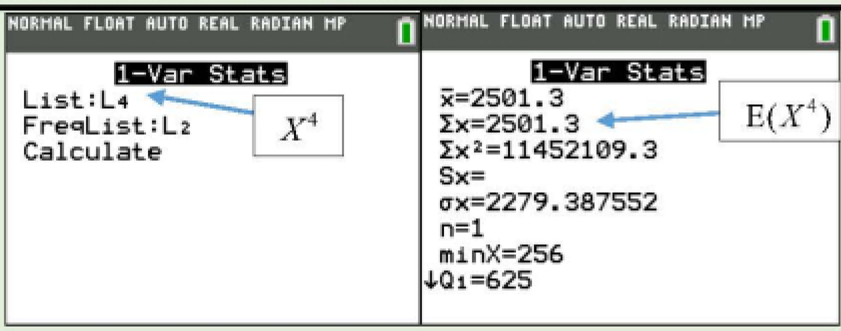
Select Calculate and press [enter].

To compute $\mathrm{E}(X^4)$, at the List option, key in L_4 and calculate.

[Note: You can input L_2 by pressing [2nd] [2], L_3 by pressing [2nd] [3] and L_4 by pressing [2nd] [4]]

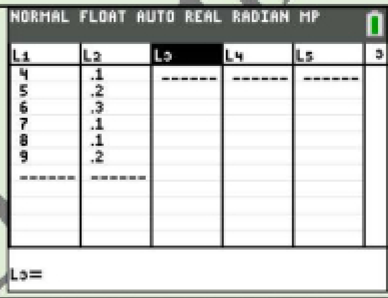




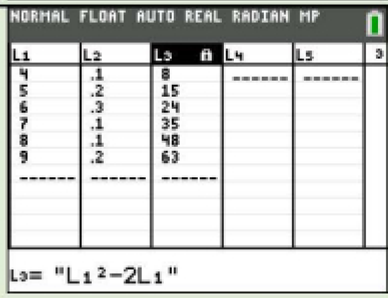


We may also compute the expectation and variance of $M = X^2 - 2X$ in Example 14(iii) directly with the graphing calculator to check the answer as follows (Note that you are to show clear workings although you can use the graphing calculator to compute):

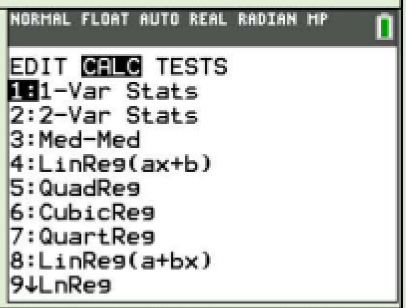
1. Press [stat] and select **1:Edit.**
Key in the values of x in list L_1 and the corresponding probabilities $\mathrm{P}(X = x)$ in list L_2.



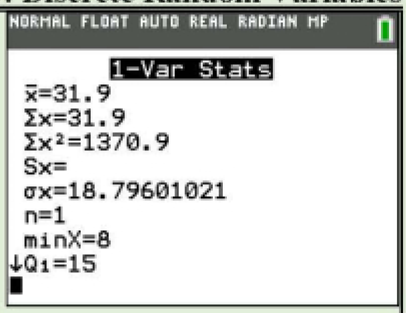
2. Navigate to the list L_3 header.
Press [alpha] [+] to key in " and then press [2nd] [1] [x²] [−] [2] [2nd] [1] to key in $L_1^2 - 2L_1$.
Press [enter] and the values for $X^2 - 2X$ will be calculated at once under list L_3.
[Note: Keying in " before $L_1^2 - 2L_1$ retains its formula. This is useful when you edit the values in L_1, the corresponding values in L_3 will be changed immediately.]



3. Press [stat]. Under the **CALC** tab, select **1: 1-Var Stats.**



4. At the List option, key in L_3.
At the FreqList option, key in L_2.
Select Calculate and press [enter].

5.	$\bar{x}$ refers to the mean i.e. $\mathrm{E}\left(X^2-2X\right)$. σ_x refers to the standard deviation i.e $\sqrt{\mathrm{Var}\left(X^2-2X\right)}$. When probabilities are entered in L_2, Σx and Σx^2 give the value of $\mathrm{E}\left(X^2-2X\right)$ and $\mathrm{E}\left(\left(X^2-2X\right)^2\right)$ respectively. Hence $\mathrm{E}(M)=\mathrm{E}\left(X^2-2X\right)=31.9$, $\mathrm{E}(M^2)=\mathrm{E}\left(\left(X^2-2X\right)^2\right)=1370.9$, $\mathrm{Var}(M)=\mathrm{Var}\left(X^2-2X\right)=(18.796)^2$	

Self-Review 1 Solution:

x	5	6	7
$\mathrm{P}(X=x)$	$\left(\frac{1}{20}\right)\left(\frac{1}{20}\right)=\frac{1}{400}$	$\left(\frac{1}{20}\right)\left(\frac{19}{20}\right)\times 2=\frac{19}{200}$	$\left(\frac{19}{20}\right)\left(\frac{19}{20}\right)=\frac{361}{400}$

Self-Review 2 Solution:

(a) $\mathrm{E}(X)=(-1)\left(\frac{1}{5}\right)+(3)\left(\frac{2}{5}\right)+(8)\left(\frac{2}{5}\right)=\frac{21}{5}$

(b) $\mathrm{E}(X^2)=(-1)^2\left(\frac{1}{5}\right)+(3)^2\left(\frac{2}{5}\right)+(8)^2\left(\frac{2}{5}\right)=\frac{147}{5}$

(c) $\mathrm{E}(|X|)=\left(\frac{1}{5}\right)+3\left(\frac{2}{5}\right)+8\left(\frac{2}{5}\right)=\frac{23}{5}$

(d) $\mathrm{E}\left(\frac{1}{X}\right)=-\left(\frac{1}{5}\right)+\frac{1}{3}\left(\frac{2}{5}\right)+\frac{1}{8}\left(\frac{2}{5}\right)=-\frac{1}{60}$

Self-Review 3 Solution:

$p_1=\frac{4}{7}$; $p_2=\left(\frac{3}{7}\right)\left(\frac{4}{6}\right)=\frac{2}{7}$; $p_3=\left(\frac{3}{7}\right)\left(\frac{2}{6}\right)\left(\frac{4}{5}\right)=\frac{4}{35}$; $p_4=\left(\frac{3}{7}\right)\left(\frac{2}{6}\right)\left(\frac{1}{5}\right)\left(\frac{4}{4}\right)=\frac{1}{35}$

Using G.C., expectation and variance of number of tins are 1.6 and 0.64 respectively.