



Chapter S2A: Discrete Random Variables

SYLLABUS INCLUDES

- Concepts of discrete random variables, probability distributions, expectations and variances

PRE-REQUISITES

- Probability
- Descriptive Statistics (See **Appendix**)

CONTENT

- 1 Random Variables
 - 1.1 Definition
 - 1.2 Discrete versus Continuous Random Variables
 - 1.3 Probability Distribution of a Discrete Random Variable
 - 1.4 Expectation of a Discrete Random Variable
 - 1.5 Relationship between Expectation and Sample Mean
 - 1.6 Independent Random Variables
 - 1.7 Functions of a Discrete Random Variable
 - 1.8 Variance and Standard Deviation of a Discrete Random Variable
 - 1.9 Linear Combination of Independent Random Variables
 - 1.10 Miscellaneous examples

Appendix:

- A A Simple Application of Statistics in Insurance
- B Descriptive Statistics (Assumed knowledge)

INTRODUCTION

It often happens in probability that the events we are interested in involve counting or measuring something.

When we toss two dice, what are the possible outcomes we can measure?



- Number of '1's obtained
- Number of even outcomes
- Sum of two dice, etc.

For instance, when tossing a pair of dice, we are often interested in the sum of the two dice and not overly concerned over the value of each individual die. That is, we may be only interested in knowing that the sum is 7 and not whether the actual outcome is $\{1,6\}$, $\{2,5\}$, $\{3,4\}$, $\{4,3\}$, $\{5,2\}$ or $\{6,1\}$.

Also in coin-flipping, we may be only interested in the number of heads obtained and not be too concerned about the actual head-tail sequences that result.

These quantities of interests, or more formally, these real value functions defined on the sample space, are known as **random variables**.

1 RANDOM VARIABLE

1.1 Definition

A **random variable** is a quantity that takes different numerical values according to the result of a random experiment whose outcome cannot be predicted exactly.

For example, consider the experiment of tossing a fair coin 3 times and observe the outcome. The sample space Ω may be represented by

$$\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}.$$

Let's denote the number of heads obtained when tossing the fair coin 3 times by X .



- The possible values X can take are 0, 1, 2 and 3 (**variable**).
- The value it assumes is subject to chance (**random**), so X is a random variable.
- In general, $P(X = x)$ refers to the probability of the random variable X assuming a specific value x .
- Eg. $P(X = 2)$ refers to the probability of obtaining 2 heads in the 3 tosses.



Important!

We use capital letters to denote random variables, eg. X , Y , and lower case letters for the values they take, eg. x and y .
Avoid using Z , B , P and N .

1.2 Discrete versus Continuous Random Variables

A **discrete random variable** takes on a *countable*¹ number of possible values while a **continuous random variable** can take any value in a given range.

Which of the following are discrete random variables and which are continuous?

- The number of phone calls a person picks up in a minute
- The life span of a light-bulb manufactured by a factory
- The number of students who are late for school on a particular day
- The distance travelled by a car in a week
- The height of a student from the Class of 2020

¹ A countable set is one whose elements are in 1-1 correspondence with the positive integers.

In general, random variables that represent the number of occurrences of an event or objects are most suited for discrete random variables while those that measure spatial dimensions, time, temperature, weight etc, are best described by continuous random variables.

Very often, random variables are associated with their **probability distributions** (for discrete random variables) or probability density functions (for continuous random variables), as well as **expectation** and **variance**. In this chapter, we will focus on discrete random variables and their associated probability distributions, expectation and variance.

1.3 Probability Distribution of a Discrete Random Variable

For a random experiment,

Let S be the sample space of the experiment (e.g. outcomes of tossing 3 coins).

Let X be the random variable defined to take values of the experiments (e.g. number of heads). A table or formula giving the values of $P(X = x)$ for every x in S is called the **probability distribution** of X .

For the experiment of tossing a fair coin 3 times where X is the number of heads obtained, the **probability distribution** of X is as follows

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

Observe that the probability distribution of X satisfy the following:

- $0 \leq P(X = x) \leq 1$ for all x in S .
- $\sum_{x \in S} P(X = x) = 1$ where the summation is over all values of x in S .

1.4 Expectation of a Discrete Random Variable

The **expectation** (or **mean**, or **expected value**) of a discrete random variable X is the weighted average of the possible values that X can take on, each value being weighted by its corresponding probability.

It is denoted by $E(X)$ or μ and is a measure of the ‘centre’ of the probability distribution of X .

Formally, if X is a discrete random variable taking values from a set S , then

$$E(X) = \sum_{x \in S} xP(X = x).$$

Example 1

Let X be the random variable representing the sum of scores obtained when 2 fair dice are thrown. State the smallest set S that contains all the possible values of X . Tabulate the probability distribution of X . Find $E(X)$.

Solution:

$$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

$d_1+d_2 \backslash d_1$ d_2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Since each of the 36 possibilities are equally likely to occur, they each carry a probability of $\frac{1}{36}$.

Therefore, the probability distribution of X is

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36} = \frac{1}{18}$	$\frac{3}{36} = \frac{1}{12}$	$\frac{4}{36} = \frac{1}{9}$	$\frac{5}{36}$	$\frac{6}{36} = \frac{1}{6}$	$\frac{5}{36}$	$\frac{4}{36} = \frac{1}{9}$	$\frac{3}{36} = \frac{1}{12}$	$\frac{2}{36} = \frac{1}{18}$	$\frac{1}{36}$

$$\text{Thus, } E(X) = \sum_{x=2}^{12} x P(X = x) = 7$$

Example 2

The probability distribution of a discrete random variable is given by

$$P(Y = y) = \begin{cases} \frac{1}{10}, & \text{if } y = 0, \\ cy^2, & \text{if } y = 1, 2, 3, 4. \end{cases}$$

Given that c is a constant, find the value of c . With this value of c , find $E(Y)$.

Solution:

As
$$\sum_{y=0}^4 P(Y = y) = 1$$

$$\Rightarrow \frac{1}{10} + c + 4c + 9c + 16c = 1$$

$$\Rightarrow 30c = \frac{9}{10}$$

$$\Rightarrow c = \frac{9}{300}$$

$$\Rightarrow c = \frac{3}{100}$$

$$E(Y) = \sum_{y=0}^4 yP(Y = y)$$

$$= 0\left(\frac{1}{10}\right) + \sum_{y=1}^4 \left(\frac{3}{100}y^3\right)$$

$$= 3$$

Recall GC usage



from Summation.

1.5 Relationship between Expectation and Sample Mean

Consider the number of heads obtained when we toss 3 fair coins. Denoting the number of heads obtained by X , the possible values X can take are 0, 1, 2 or 3. By working out the values of $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ and $P(X = 3)$, the probability distribution of X is presented in the following table.

x	0	1	2	3
$P(X = x)$	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ $= 0.125$	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times 3 = \frac{3}{8}$ $= 0.375$	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times 3 = \frac{3}{8}$ $= 0.375$	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ $= 0.125$

Based on the probability distribution,

$$E(X) = \sum_{x=0}^3 xP(X = x) = 0(0.125) + 1(0.375) + 2(0.375) + 3(0.125) = 1.5$$

However, if we toss the 3 coins 50 times and observe the number of heads obtained in each of the tosses, the data obtained can be represented in a table, known as a *frequency distribution*. From the experiment, we *may* end up with a data set as shown, which also includes the relative frequency distribution.

X	0	1	2	3
Frequency, f	5	20	18	7
Relative frequency	$\frac{5}{50} = 0.1$	$\frac{20}{50} = 0.4$	$\frac{18}{50} = 0.36$	$\frac{7}{50} = 0.14$

Based on the above sample, mean number of heads (or sample mean) obtained in each toss is

$$\frac{\sum fx}{\sum f} = \frac{0(5) + 1(20) + 2(18) + 3(7)}{50} = 0\left(\frac{5}{50}\right) + 1\left(\frac{20}{50}\right) + 2\left(\frac{18}{50}\right) + 3\left(\frac{7}{50}\right) = 1.54$$

If we increase the number of tosses to n times where n is large, we would expect the result for the relative frequency distribution to be close to the result for the probability distribution. (Refer to Appendix A in lecture notes of Chapter S1B: Probability, the empirical approach suggests that

$$\frac{f}{\sum f} = \frac{f}{n} \text{ tends to the actual probability of the event as } n \rightarrow \infty)$$

Then the expected value of X can be well approximated by the sample mean since the result for the relative frequency distribution would be close to the result for the probability distribution. We can also observe that the definition of the expected value (or mean) for a discrete random variable using a probability distribution is analogous to how we compute the sample mean for a set of data represented by a frequency distribution.

From the above, we understand intuitively why the *expected value* of a discrete random variable is the mean i.e. mean of X would be given by $E(X) = \sum x P(X = x)$.

1.6 Independent Random Variables

We have discussed what it means for two events to be independent. The concept of independent random variables is actually more complicated and requires the notion of joint probability distributions. However, there is a need for you to have at least an intuitive idea of the concept, as it will help simplify calculations.

Let X and Y be two discrete random variables taking on possible values $x_1, x_2, \dots$ and $y_1, y_2, \dots$ respectively. The random variables X and Y are said to be **independent** if for all i and j ,

$$P(X = x_i \text{ and } Y = y_j) = P(X = x_i)P(Y = y_j).$$

You should see the similarity of this definition with that of independence of events introduced earlier.

Example 3

A computer generates a random variable X whose probability distribution is given in the following table:

x	0	2	4
$P(X = x)$	0.4	0.3	0.3

Find $E(X)$.

Two independent observations of X are denoted by X_1 and X_2 . Let $Y = X_1 + X_2$. Show that $P(Y = 6) = 0.18$.

Tabulate the probability distribution of Y and find $E(Y)$. Comment on the relationship between the values of $E(Y)$ and $E(X)$.

Solution:

$$E(X) = \sum x P(X = x) = 0 \times 0.4 + 2 \times 0.3 + 4 \times 0.3 = 1.8$$

$$Y = X_1 + X_2$$

$$\begin{aligned} P(Y = 6) &= P(X_1 + X_2 = 6) \\ &= P(X_1 = 2 \text{ \& } X_2 = 4, X_1 = 4 \text{ \& } X_2 = 2) \\ &= 2(0.3 \times 0.3) \\ &= 0.18 \text{ (shown)} \end{aligned}$$

How do we obtain these values?

Y can take the values $\{0, 2, 4, 6, 8\}$.

$$P(Y = 0) = P(X_1 = 0 \text{ \& } X_2 = 0) = 0.4 \times 0.4 = 0.16$$

$$\begin{aligned} P(Y = 2) &= P(X_1 = 2 \text{ \& } X_2 = 0, X_1 = 0 \text{ \& } X_2 = 2) \\ &= (0.3 \times 0.4) + (0.4 \times 0.3) = 0.24 \end{aligned}$$

$$\begin{aligned} P(Y = 4) &= P(X_1 = 4 \text{ \& } X_2 = 0, X_1 = 2 \text{ \& } X_2 = 2, X_1 = 0 \text{ \& } X_2 = 4) \\ &= 2(0.3 \times 0.4) + (0.3 \times 0.3) = 0.33 \end{aligned}$$

$$P(Y = 8) = P(X_1 = 4 \text{ \& } X_2 = 4) = 0.3 \times 0.3 = 0.09$$

y	0	2	4	6	8
$P(Y = y)$	0.16	0.24	0.33	0.18	0.09

$$E(Y) = \sum y P(Y = y) = 0 + 0.48 + 1.32 + 1.08 + 0.72 = 3.6$$

It is noted that $E(Y) = E(X_1 + X_2) = 2E(X)$

Is this a general result?

1.7 Functions of a Discrete Random Variable

A function of a random variable is itself a random variable. Thus if X is a random variable, then functions such as $\cos X$, X^2 , $2X + 4$, $(X - E(X))^2$ are also random variables. If, in addition, X is a *discrete* random variable, then every function of X is a *discrete* random variable.

The expectation of $g(X)$, where g is a function of X , can be obtained as follows:

$$E(g(X)) = \sum_{x \in S} g(x) P(X = x).$$

In particular,

$$E(X^2) = \sum_{x \in S} x^2 P(X = x).$$

From example 3, the probability distribution of the random variable X is given in the following table:

x	0	2	4
$P(X = x)$	0.4	0.3	0.3

To calculate $E(X^2)$, we can add a row showing the values of X^2 .

x	0	2	4
x^2	0	4	16
$P(X = x)$	0.4	0.3	0.3

$$E(X^2) = \sum x^2 P(X = x) = 0 \times 0.4 + 4 \times 0.3 + 16 \times 0.3 = 6$$

To calculate $E(2X + 1)$, we can add a row showing the values of $2X + 1$.

x	0	2	4
$2x + 1$	1	5	9
$P(X = x)$	0.4	0.3	0.3

$$E(2X + 1) = 1 \times 0.4 + 5 \times 0.3 + 9 \times 0.3 = 4.6$$

In general we have the following properties:

Let X and Y be random variables and a and b be constants.
We have

- (1) $E(a) = a$
- (2) $E(aX) = a E(X)$
- (3) $E(aX \pm b) = aE(X) \pm b$

If X and Y are random variables, then

- (4) $E(aX \pm bY) = aE(X) \pm bE(Y)$

Proof for (1): $E(a) = \sum a P(X = x)$

$$= a \sum P(X = x)$$

$$= a(1)$$

$$= a$$

Proof for (3): $E(aX + b) = \sum (ax + b) P(X = x)$

$$= a \sum x P(X = x) + b \sum P(X = x)$$

$$= a E(X) + b$$

1.8 Variance and Standard Deviation of a Discrete Random Variable

The **variance** of a discrete random variable X is a measure of the *expected spread of a random variable about its mean*. It is defined as the expectation of $(X - \mu)^2$, where $\mu = E(X)$ and is denoted by **Var**(X) or σ^2 .

In the case of a discrete random variable, the variance is given by

$$\begin{aligned} \text{Var}(X) &= E(X - \mu)^2 \\ &= \sum (x - \mu)^2 P(X = x) \end{aligned}$$

The **standard deviation** of X , denoted by σ , is defined as $\sqrt{\text{Var}(X)}$.

Although the variance of X , by definition, is given by $E[(X - \mu)^2]$, for computational purposes, we often use the following equivalent equation,

$$\begin{aligned} \text{Var}(X) &= E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

Proof:

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

In a similar way, one would be able to see that the definition of the variance for a discrete random variable using a probability distribution is analogous to how we compute the variance for data represented by a frequency distribution.

For a discrete random variable, $\text{Var}(X) = E(X^2) - \mu^2 = \sum x^2 P(X=x) - \mu^2$

while variance is $\frac{1}{\sum f} \left(\sum fx^2 \right) - (\bar{x})^2$ for a set of observed data.

Consider the probability distributions of the random variables X and Y .

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

y	1	2	3	4	5
$P(Y=y)$	$\frac{1}{50}$	$\frac{4}{50} = \frac{2}{25}$	$\frac{40}{50} = \frac{4}{5}$	$\frac{4}{50} = \frac{2}{25}$	$\frac{1}{50}$

By symmetry, $E(X) = E(Y) = 3$

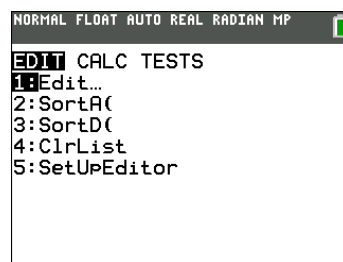
$\text{Var}(X)$ $= E(X^2) - [E(X)]^2$ $= \left(1 \times \frac{1}{5} + 4 \times \frac{1}{5} + 9 \times \frac{1}{5} + 16 \times \frac{1}{5} + 25 \times \frac{1}{5} \right) - 3^2$ $= 11 - 9 = 2$	$\text{Var}(Y)$ $= E(Y^2) - [E(Y)]^2$ $= \left(1 \times \frac{1}{50} + 4 \times \frac{4}{50} + 9 \times \frac{40}{50} + 16 \times \frac{4}{50} + 25 \times \frac{1}{50} \right) - 3^2$ $= 9.32 - 9 = 0.32$
---	--

Note:

Even though X and Y have the same expectations, their variances are different. **Var(X)** is large as the observations of X are widely scattered from the mean. **Var(Y)** is small as observations of Y tend to cluster close to the mean.

We can use the GC to help check our answers:

- Press **stat** and select **1:Edit...**
Enter the data into **L₁**, **L₂**, **L₃**.



2. Press stat and scroll to the right to select. CALC . Select 1: 1-Var Stats .	
3. Enter L1 under List . Enter L2 (for X) under FreqList . Scroll down to select Calculate .	
4. Press enter to obtain the following screen. $E(X) = \bar{x} = 3, E(X^2) = 11$ $Var(X) = E(X^2) - [E(X)]^2$ $= 11 - 3^2 = 2 = (\sigma_x)^2$	
5. Enter L3 (for Y) under FreqList . Scroll down to select Calculate .	
6. Press enter to obtain the following screen. $E(Y) = \bar{y} = 3, E(Y^2) = 9.32$ $Var(Y) = E(Y^2) - [E(Y)]^2$ $= 9.32 - 3^2$ $= 0.32 = (\sigma_y)^2$	

We have the following properties:

Let X and Y be random variables and a and b be constants.

We have

$$(1) \quad \text{Var}(a) = 0$$

$$(2) \quad \text{Var}(aX) = a^2 \text{Var}(X)$$

$$(3) \quad \text{Var}(aX \pm b) = a^2 \text{Var}(X)$$

If X and Y are *independent* random variables, then

$$(4) \quad \text{Var}(aX \pm bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y)$$

Note: The variance of a constant is zero. This is to be expected as a constant has no variation!

1.9 Linear Combination of Independent Random Variables

Given a random variable X , the probability distribution of $2X$ and $X_1 + X_2$, where X_1 and X_2 are two independent observations of X , are totally different. This is illustrated using the example below, and we will return to this important concept again in Chapter 3.

Example 4

When a tetrahedral die is thrown, the number on the face on which it lands, X , has the following probability distribution

x	1	2	3	4
$P(X=x)$	0.25	0.25	0.25	0.25

It is given that $E(X) = 2.5$ and $\text{Var}(X) = 1.25$.

- Tabulate the probability distribution of D , the random variable representing ‘double the number on which the die lands’. Find $E(D)$ and $\text{Var}(D)$.
- Tabulate the probability distribution function of S , the random variable representing ‘the sum of the two numbers when the die is thrown twice independently’. Find $E(S)$ and $\text{Var}(S)$.

Comment on the relationship between the values of

- $E(D)$ and $E(S)$;
- $\text{Var}(D)$ and $\text{Var}(S)$.

Solution:

(i) $D = 2X$

d	2	4	6	8
$P(D = d)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

$$E(D) = \frac{1}{4}(2 + 4 + 6 + 8) = 5 \text{ [or by symmetry]}$$

$$E(D^2) = \frac{1}{4}(4 + 16 + 36 + 64) = 30$$

$$\text{Var}(D) = E(D^2) - [E(D)]^2 = 30 - 5^2 = 5$$

(ii) $S = X_1 + X_2$

$x_1 + x_2$ $x_1 \backslash x_2$	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

Probability distribution of S

s	2	3	4	5	6	7	8
$P(S = s)$	$\frac{1}{16}$	$\frac{2}{16} = \frac{1}{8}$	$\frac{3}{16}$	$\frac{4}{16} = \frac{1}{4}$	$\frac{3}{16}$	$\frac{2}{16} = \frac{1}{8}$	$\frac{1}{16}$

$$E(S) = \frac{1}{16}(2 + 3(2) + 4(3) + 5(4) + 6(3) + 7(2) + 8) = 5 \text{ [or by symmetry]}$$

$$E(S^2) = \frac{1}{16}(2^2 + 3^2(2) + 4^2(3) + 5^2(4) + 6^2(3) + 7^2(2) + 8^2) = 27.5$$

$$\text{Var}(S) = E(S^2) - [E(S)]^2 = 27.5 - 5^2 = 2.5$$

(a) The means of the two distributions, D and S are the same, i.e. $E(D) = E(S)$.(b) The variances are **not** the same, with the random variable D having the greater variance.

Important Results:

Notice that $E(D) = E(2X) = 2E(X)$
 $E(S) = E(X_1 + X_2) = E(X_1) + E(X_2) = 2E(X)$

but $\text{Var}(D) = \text{Var}(2X) = 2^2 \text{Var}(X) = 4\text{Var}(X)$
 $\text{Var}(S) = \text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) = 2\text{Var}(X)$

Question:

Is $2X$ the same as $X_1 + X_2$? Why?

$2X$ is twice the observation of X but $X_1 + X_2$ is the sum of two observations of X .

Examples:

- Two identical jars, Jar 1 and Jar 2, each with 5 black marbles and 4 white marbles in it.
 X_k : the number of black marbles drawn from 3 draws, without replacement, from Jar k
 $X_1 + X_2$: the total number of black marbles drawn from Jars 1 and 2.

X : the number of black marbles drawn from 3 draws, without replacement, from a jar.
 $2X$: twice the number of black marbles drawn from a jar.

- $2X$ denotes ‘twice the number of students present in school on a particular day’ while $X_1 + X_2$ denotes the sum of the number of students present on two different days.

In general,

$$E(nX) = nE(X)$$

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = nE(X)$$

but $\text{Var}(nX) = n^2 \text{Var}(X)$
 $\text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = n \text{Var}(X)$

1.10 MISCELLANEOUS EXAMPLES

Example 5 MJCPrelim/9758/2018/02/Q6

A bag contains five balls numbered 1 to 5. A game consists of a player taking two balls from the bag consecutively, with replacement. The number on each ball is noted down. His score X is found by taking the absolute value of the difference between the two numbers.

- (i) Obtain the probability distribution of X .
- (ii) Find $E(X)$ and $\text{Var}(X)$.

A player pays \$ a to play this game. He wins an amount (in dollars) corresponding to twice his score. Determine the range of values of a if the player is expected to make a loss.

Solution:

- (i) Let X denote the absolute value of the difference between the two numbers.

X	1	2	3	4	5
1	0	1	2	3	4
2	1	0	1	2	3
3	2	1	0	1	2
4	3	2	1	0	1
5	4	3	2	1	0

The probability distribution of X is as follows:

x	0	1	2	3	4
$P(X = x)$	$\frac{5}{25} = \frac{1}{5}$	$\frac{8}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{2}{25}$

- (ii) Using GC, $E(X) = 1.6$, $\text{Var}(X) = 1.44$

Amount that he wins = $2X$

If the player is expected to make a loss,

$$E(2X - a) < 0$$

$$2(1.6) - a < 0$$

$$\therefore a > 3.2$$

Example 6 TPJCPrelim/9758/2018/02/Q5

An unbiased die has three faces painted red, two faces painted green and one face painted blue. A red face has a score of 1 point, a green face has a score of 2 points and a blue face has a score of 3 points.

Two such dice are thrown and the sum of their scores is denoted by X .

(i) Show that $P(X = 4) = \frac{5}{18}$ and find the probability distribution of X .

(ii) Find $E(X)$ and show that $\text{Var}(X) = \frac{10}{9}$.

Suppose now a red face has a score of 3 points, a green face has a score of 2 points and a blue face has a score of 1 point.

(iii) Deduce the expectation and variance of the sum of scores obtained from a throw of two such dice.

Solution:

(i)

$$P(X = 2) = P(\text{red}) \times P(\text{red}) = \frac{3}{6} \times \frac{3}{6} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$$P(X = 3) = P(\text{red}) \times P(\text{green}) + P(\text{green}) \times P(\text{red})$$

$$= \frac{3}{6} \times \frac{2}{6} + \frac{2}{6} \times \frac{3}{6} = \frac{1}{2} \times \frac{1}{3} \times 2 = \frac{1}{3}$$

$$P(X = 4) = P(\text{red}) \times P(\text{blue}) + P(\text{blue}) \times P(\text{red}) + P(\text{green}) \times P(\text{green})$$

$$= \frac{3}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{3}{6} + \frac{2}{6} \times \frac{2}{6} = \frac{1}{2} \times \frac{1}{6} \times 2 + \frac{1}{3} \times \frac{1}{3} = \frac{5}{18}$$

$$P(X = 5) = P(\text{green}) \times P(\text{blue}) + P(\text{blue}) \times P(\text{green})$$

$$= \frac{2}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{2}{6} = \frac{1}{3} \times \frac{1}{6} \times 2 = \frac{1}{9}$$

$$P(X = 6) = P(\text{blue}) \times P(\text{blue})$$

$$= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

Probability distribution of X :

x	2	3	4	5	6
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{5}{18}$	$\frac{1}{9}$	$\frac{1}{36}$

(ii)

$$\begin{aligned}E(X) &= \sum_{x=2}^6 x \cdot P(X=x) \\&= 2 \times \frac{1}{4} + 3 \times \frac{1}{3} + 4 \times \frac{5}{18} + 5 \times \frac{1}{9} + 6 \times \frac{1}{36} \\&= \frac{10}{3}\end{aligned}$$

$$\begin{aligned}E(X^2) &= \sum_{x=2}^6 x^2 \cdot P(X=x) \\&= 2^2 \times \frac{1}{4} + 3^2 \times \frac{1}{3} + 4^2 \times \frac{5}{18} + 5^2 \times \frac{1}{9} + 6^2 \times \frac{1}{36} \\&= \frac{110}{9}\end{aligned}$$

$$\begin{aligned}\text{Var}(X) &= E(X^2) - [E(X)]^2 \\&= \frac{110}{9} - \left(\frac{10}{3}\right)^2 \\&= \frac{10}{9}\end{aligned}$$

(iii)

Let Y be the sum of scores obtained from a throw of two dice where a red face has a score of 3 points, a green face has a score of 2 points and a blue face has a score of 1 point.

Then $Y = 8 - X$.

$$E(Y) = E(8 - X) = 8 - E(X) = 8 - \frac{10}{3} = \frac{14}{3}$$

$$\text{Var}(Y) = \text{Var}(8 - X) = \text{Var}(8) + \text{Var}(X) = \text{Var}(X) = \frac{10}{9}$$

Appendix A: A Simple Application of Statistics in Insurance

Let's say on average a house is worth \$600 000. A house owner could lose the house as the result of the action of fire, flood or accident. The chance of this happening is small, say $\frac{1}{500}$ in any one year.



An insurance company which decided to undertake the risk charges a premium to cover the losses if any. If the company set this premium at \$1,300 per annum and took over the risk of loss on 100,000 houses, then it would have an income of \$130 million. On average, 200 of those 100,000 houses insured would burn down, and the company would pay out \$120 million in compensation. The company is left with \$10 million for operating expenses and profit, while the individual house owners are freed from uncertainty.

The important feature from the point of view of the insurance company is the long-run average of the sums of money they must pay out. This is called the expectation of loss. In this illustration, it can be seen that the average loss is \$1,200 per house issued.

Expectation of loss	=	Size of loss × Probability of that loss
E (loss)	=	\$600 000 × $\frac{1}{500}$
	=	\$1200

Now, the house owner has exchanged the possibility of losing \$600 000 for the certainty of losing \$1,300. From his point of view,

Expectation of loss	=	Size of loss × Probability of that loss
E (loss)	=	\$1300 × 1
	=	\$1300

Before insuring against the possible loss, the expected loss was \$1,200, so the house owner has paid a little extra to remove the uncertainty. In a well-developed market, competition helps to keep the premiums down.

Let's review some concepts on Descriptive Statistics.

Appendix B:

1 MEASURES OF CENTRAL TENDENCY (AVERAGE)

We will discuss the 3 most popular ones, the mode, median and mean.

1.1 The **Mode** of a set of data is the value which occurs most frequently in the set of data.

1.2 The **Median** (also known as the 50th percentile) of a set of observations arranged in order of magnitude is

- (i) the middle value if the number of observations is odd; and
- (ii) the average of the 2 middle values if the number of observations is even.

It is the value below which half of the distribution falls.

- E.g.
- (i) The median of 7, 7, 2, 3, 4, 2, 7, 9 and 31 is 7.
 - (ii) The median of 36, 41, 27, 32, 29, 38, 39 and 43 is 37.

Note:

Other types of percentiles that are commonly computed are

1. The lower quartile Q_1 (also known as the 25th percentile), which is the value below which $\frac{1}{4}$ of the observations lie.
2. The upper quartile Q_3 (also known as the 75th percentile), which is the value below which $\frac{3}{4}$ of the observations lie.

1.3 Mean

(i) Ungrouped Data

Given a set of n observations $x_1, x_2, \dots, x_n$, the mean is

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum x_i.$$

(ii) Grouped Data

If the observations $x_1, x_2, \dots, x_n$ occur $f_1, f_2, \dots, f_n$ times respectively, the mean $\bar{x}$ is

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}.$$

Note:

1. There is no "correct" average as each has its uses.
2. In statistics, the most important average is the mean because of its sensitivity to small changes in data and ease of computation.
3. The mean is strongly affected by extreme values and in such cases the median is a more suitable average.

2 MEASURES OF DISPERSION / SPREAD

Consider the 3 sets of numbers:

- (a) 7, 7, 7, 7, 7
- (b) 4, 6, 6.5, 7.2, 11.3
- (c) -193, -46, 28, 69, 177

They all have a mean of 7 but the spread of each data set is different.

There is no variability in (a), while the numbers in (c) are obviously much more spread out than those in (b).

There are various ways of measuring the dispersion / spread of data.

2.1 Range = Largest observation – smallest observation.

2.2 Interquartile range = $Q_3 - Q_1$

2.3 Variance and Standard Deviation

(i) Ungrouped Data

Given a set of n observations $x_1, x_2, \dots, x_n$, the variance is

$$\frac{1}{n} \sum (x_i - \bar{x})^2 = \frac{1}{n} \sum x_i^2 - (\bar{x})^2.$$

(ii) Grouped Data

If the observations $x_1, x_2, \dots, x_n$ occur $f_1, f_2, \dots, f_n$ times respectively, the variance is

$$\frac{1}{n} \sum f_i (x_i - \bar{x})^2 = \frac{1}{n} \sum f_i x_i^2 - (\bar{x})^2.$$

The standard deviation is given by $\sqrt{\text{variance}}$.

Note:

Standard deviations are useful when comparing sets of data. In general,

- a lower standard deviation indicates that the observations tend to cluster close to the mean (smaller variability/dispersion/spread), and
- a higher standard deviation suggests that the observations are widely scattered from the mean (greater variability/dispersion/spread).

Example 1

Two machines A and B are used to pack biscuits. A random sample of ten packets was taken from each machine and the mass of each packet (to the nearest gram) was measured and noted. Find the mean and standard deviation of the masses for each of the two samples. Comment on the reliability of the two machines.

Machine A (x_A)	196, 198, 198, 199, 200, 200, 201, 201, 202, 205
Machine B (x_B)	192, 194, 195, 198, 200, 201, 203, 204, 206, 207

Solution:

For machine A,

$$\text{sample mean mass of a packet, } \bar{x}_A = \frac{1}{10} \sum x_A = \frac{2000}{10} = 200 \text{ g}$$

$$\begin{aligned} \text{Variance} &= \frac{1}{n} \sum x_A^2 - (\bar{x}_A)^2 \\ &= \frac{1}{10} (400056) - (200)^2 \\ &= 5.6 \text{ g}^2 \end{aligned}$$

$$\text{Standard deviation} = \sqrt{5.6} = 2.37 \text{ g (3 s.f.)}$$

For machine B,

sample mean mass of a packet,

$$\bar{x}_B = \frac{1}{10} \sum x_B = \frac{2000}{10} = 200 \text{ g}$$

$$\begin{aligned} \text{Variance} &= \frac{1}{n} \sum x_B^2 - (\bar{x}_B)^2 \\ &= \frac{1}{10} (400240) - (200)^2 \\ &= 24 \text{ g}^2 \end{aligned}$$

$$\text{Standard deviation} = \sqrt{24} = 4.90 \text{ g (3 s.f.)}$$

Machine A has less variability, indicating that it is more reliable than machine B.

Example 2

The frequency distribution of ages of a group of salesman under 30 is given as follows:

Age, x	21	22	23	24	25	26	27	28
Frequency, f	3	3	4	6	7	4	2	1

Calculate the mean and standard deviation.

Solution:

$$\text{Mean age, } \bar{x}, \text{ is } \frac{\sum fx}{\sum f} = \frac{726}{30} = 24.2 \text{ years}$$

$$\text{Variance, Var}(X) \text{ is } \frac{1}{\sum f} \left(\sum fx^2 \right) - (\bar{x})^2 = \frac{1}{30} (17668) - 24.2^2 \approx 3.2933$$

$$\therefore \text{Standard deviation} = \sqrt{\text{Var}(X)} = 1.81 \text{ (3sf)}$$