Department of Mathematical Sciences

Bachelor of Science

(Mathematical Sciences Education)

Year 2

Probability & Statistics

Random variables

By the end of this lesson you must be able to

- Understand what a discrete random variable is
- Determine whether a random variable is discrete given a context
- Represent the probability distribution of a discrete random variable using a table or probability mass function, or a graph
- Find unknown values given the table or function of a discrete random variable
- Determine if a function or table for a discrete random variable is valid or not
- Find a probability table or function given a context
- Calculate probabilities of events given a probability distribution

Random variables

- A random variable is a numerical description of the outcome of an experiment.
- A random variable is also known as stochastic variable.
- A random variable can be classified as being either discrete or continuous.

Discrete Random Variables

A random variable X is said to be discrete if it can assume only a finite or countably infinite number of distinct values.

For example, an experiment of cars arriving at a Chichiri shopping mall between 8:00am and 10:00am on Friday.

- The random variable of interest is $x =$ the number of cars arriving between 8:00am and 10:00am on Friday.
- The possible values of x are $0, 1, 2, \ldots$ So, x is a discrete random variable.

The Probability distribution of a discrete random variable

- The probability distribution describes how probabilities are distributed over the values of a random variable.
- For example, Let X denote any one of the six possible values that could be observed on the upper face when a die is tossed. After the die is tossed, the number actually will be denoted by the symbol x. Note X is the random variable but the specific observed value, x, is not random.
- For a discrete random variable, x, the probability distribution is defined by a probability mass function (p.m.f), denoted by $P(X=x)$ *or* $f(x)$. This function provides the probability for each value of the random variable.
- In the development of the probability function for a discrete random variable, two conditions must be satisfied:

  (1) $f(x)$ must be nonnegative for each value of the random variable, and

  (2) the sum of the probabilities for each value of the random variable must equal one.

Example 1

In an experiment, two fair coins are tossed once. Let X denote the number of heads observed Assuming that the two faces of a coin are named heads and tails, find the probability distribution of X.

Solution

Let X be the number of heads observed when two coins are tossed once.

The possible outcomes are TT, HT, TH, HH.

We can present this information in the Table that follows:

Outcome	**TT**	**HT**	**TH**	**HH**
Number of Heads	0	1	1	2
Probability	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

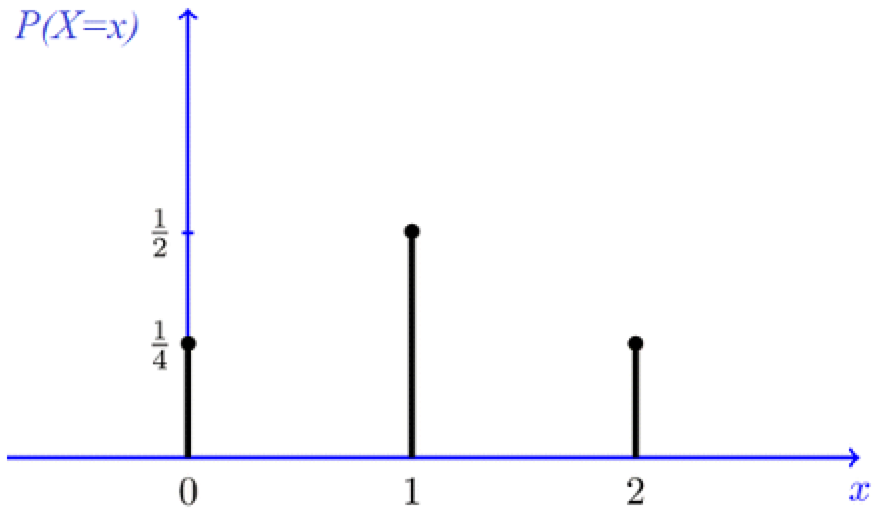
Example 1… cont'd

Note that in this experiment X can be 0, 1, or 2 since it counts the number of heads in an outcome.

We can represent a probability distribution using a **Table** and a **Graph** as follows:

X	0	1	2
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(a) Table



(b) Graph

Cumulative distribution

In a cumulative distribution, the probabilities of the outcomes of an experiment are summed up to a certain value to give a cumulative probability. It is a distribution containing cumulative probabilities.

Example

A probability distribution of a discrete random variable X is given in the table. Use it to develop a cumulative distribution.

x	1	2	3	4	5
$P(X=x)$	0.05	0.4	0.3	0.15	0.1

Cumulative distribution…cont'd

- For a discrete random variable X a cumulative distribution function is given by $F(x)$ where $F(x) = P(X \le x)$.

$$F(1) = P(X \le 1) = 0.05$$

$$F(2) = P(X \le 2) = P(X = 1) + P(X = 2) = 0.05 + 0.4 = 0.45$$

- The same pattern is followed to obtain

$$F(3) = P(X \le 3) = 0.05 + 0.4 + 0.3 = 0.75$$

$$F(4) = P(X \le 4) = 0.9$$

$$F(5) = P(X \le 5) = 1$$

Cumulative distribution…cont'd

- Therefore, the cumulative distribution is

x	1	2	3	4	5
$F(x)$	0.05	0.45	0.75	0.9	1

- In general, for ad discrete random variable X, the cumulative distribution function is given by $F(x)$ where $F(x) = P(X \leq x)$.

The expected value (mean) of a discrete random variable

- For a discrete random variable, the expected value, usually denoted as μ or $E(X)$, is calculated using:

$$\mu = E(X) = \sum xP(X = x) \quad \text{or} \quad \mu = E(X) = \sum xf(x)$$

since $f(x) = P(X = x)$.

- The formula means that we multiply each value, x, by its respective probability, $f(x)$, and then add them all together.
- You can observe that it simply is an average value but weighted by the likelihood (probability) of the value.

Example 2

A random variable X has probability distribution as shown. Find the value of h and the expected Value of X.

x	0	1	2
$f(x)$	0.25	h	0.25

Solution

We know that the sum of all probabilities in an experiment add up to 1.

Hence, $$0.25 + h + 0.25 = 1$$

$$h = 1 - 0.5 = 0.5$$

The Expected value: $$E(X) = \sum x f(x)$$

$$= (0 \times 0.25) + (1 \times 0.5) + (2 \times 0.25) = 1$$

Example 3

In a family with two children, find the mean number of children who will be girls.

Let the random variable X be the number of girls.

Solution

- In this family we could have no girls (if both children are boys); one girl and one boy; or both would be girls (no boy).
- Since X counts the number of girls in this family then it can assume the following values: 0, 1, or 2.
- The sample space would be $S = \{BB, BG, GB, GG\}$ *if* $B = boy$ *and* $G = girl$.

Example 3…cont'd

- The probability distribution for X is represented as

x	0	1	2
$f(x)$	0.25	0.5	0.25

$$E(X) = \sum xf(x) = (0 \times 0.25) + (1 \times 0.5) + (2 \times 0.25) = 1$$

The mean number of children who will be girls is 1.

Some properties of the expected value of a discrete random variable

- If X is a random variable and a is a constant, then

  ✓ $E(a) = a$

  ✓ $E(aX) = aE(X)$

- If a and b are constants then

  ✓ $E(aX + b) = aE(X) + b$

- If $X_1, X_2, X_3, \ldots, X_n$ are random variables then

  ✓ $E(X_1 + X_2 + X_3 + \ldots + X_n) = E(X_1) + E(X_2) + E(X_3) + \ldots E(X_n)$

The variance of a discrete random variable

- The variance of a discrete random variable is given by

$$\sigma^2 = Var(X) = \sum (x-\mu)^2 f(x)$$

- However, there is an easier form of this formula that we can use.

$$\sigma^2 = Var(X) = \sum x^2 f(x) - \mu^2$$

- The formula means that first, we sum the square of each value times its probability then subtract the square of the mean. We can also present the formula for variance as follows:

$$\sigma^2 = Var(X) = E(X^2) - \{E(X)\}^2 = E(X^2) - \mu^2$$

- The standard deviation of a random variable, X, is the square root of the variance.

$$\sigma = \sqrt{Var(X)} = \sqrt{\sigma^2}$$

Some properties of the variance of a discrete random variable

- If X is a random variable and a is a constant, then
  - ✓ $V(a) = 0$
  - ✓ $Var(aX) = a^2 Var(X)$
- If a and b are constants then
  - ✓ $Var(aX + b) = a^2 Var(X)$
- If $X_1, X_2, X_3, \ldots, X_n$ are random variables then
  - ✓ $Var(X_1 + X_2 + X_3 + \ldots + X_n) = Var(X_1) + Var(X_2) + Var(X_3) + \ldots + Var(X_n)$
- If X and Y random variables, then
  - ✓ $Var(X - Y) = Var(X) + Var(Y)$

Example 4

The random variable X has a probability distribution as shown in the table

x	1	2	3	4	5
$P(X=x)$	0.1	0.3	0.2	0.3	0.1

Calculate

(a) $\mathrm{Var}(X)$

(b) Standard deviation

Example 4…cont'd

(a) The variance is calculated by $Var(X) = E(X^2) + \{E(X)\}^2$

Now, we need to find $E(X^2)$

$$E(X^2) = (1^2 \times 0.1) + (2^2 \times 0.3) + (3^2 \times 0.2) + (4^2 \times 0.3) + (5^2 \times 0.1) = 10.4$$

Then let us find $E(X)$

$$E(X) \;\; = (1 \times 0.1) + (2 \times 0.3) + (3 \times 0.2) + (4 \times 0.3) + (5 \times 0.1) = 3$$

So, since we have calculated $E(X^2)$ and $E(X)$ then we can now find $Var(X)$ as follows:

$$Var(X) = E(X^2) + \{E(X)\}^2 \;\; = 10.4 - (3)^2 = 1.4$$

(b) The standard deviation is found by

$$\sigma = \sqrt{Var(X)} = \sqrt{1.4} = 1.18$$

Conclusion

- A **random variable** is a variable whose value is a numerical outcome of a random phenomenon.
  - ❖ A random variable is denoted with a capital letter
  - ❖ The probability distribution of a random variable X tells what the possible values of X are and how probabilities are assigned to those values
  - ❖ A random variable can be discrete or continuous
- A **discrete random variable** X has a countable number of possible values.

References

Bluman, A. G. (2012). *Elementary Statistics, A step by step Approach* (8th ed.). McGraw Hill Publications.

Crawshaw, J. (2001). *A concise course in A-level statistics: With worked examples* (4th ed.). Stanley Thornes

Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2007). *Mathematical statistics with Applications*. Duxbury