STAT211: Business Statistics

M4: Probability

L3: Introduction to Probability Distributions

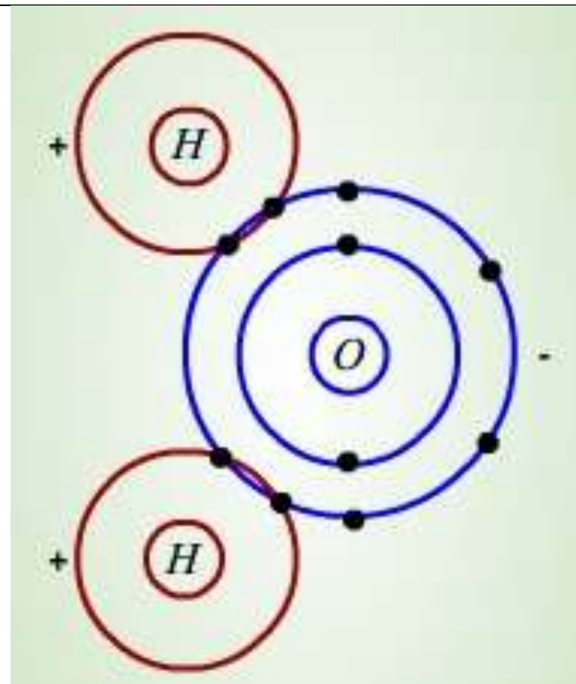
Learning Outcome

By the end of this lecture, you will be able to:

- Define a random variable
- Calculate the probability distribution of a random variable
- Compute the expected value and standard deviation for a discrete random variable or for a linear transformation of it

Introduction

Which of the two is a example of random experiment?

	
Tossing a coin	Mixing hydrogen and oxygen to create water

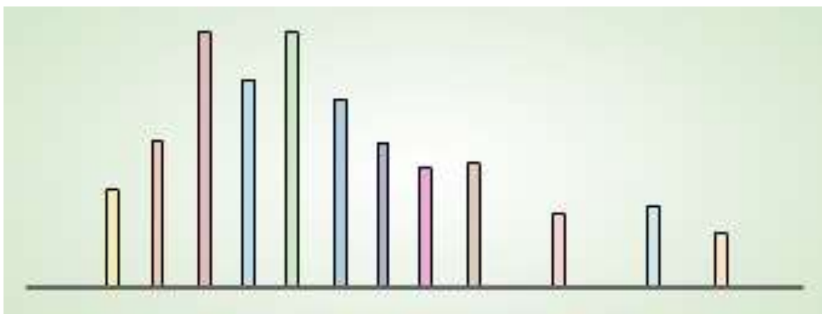
Random Variable

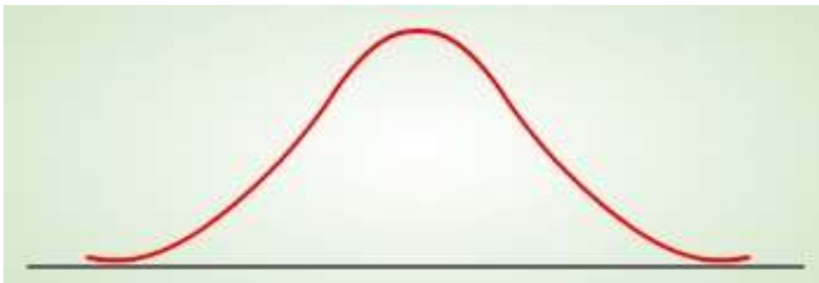
A random variable is real-valued function, defined on the sample space of a random experiment that assigns real number to each elementary event in the sample space. There are two types of random variables, discrete and continuous.

Dice

<table>
<tr>
<td>Discrete random variable</td>
<td>A discrete random variable assumes countable values.
Examples:

The number of absentees
The number of days to maturity of small projects, and so on


Discrete random variable</td>
</tr>
</table>

Continuous random variable	A continuous random variable assumes any real number in interval of real numbers, such as: • The annual revenue of a certain company • The percentage of government funds to a certain university  Continuous random variable

Random Variable > Example

Consider a random experiment to select a committee of 3 members from a group of 3 managers, 7 clerks, and 10 sales persons.

If X is a random variable that represents the number of clerks in the selected committee, then the possible values of X are $\{0, 1, 2, 3\}$.

Therefore, X is a discrete random variable.

Example

There are six questions on a multiple-choice quiz. Let Y represent the number of questions a student answers correctly.

1. Is Y a continuous or a discrete random variable?
2. What are the possible values of Y?

Solution:

Y is counting the number of correctly answered questions. Then,

1. Y is a discrete random variable.

2. The possible values for $Y = \{0, 1, 2, 3, 4, 5, 6\}$.

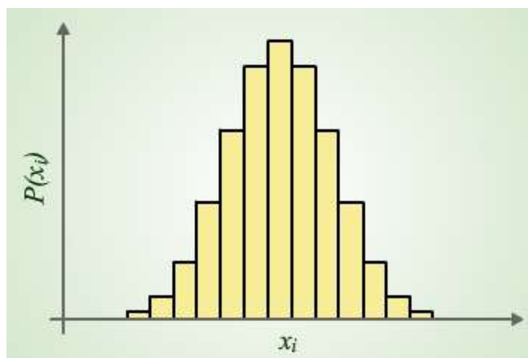
Discrete Probability Distribution

The probability distribution of a discrete random variable is a list of all possible values of the random variable along with their respective probabilities or a formula that calculates the probability of any possible value of the random variable.

Discrete probability distribution is a list of all possible $(x_i, P(x_i))$ pairs, where:

x_i = Value of the random variable X, i.e.an outcome

$P(x_i)$ = Probability associated with the value x_i



Note:

1. x_i's are mutually exclusive.
2. x_i's are collectively exhaustive.
3. $0 \le P(x_i) \le 1$ for all x_i.

$$\sum_{all\ i} P(x_i) = 1$$

Example

Consider rolling a die twice. Let X represents the number of times the digit 4 comes up. Write the sample space and the probability distribution of X.

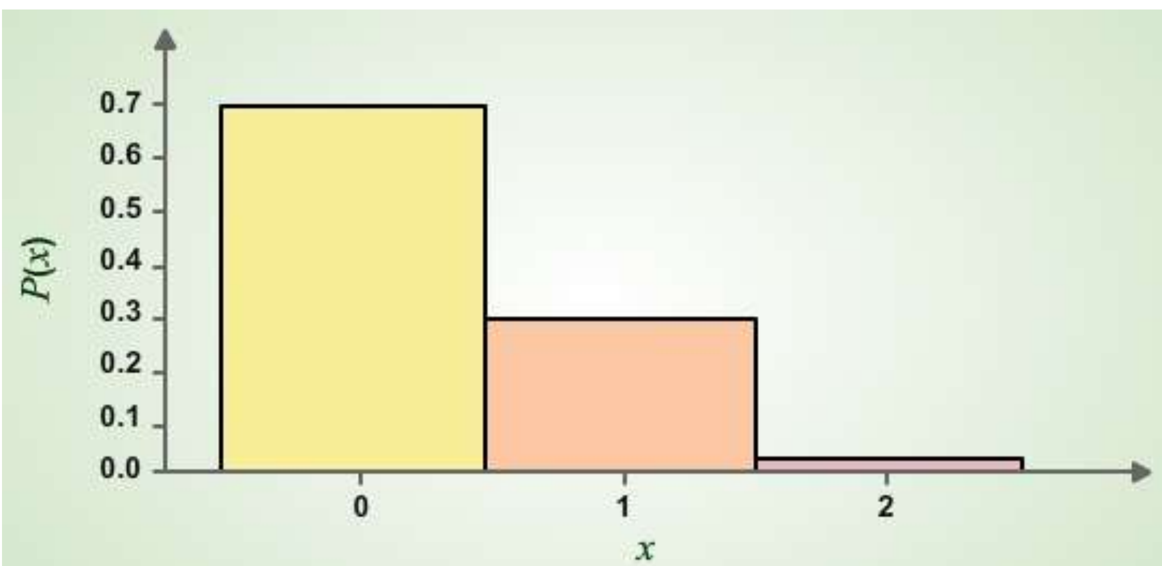
Solution:

$$S = \{(1,1), (1,2), \dots, (1,6), (2,1), (2,2), \dots, (2,6), \dots, (6,1), (6,2), \dots, (6,6)\}$$

X is a discrete random variable with the values $X = \{0, 1, 2\}$ and their corresponding probabilities as shown in the table.

X*= *x	**Elementary events**	**# E.E**	***P*(*x*)**
0	All the others	25	25/36 = 0.69
1	(1,4),(2,4),(3,4),(4,1),(5,4),(6,4), (4,2),(4,3),(4,5), (4,6)	10	10/36 = 0.28
2	(4,4)	1	1/36 = 0.028
Total		**36**	**36/36**

Moreover, the probability distribution can be displayed graphically in the probability histogram given.



Histogram of X

Example

The lab in your department contains 15 computers of which 3 are defective. Two of these computers are randomly chosen and inspected. Let U denote the number of defective computers. Find the probability distribution of U.

Lab

Solution:

Since U is counting the number of defective computers then,

1. U is a discrete random variable.
2. The possible values for $U = \{0, 1, 2\}$.
3. Let D be defective and N be non-defective, then the probability distribution of U is:

$U = u$	**Elementary events**	**# E.E**	$P(x)$
0	*NN*	1	(12/15)(11/14) = 22/35
1	*DN*, *ND*	2	2*(3/15)(12/14) = 12/35
2	*DD*	1	(3/15)(2/14) = 1/35
Total		**4**	**1**

Example

Suppose we are about to learn the sexes of the three children of a certain family. Let V denote the number of females in this family. Write the values of V and find the probability distribution of V.

Solution:

Since V is counting the number of female children then,

1. V is a discrete random variable.
2. The possible values for $V = \{0, 1, 2, 3\}$.
3. Let F be Female and M be Male, then the probability distribution of V is:

$V = v$	Elementary events	# E.E	$P(v)$
0	*MMM*	1	1 / 8 = 0.125
1	*FMM, MFM, MMF*	3	3 / 8 = 0.375
2	*FFM, FMF, MFF*	3	3 / 8 = 0.375
3	*FFF*	1	1 / 8 = 0.125
Total		**8**	**8 / 8 = 1**

Discrete Random Variable > Summary Measures

The summary measures of a discrete random variable can be:

Expected value of a discrete distribution	It is a weighted average. If X is a discrete random variable that takes the values $x_1, x_2, x_3, \ldots, x_n$, then the expected value of X is defined by: $$\mu_X = E(X) = \sum_{All\ i} x_i P(x_i)$$ where: $E(X)$ =Expected value of the random variable.

Variance of a discrete distribution	If X is a discrete random with expected value μ_x, then the variance of X is defined by: $$\sigma_X^2 = Var(X) = E(X-\mu_x)^2 = \sum [x-\mu_x]^2 P(x_i)$$ $$= \sum_{All\ x} x_i^2 P(x_i) - [\mu_x]^2 = \sigma_X^2 = E(X^2) - (E(X))^2$$ where: $E(X)$ = Expected value of the random variable. Note: that the standard deviation is given by: $SD(X) = \sqrt{Var(X)}$.

Properties of the Expected Value and the Variance

The properties of the expected value and the variance can be examined in two cases:

Case 1	For any random variable X with expected value $E(X)$, variance $Var(X)$, and any constant c, then 1. $E(cX) = cE(X)$ 2. $E(X + c) = E(X) + c$ 3. $Var(cX) = c^2 Var(X)$ 4. $Var(X + c) = Var(X)$
Case 2	If X and Y are two random variables, and a and b are two constants, then 1. $E(aX \pm bY) = aE(X) \pm bE(Y)$ 2. If X and Y are independent then $$Var(aX \pm bY) = a^2 Var(X) \pm b^2 Var(Y)$$ Note that: 1. If $X_1, X_2, X_3, \ldots, X_k$ are k randorn variables, then $$E\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} E(X_i)$$

	2. If $X_1, X_2, X_3, ..., X_k$ are independent random variables, then $$Var\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} Var(X_i)$$

Example

Consider a random variable X that takes on the values 1 or 0 with probability p and $1 - p$ respectively. Find $E(X)$ and $Var(X)$.

Solution:

The probability distribution for X can be shown in the table.

X	0	1	**Total**
***P*(*x*)**	$1 - p$	p	1

$$\mu_X = E(X) = \sum_{i=0}^{1} x_i P(x_i) = 1(p) + 0(1-p) = p$$

$$\sigma_X^2 = Var(X) = E(X^2) - E(X)^2 = \left[\sum_{i=0}^{1} x_i^2 P(x_i)\right] - [\mu_x]^2$$

$$= [0^2(1-p) + 1^2(p)] - p^2 = p - p^2 = p(1-p)$$

Example

Consider the given discrete probability distribution:

y	-2	1	3	6
***P* (*y*)**	0.13	0.12	0.15	0.6

1. Calculate the expected value of Y.
2. Calculate the variance of Y.
3. Let the random variable $W = Y + 7$, calculate the expected value and the variance of W.
4. Let the random variable $V = -7Y$, calculate the expected value and the variance of V.

Solution:

y	-2	1	3	6	**Sum**
$P(y)$	0.13	0.12	0.15	0.6	1
$yP(y)$	-0.26	0.12	0.45	3.6	3.91
$y^2P(y)$	0.52	0.12	1.35	21.6	23.59

1.
$$\mu_Y = E(Y) = \sum_{All\ i} y_i P(y_i)$$
$$= 3.91$$

2.
$$E(Y^2) = \sum_{All\ i} y_i^2 P(y_i)$$
$$= 23.59$$
$$\sigma_Y^2 = Var(Y) = E(Y^2) - E^2(Y)$$
$$= 23.59 - (3.91)^2$$
$$= 8.3019$$

3.
$$E(W) = E(Y) + 7$$
$$= 3.91 + 7$$
$$= 10.71$$
$$Var(W) = Var(Y) = 8.3019$$

4.
$$E(V) = -7E(Y)$$
$$= -7(3.91)$$
$$= -27.37$$
$$Var(V) = (-7)^2\, Var(Y)$$
$$= 49(8.3019)$$
$$= 406.7931$$

Recap

In this lecture, you have learned that:

- Random variable is a real-valued function defined on the sample space of a random experiment that assigns a real number to each elementary event in the sample space. The two types are:
    - Discrete
    - Continuous
- The probability distribution of a discrete random variable is a list of all possible values of the random variable along with their respective probabilities or a formula
- The expected value of a discrete random variable is:

$$\mu_x = E(X) = \sum_{All\ i} x_i P(x_i)$$

- The variance of a discrete random variable is:

$$\sigma_X^2 = Var(X) = E(X-\mu_x)^2 = \sum [x-\mu_x]^2 P(x_i)$$
$$= \sum_{All\ x} x_i^2 P(x_i) - [\mu_x]^2 = E(X^2) - (E(X))^2$$