1. Introduction

1.1 Review of Some Statistical Concepts

1.1.1 Probability and Random Variable

Probability is the possibility that any event will occur, given some specific sample space.

Let A be the event occurring in the given sample space and $P(A)$ be the probability that A will happen. Then, $P(A)$ is defined as;

$$P(A) = \frac{\text{the number of times the event A will occur}}{\text{the number of all possible outcomes in sample space}} \tag{1.1}$$

For instance, to draw one card from the standard 52-card deck, let A be the event that the rank of card is 2. Times the event will occur is 4 and the amount of all possible outcomes is 52; hence, the probability of A is $\frac{4}{52}$ or $\frac{1}{13}$.

Some properties of probability are;

1. $0 \leq P(A) \leq 1$

2. If A, B and C are exhaustive set, then,

$$P(A) + P(B) + P(C) = 1$$

1. Introduction

1.1 Review of Some Statistical Concepts

1.1.1 Probability and Random Variable

Probability is the possibility that any event will occur, given some specific sample space.

Let A be the event occurring in the given sample space and $P(A)$ be the probability that A will happen. Then, $P(A)$ is defined as;

$$P(A) = \frac{\text{the number of times the event A will occur}}{\text{the number of all possible outcomes in sample space}} \tag{1.1}$$

For instance, to draw one card from the standard 52-card deck, let A be the event that the rank of card is 2. Times the event will occur is 4 and the amount of all possible outcomes is 52; hence, the probability of A is $\frac{4}{52}$ or $\frac{1}{13}$.

Some properties of probability are;

1. $0 \leq P(A) \leq 1$

2. If A, B and C are exhaustive set, then,

$P(A) + P(B) + P(C) = 1$

3. If A, B and C are mutually exclusive, then,

$$P(A+B+C) = P(A) + P(B) + P(C)$$

Suppose that the results of an experiment are in the form of value, the variable, whose value is determined by one of those results, is known as **Random Variable**. Random variable can be either **discrete** or **continuous value**.

For discrete random variable, the example is the sum of the values on the face of two dice, when rolling two dice once. In other word, the obtained sum will range from 2 to 12, and it is impossible to get 2.5 or 3.5.

For continuous random variable, the example is the height of the high-school student, constricted to the range from 160 to 180 centimetres. It can be seen that the value of the height need not be the integers and can take the value of 160.5 or 160.52 centimetres.

These two distinct characteristics of random variable enable us to classify them into different probability density functions, which would be stated in section 1.4.

Random variables are usually denoted by the capital letter X, Y, Z, and so on, and the values taken by them are denoted by small letters x, y, z

1.2 Probability Density Function

As the value of random variable depends on an experiment, the **probability density function** would portray the overall image of possible random results. The type of the probability density function relies on the characteristics of the random variable. In this section, many important types are discussed.

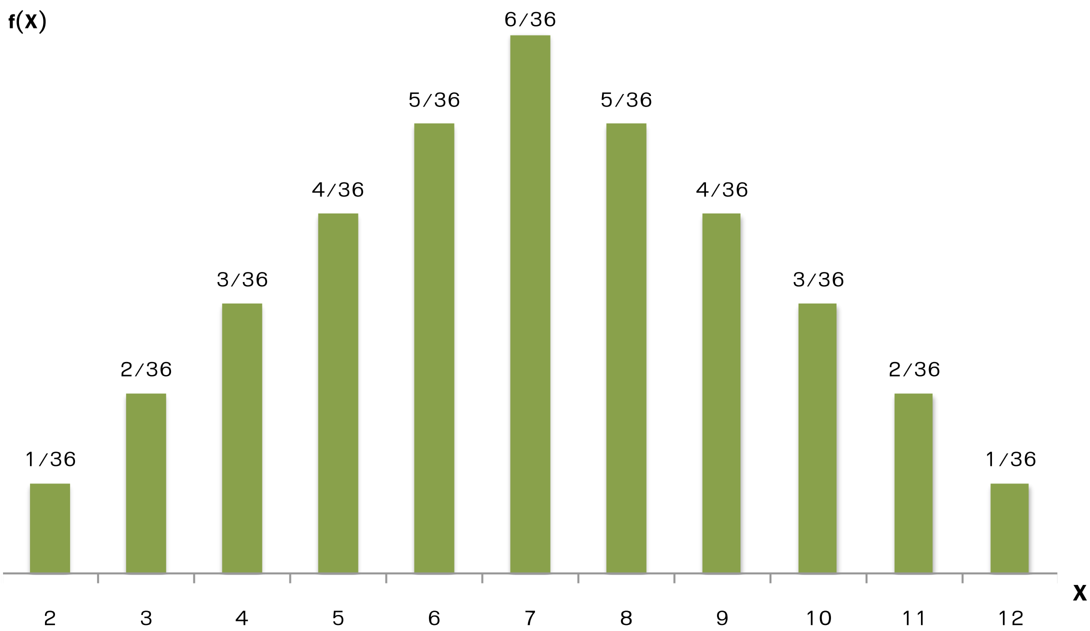
1.2.1 Probability Density Function for Discrete Random Variable

Let X be the discrete random variable with the value $x_1, x_2, \ldots, x_n$ and we get,

$$\begin{array}{rcl} f(x) & = & P(X = x_i) \quad \text{for} \quad i = 1,2,...,n \\ f(x) & = & 0 \quad \text{for} \quad x \neq x_i \end{array}$$

Example: Let X be random variable of the sum of values on the face of two dices. The value might be 2 or 12, that is the value from both rolling round is 1 or 6, respectively. The Figure 1 summarizes all possible results#

Figure 1: Probability Density function of the Sum of Values on the Side of the Dice, Obtained from Rolling the Dice Twice



Example

In a throw of two dice, the random variable X, the sum of the numbers shown on two dice can take one of the 11 values

2nd dice \ 1st dice	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

$f(x) = P(X = x_i)$

$0 \leq f(x_i) \leq 1$

$\sum f(x_i) = 1$

x : 2 3 4 5 6 7 8 9 10 11 12

$f(x)$: $\frac{1}{36}$ $\frac{2}{36}$ $\frac{3}{36}$ $\frac{4}{36}$ $\frac{5}{36}$ $\frac{6}{36}$ $\frac{5}{36}$ $\frac{4}{36}$ $\frac{3}{36}$ $\frac{2}{36}$ $\frac{1}{36}$

$$0 \le f(x_i) \le 1$$

$$\sum f(x_i) = 1$$

Tossing 2 coins

2nd \ 1st	H	T
H	HH	HT
T	TH	TT

{HH, HT, TH, TT}

outcome X => #H (number of head) x_i	$f(x_i)$
0	$\frac{1}{4}$
1	$\frac{1}{2}$
2	$\frac{1}{4}$

$\sum f(x) = 1$

1.2.2 Probability Density Function for Continuous Random Variable

Let X be the continuous random variable. The probability density function of X satisfies the three following conditions.

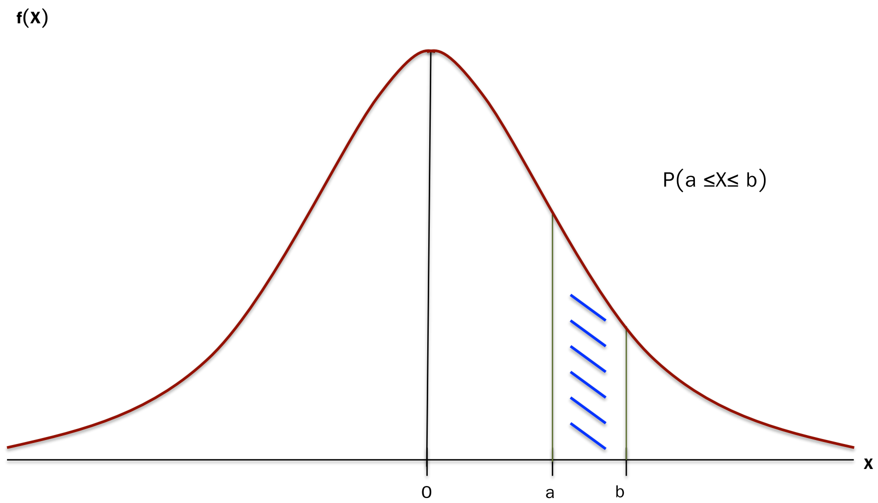
Discrete

$\sum f(x) = 1$

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$ continuous RV
3. $\int_{a}^{b} f(x)dx = P(a \leq x \leq b)$

Figure 2 exhibits the probability density function for the continuous random variable, where the area under the curve represents the probability that the variable will lay on that range. Specifically, $P(a \leq X \leq b)$ means the probability that X will take the value between a and b.

Figure 2: Probability Density Function for Continuous Random Variable



E.g. $f(x) = \frac{1}{9}x^2, \quad 0 \leq x \leq 3$

$$\int_0^3 \frac{1}{9}x^2 dx = 1$$

$$\frac{1}{27}x^3 \Big|_0^3 = 1$$

1.2.3 Joint Probability Density Function

In this section, only **joint probability density function** for discrete variable is discussed. Let X and Y be discrete random variables. The joint probability density function, identifying the probability that X and Y happen simultaneously, is written as,

$$f(X,Y) = P(X = x \text{ and } Y = y)$$

Example: The following table explains the joint probability density function.

Table 1: The table illustrating the joint probability density function of X and Y

		X		
		-1	0	1
	1	0.11	0.08	0.05
Y	2	0.09	0.05	0.03
	3	0.35	0.07	0.17

According to the table 1, the probability that random variable X will be 0 and random variable Y will be 3 is 0.07 or 7 percent. In mathematical term, it can be written as $f(X = 0, Y = 3) = 0.07$.

1 $f(X=-1, Y=1) =$

2 $f(X=0, Y=2) =$

3 $f(X=1, Y=3) =$

1.2.4 Marginal Probability Density Function

The above joint probability density function $f(X,Y)$ shows the joint distribution of two variables. On the other hand, **marginal probability density function** with respect to joint probability function, displays the probability density function of single variable like $f(X)$, $f(Y)$, which can be derived from;

$$\begin{aligned} f(X) &= \textstyle\sum_Y f(X,Y) \quad \text{called} \quad \text{marginal PDF of X} \\ f(Y) &= \textstyle\sum_X f(X,Y) \quad \text{called} \quad \text{marginal PDF of Y} \end{aligned}$$

where $\sum_Y$ or $\sum_X$ means the summation of probability over all values of X and Y respectively.

Example: According to Table 2 above, marginal PDF of X is obtained from

$$\begin{aligned} (X=-1) &= \\ &= \\ &= \\ &= \\ f(X=0) &= \textstyle\sum_Y f(X=0,Y) \\ &= f(X=0,Y=1)+f(X=0,Y=2)+f(X=0,Y=3) \\ &= 0.08+0.05+0.07 \\ &= 0.20 \\ f(X=1) &= \textstyle\sum_Y f(X=1,Y) \\ &= f(X=1,Y=1)+f(X=1,Y=2)+f(X=1,Y=3) \\ &= 0.05+0.03+0.17 \\ &= 0.25 \end{aligned}$$

		X		
		-1	0	1
Y	1	0.11	0.08	0.05
	2	0.09	0.05	0.03
	3	0.35	0.07	0.17

Marginal PDF of Y is obtained from

$$
\begin{aligned}
f(Y=1) &= \\
&= \\
&= \\
&= \\
f(Y=2) &= \textstyle\sum_X f(X,Y=2) \\
&= f(X=-1,Y=2)+f(X=0,Y=2)+f(X=1,Y=2) \\
&= 0.09+0.05+0.03 \\
&= 0.17 \\
f(Y=3) &= \textstyle\sum_X f(X,Y=3) \\
&= f(X=-1,Y=3)+f(X=0,Y=3)+f(X=1,Y=3) \\
&= 0.35+0.07+0.17 \\
&= 0.59
\end{aligned}
$$

		X		
		−1	0	1
Y	1	0.11	0.08	0.05
	2	0.09	0.05	0.03
	3	0.35	0.07	0.17

According to the calculation above, the result can be summarized into Table 2.

Table 2 shows joint probability of random variable X and Y

		X -1	0	1	
	1	0.11	0.08	0.05	$f(Y=1)$ =
Y	2	0.09	0.05	0.03	$f(Y=2)$ =
	3	0.35	0.07	0.17	$f(Y=3)$ =
		$f(X=-1)$ =	$f(X=0)$ =	$f(X=1)$ =	$f(X)=$ $f(Y)=$

1.2.5 Conditional Probability Density Function

Conditional probability density function is the probability of one event given that some events have already occurred. The function is written as,

$$f(X|Y) = P(X=x|Y=y)$$

This function can be obtained from the joint probability density function through,

$$f(X|Y) = \frac{f(X,Y)}{f(Y)}$$

Example: According to Table 2, find $f(X=1|Y=2)$ and $f(Y=2|X=0)$

$f(x|y) = \frac{f(x,y)}{f(y)}$ Conditional PDF of X

$f(y|x) = \frac{f(x,y)}{f(x)}$ Conditional PDF of Y