Universiti Malaysia PAHANG
Engineering • Technology • Creativity

DUM 2413 STATISTICS & PROBABILITY

CHAPTER 4
DISCRETE PROBABILITY DISTRIBUTIONS

PREPARED BY:
DR. CHUAN ZUN LIANG; DR. NORATIKAH ABU; DR. SITI ZANARIAH SATARI
FACULTY OF INDUSTRIAL SCIENCES & TECHNOLOGY
chuanzl@ump.edu.my; atikahabu@ump.edu.my; zanariah@ump.edu.my

EXPECTED OUTCOMES

- Able to determine the expected value, standard deviation and variance of a discrete random variable
- Able to distinguish between the binomial and Poisson distributions
- Able to solve the application problems, which involved the binomial and Poisson distributions
- Able to identify the relationship between the binomial and Poisson distributions

CONTENT

4.1
DISCRETE RANDOM VARIABLE AND PROBABILITY DISTRIBUTION

RECALLED TO CHAPTER 1

VARIABLE

Any characteristic or property of the population of interest
E.g. : The average monthly income per household in Malaysia, level of education and age of residents in a community, etc.

QUANTITATIVE OR NUMERICAL

* Measures on the numeric scale
* Resulted a numeric response
E.g.: What is your age?
The answer is a numeric.

…VE OR…

* Measu… …scale
* Resulted… …onse
E…
…a non-nu…

IT IS RELATED TO RANDOM VARIABLE

DISCRETE

* The numeric responses which arises from *counting process*
E.g.: How many balls in the basket?

CONTINUOUS

* The numeric response which arises from a *measuring process*
E.g.: What is your weight and height?

RANDOM VARIABLE

RANDOM VARIABLE

DEFINITION: If *X* is a ***real-valued*** function defined in a ***sample space***, ***S*** in probability measure, therefore *X* is known as ***a random variable***.

TYPE

DISCRETE RANDOM VARIABLE

DISCUSSED IN THIS CHAPTER
(“=” equal signs is matters)

CONTINUOUS RANDOM VARIABLE

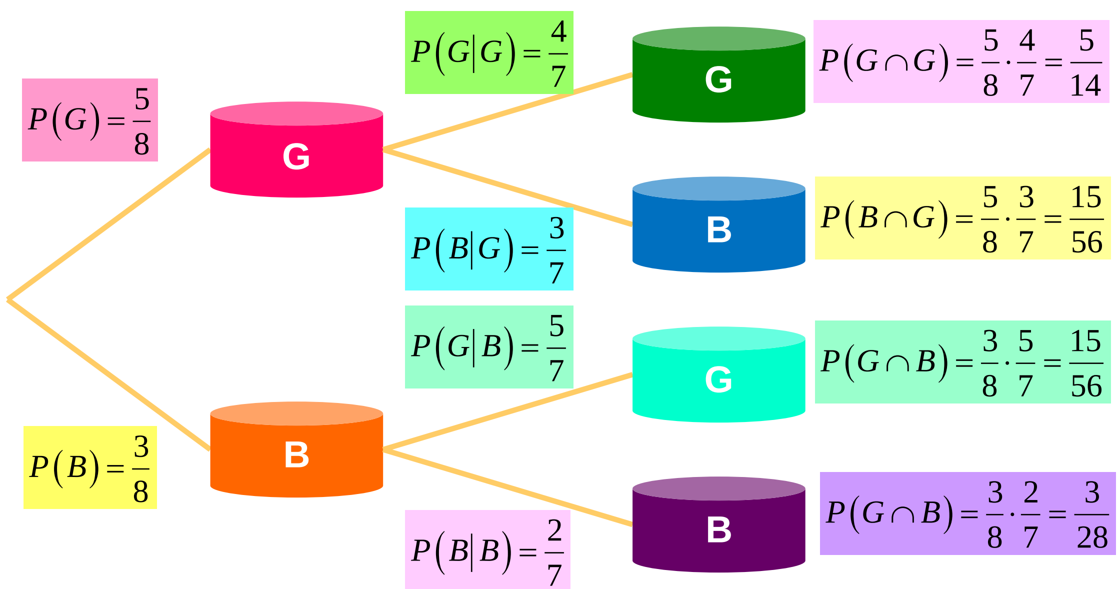
DISCUSSED IN CHAPTER 5
(“=” equal signs don’t matters)

YouTube: https://www.youtube.com/watch?v=lHCpYeFvTs0

EXAMPLE 4.1

Jenny randomly selects *two socks* in succession without replacement from a box, which the box contains *five green (G) socks* and *three blue (B) socks*. Based on the given information, list all the elements of the sample space and their corresponding probability. Also, denotes that a random variable, X represents the *number of green socks* are selected, hence give the corresponding probability of X.

SOLUTION



Elements	Probability
GG	$P(X=2)={}^{5}/_{14}$
GB	$P(X=1)={}^{15}/_{28}$
BG	
BB	$P(X=0)={}^{3}/_{28}$

PROBABILITY DISTRIBUTIONS vs. DISTRIBUTION FUNCTION

PROBABILITY DISTRIBUTIONS

If X is a ***discrete random variable***, which have a function, $f(x) = P(X = x)$ for each x within the range of X is known as the ***probability distribution*** of X.

PROBABILITY DISTRIBUTIONS

where the value $f(x) = P(X = x)$ ***should be satisfy the conditions***

(i) $f(x) \geq 0$ (ii) $\sum_{x} f(x) = 1$

DISTRIBUTION FUNCTION

If X is a ***discrete random variable***, the function given by

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad \text{for } -\infty < x < \infty$$

where $f(t)$ is the ***value of the probability distribution*** X of at t, is called the ***distribution function*** (***cumulative distribution***) of X.

DISTRIBUTION FUNCTION

where the values $F(x)$ ***should be satisfy the conditions***

1. $F(-\infty) = 0$ and $F(\infty) = 1$
2. if $a < b$, then $F(a) \leq F(b)$ for any real number a and b.

EXAMPLE 4.2

Determine whether ***the given function is permitted to serve as the probability distribution of a discrete random variable, X*** **with the range** $x = 1,\ 2,\ 3,\ 4.$

(i) $f(1) = 0.25,\ f(2) = 0.75,\ f(3) = 0.25$ and $f(4) = -0.25.$

(ii) $f(1) = 0.15,\ f(2) = 0.27,\ f(3) = 0.29$ and $f(4) = 0.29.$

(iii) $f(1) = \frac{1}{19},\ f(2) = \frac{10}{19},\ f(3) = \frac{2}{19}$ and $f(4) = \frac{5}{19}.$

SOLUTION

i

Not permissible. This is due to $f(4) = P(X = 4) < 0.$

ii

Permissible. This is due to

(i) $f(x) \ge 0$ **(ii)** $\sum_{x \in S} f(x) = f(1) + f(2) + f(3) + f(4) = 1.$

iii

Not permissible. This is due to $\sum_{x \in S} f(x) = f(1) + f(2) + f(3) + f(4) \neq 1.$

EXAMPLE 4.3

Find the value of constant c if the given function $f(x)$ is satisfies the conditions of ***a probability distribution*** for ***a random variable, X***.

(i) $f(x)=\dfrac{x}{c}, \quad x=1,\ 2,\ 3,\ 4.$

(ii) $f(x)=cx, \quad x=1,\ 2,\ 3,\ 4,\ 5.$

(iii) $f(x)=c(x+1)^2, \quad x=0,\ 1,\ 2,\ 3.$

SOLUTION

i

$$f(1)=\frac{1}{c};\ f(2)=\frac{2}{c};\ f(3)=\frac{3}{c};\ f(4)=\frac{4}{c}$$

Therefore, $\sum_{x\in S} f(x)=f(1)+f(2)+f(3)+f(4)=\frac{1}{c}+\frac{2}{c}+\frac{3}{c}+\frac{4}{c}=1\Rightarrow c=10$

ii

$$f(1)=c;\ f(2)=2c;\ f(3)=3c;\ f(4)=4c;\ f(5)=5c$$

Therefore,

$$\sum_{x\in S} f(x)=f(1)+f(2)+f(3)+f(4)+f(5)=c+2c+3c+4c+5c=1\Rightarrow c=\frac{1}{15}$$

iii

$$f(0)=c;\ f(1)=4c;\ f(2)=9c;\ f(3)=16c$$

Therefore,

$$\sum_{x\in S} f(x)=f(0)+f(1)+f(2)+f(3)=c+4c+9c+16c=1\Rightarrow c=\frac{1}{30}$$

EXAMPLE 4.4

Determine whether *the given function is permitted to serve as the probability distribution of a discrete random variable, X* with the range $x = 1,\ 2,\ 3 \text{ and } 4.$

(i) $F(1) = 0.3,\ F(2) = 0.5,\ F(3) = 0.8 \text{ and } F(4) = 1.2.$

No permissible. This is due to $F(4) > 1.0.$

$$P(X = 2) = P(X \le 2) - P(X \le 1) = F(2) - F(1)$$
$$= 0.4 - 0.5$$
$$= -0.1$$

This is violated the axiom of probability

(ii) $F(1) = 0.5,\ F(2) = 0.4,\ F(3) = 0.7 \text{ and } F(4) = 1.0.$

No permissible. This is due to $(F(2) = 0.4) < (F(1) = 0.5).$

(iii) $F(1) = 0.25,\ F(2) = 0.61,\ F(3) = 0.83 \text{ and } F(4) = 1.00.$

Permissible.

EXAMPLE 4.5

Find the *distribution function* of the random variable, X that has the *probability distribution*

$$f(x)=\frac{x}{15} \quad \text{for } x=1,\ 2,\ 3,\ 4,\ 5$$

SOLUTION

$f(1)=\frac{1}{15}$

$f(2)=\frac{2}{15}$

$f(3)=\frac{3}{15}$

$f(4)=\frac{4}{15}$

$f(5)=\frac{5}{15}$

$$F(x)=\begin{cases} 0, & x<1 \\ 1/15, & 1\le x<2 \\ 1/5, & 2\le x<3 \\ 2/5, & 3\le x<4 \\ 2/3, & 4\le x<5 \\ 1, & x\ge 5 \end{cases}$$

$F(1)=f(1)=1/15$

$F(1)=f(1)+f(2)=1/5$

$F(1)=f(1)+f(2)+f(3)=2/5$

$F(1)=f(1)+f(2)+f(3)+f(4)=2/3$

$F(1)=f(1)+f(2)+f(3)+f(4)+f(5)=1$

EXERCISE 4.1

Determine whether *the given function is permitted to serve as the probability distribution of a discrete random variable, X* with the range $x = 1,\ 2,\ 3,\ 4.$

(i) $f(x) = \dfrac{x-2}{5}, \quad x = 1,\ 2,\ 3,\ 4,\ 5.$

(ii) $f(x) = \dfrac{x^2}{30}, \quad x = 0,\ 1,\ 2,\ 3,\ 4.$

(iii) $f(x) = \dfrac{1}{5}, \quad x = 0,\ 1,\ 2,\ 3,\ 4,\ 5.$

SOLUTION

i

$$f(1)=\frac{1-2}{5}=-\frac{1}{5};\ f(2)=\frac{2-2}{5}=0;\ f(3)=\frac{3-2}{5}=\frac{1}{5};\ f(4)=\frac{4-2}{5}=\frac{2}{5};\ f(5)=\frac{5-2}{5}=\frac{3}{5}$$

Not permissible. This is due to $\left(f(1)=-\frac{1}{5}\right)<0.$

ii

$$f(0)=\frac{0^2}{30}=0;\ f(1)=\frac{1^2}{30}=\frac{1}{30};\ f(2)=\frac{2^2}{30}=\frac{2}{15};\ f(3)=\frac{3^2}{30}=\frac{3}{10};\ f(4)=\frac{4^2}{30}=\frac{8}{15}$$

Permissible. This is due to

(i) $f(x)\geq 0$ (ii) $\sum_{x\in S} f(x)=f(0)+f(1)+f(2)+f(3)+f(4)=1.$

iii

$$f(0)=\frac{1}{5};\ f(1)=\frac{1}{5};\ f(2)=\frac{1}{5};\ f(3)=\frac{1}{5};\ f(4)=\frac{1}{5};\ f(5)=\frac{1}{5}$$

Not permissible. This is due to

$$\sum_{x\in S} f(x)=f(0)+f(1)+f(2)+f(3)+f(4)+f(5)\neq 1.$$

EXERCISE 4.2

A random variable, X has the distribution function

$$F(x)=\begin{cases}0 & \text{for } x<-1\\ 1/4 & \text{for } -1\le x<1\\ 1/2 & \text{for } 1\le x<3\\ 3/4 & \text{for } 3\le x<5\\ 1 & \text{for } x\ge 5\end{cases}$$

Find

(i) $P(X\le 3)$ (ii) $P(X=3)$ (iii) $P(X<3)$

(iv) $P(X\ge 1)$ (v) $P(-0.4<X<4)$ (vi) $P(X=5)$

i $P(X\le 3)=\frac{3}{4}$

iii $P(X<3)=\frac{1}{2}$

v $P(-0.4<X<4)=P(X\le 3)-P(X\le -1)$
$=\frac{3}{4}-\frac{1}{4}=\frac{1}{2}$

ii $P(X=3)=P(X\le 3)-P(X\le 1)$
$=\frac{3}{4}-\frac{1}{2}=\frac{1}{4}$

iv $P(X\ge 1)=1-P(X\le -1)$
$=1-\frac{1}{4}=\frac{3}{4}$

vi $P(X=5)=P(X\le 5)-P(X\le 3)$
$=1-\frac{3}{4}=\frac{1}{4}$

4.1
MEAN AND VARIANCE OF A DISCRETE RANDOM VARIABLE

MEAN AND VARIANCE

MEAN

Let X be a *discrete random variable* with *probability function* $f(x) = P(X = x)$
Then the *expected value* of X denoted by $\mathrm{E}(X)$, defined by

$$\mu = \mathrm{E}(X) = \sum_{x \in S} x \cdot f(x) = \sum_{x \in S} x \cdot P(X = x)$$

VARIANCE

The *variance* of a *random variable* X is defined by

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

$$= \left(\sum_{x \in S} x^2 \cdot f(x) \right) - \mu^2 = \left(\sum_{x \in S} x^2 \cdot P(X = x) \right) - \mu^2$$

SOME PROPERTIES OF MEAN AND VARIANCE

Let a and b be a constants. Then the following results hold:

(i) $\mathrm{E}(a) = a$ **(ii)** $\mathrm{E}(aX) = a\,\mathrm{E}(X)$

(iii) $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$. In particular, $\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X)$, $\mathrm{Var}(b) = 0$.

EXAMPLE 4.6

A *random variable, X* have the *probability distribution*

$$f(x)=\frac{(|x|+1)^2}{9},\qquad x=-1,0,1$$

Determine $\mathrm{E}(X)$, $\mathrm{E}(X^2)$, and $\mathrm{E}(3X^2-2X+4)$.

SOLUTION

i

In general,

$$\mathrm{E}(X)=\sum_{x\in S} x\cdot f(x)$$

Therefore,

$$\mathrm{E}(X)=(-1)\left(\frac{(|-1|+1)^2}{9}\right)+(0)\left(\frac{(|0|+1)^2}{9}\right)+(1)\left(\frac{(|1|+1)^2}{9}\right)=0$$

ii

In general,

$$\mathrm{E}(X^2)=\sum_{x\in S} x^2\cdot f(x)$$

Therefore,

$$\mathrm{E}(X^2)=(-1)^2\left(\frac{(|-1|+1)^2}{9}\right)+(0)^2\left(\frac{(|0|+1)^2}{9}\right)+(1)^2\left(\frac{(|1|+1)^2}{9}\right)=\frac{8}{9}$$

iii

$$\mathrm{E}(3X^2-2X+4)=3\mathrm{E}(X^2)-2\mathrm{E}(X)+\mathrm{E}(4)=3\left(\frac{8}{9}\right)-2(0)+4=\frac{20}{3}$$

EXAMPLE 4.7

Find the mean and variance for the discrete probability distribution below.

(i) $f(x)=\frac{1}{5}, \quad x=5,\ 10,\ 15,\ 20,\ 25.$

(ii) $f(x)=1, \quad x=5.$

(iii) $f(x)=\frac{4-x}{6}, \quad x=1,\ 2,\ 3.$

SOLUTION

i

Mean

$$\mu=\mathrm{E}(X)=\sum_{x\in S}x.f(x)$$
$$=\left(5\cdot\frac{1}{5}\right)+\left(10\cdot\frac{1}{5}\right)+\left(15\cdot\frac{1}{5}\right)+\left(20\cdot\frac{1}{5}\right)+\left(25\cdot\frac{1}{5}\right)$$
$$=15$$

Variance

$$\sigma^2=\mathrm{Var}(X)=\mathrm{E}(X^2)-(\mathrm{E}(X))^2$$
$$=\left(\left(5^2\cdot\frac{1}{5}\right)+\left(10^2\cdot\frac{1}{5}\right)+\left(15^2\cdot\frac{1}{5}\right)+\left(20^2\cdot\frac{1}{5}\right)+\left(25^2\cdot\frac{1}{5}\right)\right)-(15)^2$$
$$=50$$

ii

Mean

$$\mu = \mathrm{E}(X) = \sum_{x \in S} x.f(x)$$
$$= 5 \cdot 1$$
$$= 5$$

ii

Variance

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$
$$= (5^2 \cdot 1) - (5)^2$$
$$= 0$$

iii

Mean

$$\mu = \mathrm{E}(X) = \sum_{x \in S} x.f(x)$$
$$= \left(1 \cdot \frac{4-1}{6}\right) + \left(2 \cdot \frac{4-2}{6}\right) + \left(3 \cdot \frac{4-3}{6}\right)$$
$$= \frac{5}{3}$$

iii

Variance

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$
$$= \left(\left(1^2 \cdot \frac{4-1}{6}\right) + \left(2^2 \cdot \frac{4-2}{6}\right) + \left(3^2 \cdot \frac{4-3}{6}\right)\right) - \left(\frac{5}{3}\right)^2$$
$$= \frac{5}{9}$$

EXERCISE 4.3

A discrete *random variable, X* represents *the larger value when a pair of four-sided dice is rolled simultaneously*, where the *probability mass function* of *X* is

$$f(x) = \frac{2x-1}{16}, \quad x = 1,\ 2,\ 3,\ 4.$$

Find the *mean*, *variance* and *standard deviation* of *X*.

SOLUTION

In general,

$$\mathrm{E}(X) = \sum_{x \in S} x \cdot f(x)$$

$$\mathrm{E}(X) = \left(1 \cdot \frac{2(1)-1}{16}\right) + \left(2 \cdot \frac{2(2)-1}{16}\right) + \left(3 \cdot \frac{2(3)-1}{16}\right) + \left(4 \cdot \frac{2(4)-1}{16}\right)$$

$$= \frac{25}{8}$$

In general,

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

$$\mathrm{Var}(X) = \left(\left(1^2 \cdot \frac{2(1)-1}{16}\right) + \left(2^2 \cdot \frac{2(2)-1}{16}\right) + \left(3^2 \cdot \frac{2(3)-1}{16}\right) + \left(4^2 \cdot \frac{2(4)-1}{16}\right)\right) - \left(\frac{25}{8}\right)^2$$

$$= \frac{55}{64}$$

Standard Deviation, $\sigma = \sqrt{\mathrm{Var}(X)}$

$$= \sqrt{\frac{55}{64}}$$

$$= 0.9270$$

EXERCISE 4.4

Given $E(X+4)=10$ and $E\left[(X+4)^2\right]=116.$ Determine

(i) $\mathrm{Var}(X+4)$ (ii) $\mathrm{E}(X)$ (iii) $\mathrm{Var}(X)$

SOLUTION

i

$$\mathrm{Var}(X+4)=\mathrm{E}\left[(X+4)^2\right]-\left[\mathrm{E}(X+4)\right]^2$$
$$=116-10^2$$
$$=16$$

ii

$$\mathrm{E}(X+4)=\mathrm{E}(X)+\mathrm{E}(4)=10$$
$$\mathrm{E}(X)+4=10$$
$$\therefore \mathrm{E}(X)=6$$

iii

$$\mathrm{Var}(X+4)=\mathrm{Var}(X)+\mathrm{Var}(4)=16$$
$$\mathrm{Var}(X)+0=16$$
$$\therefore \mathrm{Var}(X)=16$$

EXERCISE 4.5

An insurance company will be *paid RM200 to a patient for each of the first two days in the hospital* and *RM 100 for each day after the first two days.* Let a *random variable, X* represents *the number of days a patient needs to be in the hospital*, where the *probability distribution of X* is defined as below.

$$f(x)=\frac{5-x}{10}, \qquad x=1,\ 2,\ 3,\ 4.$$

Based the information given above, determine the *expected payment* for the hospitalization.

SOLUTION

$$\begin{aligned}E(\text{Payment}) &= \underbrace{200\left(1\cdot\frac{5-1}{10}\right)}_{x=1}+\underbrace{200\left(2\cdot\frac{5-2}{10}\right)}_{x=2}+\underbrace{200\left(2\cdot\frac{5-3}{10}\right)+100\left(1\cdot\frac{5-3}{10}\right)}_{x=3}+\underbrace{200\left(2\cdot\frac{5-4}{10}\right)+100\left(2\cdot\frac{5-4}{10}\right)}_{x=4}\\ &= \text{RM}\,360\end{aligned}$$

THANK YOU

END OF CHAPTER 4 (PART 1)