Chapter 2 Random Variables

§2.1 Concept of Random Variables

I. Motivation:

1. An example: (Again!!)

Experiment: A coin is thrown 3 times.

$\Omega = \{HHH; HHT; HTH; HTT; THH; THT; TTH; TTT\}$

Now, perform the experiment and the total number of heads (appeared) is observed and recorded.

If outcome is $\{HHH\} \hat{=} \omega_1$, then 3.

If ..... $\{HHT\} \hat{=} \omega_2$ then 2.

— —— — . $\{HTH\} \hat{=} \omega_3$. then 2.

. . . — — — — .

If outcome is $\{TTH\} \hat{=} \omega_7$ then 1.

If outcome is $\{TTT\} \hat{=} \omega_8$ then 0.

If we let X denote the "total number of heads", then we can see:

① X is a "variable" (the values of X are numbers and it can take different values.

② The value of X depends upon the outcome, i.e. depends upon the particular ω in the sample space Ω.

In short, X can be viewed as a function from Ω to the real number.

③ Since the outcome is random, and so the value of X is also random.

So, essentially, X is a random number

We usually call "such random number X" as "random variable".

II. Definition:

1. Def.: A random variable is a real-valued function defined on the points of a sample space.

2. Notes:

① First a random experiment with all possible outcomes: sample space Ω.
then for each ω in Ω (a point in Ω) we assign a real value.
Thus a real-valued "function" on Ω.

② Since the random variable X is a function on Ω, so, often write it as $X(\omega)$.

3. Notations:

① "Random Variable" → "r.v"

② Use capital letters X, Y, Z etc (or $X(w)$, $Y(w)$, $Z(w)$) to denote random variables

③ Return to our example (P2.1)

We see that "the total number of heads", X is a random variable.

Question: For what outcomes, the total number of heads is 2 (exactly 2 !!)

Ans: {HHT; HTH; THH}

This is an event (∵ a subset of Ω !)

denote by A = {HHT; HTH; THH}

The event A is formed by such outcome for which the number of heads is exactly 2, i.e.

$$A = \{\omega \in \Omega;\ X(\omega) = 2\}$$

$$= \{HHT;\ HTH;\ THH\}$$

Since A is an event, then we can consider its probability, i.e.

$$P(A) = P(\{\omega \in \Omega;\ X(\omega) = 2\})$$

Usually simply written as

$$P(X(\omega) = 2) \quad \text{or} \quad P(X = 2).$$

Now what's the meaning of

$$P(X \leq 2) \quad \text{or} \quad P(X < 2)\ ?$$

the former is $P(B)$ where

$$B = \{HHT;\ HTH;\ HTT;\ THH;\ THT;\ TTH;\ TTT\}$$

§2.2 Discrete Random Variables

I. Concept:

1. Example: Recall the example in §2.1
   (Keep this example in mind !!)

Experiment: A fair coin is thrown 3 times.

$$\Omega = \{HHH;\ HHT;\ HTH;\ HTT;\ THH;\ THT;\ TTH;\ TTT\}$$
$$= \{\omega_1;\ \omega_2;\ \omega_3;\ \omega_4;\ \omega_5;\ \omega_6;\ \omega_7;\ \omega_8\}$$

Let X denote the number of heads, then X is a random variable. Indeed.

$X(\omega_1) = 3$; $X(\omega_2) = 2$; $X(\omega_3) = 2$

$X(\omega_4) = 1$; $X(\omega_5) = 2$; $X(\omega_6) = 1$

$X(\omega_7) = 1$; $X(\omega_8) = 0$;

We can say, the r.v. X takes the values of $\{0; 1; 2; 3\}$.

These values are called the possible values of the r.v. X.

For the above r.v. Z, it can take only a <u>finite number</u> of values.

(The number of possible values is finite)

We call it a discrete random variable.

(In our example, there are 4 possible values i.e. $\{0; 1; 2; 3\}$).

2. Definition: A random variable is called discrete r.v. if it can take only a finite

or at most a countably infinite number of values.

3. Note: "countably infinite number of values" means "all the possible values can be written as a sequence."

For example, non-negative integers $\{0, 1, 2, 3, 4, \dots\}$

4. Notation: For a r.v. X, we shall use the lowercase letter x to denote the possible values.

Hence, for a discrete r.v. X, all the possible values can be rewritten as

$$\{x_1; x_2; \cdots, x_n\} \quad \text{(finite many.)}$$

or $$\{x_1; x_2; \cdots, x_n; \cdots\} \quad \text{(a sequence!)}$$

Note that, here, x_1, say, is a particular real number, it is not a random variable (The random variable is X.)

In the above example, the random variable X is the total number of heads, so the possible values of X are $\{0, 1, 2, 3\}$.

Here we may write $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, $x_4 = 3$.

Now, here x_3, say, is just the real number 2. A particular real number 2 can not, of course, be called a random variable. We have mentioned the meaning of

$P(X=2)$, (Recall here!!)

Now, since $x_3 = 2$, the meaning of

$P(X = x_3)$ should be clear.

II. Probability mass function:

1. Example: Again, for the above example, since the coin is fair, so all the 8 different outcomes have the same probabity. (i.e. $\frac{1}{8}$) (Equally likely case!!)

Hence,

$P(X=0) = \frac{1}{8}$; $(X=0) = \{TTT\} = \{w_8\}$

$P(X=1) = \frac{3}{8}$; $(X=1) = \{w_4, w_6, w_7\}$

$P(X=2) = \frac{3}{8}$; $(X=2) = \{w_2, w_3, w_5\}$

$P(X=3) = \frac{1}{8}$; $(X=3) = \{HHH\} = \{w_1\}$

Now, we let ("$\triangleq$" means "defined as")

$$p(0) \triangleq P\{X = 0) = \frac{1}{8}$$

$$p(1) \triangleq P(X = 1) = \frac{3}{8}$$

$$p(2) \triangleq P(X = 2) = \frac{3}{8}$$

$$p(3) \triangleq P(X = 3) = \frac{1}{8}$$

i.e. in general, let (Lowercase letter $p(\cdot)$)

$$p(x_i) \triangleq P(X = x_i),$$

we then get a function p.

This function p is called the Probability mass function of the random variable X.

Note that this function p is a real function in the ordinary meaning, i.e. $p: R \to R$.

Essentially, this function can be rewritten as

x_i	0	1	2	3
$p(x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Note also that, this function p has the property that

$\forall x_i, \quad p(x_i) \geq 0$

and $\sum_{i=1}^{4} p(x_i) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$

2. Definition: Suppose X is a discrete random variable and all the possible values of X are $x_1, x_2, \cdots$.
(finite or countably infinite number of values)
Then the function p defined by

$$p(x_i) = P(X = x_i) \qquad (2.2.1)$$

is called the probability mass function of the random variable X.

"Probability Mass Function" will be denoted by "p.m.f"

3. Meaning:

The p.m.f of a r.v. X tells us two things:

① all the possible values;

② the probability of taking each value,

and thus everything about the r.v. X.

Namely, "p.m.f" contains all the information about the r.v.

4. Properties:

Let $p(x_i)$ $(i=1,2,\cdots)$ be the p.m.f of r.v. X,

then (i) $p(x_i) \geq 0$ for all x_i (2.2.2)

(ii) $\sum_i p(x_i) = 1$ (2.2.3)

The proof is easy. Note that by (i) and (ii) we also have, $p(x_i) \leq 1$ for all x_i.

III. Cumulative Distribution Function

1. Example (again): A fair coin is tossed 3 times

X: the total number of heads

Then the p.m.f is given by

x_i	0	1	2	3
$p(x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

i.e. all the possible values: $\{0, 1, 2, 3\}$

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

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

Now, for any real value x, we consider

$P(X \le x)$ (Recall the meaning !!)

⌜Recall:

$$P(X \le x) = P(\{\omega \in \Omega;\ X(\omega) \le x\})$$

for example

$$P(X \le 5.4) = P(\{\omega \in \Omega;\ X(\omega) \le 5.4\})$$

Suppose

① $x = -2$, say. then $\{\omega \in \Omega;\ X(\omega) \le -2\} = \phi$

(Impossible event, since all the possible values are $\{0, 1, 2, 3\}$!!)

$\Rightarrow P(X \le -2) = P(\phi) = 0$

② $x = -0.34$, say, similarly

$P(X \le -0.34) = 0$

③ $x = 0$, say. then $\{\omega \in \Omega;\ X(\omega) \le 0\}$

$= \{\omega \in \Omega;\ X(\omega) = 0\}$

(Again. all the possible values: $0, 1, 2, 3$)

$\Rightarrow P(X \le 0) = P(X = 0) = \frac{1}{8}$

④ $x = 0.503$, say.

then $\{X \le 0.503\} = \{X(\omega) = 0\}$

$\Rightarrow P(X \le 0.503) = P(X = 0) = \frac{1}{8}$

⑤ Similarly, $P(X \le 0.9999) = \frac{1}{8}$

⑥ $x = 1$, say, then

$$\{X \le 1\} = \{X = 0\} \cup \{X = 1\}$$ $\leftarrow$ or!!

$$\Rightarrow: P(X \le 1) = P(\{X=0\} \cup \{X=1\})$$

$$= P(X=0) + P(X=1)$$ (Think why!!)

$$= \frac{1}{8} + \frac{3}{8} = \frac{4}{8} = \frac{1}{2}$$

Similarly,

⑦ $P(X \le 1.5) = \frac{1}{2}$

⑧ $P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$$

⑨ $P(X \le 2.99) = P(X=0) + P(X=1) + P(X=2) = \frac{7}{8}$

⑩ $P(X \le 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

⑪ $P(X \le 1093) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 1$

actually, for any $x \ge 3$, we have $P(X \le x) = 1$.

P2.20
P2.16

Thus for any real value x, we can get a corresponding value by $P(X \le x)$. thus we get another function. We denote it as

$$F(x) = P(X \le x)$$

This function $F(x)$ (again the ordinary meaning) is called the "Cumulative Distribution Function" of the random variables.

2. Definition: Suppose X is a random variable, the cumulative distribution function F of the random variable X is defined for all real numbers x, $(-\infty < x < +\infty)$, by

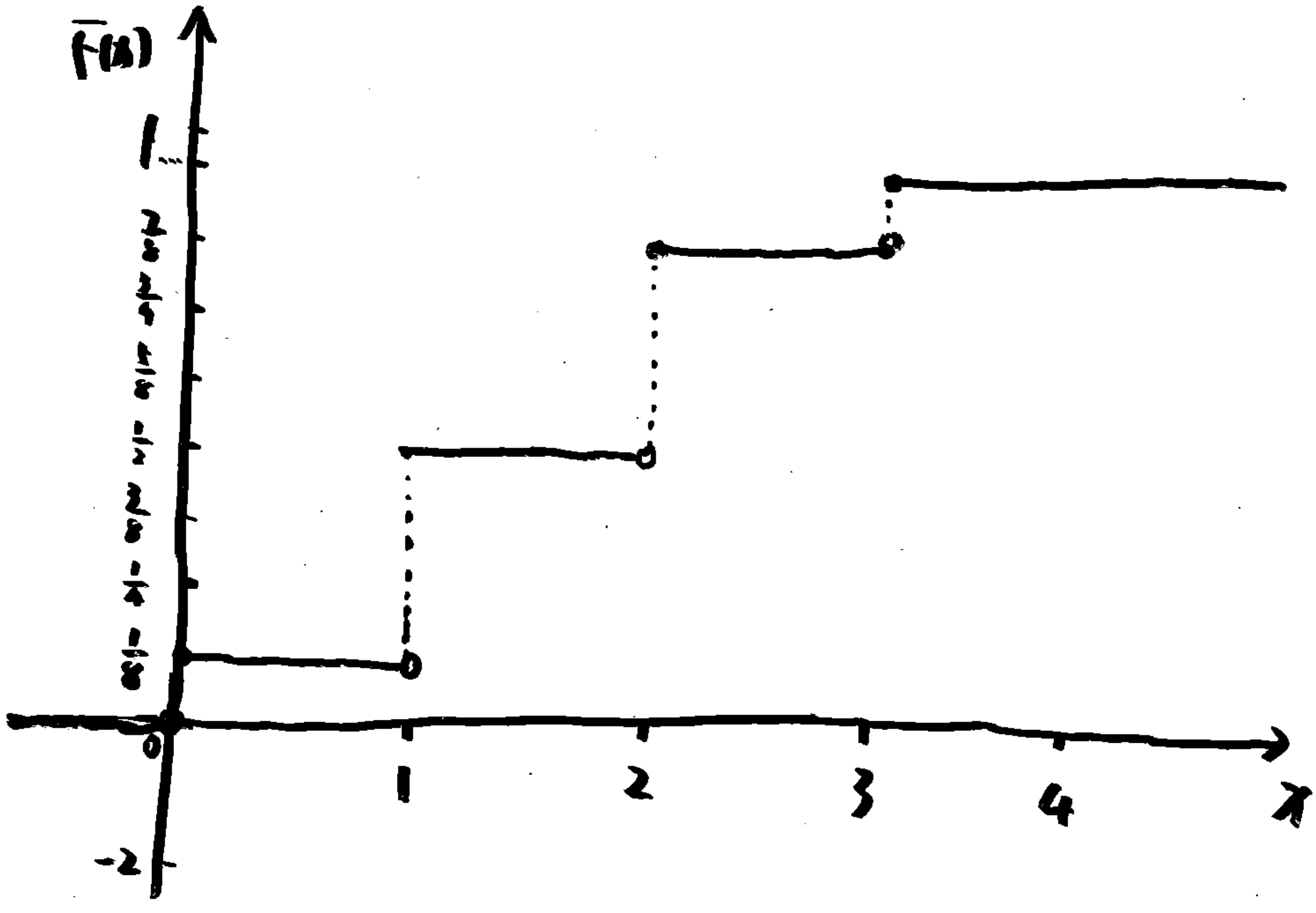
$$F(x) = P(X \le x) \qquad (2.2.4)$$

3. Graph:

For the example above, $F(x)$ is actually

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{8} & \text{if } 0 \le x < 1 \\ \frac{1}{2} & \text{if } 1 \le x < 2 \\ \frac{7}{8} & \text{if } 2 \le x < 3 \\ 1 & \text{if } x \ge 3 \end{cases}$$

So, the graph of $F(x)$ is



We can see that $F(x)$ is a "Step Function":
it is non-negative ; increasing and
it jumps at 0 (with jump $\frac{1}{8}$), 1 (with jump $\frac{3}{8}$)
2 (with jump $\frac{3}{8}$) and 3 (with jump $\frac{1}{8}$),
i.e. it jumps wherever $p(x)>0$ and that
the jump at x_i is $p(x_i)$. Also
$\boxed{0 \leq F(x) \leq 1 \ \forall x, \text{ and } \lim_{x\to-\infty} F(x)=0, \ \lim_{x\to+\infty} F(x)=1}$
The similar properties hold true
for other discrete random variables. Also, the
behaviour of the graph ...
(But, possibly for countably infinite number
of jumps !). In particular, again:
$0 \leq F(x) \leq 1$ for all x; $F(x)$ is an increasing function;
$F(x) \to 0 \ (x \to -\infty)$; $F(x) \to 1 \ (x \to +\infty)$.

IV. Relation Between "p.m.f" and "c.d.f":

Recall: for a discrete random variable X,

$$\text{p.m.f:} \quad p(x_i) = P(X = x_i)$$

$$\text{c.d.f:} \quad F(x) = P(X \le x)$$

There exists a close link between them.

For simplicity, let's just consider the case where all the possible values of the r.v. X are non-negative values $\{0, 1, 2, 3, 4, 5, \ldots\}$.

Then

1. if we know p.m.f. i.e $p(0)$, $p(1)$ ...

$\Rightarrow$: c.d.f.

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i).$$

for example,

$$F(17.1) = p(0) + p(1) + \cdots + p(17) = \sum_{k=0}^{17} p(k)$$

$$F(14) = p(0) + \cdots + p(14) = \sum_{k=0}^{14} p(k)$$

2. If we know c.d.f, i.e.

$F(x)$ is known for each $x \in (-\infty, +\infty)$

$\Rightarrow$: $p(x_i)$

(This is usually the case in application!!)

Suppose we want to get

$$p(17) = P(X=17) \text{ , say.}$$

then $p(17) = P(X=17)$

$$= P(X \le 17) - P(X \le 16)$$

$$= F(17) - F(16)$$

Similarly, for example,

$$p(304) = F(304) - F(303)$$

The reason for the above key step is, say

$(X \le 17) = (X \le 16) \cup (X = 17)$

$\Rightarrow$. $P(X \le 17) = P(X \le 16) + P(X=17)$ (Think why!!)

$\Rightarrow$: $P(X=17) = P(X \le 17) - P(X \le 16)$

i.e. $p(17) = F(17) - F(16)$

§2.3. Examples of Discrete Random Variables

I. Bernoulli Random Variables:

1. Def. A Bernoulli random variable takes on only two values: 0 and 1, with probabilities $1-p$ and p, respectively, where

$$0 < p < 1 \qquad (2.3.1)$$

2. P.m.f: Bernoulli random variable X:

x_i	0	1
$p(x_i)$	$1-p$	p

i.e.

$$p(0) = P(X=0) = 1-p \qquad (2.3.2)$$

$$p(1) = P(X=1) = p \qquad (2.3.3)$$

3. Note: The p.m.f of Bernoulli random variable X satisfies: $p(0) \geq 0$, $p(1) \geq 0$ (See 2.3.1)

and $p(0) + p(1) = 1$

II. The Binomial Distribution:

1. Definition: A random variable is called a Binomial random variable, (or, more often, the r.v. X obeys the Binomial distribution) if all the possible values of X are

$$\{0, 1, 2, 3, \cdots, n\}$$

where n is a positive integer ($n \geq 1$) and that

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{2.3.4}$$

where

$$0 < p < 1 \tag{2.3.5}$$

(Trivial, if $p = 0$ or $p = 1$)

2. p.m.f:

Usually we let $q = 1 - p$

(and since $0 < p < 1$ we also have $0 < q < 1$)

x_i	0	1	2	$\cdots$	n
$p(x_i)$	q^n	$\binom{n}{1} p q^{n-1}$	$\binom{n}{2} p^2 q^{n-2}$	$\cdots$	p^n

i.e.

$$p(k) = \binom{n}{k} p^k q^{n-k} = P(X=k) \quad (2.3.6)$$

where $k = 0, 1, 2, \cdots, n$

and $0 < p < 1$, $0 < q < 1$, $p + q = 1$ (2.3.7)

3. Notes:

① By (2.3.6) we again have

$p(k) \geq 0$ for all $k = 0, 1, \cdots, n$

$$\sum_{k=0}^{n} p(k) = 1 \quad (2.3.8)$$

Indeed, $\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$.

② If $n=1$ we return to Bernoulli r.v.

4. Notation:

For a Binomial r.v. X, the distribution depends upon two parameters: n and p. so, usually we write it as $B(n, p)$.

Thus:

"X is a Binomial Random Variable with parameters n and p"

$\Leftrightarrow$ "X obeys the Binomial distribution with parameters n and p"

$\Leftrightarrow$ " $X \sim B(n, p)$ "

$\Leftrightarrow$ " $p(k) = P(X=k) = \binom{n}{k} p^k q^{n-k}$ "

$k = 0, 1, 2, \cdots, n$

III. Poisson Distribution.

1. Definition: We say that the random variable X obeys the Poisson distribution if all the possible value of X are non-negative integers $\{0, 1, 2, 3 \cdots\}$ and the probability mass function takes the form of

$$P_K = P\{X=K\} = e^{-\lambda} \cdot \frac{\lambda^K}{K!} \qquad (2.3.9)$$

$(K = 0, 1, 2, 3, \cdots)$

where $\lambda > 0$ is a constant.

Note that, again, we have

$$P_K \geq 0 \quad (\forall K), \quad \text{and} \quad \sum_{K=0}^{\infty} P_K = 1. \qquad (2.3.10)$$

Indeed, $\sum_{K=0}^{\infty} P_K = \sum_{K=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^K}{K!} = e^{-\lambda} \cdot \sum_{K=0}^{\infty} \frac{\lambda^K}{K!}$

$$= e^{-\lambda} \cdot e^{\lambda} = e^0 = 1$$

2. Applications:

The most important discrete random variable

① Number of telephone calls received in a time interval.

② The number of radioactive particles observed in a time interval,
(or, the number of other kind of particles ...)

③ Queueing theory: the length of some kind of queue.

④ Traffic studies: the number of cars arrived in a junction.

⑤ - - - -

⑥ - - - - -

3. Parameter: The constant $\lambda > 0$ in (2.3.9) is called the parameter of the Poisson distribution.

The meaning of λ will be clear later.

(Also called: the mean; density; etc)

4. Notation:

Note that Poisson distribution depends upon the parameter λ.

" X obeys the Poisson distribution with parameter λ "

$\Leftrightarrow$ " The r.v. X has a Poisson distribution with parameter λ "

$\Leftrightarrow$ " $X \sim \text{Poisson}(\lambda)$ "

5. c.d.f and the table:

If $X \sim Poisson(\lambda)$, then the c.d.f of X:

$$F_X(x) = P(X \le x) = \sum_{k \le x} P\{X = k\}$$

$$= \sum_{k \le x} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

for example,

$$F_X(8) = \sum_{k \le 8} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k=0}^{8} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=0}^{8} \frac{\lambda^k}{k!}$$

$$= e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots + \frac{\lambda^8}{8!}\right).$$

Also, say

$$F_X(3.52) = \sum_{k \le 3.52} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k=0}^{3} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$= e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right).$$

The value of $F(x)$ can be obtained by checking the table.

See Table III Pages 610 — 612.