§ 3.2 Probability Distributions

In the previous examples, we saw how the prob. measure on the sample space determined the prob. with which the r. v. assumed different values.

e.g. the example with the two dice.

x	$P(X=x)$
2	$1/36$
3	$2/36$
4	$3/36$
5	$4/36$
6	$5/36$
7	$6/36$
8	$5/36$
9	$4/36$
10	$3/36$
11	$2/36$
12	$1/36$

One can easily check that the function

$$f(x) = \frac{6 - |x-7|}{36}, \quad x = 2, 3, \ldots, 12$$

gives the correct probabilities.

<u>Defn 3.2</u> If X is a discrete r.v., the fn given by $f(x) = P(X=x)$ for each x within the range of X is called the <u>probability distribution function</u> (PDF) of X.

From the postulates of prob., we get

<u>Thm 3.1</u> A fn f can serve as the prob. distr. of a discrete r.v. iff its values $f(x)$ satisfy

1) $f(x) \geq 0$ for each x within the domain of f;

2) $\sum_{x \in dom(f)} f(x) = 1$.

Ex. Find a formula for the prob. distr of the total no. of heads obtained in 4 tosses of a balanced coin.

From prev. ex.

$$P(X=0) = \frac{1}{16}$$

$$P(X=1) = \frac{4}{16}$$

$$P(X=2) = \frac{6}{16}$$

$$P(X=3) = \frac{4}{16}$$

$$P(X=4) = \frac{1}{16}$$

Note $\binom{4}{0} = 1$, $\binom{4}{1} = 4$, $\binom{4}{2} = 6$, $\binom{4}{3} = 4$, $\binom{4}{4} = 1$,

So

$$f(x) = \frac{\binom{4}{x}}{16}, \quad x = 0, \ldots, 4$$

will do.

Ex. Can the fn

$$f(x) = \frac{x+2}{25}, \quad x = 1, 2, \ldots, 5$$

be the prob. dist. fn of a discrete r.v.?

Soln. Easy to see that $f(x) \geq 0$ for $x = 1, \ldots, 5$, so the first requirement of Thm 3.1 is met.

For the second requirement

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

$$= \frac{3}{25} + \frac{4}{25} + \frac{5}{25} + \frac{6}{25} + \frac{7}{25}$$

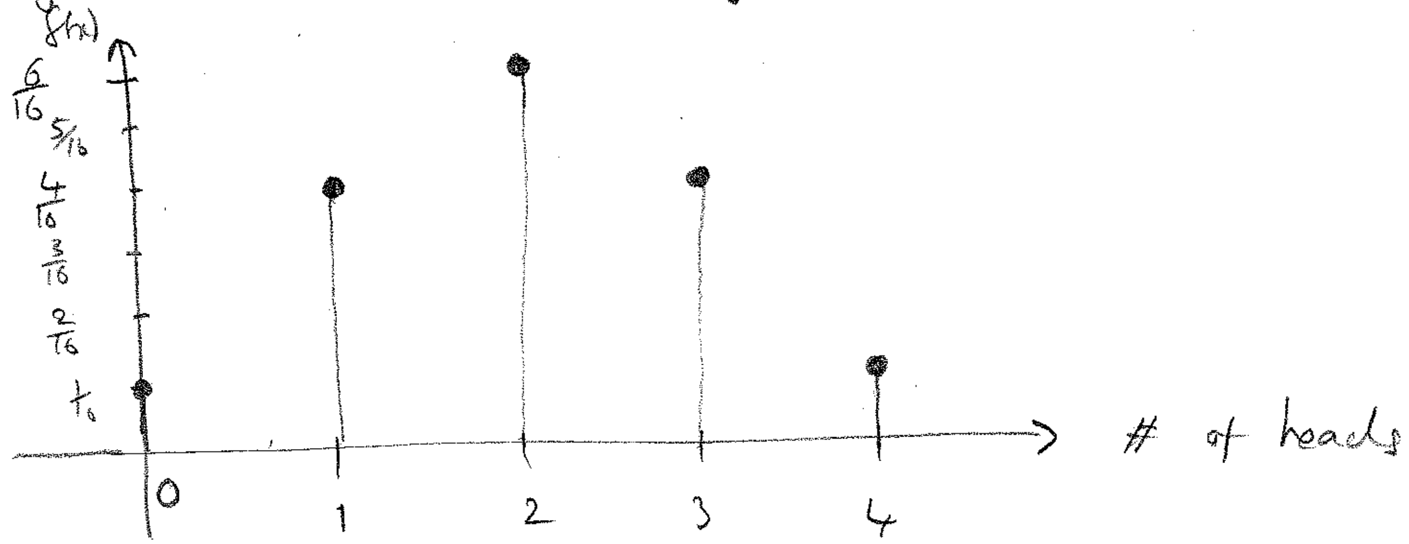
$$= 1,$$

so this requirement is also met.

Thus by Thm 3.1, f could be the prob. distr fn of a discrete r.v.

Prob. distr. fns are often represented as graphs.

e.g. for the coin tossing example



The book discusses other ways, but for discrete r.v.'s this sort of picture is the best.

Often we wish to concentrate on the probabilities of the events $\{X \leq x\}$ which leads to the following defn.

<u>Defn.</u> 3.3 If X is a discrete r.v., the fn given by

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

where $f(t)$ is the prob. distr. fn of X, is called the cumulative distribution fn (CDF) of X.

We immediately find.

<u>Theorem 3.2</u> The values $F(x)$ of the CDF of a discrete r.v. X satisfy the conditions

1. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$;

2. if $a < b$, then $F(a) \leq F(b)$ for any real nos. a, b.

<u>Ex.</u> Balanced coin tossed 4 times (again).

Had $f(0) = \frac{1}{16}$, $f(1) = \frac{4}{16}$, $f(2) = \frac{6}{16}$, $f(3) = \frac{4}{16}$, $f(4) = \frac{1}{16}$

Then

$$F(0) = f(0) = \frac{1}{16}$$

$$F(1) = f(0) + f(1) = \frac{5}{16}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16}$$

$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16}$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1.$$

Thus, the CDF is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{16}, & 0 \le x < 1 \\ \frac{5}{16}, & 1 \le x < 2 \\ \frac{11}{16}, & 2 \le x < 3 \\ \frac{15}{16}, & 3 \le x < 4 \\ 1, & x \geq 4. \end{cases}$$

Ex. Find the CDF of the r.v. for the socks example (5 Br, 3 Gr, choose 2 without repl., $W = \#$ brown chosen).

From the table

$$P(W=0) = \frac{3}{28} = f(0)$$

$$P(W=1) = \frac{15}{56} + \frac{15}{56} = \frac{15}{28} = f(1)$$

$$P(W=2) = \frac{5}{14} = f(2)$$

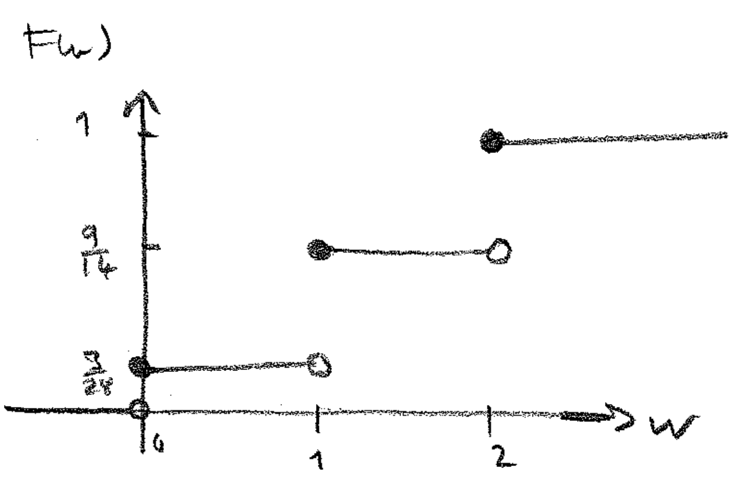
Thus

$$F(0) = f(0) = \frac{3}{28}$$

$$F(1) = f(0) + f(1) = \frac{9}{14}$$

$$F(2) = f(0) + f(1) + f(2) = \underline{1}.$$

Graph of F looks like



Note that it is cts from the $\underline{\text{right}}$ (but not the left) which is a general feature of CDF's.

Can also go the other way and get the PDF from the CDF.

<u>Theorem 3.3</u> If the range of a r.v. consists of the values $x_1 < x_2 < \cdots < x_n$, then

$$f(x_1) = F(x_1) \text{ and}$$

$$f(x_i) = F(x_i) - F(x_{i-1}), \quad i = 2, \ldots, n.$$

<u>Ex.</u> If the CDF of a r.v. X is given by

$$F(x) = \begin{cases} 0, & x < 2 \\ 1/36, & 2 \le x < 3 \\ \frac{3}{36}, & 3 \le x < 4 \\ \frac{6}{36}, & 4 \le x < 5 \\ \frac{10}{36}, & 5 \le x < 6 \\ \frac{15}{36}, & 6 \le x < 7 \\ \frac{21}{36}, & 7 \le x < 8 \\ \frac{26}{36}, & 8 \le x < 9 \\ \frac{30}{36}, & 9 \le x < 10 \\ \frac{33}{36}, & 10 \le x < 11 \\ \frac{35}{36}, & 11 \le x < 12 \\ 1, & 12 \le x \end{cases},$$

find the PDF of this r.v.

Soln. By Thm 3.3

$$f(2) = \frac{1}{36}, \quad f(3) = \frac{3}{36} - \frac{1}{36} = \frac{2}{36},$$

$$f(4) = \frac{6}{36} - \frac{3}{36} = \frac{3}{36}, \quad f(5) = \frac{10}{36} - \frac{6}{36} = \frac{4}{36},$$

$$\vdots$$

$$f(12) = 1 - \frac{35}{36} = \frac{1}{36}.$$

Recognise this as the pdf of the r.v. which was the sum of the numbers on the faces of the red and green dice from earlier.

Basically

PDF $\longrightarrow$ CDF is integration

CDF $\longrightarrow$ PDF is differentiation.

More about this later!