DISCRETE RANDOM VARIABLES (Chapter 3)

A *random variable* (rv) is a numerical variable whose value depends on the outcome of an experiment.

Notes: A random variable must be a number; it cannot be a letter, say.

More precisely, a *random variable* is a "real-valued function for which the domain is a sample space" (Definition 2.12 on page 76 in the text).

Example 1 A coin is tossed twice and the sequence of H's and T's is observed.

(Recall that an experiment consists of two things: what is *done*, and what is *observed*.)

Let Y be the number of H's which come up.
Show that Y is a random variable.

The experiment here has 4 possible outcomes: TT, TH, HT, TT.
Also: $Y = 0$ if the outcome is TT
$Y = 1$ if the outcome is TH or HT
$Y = 2$ if the outcome is HH.

We see that Y is a numerical variable whose value depends on the outcome of an experiment. Therefore Y is a random variable.

A random variable is *discrete* if its possible values are either
finite or *countably infinite* in number (or in other words, if those values can be *listed*).

For instance, Y in Ex. 1 is a discrete random variable (because $\{0,1,2\}$ is a finite set).

Notes: If the possible values of a rv Y are 0,3,6,..., then Y is *discrete*.

If the possible values of Y are all the real numbers in the set $(3,7)$,
then Y is not discrete; in this case we say that Y is a *continuous* rv (see Ch 4).

There is another type of random variable which is neither discrete nor continuous
but a combination of both types; such a r.v. is said to have a *mixed* distribution.
This may be covered later in Chapter 4 (e.g. see Example 4.19).

The probability that a discrete random variable Y takes on a particular value y is the sum of the probabilities of all sample points in the sample space S that are associated with y.

We write this probability $P(Y = y)$.

The *probability distribution (pr dsn)* of a discrete random variable Y is any information which provides $P(Y = y)$ for each possible value y of Y.

This information may take the form of a list, table, function (formula) or graph.

Example 2 Find the pr dsn of Y in Example 1.

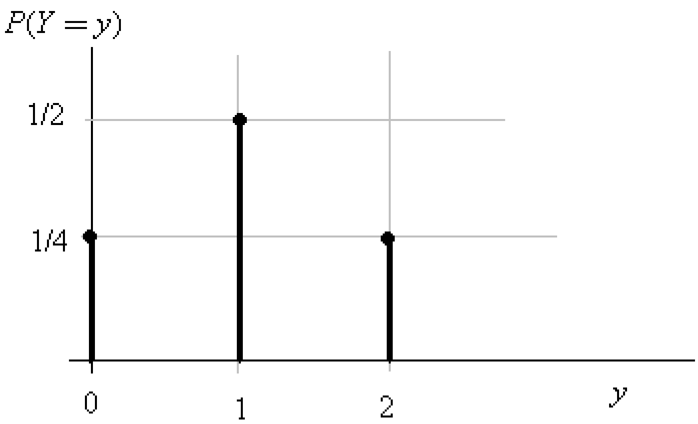
$P(Y = 0) = P(\text{TT}) = 1/4$
$P(Y = 1) = P(\text{TH}) + P(\text{HT}) = 1/4 + /14 = 1/2$
$P(Y = 2) = P(\text{HH}) = 1/4.$

The above is a *list*. Y's probability distribution can also be presented in other ways. For example:

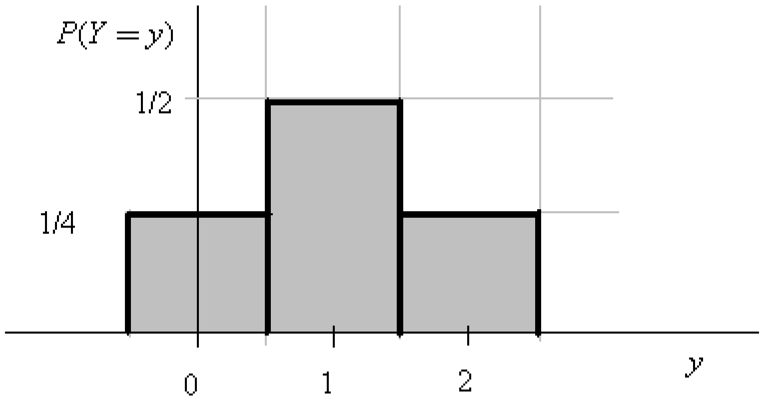
table

y	$P(Y = y)$
0	1/4
1	1/2
2	1/4

graph



another graph



formula $$P(Y=y)=\begin{cases}1/4, & y=0,2\\ 1/2, & y=1\end{cases}$$

another formula $$P(Y=y)=\frac{1}{2^{1+(y-1)^2}},\quad y=0,1,2$$

yet another formula $$P(Y=y)=\frac{1}{2(1+|y-1|)},\quad y\in\{0,1,2\}$$

Notation and terminology

It is conventional to denote rv's by upper case letters (eg, *Y*, *X*, *U*) and possible values of those rv's by the corresponding lower case letters (eg, *y*, *x*, *u*).

P(*Y* = *y*) is called the *probability density function (pdf)* of *Y*,
and is often written *p*(*y*) (or $p_Y(y)$ or $f(y)$ or $f_Y(y)$).

P(*Y* = *y*) may also be called the *density function*, the *density*,
or the *probability mass function (pmf)*.

Two properties of discrete pdf's

1. $0\le p(y)\le 1$ for all *y*.

(Every value of the pdf must be at least 0 and no more than 1.)

2. $\sum_y p(y)=1$.

(All the values of the pdf must add up to 1.)

Example 3 A coin is repeatedly tossed until the first head comes up.
Let Y be the number of tosses.
Derive the pdf of Y, and check that it satisfies the two properties of discrete pdf's.

$P(Y = 1) = P(\mathrm{H}) = 1/2$
$P(Y = 2) = P(\mathrm{TH}) = (1/2)(1/2)$
$P(Y = 3) = P(\mathrm{TTH}) = (1/2)(1/2)(1/2)$, etc.

Thus Y has pdf $p(y) = \begin{cases} 1/2, & y = 1 \\ 1/4, & y = 2 \\ 1/8, & y = 3 \\ \text{etc} & \end{cases}$

(Equivalently, $p(y) = 1/2^y, y = 1, 2, 3, ...$)

We observe that Property 1 is satisfied, since 1/2, 1/4, 1/8, ... are all between 0 and 1.

Also, $\sum_y p(y) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1$. Thus Property 2 is also satisfied.

Notes: Y is a discrete rv in this example because {1,2,3,...} is a *countably infinite* set (its elements can be *listed*).

A pdf uniquely defines a rv or pr dsn.
Thus a rv can't have 2 or more different pdf's.

Not all functions are valid pdf's.
For example, consider $p(y) = 2/3$, $y = 6,10$.
This function is not a pdf, since $2/3 + 2/3 \neq 1$.

Some important discrete probability distributions

The binomial distribution

This has to do with experiments which involve doing something several times, independently, and observing the number of 'successes'.

Example 4 A die is rolled 7 times. Let *Y* be the number of 6's which come up. Find *Y*'s pdf.

$P(Y = 3)$ = P(Three 6's and four non-6's, in any order)
= $P(6660000) + P(6606000) + \dots + P(0000666)$
123 124 567 <--- 'position numbers'
(0 here stands for 1,2,3,4 or 5 coming up, ie a 'non-6')

$$= \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 + \left(\frac{1}{6}\right)^2\frac{5}{6}\frac{1}{6}\left(\frac{5}{6}\right)^3 + \dots + \left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)^3 \qquad (\binom{7}{3} \text{ identical terms})$$

$$= \binom{7}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 \qquad (= 0.0781)\ .$$

Similarly, $P(Y = 2) = \binom{7}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^5$ (= 0.2344) , etc.

We see that $p(y) = \binom{7}{y}\left(\frac{1}{6}\right)^y\left(\frac{5}{6}\right)^{7-y}, \quad y = 0,\dots,7$.

We say that *Y* has a binomial distribution (with parameters 7 and 1/6).

A random variable *Y* has the *binomial distribution* with parameters *n* and *p* if its pdf is of the form

$$p(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y = 0,\dots,n \quad (n = 1,2,3,...;\ 0 \le p \le 1)\ .$$

We write $Y \sim \text{Bin}(n,p)$ and $p(y) = p_{\text{Bin}(n,p)}(y)$.

We call *n* the *number of trials* and *p* the *probability of success*.

Note: Don't confuse the two uses of *p* here: *p* in $p(y)$ refers to a function
p in p^y refers to a constant, etc.

Check that the binomial pdf is proper:

$$\sum_{y=0}^{n}\binom{n}{y}p^y(1-p)^{n-y} = (p+(1-p))^n = 1 \quad \text{(Property 2).}$$

Note: We have here used the *binomial theorem*: $(a+b)^m = \sum_{x=0}^{m}\binom{m}{x}a^x b^{m-x}$

(which can be proved by indiction).

For example, $(a+b)^2 = \binom{2}{0}a^0b^{2-0} + \binom{2}{1}a^1b^{2-1} + \binom{2}{2}a^2b^{2-2} = b^2 + 2ab + a^2$.

Example 5 A coin is going to be tossed 10 times.
Find the probability that 3 heads come up.

Let Y be the number of heads that come up.
Then Y ~ Bin(10,0.5), with pdf

$$p(y) = \binom{10}{y}0.5^y(1-0.5)^{10-y} = \binom{10}{y}\frac{1}{2^{10}}, \quad y = 0,\ldots,10.$$

So $P(Y=3) = p(3) = \binom{10}{3}\frac{1}{2^{10}} = 0.117$.

Alternatively, we can make use of tables.

On page 839 of the text we find that: $P(Y \le 3) = 0.172$ (= p(0) + p(1) + p(2) + p(3))
$P(Y \le 2) = 0.055$ (= p(0) + p(1) + p(2)).

Hence $P(Y=3)$ $(= p(3))$ $= P(Y \le 3) - P(Y \le 2) = 0.172 - 0.055 = 0.117$.

The Bernoulli distribution (a special case of the binomial)

A random variable Y has the *Bernoulli distribution* with parameter p if it has the binomial distribution with parameters 1 and p.

The pdf of Y is then $p(y) = \begin{cases} p, & y=1 \\ 1-p, & y=0 \end{cases}$ $(0 \le p \le 1)$.

We write Y ~ Bern(p) (or Y ~ Bin(1,p)) and $p(y) = p_{\text{Bern}(p)}(y)$.

We call p the *probability of success*, as before.

Example 6 A coin is tossed once. Let *Y* be the number of heads that come up. What is *Y*'s distribution? Write down *Y*'s pdf.

Y ~ Bern(0.5), and *Y*'s pdf is $p(y) = 0.5$, $y = 0,1$.

Notes: We may call *Y* the *indicator variable* for the event that heads come up and write this rv as $Y = I(A)$, where A = "Heads come up".
Here, $I(.)$ denotes the *standard indicator function*.

All indicator variables have a Bernoulli distribution.

A binomial experiment consists of several *Bernoulli trials*.
A binomial rv can be thought of as the sum of several Bernoulli rv's.

Eg, we toss two coins. Let *Y* = number of heads that come up.
Then $Y = Y_1 + Y_2$, where Y_i = number of heads on *i*th toss (must be 1 or 0)

$$= I(\text{heads on } i\text{th toss})$$

$$= \begin{cases} 1 & \text{if heads comes up on the } i\text{th toss} \\ 0 & \text{if tails comes up on the } i\text{th toss.} \end{cases}$$

Here $Y_1, Y_2 \sim iid\ Bern(1/2)$ and $Y \sim Bin(2,1/2)$.

The geometric distribution

Example 7 A die is rolled repeatedly until the first 6 comes up. Find the pdf of *Y*, the number of rolls.

$P(Y = 3) = P(\text{Two non-6's and then a 6}) = P(006) = (5/6)^2(1/6)$ (= 0.0231).
Similarly, $P(Y = 4) = (5/6)^3(1/6)$ (= 0.00386), etc.
We see that $p(y) = (5/6)^{y-1}(1/6), \quad y = 1,2,3,\ldots$

We say that *Y* has a geometric distribution (with parameter 1/6).

A random variable *Y* has the *geometric distribution* with parameter *p* if its pdf is of the form

$$p(y) = (1-p)^{y-1}p, \quad y = 1,2,3,\ldots \quad (0 \le p \le 1).$$

We write Y ~ Geo(p) and $p(y) = p_{\text{Geo}(p)}(y)$.

We call *p* the *probability of success*, as before.

Check that the geometric pdf is proper:

$$\sum_{y=1}^{\infty}(1-p)^{y-1}p = p\sum_{x=0}^{\infty}(1-p)^{x} \qquad \text{(put } x = y - 1\text{)}$$

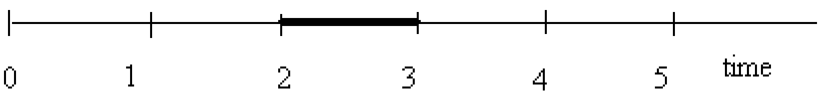
$$= p \times \frac{1}{1-(1-p)} = 1 \quad \text{(Property 2).}$$

The geometric distribution can be used to model *waiting times*.

Example 8 Suppose that the probability of an engine malfunctioning during any 1-hr period is 0.02. Find the pr that the engine will survive 2 hours.

Let Y be the number of 1-hr periods until the first malfunction (including the 1-hr period in which that malfunction occurs). Then $Y \sim \text{Geo}(p)$, where $p = 0.02$.

Eg: A malfunction in the 3rd 1-hr period means that $Y = 3$.



$$P(Y > 2) = \sum_{y=3}^{\infty} q^{y-1}p \qquad \text{where } q = 1 - p = 0.98$$

$$= pq^2\sum_{y=3}^{\infty} q^{y-1-2} \qquad = pq^2\sum_{x=0}^{\infty} q^{x} \qquad \text{(after putting } x = y - 3\text{)}$$

$$= pq^2\frac{1}{1-q} = q^2 = 0.98^2 = 0.9604.$$

Note: This may be calculated more simply as:

$$\begin{aligned} P(Y > 2) &= 1 - P(Y \le 2) \\ &= 1 - \{p(1) + p(2)\} \\ &= 1 - (p + qp) = q^2 = 0.9604\,, \end{aligned}$$

or even more simply as:

$$\begin{aligned} P(Y > 2) &= P(\text{Survive 2 hours}) \\ &= P(\text{No failures in the first 2 hours}) \\ &= q^2 = 0.9604. \end{aligned}$$

The hypergeometric distribution

This has to do with sampling objects from a box, without replacement, and observing how many have a certain characteristic.

Example 9 A box has 9 marbles, of which 3 are white and 6 are black. You randomly select 5 marbles from the bag (without replacement). Find the pdf of Y, the number of white marbles amongst the selected 5.

Number the 9 marbles 1,2,...,9, with the first 3 being white and the last 6 black. Then the sample points may be represented by writing 12345, 12346, ..., 56789.

Note: We don't write 13245, because this represents the same sample point as 12345. I.e., the distinct sample points correspond to strings of numbers in increasing order.

We see that the number of sample points is $n_S = \binom{9}{5}$.

The sample points associated with the event $Y = 2$ are **<u>12</u>**456, **<u>12</u>**457, ..., **<u>23</u>**789, and the number of these is $n_2 = \binom{3}{2}\binom{6}{3}$.

Note: We require 2 numbers to be from 1,2,3, and the other 3 from 4,5,6,7,8,9.

Hence $P(Y = 2) = \frac{n_2}{n_S} = \frac{\binom{3}{2}\binom{6}{3}}{\binom{9}{5}}$ $\left(= \frac{3(20)}{126} = \frac{10}{21} = 0.4762\right)$.

Similarly, $P(Y = 1) = \frac{n_1}{n_S} = \frac{\binom{3}{1}\binom{6}{4}}{\binom{9}{5}}$ $\left(= \frac{3(15)}{126} = \frac{5}{14} = 0.3571\right)$, etc.

We see that Y has pdf $p(y) = \frac{\binom{3}{y}\binom{6}{5-y}}{\binom{9}{5}}, \quad y = 0,1,2,3$.

We say that Y has a hypergeometric distribution (with parameters 9, 3 and 5).

A random variable Y has the *hypergeometric distribution* with parameters N, r and n if its pdf is of the form

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, \quad y = 0,\ldots,r \quad \text{subject to} \quad 0 \le n-y \le N-r$$

$(N = 1,2,3,...;\ r = 1,2,...,N;\ \ n = 1,2,...,N)$.

Note: The number of black balls sampled, $n - y$, can't be less than 0 or more than the total number of black balls in the bag, $N - r$.

We write $Y \sim \text{Hyp}(N,r,n)$ and $p(y) = p_{\text{Hyp}(N,r,n)}(y)$.

We may call N "the number of balls" (parameter), r "the number of white balls", and n "the number of sampled balls".

Example: There are 10 men and 15 women in a room.
8 persons are chosen randomly to form a committee.
Find the probability that the committee contains 6 women.

Solution: Let Y = number of women on the committee.
Then $Y \sim \text{Hyp}(25,15,8)$
and so $p(6) = C(15,6)C(10,2)/C(25,8) = 0.2082$.