CHAPTER 2

Discrete random variables

2.1 Introduction

Consider throwing two fair coins. We know that

P(2 heads) = $\frac{1}{4}$, P(1 head, 1 tail) = $\frac{1}{2}$, P (2 tails) = $\frac{1}{4}$

if we write the quantity X for the number of heads we have

P(X = 2) = $\frac{1}{4}$, P(X = 1) = $\frac{1}{2}$, P (X = 0) = $\frac{1}{4}$

X is an example of a *random variable.* It expresses the result of the experiment as a number. More formally we make the following definition

Definition: A random variable *X* is a numerically valued function defined on the sample space, Ω.

$$X : \Omega \to R$$

We say that *X* is a discrete random variable if it can take only a countable set of values, i.e. integer or rational values. We say *X* is a continuous random variable if it can take a continuous set of values, i.e. a subset of the reals.

Note that a random variable is usually written with a capital letter, *X* for example, the realised value of the random variable is written as a lower case letter, z for example. In this chapter we shall consider discrete random variables.
We next define what we mean by a discrete probability distribution.

Definition: if we have a discrete random variable *X* taking values $x_1, x_2, ..., x_n$ with probabilities $p_1, p_2, ..., p_n$ where

$$p_1 + p_2 + p_3 + ... + p_n = 1 \qquad p_i > 0 \ , \ \forall i$$

then this defines a discrete probability distribution for X Although we have written the random variable *X* as taking a finite set of values in this definition, it also holds for an *X* which takes an infinite countable set of values, e.g. all non-negative integers.
We may write $P(X = x_i)$ as p_i. This is sometimes referred to as the probability *function* for *X*.

Example : Two fair dice are thrown. Let X be the sum of the values on the upturned faces. Find the probability distribution for X.
Note that X =.2 if and only if both dice show 1. Thus $P(X = 2) = \frac{1}{36}$
Also, X = 3 if and only if one die shows 1 and the other 2. Thus $P(X = 3) = \frac{1}{36}$

Similarly we find that P(X=4)= $\frac{3}{36}$, P(X=5)= $\frac{4}{36}$, P(X=6)= $\frac{5}{36}$, P(X=7)= $\frac{6}{36}$,
P(X=8)= $\frac{5}{36}$,P(X=9)= $\frac{4}{36}$,P(X=10)= $\frac{3}{36}$,P(X=*11)*= $\frac{2}{36}$ *and P(X=12)*= $\frac{1}{36}$
Note that the sum of the probabilities is one and all are positive so this is a valid probability distribution.

Example: The discrete random variable *X* has probability function given by
$P(X=x) = cx^2$, x=1,2,3,4. Find c.

We know that

$$\sum_{x=1}^{4} P(X = x) = 1$$

So c + 4c + 9c + 16c =1 and hence c = $\frac{1}{30}$

2.2 Mean and Variance of a Random Variable

We begin with the definition of a mean of a random variable.

Definition: Suppose X is a discrete random variable taking values $x_1, x_2, \ldots, x_n$, with probabilities $p_1, p_2, p_3, \ldots, p_n$ then the mean or expected value of X, written as μ or E[X] is given by

$$\mu = E(X) = \sum_{i=1}^{n} p_i x_i$$

If *X* takes an infinite number of values the sum is taken over all values of i.
To justify this definition, suppose we had a sample of *x* values where *z* occurs with frequency *f*. Then the sample mean would be

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \sum \left(\frac{f_i}{\sum f_i} \right) x_i$$

In the limit if we collect enough data $f_i \big/ \sum f_i$ tends to *p*.
Example: A die is thrown, what is the mean (or expected) score?

$$E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6}$$

Note that the expected value of a random variable is not necessarily a value the random variable can take. The expected score when we throw a. fair die is 7/2, but a die cannot take this value. Think of the expected value or mean of a random variable as a measure of where the distribution is centred around.
The expectation of any *function* of a random variable, *g(X)* say, is defined in a similar way.

Definition : If X is a discrete random variable then the expectation of *X* is given by

$$E[g(X)] = \sum_{i=1}^{n} p_i g(x_i)$$

We can also define the variance and standard deviation of a random variable.

Definition : If X is a discrete random variable then its variance, written *Var[X]* is defined by

$$Var[X] = \sum_{i=1}^{n} p_i (x_i - \mu)^2$$

The standard deviation of X is the positive square root of the variance of *X*.
By multiplying out the bracket it is straightforward to see that the variance is given by
$Var[X] = (\sum p_i x_i^{\,2}) - \mu^2$ or $Var[X] = E[X^2] - (E[X])^2$
Example A die is thrown, what is the variance of the score?

$$E[X^2] = 1^2 \times \frac{1}{6} + 2.^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = \frac{91}{6}$$

$$V[X] = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

The variance gives an idea of how spread out the distribution is.

2.2 The Binomial distribution

We saw in section 3.3 that if we have n Bernoulli trials with probability of a success equal to p then the probability of r successes is given by the binomial probability

$$\binom{n}{r} p^r (1-p)^{n-r} \quad r = 0, 1, 2, \ldots, n.$$

Thus if we consider the random variable X which is the number of successes in n Bernoulli trials, then

$P(X = r)$ is given by the binomial probability. We say that X has a binomial distribution with parameters n and p.

Table (1) of gives the probability of r or more successes in vi independent trials with the probability of success p. For example if we wanted the probability of obtaining 23 or more heads in 50 tosses of a fair coin we find the answer is 0.76006.

Example:

Suppose that 5% of the articles made. by a factory are defective. What is the probability of finding 1 defective in a sample of 10 from a very large batch. Since it is a large batch we may treat this as sampling without replacement and the number of defectives, X, will have a binomial distribution with vi =10 and

p = 0.05. Thus

$$P(X = 1) = \binom{10}{1} 0.05 \times 0.95^{19} = 0.315$$

We can also find this quantity from the tables,

$$P(X = 1) = P(X \geq 1) - P(X \geq 2) = 0.40126 - 0.08614 = 0.31512$$

The tables are only given for some values of n and p so are not always useful, but you should know how to use them. Note that although p is only given up to 0.5, we can always turn a problem where the probability of a 'success' is greater than 0.5 into a question about 'failures' which will have probability less than 0.5. An example of this is given next.

Example Fifty seeds were planted and it is known that the probability of any seed germinating is 0.8. Assuming that the number of seeds germinating follows a binomial distribution, using tables find the probabilities of the following events (a) exactly 40 seeds germinate,

(b) more than 12 seeds fail to germinate,

(c) more than 38 but fewer than 45 seeds germinate.

The tables only give values of p up to 0.5 so we have to convert the events to questions about failing to germinate. The chance of a seed failing to germinate is 0.2. Let X be the number of seeds that germinate and Y the number of seeds that fail to germinate so that X + Y = 50. Thus we require for (a)

$$
\begin{aligned}
P(X=10) &= P(Y=10) \\
&= P(Y \geq 10) - P(Y \geq 11) \\
&= 0.55626 - 0.41644 \\
&= 0.13982 = 0.140
\end{aligned}
$$

For (b) $Y > 12$ is the same as $Y \geq 13$ and from the tables $P(Y \geq 13) = 0.18606 = 0.186$
For (c) 38< X<45 is the same as 6 $\leq Y \leq 11$ and the probability of this is

$$
\begin{aligned}
P(Y \geq 6) - P(Y \geq 12) &= 0.95197 - 0.28933 \\
&= 0.66264 = 0.663
\end{aligned}
$$

Note that we quote the final answers to no more than 3 decimal places.
When the tables do not exist there are approximations we may use. For example it can be shown that as $n \to \infty$ and $\pi \to 0$, a binomial distribution tends to a Poisson distribution (see later). Thus if *vi* is large and p is small we may approximate a binomial random variable by a Poisson one.

We next derive the mean and variance of a random variable *X* which has a binomial distribution with parameters *n* and *p*

Proposition ***9*** If *X* has a binomial distribution with parameters n and p then
E[X]= np and *Var(X] = np(1—p).*

Proof By the definition of the mean

$$
\begin{aligned}
E(X) &= \sum_{r=0}^{n} rP(X=r) \\
&= \sum_{r=0}^{n} r\binom{n}{r} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} r\binom{n}{r} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} r.\frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} \frac{n(n-1)!}{(r-1)!(n-r)!} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} n\binom{n-1}{r-1} p^r (1-p)^{n-r} \\
&= np\sum_{r=1}^{n} r\binom{n-1}{r-1} p^{r-1} (1-p)^{n-r} \\
&= np\sum_{k=0}^{n-1} \binom{n-1}{k} p^r (1-p)^{n-1-k} \\
&= np[p + (1-p)]^{n-1} \\
&= np
\end{aligned}
$$

Recall that $Var[X] = E[X^2] - [E[X]]^2$
Now $E[X^2] = E[X^2] = E[X(X-1)] + E[X]$
Then

$$
\begin{aligned}
E[X(X-1)] &= \sum_{r=0}^{n} r(r-1)P(X=r) \\
&= \sum_{r=2}^{n} r(r-1)\frac{n!}{r!(n-r)!}p^r(1-p)^{n-r} \\
&= \sum_{r=2}^{n} \frac{n(n-1)(n-2)!}{(r-2)!(n-r)!}p^r(1-p)^{n-r} \\
&= \sum_{r=2}^{n} n(n-1)\binom{n-2}{r-2}p^r(1-p)^{n-r} \\
&= n(n-1)p^2\sum_{r=2}^{n}\binom{n-2}{r-2}p^{r-2}(1-p)^{n-r} \\
&= n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^k(1-p)^{n-2-r} \\
&= n(n-1)p^2[p+(1-p)]^{n-2} \\
&= n(n-1)p^2
\end{aligned}
$$

Thus $Var[X] = n(n-1)p.^2 + np - (np)^2 = np(1-p)$

Example: The random variable X has a binomial distribution with parameters n=100 and p=0.8. Find the mean and variance of X.
The mean $\mu =$ *np* = 80, the variance is *np (np-1)* =16

2.4 The Poisson Distribution

Suppose events occur at random at an average rate λ per minute. Examples include radioactive decay and arrivals in a queue. Then the distribution of the number of events which occur in one minute is said to have a Poisson distribution with parameter λ. If X has a Poisson distribution then

$$P[X=r] = \exp(-\lambda)\frac{\lambda^r}{r!} \quad r = 0,1,2,\ldots$$

where λ>0. Note that

$$
\begin{aligned}
\sum_{r=0}^{\infty}\exp(-\lambda)\frac{\lambda^r}{r!} &= \exp(-\lambda)\sum_{r=0}^{\infty}\frac{\lambda^r}{r!} \\
&= \exp(-\lambda)\exp(-\lambda) \\
&= 1
\end{aligned}
$$

so this is a valid probability distribution.
It can be shown that $E(X) = \lambda, \ V(X) = \lambda$

Table (2) gives the probability that a Poisson random variable with mean λ will be greater or equal to r. These table can be used in the same way as the binomial tables. For example, suppose that X is a random variable with Poisson distribution with mean 2.0. Find
(1) $P(X = 2)$ (2) $P(X \geq 3)$ (3) $P(X < 2)$

$$P(X = 2) = P(X \geq 2) - P(X \geq 3) = 0.59399 - 0.32332 = 0.27067$$
$$P(X \geq 3) = 0.32332$$
$$P(X < 2) = 1 - P(X \geq 2) = 1 - 0.59399 - 0.40611$$

A property of the Poisson distribution is that if *X* is Poisson with mean λ then *kX* is Poisson with mean kλ. this can be useful in calculating probabilities of numbers of event in a time period different to that for which information is given.

Example: The number of particles emitted by radio active source in an hour has a Poisson distribution with mean $\frac{20}{30} = \frac{2}{3}$. Therefore the probability that one particle is emitted is a two minute period is

$$\exp(-\frac{2}{3})\frac{2}{3} = 0.342$$

Read that if *X* has a binomial distribution with parameters n and p that $E(X) = np$ and $Var[X] = np(1-p)$. Now if p is small then 1-p is close to one and np(1-p) $\approx np$. This suggest that if p is small we may be able to approximate *X* by a Poisson random variable with mean *np*. So long as *p* is small (may be < 0.1) and n is large (may be >50) a binomial distributed random variable is well approximated by a Poisson random variable of mean *np*.

Example: IF *X* has a binomial distribution, *n*=100, p=0.01 then from the tables

$$P(X \geq 1) = 0.63397$$
$$P(X \geq 2) = 026424$$
$$P(X = 1) = 0.36973$$

The corresponding quantities from the Poisson tables with λ=1 are

$$P(Y \geq 1) = 0.63212$$
$$P(Y \geq 2) = 0.26424$$
$$P(Y = 1) = 0.36788$$

Example: The probability that a car has defective gearbox is 0.02. If I check the gearboxes of 140 cars what is a suitable approximation to the probability that 1 find
2 defectives (b) more than 5 defectives (c) fewer than 4 defectives

Let *X* be the number of defective gearboxes that I find. Then *X* has a binomial distribution with n=140 and p=0.02. Since n is large and p is small a Poisson random variable Y with mean $\lambda = np = 2.8$ will give a good approximation to *X* tables

$(a)\ P(Y=2)=P(Y\geq 2)-P(Y\geq 3)=0.76892-0.53055=0.238$

$(b)\ P(Y>5)=P(Y\geq 6)=0.065$

$(c)\ P(Y<4)=1-P(Y\geq 4)=1-0.30806=0.692$

2.5 The discrete uniform distribution

Consider throwing a single die. If *Y* is the score on the upturned face then $P(Y=r)=\frac{1}{6}$ for r $=1,2,3,4,5,6$. This is a special case of a distribution called *discrete uniform distribution.* In general we consider a random variable *X* which is equally likely to take any integer value between two limits , a + 1 and b say. Then

$$P(X=r)=\frac{1}{b-a} \quad r=a+1,\ a+2,\ldots b$$

The case of a die is when a = 0 and b = 0.

The discrete uniform distribution is an example where the probability generating function is not useful. Instead we find the mean and variance directly.

For the mean

$$\begin{aligned}E[X] &= \sum_{r=a+1}^{b} r.\frac{1}{b-a}\\ &= \frac{1}{b-a}\sum_{r=a+1}^{b} r\\ &= \frac{1}{b-a}\left[\sum_{r=1}^{b} r-\sum_{r=1}^{a} r\right]\\ &= \frac{1}{b-a}\left[\frac{b(b+1)}{2}-\frac{a(a+1)}{2}\right]\\ &= \frac{b^2-a^2+b-a}{2(b-a)}\\ &= \frac{a+b+1}{2}\end{aligned}$$

Note for a die we obtain $E[Y]=\frac{7}{2}$ in agreement with a result found earlier

To find variance we first find $E[X^2]$

$$E[X^2] = \sum_{r=a+1}^{b} r^2 . \frac{1}{b-a}$$
$$= \frac{1}{b-a} \sum_{r=a+1}^{b} r^2$$
$$= \frac{1}{b-a} \left[\sum_{r=1}^{b} r^2 - \sum_{r=1}^{a} r^2 \right]$$
$$= \frac{1}{b-a} \left[\frac{b(b+1)(2b+1)}{6} - \frac{a(a+1)(2a+1)}{6} \right]$$
$$= \frac{2(b^3 - a^3) + 3(b^2 - a^2) + b - a}{6(b-a)}$$
$$= \frac{2(b^2 + ab + a^2) + 3(a+b) + 1}{6}$$

Thus

$$V[X] = E[X^2] - (E[x])^2$$
$$= \frac{2(b^2 + ab + a^2) + 3(a+b) + 1}{6} - \frac{(a+b+1)^2}{4}$$
$$= \frac{(b-a)^2 - 1}{12}$$

We have it as an exercise to missing algebra.

2.6 Exercise

A manufacturing process produces components which are free from any faults with probability p. Find the probability that in a sample of size 50 from a large batch there are fewer than 4 faulty components when $p = 0.95$. Find the probability that in a sample of size 50 there are fewer than 10 faulty when $p = 0.75$.

Use the table to give a suitable approximation to the probability that $X \geq 5$ where X is binomial random variable with parameters $p = 0.05$ and $n = 400$