"JUST THE MATHS"

UNIT NUMBER

19.7

PROBABILITY 7
(The Poisson distribution)

by

A.J.Hobson

UNIT 19.7 - PROBABILITY 7

THE POISSON DISTRIBUTION

19.7.1 THE THEORY

We recall that, in a binomial distribution of n trials, the probability, P_r, that an event occurs exactly r times out of a possible n is given by

$$P_r = \frac{n!}{(n-r)!r!}p^r q^{n-r},$$

where p is the probability of success in a single trial and $q = 1 - p$ is the probability of failure.

Now suppose that n is very large compared with r and that p is very small compared with 1.

Then,

(a)

$$\frac{n!}{(n-r)!} = n(n-1)(n-2).....(n-r+1) \simeq n^r.$$

ILLUSTRATION

If $n = 120$ and $r = 3$, then

$$\frac{n!}{(n-r)!} = \frac{120!}{117!} = 120 \times 119 \times 118 \simeq 120^3.$$

(b)

$$q^r = (1-p)^r \simeq 1,$$

so that

$$q^{n-r} \simeq q^n = (1-p)^n.$$

We may deduce that

$$P_r \simeq \frac{n^r p^r (1-p)^n}{r!} = \frac{(np)^r (1-p)^n}{r!}$$

or

$$P_r \simeq \frac{(np)^r}{r!} \left[1 - np + \frac{n(n-1)}{2!} p^2 - \frac{n(n-1)(n-2)}{3!} p^3 + \ldots \right]$$

$$\simeq \frac{(np)^r}{r!} \left[1 - np + \frac{(np)^2}{2!} - \frac{(np)^3}{3!} + \ldots \right].$$

Hence,

$$P_r \simeq \frac{(np)^r}{r!} e^{-np}.$$

The number, np, in this formula is of special significance, being the average number of successes to be expected a single set of n trials.

If we denote np by μ, we obtain the **"Poisson distribution"** formula,

$$P_r \simeq \frac{\mu^r e^{-\mu}}{r!}.$$

Notes:

(i) Although the formula has been derived from the binomial distribution, as an approximation, it may also be used in its own right, in which case we drop the approximation sign.

(ii) The Poisson distribution is more use than the binomial distribution when n is a very large number, the binomial distribution requiring the tedious evaluation of its various coefficients.

(iii) The Poisson distribution is of particular use when the average frequency of occurrence of an event is known, but not the number of trials.

EXAMPLES

1. The number of cars passing over a toll-bridge during the time interval from 10a.m. until 11a.m. is 1,200.

   (a) Determine the probability that not more than 4 cars will pass during the time interval from 10.45a.m. until 10.46a.m.

   (b) Determine the probability that 5 or more cars pass during the same interval.

   Solution

   The number of cars which pass in 60 minutes is 1200, so that the average number of cars passing, per minute, is $20 = \mu$.

   (a) The probability that not more than 4 cars pass in a one-minute interval is the sum of the probabilities for 0,1,2,3 and 4 cars.

   That is,

   $$\sum_{r=0}^{4} \frac{(20)^r e^{-20}}{r!} =$$

   $$\left[\frac{(20)^0}{0!} + \frac{(20)^1}{1!} + \frac{(20)^2}{2!} + \frac{(20)^3}{3!} + \frac{(20)^4}{4!} \right] e^{-20} =$$

   $$8221 e^{-20} \simeq 1.69 \times 10^{-5}.$$

   (b) The probability that 5 or more cars will pass in a one-minute interval is the probability of failure in (a). In other words,

   $$1 - \sum_{r=0}^{4} \frac{(20)^r e^{-20}}{r!} =$$

   $$1 - 8221 e^{-20} \simeq 0.99998$$

2. A company finds that, on average, there is a claim for damages which it must pay 7 times in every 10 years. It has expensive insurance to cover this situation.

   The premium has just been increased, and the firm is considering letting the insurance lapse for 12 months as it can afford to meet a single claim.

Assuming a Poisson distribution, what is the probability that there will be at least two claims during the year ?

Solution

Using $\mu = \frac{7}{10} = 0.7$, the probability that there will be at <u>most</u> one claim during the year is given by

$$P_0 + P_1 = e^{-0.7} + e^{-0.7}0.7 = e^{0.7}(1 + 0.7).$$

The probability that there will be at <u>least</u> two claims during the year is given by

$$1 - e^{-0.7}(1 + 0.7) \simeq 0.1558$$

3. There is a probability of 0.005 that a welding machine will produce a faulty joint when it is operated. The machine is used to weld 1000 rivets. Determine the probability that at least three of these are faulty.

Solution

First, we have $\mu = 0.005 \times 1000 = 5$.

Hence, the probability that at <u>most</u> two will be faulty is given by

$$P_0 + P_1 + P_2 = e^{-5}\left[\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!}\right].$$

That is,

$$e^{-5}[1 + 5 + 12.5] \simeq 0.125$$

Hence, the probability that at least three will be faulty is approximately

$$1 - 0.125 = 0.875$$

19.7.2 EXERCISES

1. The probability that a glass fibre will shatter during an experiment is believed to follow a Poisson distribution.

   In a certain apparatus, it was found that, on average, 7 glass fibres shattered.

   Determine the probability that, during a single demonstration of the experiment,

   (a) two glass fibres will shatter;
   (b) at least one glass fibre will shatter.

2. A major airline operates 350 flights per day throughout the world. The probability that a flight will be delayed for more than one hour, for any reason, is 0.7%. If more than four flights suffer such delays in one day, the implications for route organisation and crewing become serious. Calling such a day a "bad day", determine the probabilities that

   (a) any particular day is a bad day;
   (b) at most two bad days occur in one week;
   (c) more than 50 bad days occur in a year of 365 days.

   State your answers correct to three significant figures.

3. It is known that 3% of bolts made by a certain machine are defective. If the bolts are packaged in boxes of 50, determine the probability that a given box will contain 4 defectives.

4. If 0.04% of cars break down while driving through a certain tunnel, determine the probability that at most 2 break down out of 2000 cars entering the tunnel on a given day. State your answer correct to three significant figures.

19.7.3 ANSWERS TO EXERCISES

1. (a) 2.28×10^{-3}; (b) 0.0676

2. (a) 0.102; (b) 0.128; (c) 0.011

3. 0.047 using $\mu = 1.5$

4. 0.952 using $\mu = 0.8$