Lecture - 5

~~Chapters~~: Binomial & Poission Distributions

In this chapter we seek to study some specialized probability distributions, which has a lot of applicability. In this chapter in particular we shall focus on Binomial and Poission distributions which are important distributions for discrete random variables. For a discrete random variable X, ~~its~~ its probability distribution is simply an algebraic expression of its probability mass function. Of course special distributions are for particular type of random variables, describing special situations.

Section 1: Binomial Distributions

Suppose we make repeated coin tossing. Of course a fair coin is used. Suppose we toss a coin 100 times and ask the question that what is the probability that head appears 25 times? The Binomial Distribution answers such questions.

Binomial Distribution deals with repeated trials and each trial has ~~two~~ only two possible outcomes, which mark as success or failure. Each such trial is often called a Bernoulli trial. For example when we toss a coin for example we can consider the ~~appearan~~ appearance of head (H) as success and appearance of tail as failure. If we denote ~~sucess~~ success as S and failure as F, then any repeated random trial with ~~#~~ n trials ~~looks li~~ may look like the following cases

SSFF...FSSS.....S

So there could be k successes out of n trials. We ask ~~our selve~~ ourselves what is the probability of such an event if each success has a fixed probability p. Note when n trials are carried out k successes can appear in $\binom{n}{k}$ ways.

Thus if X denotes the random variable ~~of~~ denoting the number of succen, then

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

~~In fact S denot.~~

So the Binomial p.m.f is usually written as

$$\boxed{f_X(x) = \binom{n}{x} p^{x} q^{n-x}, \quad \text{where } q = 1-p}$$

Let us first check that f_X is indeed a probability mass function.

i) $f_X(x) \geq 0$, is of course clear as $p \geq 0$

ii) $$\sum_{x=1}^{n} f_X(x) = \sum_{x=1}^{n} \binom{n}{x} p^{x} q^{n-x}$$

$$= (p+q)^{n} = 1.$$ (Last step uses the Binomial Theorem)

So why it is called a Binomial distribution is clear.
Before providing ~~and~~ any example of a numerical problem let us compute two important measures associated with a random variable, namely expectation and variance.
Thus if X is a binomial random variable, then we will compute $E(X)$ & $Var(X)$. There are two approaches to it. We shall choose the approach using the moment generating function. Now the mgf of the binomial random variable is

$$m_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^{x} q^{n-x}$$

$$= (pe^{t} + q)^{n}.$$

Thus

$$\boxed{m_X(t) = \left(pe^t + q\right)^n}$$

$\therefore \quad m'_X(t) = n\left(pe^t + q\right)^{n-1} pe^t$

$\therefore \quad m'_X(0) = n\left(p+q\right)^{n-1} p$

$m'_X(0) = np \quad (\text{as } p+q=1)$

Hence $\quad \mu_X = E(X) = m'_X(0) = np.$

Now to compute the variance we compute $m''_X(0)$. So

$$m''_X(t) = n(n-1)\left(pe^t+q\right)^{n-2} pe^t\, pe^t + n\left(pe^t+q\right)^{n-1} pe^t$$

$$\therefore \quad m''_X(0) = n(n-1)\left(p+q\right)^{n-2} \cancel{(pe^t)}\, p^2 + n\left(p+q\right)^{n-1} p.$$

$$\therefore \quad m''_X(0) = n(n-1)p^2 + np$$

$$= (n^2 - n)p^2 + np$$

$$= n^2p^2 - np^2 + np$$

$$= n^2p^2 + np - np^2$$

$$= n^2p^2 + np(1-p)$$

$$= n^2p^2 + npq$$

$\therefore \quad E(X^2) = n^2p^2 + npq$

$\therefore \quad Var(X) = E(X^2) - [E(X)]^2 = n^2p^2 + npq - n^2p^2$

$$\therefore \quad \boxed{Var(X) = npq}$$

The whole idea of using m.g.f might appear counter-intuitive as if something is forced on to us, without motivation and does not seem to have immediate ~~motivation~~ connect with the definition. Of course we can definitely compute $E[X]$ and $Var(X)$ of a binomial random variable, using brute force, which we demonstrate below.

Brute force computation:

$$\mu_X = E(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x \cdot \frac{n!}{x!\,(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x q^{n-x}$$

$$\therefore \quad \mu_x = np \sum_{x=0}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} q^{n-x}$$

$$= np \sum_{x=0}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np\,(p+q)^{n-1}$$

$$= np.$$

However computing $E(X^2)$ using brute force technique is not so simple. There is a small trick. This is as follows.

$$E(X^2) = E(X^2 - X + X)$$

$$= E(X^2 - X) + E(X)$$

$$= E(X(X-1)) + E(X).$$

$$E\left(X(X-1)\right) = \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^{x} q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\frac{n!}{x!\,(n-x)!} p^{x} q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\frac{n!}{x(x-1)(x-2)!\,(n-x)!} p^{x} q^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n!}{(x-2)!\,(n-x)!} p^{x} q^{n-x}$$

$$= n(n-1)p^{2}\sum_{x=0}^{n} \frac{(n-2)!}{(x-2)!\,(n-2)!} p^{x-2} q^{n-x}$$

$$= n(n-1)p^{2}(p+q)^{n-2}$$

$$= n(n-1)p^{2}$$

$$\therefore \quad E(X^{2}) = n(n-1)p^{2} + np.$$

Hence Var(X) is given as

$$Var(X) = E(X^{2}) - [E(X)]^{2}$$

$$= n(n-1)p^{2} + np - n^{2}p^{2}$$

$$= n^{2}p^{2} - np^{2} + np - n^{2}p^{2}$$

$$= np(1-p) = npq.$$

The square root of the ~~Vart~~ variance of X is called the standard deviation of X. $\sigma_X = \sqrt{Var(X)}$. Let us now look into a numerical example, where we can see the application of binomial distribution.

Example 5.1: A jumbo-jet, say Boeing-747-400 has four engines, the operate independently. Let each engine has a probability ~~p~~ q of failure and for a succesful flight at least two engines should work. What is the probability that ~~we will~~ the flight will be succesful.

Solution: The probability of engine not failing is p

Let X be the number of engines working during the flight

We have to find the probability that $X \geq 2$, ie

$$P(X \geq 2) = 1 - P(X < 2)$$

Of course X follows a binomial distribution with $n = 4$ & probability of succes p. Thus

$$P(X \geq 2) = 1 - (P(X=0) + P(X=1))$$

$$= 1 - \left(\binom{4}{0} p^0 q^{4-0} + \binom{4}{1} p^1 q^{4-1} \right)$$

$$= 1 - \left[q^4 + 4pq^3 \right]$$

$$\therefore \boxed{P(X \leq 2) = 1 - q^4 - 4pq^3} \quad \square$$

Observe that if X is a binomial random variable, then its distribution depends on two parameters, the number of trials n and the probability of succes p. So often one write

$$\boxed{X \sim B(n, p)},$$

which means that X is a binomial random variable with parameters n & p.

Section 2: Poisson Distribution

Suppoce you are editing an essay with 5,000 words and you want to know ~~if there~~ what is the probability, that there are five spelling mistakes, given that the probability of ~~finding~~ committin the mistake is 0.005. If we consider that ~~find~~ occurence of a spelling mistake is a succes, then modelling through the Binomial distribution shows that

$$f_X(5) = \binom{5000}{5} (0.005)^5 (0.995)^{4995}$$

Of course this is a cumbersome computation. The question is can we make it simpler, We are in a situation where n is large and p small while the product np is moderate. Our question is such cases, what form would the Binomial pmf take. Trying to figure out this leads us to the Poisson distribution.

Let us assume the $np = \lambda > 0$, then setting $p = \frac{\lambda}{n}$, we can rewrite the Binomial distribution as

$$f_X(x) = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n!}{x!\,(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)(n-x)!}{x!\,(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

Thus

$$f_X(x) = \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{(x-1)}{n}\right)}{x!} \lambda^x \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

Now we need to see what happens to the right hand side when $n \to \infty$

We of course know the celebrated formula

$$\lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = e^z$$

Thus

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Further $\left(1-\frac{\lambda}{n}\right)^{-x} \to 1$ as $n \to \infty$. Hence the rhs becomes (as $n\to\infty$)

$$\boxed{\frac{\lambda^x e^{-\lambda}}{x!}}$$

This gives us a new kind of probability distribution, which we call a Poisson distribution with parameter λ.

$$\boxed{f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}}, \quad x = 0, 1, 2, 3, \ldots$$

Note that though x can take any non-negative value, in practice it is small, but even if it is large we can always use Stirlings approximation to $x!$. Let us now compute the mean and variance of the Poission ~~rand~~ random variable. In fact we write $X \sim \text{Poisson}(\lambda)$, i.e X follows a Poisson distribution with parameter λ. Let us use the m.g.f technique and leave the bruce force technique to the reader

$$m_X(t) = \sum_{x=0}^{\infty} \frac{e^{tx}\lambda^x e^{-\lambda}}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda}e^{\lambda e^t}$$

Now

$$m'_X(t) = e^{-\lambda}\lambda e^{\lambda e^t}e^t = \lambda e^{-\lambda}e^t e^{\lambda e^t}$$

$$\boxed{m'_X(0) = \mu_X = \lambda}$$

$$m_X''(t) = \lambda e^{-\lambda}\left[e^t e^{\lambda e^t} + \lambda e^t e^{\lambda e^t} e^t\right]$$

$$m_X''(0) = \lambda e^{-\lambda}\left[e^{\lambda} + \lambda e^{\lambda}\right]$$

$$\boxed{m_X''(0) = \lambda + \lambda^2}$$

Thus

$$\begin{aligned} Var(x) &= E(x^2) - (\mu_X)^2 \\ &= m_X''(0) - (\mu_X)^2 \\ &= \lambda + \lambda^2 - \lambda^2 \\ &= \lambda \end{aligned}$$

$$\boxed{Var(x) = \lambda}$$

I find this amazing, both mean and variance have the same value, λ the parameter of the Poisson distribution.

In his small book called "Facts from Figures", M. J. Moroney almost magically brings up the Poisson Distribution, for a discrete random variable X, with countable range, i.e.

$$\begin{aligned} \text{Range } X &= \{0, 1, 2, 3, \ldots\ldots\} \\ &= \mathbb{N} \cup \{0\}. \end{aligned}$$

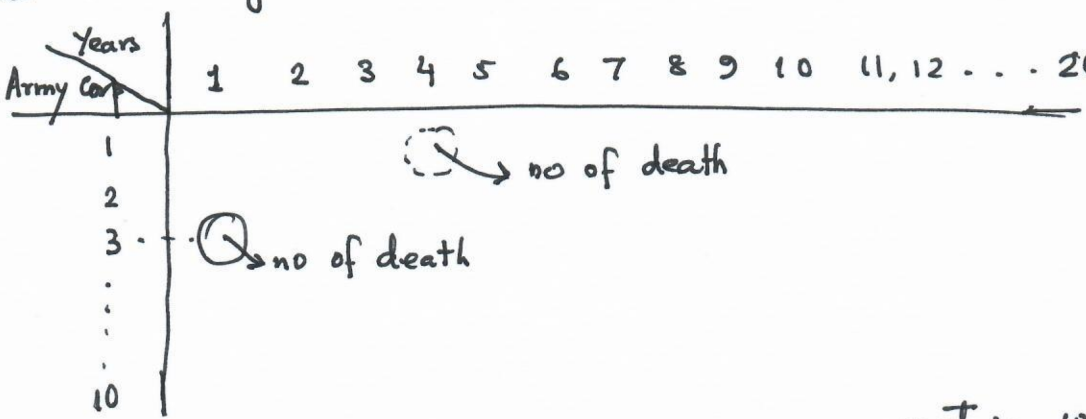
Observe that,

$$\cancel{1 = e^{-z} e^{z} = e^{-z}\left(1 + \frac{z}{1!} + \frac{z^2}{2!} + \right.}$$

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

$$1 = e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \cdots\cdots \frac{\lambda^n e^{-\lambda}}{n!} + \cdots\cdots$$

Observe that each term on the rhs is a term of the Poisson distribution for different values of x. In fact to show that how Poission distribution works effectively ~~Morney~~ Moroney tested it on ~~ha~~ a data collected by the great statistician R.A. Fisher. This a data on the death of a cavalryman getting killed by a horsekick in the course of a year. The data has been collected over twenty years from ten army corps (of Britain I guess) thus has 200 readings



So the data is represent is a 10 × 20 matrix, which has 200 entries. It can be summarized as follows

No of deaths	Frequency of occurence of such deaths [No cells in the above matrix with the given no]
0	109
1	65
2	22
3	3
4	1
5	0
6	0

$$\text{Total death} = (0\times109)+(1\times65)+(2\times22)+(3\times3)+(4\times1)+(5\times0)+(6\times0)$$

$= 122$ (in twenty years) among 200 observations

So average death per year per corps $= \frac{122}{200} = 0.61$

So $\lambda = 0.61$ in this case. Once we fix it we have $e^{-\lambda} \approx 0.543$. Let us assume that number of deaths is a random variable following Poission with the assumption that $\lambda = 0.61$ is fixed.

No of deaths	Poisson Prob.	Poisson freq = 200 × Poisson Prob	Actual
0	0.543	109	109
1	0.331	66.3	65
2	0.101	20.2	22
3	0.021	4.1	3
4	0.003	0.6	1

[Table is taken from page 98 of "Facts ~~and~~ from Figures" by M.J. ~~Morner~~ Moroney, Pelican, 1951]

So you see this data can be indeed ~~very~~ modelled very well by Poisson distributions.

"Thus statistics works"

—x—