Discrete Probability Distributions

Often, the observations generated by different statistical experiments have the same general type of behavior. Consequently, discrete r.v's associated with these experiments can be described by essentially the same probability distribution and therefore can be represented by a single formula. In fact, one needs only a handful of important prob. dist. to describe many of the discrete r.v's encountered in practice.

Such a handful of distributions describe several real life random phenomena. For instance, in a study involving testing the effectiveness of a new drug, the number of cured patients among all the patients who use the drug approximately follows a binomial distribution.

Binomial distribution

An experiment often consists of repeated trials, each with two possible outcomes that may be labelled as success or failure. For eg, testing of items as they come off an assembly line, where each trial may indicate a defective or a nondefective item. We may choose to define either outcome as a success. The process is called a Bernoulli process. Each trial is called a Bernoulli trial.

Bernoulli Process – It must possess the following properties–

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.

3. The probability of success, denoted by p, remains constant from trial to trial.

4. The repeated trials are independent.

Consider the set of Bernoulli trials where 3 items are selected at random from a manufacturing process, inspected & classified as defective or nondefective. A defective item is designated a success.

X : no. of successes assuming values 0 to 3.

Eight possible outcomes & corresponding X are.

Outcome	NNN	NDN	NND	DNN	NDD	DND	DDN	DDD
x	0	1	1	1	2	2	2	3

Since the items are selected independently & process produces 25% defectives,

$$P(NDN) = P(N)P(D)P(N) = \frac{3}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{64}$$

x	0	1	2	3
$f(x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

Binomial distribution -

The no. X of successes in n Bernoulli trials is called a binomial r.v. The probability distribution of this discrete r.v. is called the binomial distribution & it depends on the no. of trials & prob. of a success on a given trial.

Thus, for the prob. dist of X, the no. of defectives is

$$P(X=2) = f(2) = b\left(2; 3, \frac{1}{4}\right) = \frac{9}{64}$$

general

$b($

$\not{x}$

$\not{x}$

generalize the above illustration, to yield a formula for $b(x;n,p)$. ie., prob of x successes in n trials for a binomial experiment.

First consider the prob. of x successes and $n-x$ failures in a specified order. Since the trials are independent, each success occurs with prob. p & each failure with prob. $q = 1-p$.

∴ the prob. for a specified order is $p^x q^{n-x}$.

To determine the total no. of sample points in the experiment that have x successes and $n-x$ failures. This number is equal to the no. of partitions of n outcomes into two groups with x in one group & $n-x$ in the other & is written as ${}^nC_x$. Because these partitions are mutually exclusive, we add the probabilities of all different partitions to obtain the general formula, or simply multiply $p^x q^{n-x}$ by ${}^nC_x$.

* A Bernoulli trial can result in a success with prob. p and a failure with probability $q = 1-p$. Then the probability distribution of a binomial r.v. X, the no. of successes in n independent trials, is

$$P(X=x) = f(x) = b(x;n,p) = {}^nC_x \, p^x q^{n-x}, \quad x = 0,1,2,\ldots,n.$$

x & p are independent constants, known as parameters of the dist.

* for $n=3$, $p=1/4$, the prob. dist of X, the no. of defectives can be written as

$$b(x;3,\tfrac{1}{4}) = \binom{3}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x}, \quad x = 0,1,2,3.$$

Ex The prob that a certain kind of component will survive a shock test is 3/4. find the prob that exactly 2 of the next 4 components tested survive.

$$b(2;4,3/4) = {}^4C_2 \left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{4!}{2!2!}\frac{3^2}{4^4} = \frac{27}{128}$$

* $$(q+p)^n = {}^nC_0 q^n + {}^nC_1 pq^{n-1} + {}^nC_2 p^2 q^{n-2} + \dots + {}^nC_n p^n$$

$$= b(0;n,p) + b(1,n,p) + \dots + b(n;n,p)$$

Since $p+q=1$ $\sum_{x=0}^{n} b(x;n,p) = 1$, a condition that must hold for any prob. dist.

Ex The prob. that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the prob. that

a) at least 10 survive

let X be the no. of people who survive

$$P(X \geq 10) = 1 - P(X<10)$$

$$= 1 - \sum_{x=0}^{9} b(x;15,0.4) = 1-0.9662$$

$$= 0.0338$$

b) from 3 to 8 survive.

$$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x;15,0.4) = \sum_{x=0}^{8} b(x;15,0.4) - \sum_{x=0}^{2} b(x;15,0.4)$$

$$= 0.9050 - 0.0271 = 0.8779$$

c) $$P(X=5) = b(5;15,0.4)$$

$$= \sum_{x=0}^{5} b(x;15,0.4) - \sum_{x=0}^{4} b(x;15,0.4)$$

$$= 0.4032 - 0.2173$$

~~= 0.8779~~

~~$P(x=5) = b(5;15,0.4)$~~ $= \sum_{x=0}^{5} b(x;15,0.4) - \sum_{x=0}^{4} b(x;15,0.4)$

$= 0.4032 - 0.2173 = 0.1859.$

Ex A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%. The inspector picks 20 items randomly from a shipment. What is the prob. that there will be atleast one defective item among these 20 ?

let X : no. of defective devices among 20

X follow $b(x; 20, 0.03)$ distribution.

$$P(X \geq 1) = 1 - P(X=0) = 1 - b(0; 20, 0.03)$$

$$= 1 - (0.03)^0 (1-0.03)^{20-0} = 0.4562$$

Ex A and B play a game in which their chances of winning are in the ratio 3 : 2. find A's chance of winning atleast three games out of the five games played.

let p be the prob. that A wins the game.

$n = 5, \quad p = \frac{3}{5} \Rightarrow q = \frac{2}{5}$

let A wins 'x' games

$$P(X \geq 3) = b(3; 5, \tfrac{3}{5}) + b(4; 5, \tfrac{3}{5}) + b(5; 5, \tfrac{3}{5})$$

$$= {}^5C_3 \left(\tfrac{3}{5}\right)^3 \left(\tfrac{2}{5}\right)^2 + {}^5C_4 \left(\tfrac{3}{5}\right)^4 \left(\tfrac{2}{5}\right) + {}^5C_5 \left(\tfrac{3}{5}\right)^5$$

$$= \frac{3^3}{5^5} \left[{}^5C_3 \cdot 2^2 + {}^5C_4\, 3 \times 2 + {}^5C_5\, 3^2 \right]$$

$$= \frac{27}{3125} \left(40 + 30 + 9\right) = 0.68$$

$(x+y)^n = \sum_{k=0}^{n} {}^nC_k x^{n-k} y^k$

Mean and Variance of binomial distribution

$$\mu_1' = E(X) = \sum_{x=0}^{n} x f(x) = \sum x\, {}^nC_x p^x q^{n-x}$$

$${}^nC_x = \frac{n}{x}\, {}^{n-1}C_{x-1} \qquad \left({}^nC_x = \frac{n!}{(n-x)!\,x!} = \frac{n(n-1)!}{x(x-1)![n-1-(x-1)]!} = \frac{n}{x}\, {}^{n-1}C_{x-1} \right)$$

$$E(X) = np \sum_{x=1}^{n} {}^{n-1}C_{x-1} p^{x-1} q^{n-x}$$

$n-1-(x-1)$

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

$$E(X^2) = \sum_{x=0}^{n} x^2\, {}^nC_x p^x q^{n-x} = \sum_{x=0}^{n} [x(x-1)+x] \frac{n(n-1)}{x(x-1)}\, {}^{n-2}C_{x-2}\, p^x q^{n-x}$$

$$= n(n-1)p^2 \sum_{x=2}^{n} {}^{n-2}C_{x-2} p^{x-2} q^{n-x} + np$$

$n-2-(x-2)$

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

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

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

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

$$= npq$$

kth central moment

$$\mu_k = E[X - E(X)]^k, \quad E(X) = np$$

$$= \sum_{x=0}^{n} (x-np)^k\, {}^nC_x p^x q^{n-x}$$

$$\frac{d\mu_k}{dp} = \sum_{x=0}^{n} {}^nC_x \left[-nk(x-np)^{k-1} p^x q^{n-x} + (x-np)^k \{ x p^{x-1} q^{n-x} - (n-x) p^x q^{n-x-1} \} \right]$$

$$= -nk \sum_{x=0}^{n} {}^nC_x (x-np)^{k-1} p^x q^{n-x} + \sum_{x=0}^{n} {}^nC_x (x-np)^k p^x q^{n-x} \left(\frac{x}{p} - \frac{n-x}{q} \right)$$

$$= -nk\sum_{x=0}^{n}(x-np)^{k-1}p_x + \sum_{x=0}^{n}(x-np)^k p_x \frac{(x-np)}{pq}$$

$$= -nk\,\mu_{k-1} + \frac{1}{pq}\mu_{k+1}$$

$$\therefore \mu_{k+1} = pq\left(nk\,\mu_{k-1} + \frac{d}{dp}\mu_k\right)$$

Put $k = 1, 2, 3$

$$\mu_2 = pq\left(n\mu_0 + \frac{d\mu_1}{dp}\right) = npq \qquad (\because \mu_0 = 1,\ \mu_1 = 0)$$

$$\mu_3 = pq\left(2n\mu_1 + \frac{d\mu_2}{dp}\right) = pq\frac{d(npq)}{dp} = \cancel{pq\,nq}$$

$$= pq\frac{d}{dp}[\underbrace{np(1-p)}_{p-p^2}] = npq[1-2p] = npq(q-p)$$

$$\mu_4 = pq\left(3n\mu_2 + \frac{d\mu_3}{dp}\right) = pq\left[3n.\,npq + \frac{d}{dp}(np(1-p)(1-2p))\right]$$

$$= pq\left[3n^2pq + n\frac{d}{dp}(p - 3p^2 + 2p^3)\right]$$

$$= pq\left[3n^2pq + n(1 - \underbrace{6p + 6p^2}_{})\right] \quad 1-6p(1-p) \quad 1-6pq$$

$$= npq[3npq + 1 - 6pq]$$

$$= npq[1 + 3pq(n-2)]$$

Moment generating function of Binomial distribution.

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx}\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum_{x=0}^{n} {}^nC_x (e^t p)^x q^{n-x} = (q + pe^t)^n$$

Additive property of Binomial distribution

~~Let~~ $X \sim B(n_1, p_1)$ & $Y \sim B(n_2, p_2)$ ~~be in~~

$M_x(t) = (q_1 + p_1 e^t)^{n_1}$, $M_y(t) = (q_2 + p_2 e^t)^{n_2}$

$M_{x+y}(t) = M_x(t) \cdot M_y(t)$, if x & y are independent

$= (q_1 + p_1 e^t)^{n_1} (q_2 + p_2 e^t)^{n_2}$

Since it cannot be written in the form $(q + pe^t)^n$. Hence $M_{x+y}(t)$ is not a binomial variate. In other words, binomial distribution doesnot possess the additive or reproduction property.

However, if we take $p_1 = p_2 = p$ (say), then mgf is a binomial variate with parameters $(n_1 + n_2, p)$. Thus $X + Y \sim B(n_1 + n_2, p)$

Generalization – If X_i ($i = 1, 2, --, k$) are independent binomial variates with parameters (n_i, p) $i = 1, 2 --, k$, then their sum

$$\sum_{i=1}^{k} X_i \sim B\left(\sum_{i=1}^{k} n_i, p\right).$$

Ex If mean & variance of binomial distribution are 4 & 4/3 resp. Find $P(X \geq 1)$.

$E(X) = np = 4$, $npq = 4/3$ $\Rightarrow$ $q = 1/3$ and so $p = 2/3$

$n = \frac{4}{p} = 6$

$P(X \geq 1) = 1 - P(X = 0) = 1 - q^n = 1 - \left(\frac{1}{3}\right)^6 = 1 - \frac{1}{729}$

$= 0.998$

Ex A target is to be destroyed in a bombing exercise. There is 75% chance that any one bomb will strike the target. Assume that two direct hits are required to destroy the target completely. How many bombs must be dropped in order that the chance of destroying the target is $\geq 99\%$?

X = r.v representing no of bombs to be used

p = prob that bombs hit the target = $3/4$, $q = 1/4$

n = no. of bombs reqd to destroy the target

$X \sim b(x; n, 3/4)$

$$P(X \geq 2) = 0.99 \quad \text{or} \quad 1 - P(X \leq 1) \geq 0.99$$

$$1 - [P(X=0) + P(X=1)] \geq 0.99$$

$$\left(\frac{1}{4}\right)^n + n\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{n-1} \leq 0.01$$

$$1 + 3n \leq (0.01)\, 4^n$$

$$\Rightarrow \quad n = 6$$

If for an experiment, a r.v X is the no of - -

Poisson distribution 'Pwa-son'

Experiments yielding numerical values of a r.v. X, the no. of outcomes occuring during a given time interval or in a specified region are called Poisson experiments.

The given time interval may be of any length, such as a minute, a day, a week, a month or even a year. For eg: Poisson experiment can generate observations for a r.v. X representing the no. of telephone calls received per hour by an office, the no. of days school is closed due to snow during the winter or the no. of games postponed due to rain. The specified region could be a line segment, an area, a volume, or a piece of material. X might represent the no. of typing errors per page, no. of bacteria in a given culture.

more Ex: No. of defectives items out of lots produced in a large factory.

No. of vehicles passing/min through a particular traffic junction.

Poisson experiment possesses following properties -

1. The no. of outcomes occurring in one time interval or specified region of space is independent of the no. that occur in any other disjoint time interval or region.

2. The prob. that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region & does not depend on the no. of ~~succes~~ outcomes occurring outside this time interval or region.

3. The prob that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

The no. X of outcomes occurring during a Poisson experiment is called a Poisson r.v. and its prob. dist is called Poisson distribution. It can be used under the following conditions.

(i) Each trial results in mutually exclusive outcomes termed as success or a failure.

(ii) n, the no of trials is very large, i.e., $n \to \infty$

(iii) p - prob of success, is very small, i.e., $p \to 0$

(iv) $np = \lambda$, (λ : +ve real no.), is finite.

$\therefore \quad p = \frac{\lambda}{n} \quad \& \quad q = 1 - (\lambda/n)$

A r.v. X is said to follow Poisson dist. if it assumes only non-negative values & its pmf is

$$P(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!}, & x = 0, 1, 2 \ldots, \quad \lambda > 0 \\ 0 & \text{o.w.} \end{cases}$$

λ is a parameter here.

2. $\sum_{x=0}^{\infty} P(X=x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$

2. The distribution function is given by

$$F(x) = P(X \le x) = \sum_{r=0}^{x} p(r) = e^{-\lambda} \sum_{r=0}^{x} \frac{\lambda^r}{r!}, \quad x = 0, 1, 2, \ldots$$

3. Unlike in binomial dist, Poisson dist occurs when there are events which do not occur as outcomes of a definite no of trials.

4. Binomial probabilities may be computed approx. by computing the corresponding Poisson probabilities, whenever n is large and p is small.

Let X be binomially distributed with parameters n & p

Then for $r = 0, 1, 2, \ldots, n$ & $p = \lambda/n$,

$$P(X=x) = {}^nC_x \, p^x q^{n-x} = \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \left[1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

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

when n is finite

Taking $n \to \infty$,

$$\lim_{\substack{n\to\infty \\ np=\lambda=\text{finite}}} P(X=x) = \frac{\lambda^x}{x!}(1)e^{-\lambda} = \frac{e^{-\lambda}\lambda^x}{x!}$$

Which is the pmf of Poisson variate X.

Mean & variance of the Poisson distribution

$$E(X) = \sum_{x=0}^{\infty} x\, p(x) = \sum_{x=0}^{\infty} \frac{x\, e^{-\lambda} \lambda^{x}}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

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

$$E(X^2) = \sum_{x=0}^{\infty} x^2 p(x) = \sum_{x=0}^{\infty} \frac{x^2 e^{-\lambda} \lambda^{x}}{x!} = \sum_{x=0}^{\infty} \frac{[x(x-1)+x] e^{-\lambda} \lambda^{x}}{x!}$$

$$= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \underbrace{\lambda}_{\text{from } E(X)} = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$Var(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$\therefore$ for Poisson dist, mean = variance = λ.

Recurrence formula for the moments of PD

kth central moment $\mu_k = E[X - E(X)]^k$

$$= E(X-\lambda)^k$$

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

$$\frac{d\mu_k}{d\lambda} = \sum_{r=0}^{\infty} \frac{1}{x!} \left[e^{-\lambda} \left\{ -k(x-\lambda)^{k-1} \lambda^{x} + (x-\lambda)^k (x\lambda^{x-1} - \lambda^{x}) \right\} \right]$$

$$= -k\mu_{k-1} + \sum_{x=0}^{\infty} \frac{(x-\lambda)^{k+1} e^{-\lambda} \lambda^{x-1}}{x!}$$

$$= -k\mu_{k-1} + \frac{1}{\lambda}\mu_{k+1}$$

$$\mu_{k+1} = \lambda\left[k\mu_{k-1} + \frac{d\mu_k}{d\lambda} \right]$$

$\mu_0 = 1$

$\mu_1 = 0$

$$\mu_2 = \lambda\left[\mu_0 + \frac{d\mu_1}{d\lambda} \right] = \lambda$$

$$\mu_3 = \lambda\left[2\mu_1 + \frac{d\mu_2}{d\lambda} \right] = \lambda, \quad \mu_4 = \lambda\left[3\mu_2 + \frac{d\mu_3}{d\lambda} \right] = \lambda(3\lambda+1)$$

ficient of skewness

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}$, $\gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}$

kurtosis $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1}{\lambda}$, $\gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}$

Hence PD is always a skewed dist.

As $\lambda \to \infty$, $\beta_1 = 0$ & $\beta_2 = 3$

mgf.

Remark

If $X_1, X_2, \ldots, X_n$ are independent Poisson variate with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $\sum_{i=1}^{n} X_i$ is also a Poisson variate with parameters $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Let $X_1 \sim P(\lambda_1)$, $X_2 \sim P(\lambda_2)$, $\ldots$, $X_n \sim P(\lambda_n)$

Then $M_{X_i}(t) = e^{\lambda_i(e^t-1)}$, $i = 1, 2, \ldots, n$

Now
$$M_{X_1+X_2+\ldots+X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \ldots M_{X_n}(t)$$
$$= e^{\lambda_1(e^t-1)}\, e^{\lambda_2(e^t-1)} \ldots e^{\lambda_n(e^t-1)}$$
$$= e^{(\lambda_1+\lambda_2+\ldots+\lambda_n)(e^t-1)} = e^{\lambda(e^t-1)}$$

since X_i are independent & $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Hence $\sum X_i$ ~~are~~ is also a Poisson variate with parameter $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

* The converse of this result is also true., ie., if $X_1, X_2 \ldots X_n$ are independent & $\sum X_i$ ~~are~~ has a Poisson dist, then each of the r.v. has a Poisson dist.

Let X_1 & X_2 are independent r.v.

$X_1 \sim P(\lambda_1)$ & $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$ To prove $X_2 \sim P(\lambda_2)$

$M_{X_1+X_2}(t) = M_{X_1}(t)\, M_{X_2}(t)$

$e^{(\lambda_1+\lambda_2)(e^t-1)} = M_{X_2}(t)\, e^{\lambda_1(e^t-1)}$ $\Rightarrow M_{X_2}(t) = e^{\lambda_2(e^t-1)}$

$X_2 \sim P(\lambda_2)$

* The difference of two independent Poisson variate is not Poisson variate propoti

→ Mgf of PD

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\left[1+\lambda e^t + \frac{(\lambda e^t)^2}{2!} + \cdots\right]$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

Ex A manufacturer of pins knows that 5% of his products is defective. If he sells pins in boxes of 100 & guarantees that not more than 10 pins will be defective what is the approximate prob. that a box will fail to meet the guaranteed quality?

$n = 100$, p = prob of defective pin $= 5\% = 0.05$
(since p is small, we can use PD)

λ = Mean no. of defective pins $= np = 100 \times 0.05 = 5$

X : no. of defective pins in a box of 100

Prob of x defective pins $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-5}5^x}{x!}$, $x = 0, 1, 2 \ldots$

$$P(X>10) = 1 - P(X \le 10) = 1 - \sum_{x=0}^{10} \frac{e^{-5}5^x}{x!} = 1 - e^{-5}\sum_{x=0}^{10}\frac{5^x}{x!}$$

$e = 2.718$

$= 1 - 0.9863$

Ex A car hire firm has two cars, which it hires out day by day. The no. of demands for a car on each day is distributed as PD with mean 1.5
Calculate the proportion of days on which

(i) neither car is used.

X : no. of demands for a car on any day.

$\lambda = 1.5$

x

proportion of days on which there are x demands for a car is

$$P(X=x) = \frac{e^{-1.5}(1.5)^x}{x!}, \quad x=0,1,2,\ldots$$

(i) $P(X=0) = e^{-1.5} = 1 - 1.5 + \frac{(1.5)^2}{2!} - \frac{(1.5)^3}{3!} \ldots = 0.2231$

(ii) Proportion of days on which some demand is refused is

$$P(X>2) = 1 - P(X\leq 2) = 1 - [P(X=0) + P(X=1) + P(X=2)]$$

$$= 1 - e^{-1.5}\left[1 + 1.5 + \frac{(1.5)^2}{2}\right] = 1 - 0.2231 \times 3.625$$

$$= 0.19126$$

Recurrence formula for probabilities of PP

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

& $$p(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}, \quad x=0,1,2\ldots$$

$$\therefore \frac{p(x+1)}{p(x)} = \frac{\lambda}{x+1} \quad \text{or} \quad p(x+1) = \frac{\lambda}{x+1}p(x)$$

This formula is called "fitting of PD".

If we know, $p(0) = e^{-\lambda}$, then we can calculate all other prob. $p(1), p(2) \ldots$

Ex After correcting 50 pages of the proof of a book, the proof reader finds that there are, on the average 2 errors per 5 pages.

How many pages would you expect to find with 0, 1, 2, 3 & 4 errors, in 1000 pages of the first print of the book? (Given $e^{-0.4} = 0.6703$)

X . no. of errors per page

$\lambda = \frac{2}{5}$ mean errors per page

$= 0.4$

Prob of x errors per page

$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.4}(0.4)^x}{x!}, \quad x = 0, 1, 2 \ldots$

Expected no. of pgs with x errors per page in a book of $N = 1000$ pages.

$f(x) = N \times P(X=x) = 1000 \times \frac{e^{-0.4}(0.4)^x}{x!}, \quad x = 0, 1, 2 \ldots$

No. of errors per page (x)	$p(x)$ probability	Expected no. of pages $= Np(x)$
0	$p(0) = e^{-0.4} = 0.6703$	$670.3 \simeq 670$
1	$p(1) = \frac{0.4}{0+1} p(0) = 0.26812$	$268.12 \simeq 268$
2	$p(2) = \frac{0.4}{1+1} P(1) = 0.053624$	$53.624 \simeq 54$
3	$p(3) = \frac{0.4}{2+1} P(2) = 0.0071298$	$7.1298 \simeq 7$
4	$p(4) = \frac{0.4}{3+1} p(3) = 0.00071298$	$0.71298 \simeq 1$

Ex If the independent r.v's X & Y are binomially distributed resp with $n_1 = 3$, $p = 1/3$ & $n_2 = 5$, $p = 1/3$

Write the prob that $X + Y \geq 1$

$X \sim B(3, 1/3)$, $Y \sim B(5, 1/3)$

$X+Y \sim B(3+5, 1/3)$

$X+Y \sim B(8, 1/3)$

$P(X+Y=k) = {}^8C_k (1/3)^k (2/3)^{8-k}$

$P(X+Y \geq 1) = 1 - P(X+Y < k) = 1 - P(X+Y=0) = 1 - {}^8C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^8$

$= 1 - 1 \times 1 \times \left(\frac{2}{3}\right)^8$

$= 1 - \left(\frac{2}{3}\right)^8$

no. of exce

Bin $E(x)$ 9

Two dice are thrown 120 times. Find the average no. of times in which the no. on the first dice exceeds the no. on the second dice.

$n = 120$

$p = ?$

(1,1), ... (1,6)

(2,1) .... (2,6)

(3,1) ... (3,6)

⋮

(6,1) ... (6,6)

$S = \{(2,1), (3,1), (4,1) (5,1) (6,1), (3,2) (4,2) (5,2) (6,2), (4,3) (5,3) (6,3), (5,4), (6,4), (6,5)\}$

15 pts.

$P(\text{success}) = \frac{15}{36} = \frac{5}{12}$

$X \sim B(n, p) = B(120, 5/12)$

$E(x) = np = 120 \times \frac{5}{12} = 50.$

Bin

Ex Six dice are thrown 720 times. How many times do you expect atleast 3 dice to show 5 or 6?

X : No. of dice that shows 5 or 6

$n = 6$

$p = 2/6$

$X \sim B(n=6, p=2/6)$

$P(X \geq 3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$

$= 1 - \left[6C_0\left(\frac{2}{6}\right)^0\left(\frac{4}{6}\right)^6 + 6C_1\left(\frac{2}{6}\right)^1\left(\frac{4}{6}\right)^5 + 6C_2\left(\frac{2}{6}\right)^2\left(\frac{4}{6}\right)^4\right]$

$= 1 - \frac{4^4}{6^4}\left[\frac{16 \cancel{\times 36}}{36} + 6 \times \frac{2}{6} \times \frac{4}{6} + \frac{6!}{4!2!} \times \frac{4}{36} \times 1\right].$

$= 1 - \frac{4^4}{6^4}\left(\frac{4}{9} + \frac{2}{3} + \cancel{15\times} \frac{5}{3}\right] = 1 - \frac{4^4}{6^4}\left[\frac{4+6+15}{9}\right]$

$= 1 - \frac{4^4}{6^4} \times \frac{25}{9} = 1 - \frac{256 \times 25}{1296 \times 9} = 1 - \frac{16 \times 25}{81 \times 9} = 1 - \frac{400}{729} =$

324 that 64 16 81 324

Ex The average no. of phone calls/min coming into a switch board b/w 2 pm and 4 pm is 2.5. find the prob. that during one particular minute there will be (i) 4 or fewer (ii) more than 6 calls

Prob first

$\lambda = 2.5$

X : no of phone calls in a min

$X \sim P(\lambda)$ ie $X \sim P(2.5)$

(i) $P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$

$P(X \le 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$

$= e^{-\lambda} + e^{-\lambda}\lambda + \dfrac{e^{-\lambda}\lambda^2}{2!} + \dfrac{e^{-\lambda}\lambda^3}{3!} + \dfrac{e^{-\lambda}\lambda^4}{4!}$

$= e^{-\lambda}\left[1 + \lambda + \dfrac{\lambda^2}{2!} + \dfrac{\lambda^3}{3!} + \dfrac{\lambda^4}{4!}\right] = 0.912$

(ii) $P(X > 6) = 1 - P(X \le 6)$

$= 1 - [P(X \le 4) + P(X=5) + P(X=6)]$

$= 1 - 0.9858$

Geometric distribution

Prob. dist. for the no of trials required for the first success. for eg. tossing of a coin until a head occurs. To find the prob. that the first head occurs on the fourth toss.

A r.v. X {no. of trials} until the first success) is said to have a geometric dist. if it assumes non-negative values & its pmf is

$g(x, p)$

$$p(X=x) = pq^{x-1}, \quad x = 1, 2 \ldots$$

- the trials are independent Bernoulli trials
- & 'p' remains same in each trial.

* Since the successive terms constitute a geometric progress. hence the name geometric distribution.

* $\sum_{1}^{\infty} P(X=x) = \sum_{1}^{\infty} pq^{x-1} = P(1+q+q^2+\ldots) = \frac{P}{1-q} =$

Ex For a certain manufacturing process, it is known that, on an average, 1 in every 100 items is defective. What is the prob. that the fifth item inspected is the first defective found.

$x=5, \quad p=0.01$

$$g(5, 0.01) = (0.01)(0.99)^4 = 0.0096.$$

Mean & Variance of GD

$f(x) = P(X=x) = pq^{x-1}, \quad x = 1, 2 \ldots$

Expected value

$$E(X) = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} x\, p q^{x-1}$$

$$= p \sum_{1}^{\infty} x q^{x-1}$$

$$= p \sum \frac{d(q^x)}{dq}$$

$$= p \frac{d}{dq} \sum_{1}^{\infty} q^x$$

$$= p \cdot \frac{d}{dq} [q + q^2 + q^3 + \ldots] \quad \text{GD}$$

$S_n = \frac{a}{1-r}$

$$= p \frac{d}{dq} \frac{q}{1-q}$$

$$= p \left[\frac{(1-q) + q}{(1-q)^2} \right] = \frac{p}{(1-q)^2}$$

$$= \frac{1}{p}$$

$$E(X) = \frac{1}{p}$$

Variance $\quad V(X) = E(X^2) - E(X)^2$

$$= E(X(X-1) + X) - E(X)^2$$

$$= \underbrace{E(X(X-1))}_{?} + E(X) - E(X)^2$$

$$E(X(X-1)) = \sum_{1}^{\infty} x(x-1)\, p q^{x-1} = p \sum_{1}^{\infty} x(x-1) q^{x-1}$$

$$= p \sum_{1}^{\infty} \frac{d}{dq} [(x-1) q^x] = p \frac{d}{dq} \left(\sum_{1}^{\infty} (x-1) q^x \right)$$

$$= p \frac{d}{dq}\left(q^2 \sum_{2}^{\infty} (x-1) q^{x-2} \right)$$

$$= p \frac{d}{dq}\left(q^2 \sum_{2}^{\infty} \frac{d}{dq} q^{(x-1)} \right)$$

$$= p \frac{d}{dq}\left(q^2 \frac{d}{dq} (q + q^2 + q^3 + \cdots) \right)$$

$$= p \frac{d}{dq}\left(q^2 \frac{d}{dq}\left(\frac{q}{1-q} \right) \right) = p \frac{d}{dq} q^2 \left(\frac{1}{(1-q)^2} \right)$$

$$= p \left[\frac{2(1-q)^2 q + 2q^2 (1-q)}{(1-q)^4} \right]$$

$$= p(1-q) \left[\frac{2q - \cancel{2q^2} + \cancel{2q^2}}{(1-q)^4} \right]$$

$$= \frac{p \cdot 2q}{p^3} = \frac{2q}{p^2}$$

$$V(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q + p - 1}{p^2} = \frac{2q - q}{p^2}$$

$$= \frac{q}{p^2}$$

$$\boxed{\text{Var } X = \frac{q}{p^2}}$$

mgf of GD

$$M_X(t) = E(e^{tX}) = \sum_{1}^{\infty} e^{tx}\, p\, q^{x-1}$$

$$= \frac{p}{q}\sum_{1}^{\infty}(e^t q)^x = \frac{p}{q}\, e^t q\,(1-qe^t)^{-1}$$

$$= \frac{pe^t}{1-qe^t}.$$