The Binomial Distribution

References:

[1] MILLER & FREUND'S, PROBABILITY AND STATISTICS FOR ENGINEERS, NINTH EDITION ,Richard A. Johnson, Pearson Ed., 2018

[2] http://www.universityofcalicut.info/SDE/Probability_distribn_4june2015.pdf

Introduction

We are interested in the probability of getting x **successes** in n **trials**, or, in other words, x successes and $n - x$ failures in n attempts.

……………………………
A series of trials which we refer to as **Bernoulli trials** if the following assumptions hold:

1. **There are only two possible outcomes for each trial** (arbitrarily called "success" and "failure," without inferring that a success is necessarily desirable).
2. **The probability of success is the same for each trial**.
3. **The outcomes from different trials are independent**.

………………………………………

Example: Drivers stopped at a roadblock will be checked for failure to wear a seatbelt. Can it be treated as Bernoulli trials?

Solution. There are only two outcomes, and we call not wearing a seatbelt a success. (Success in this context does not mean success in life.)

If all cars are treated alike, their drivers would all have the same probability of not wearing a seatbelt.

If drivers are grouped by age, you may need different probabilities for persons under 20 than for those 30 to 40 years old. Then you would not have Bernoulli trials.

The results on seatbelt wear, for different drivers, should be independent.

There is no obvious common cause.

If someone caught without a seatbelt were to inform oncoming cars about the checkpoint, that would introduce dependence. //

…………………..
In the Bernoulli trials, let us consider the additional assumption that the number of trials is fixed in advance, that is,

4. **There are a fixed number *n* of Bernoulli trials conducted**.

Binomial Distribution
Let X be the random variable that equals the number of successes in n trials.

To obtain probabilities concerning X, we proceed as follows:

If p and $1 - p$ are the probabilities of success and failure on any one trial, then the probability of getting x successes and $n-x$ failures, in some specific order, is

$$p^x(1-p)^{n-x}.$$

Clearly, in this product of p's and $(1-p)$'s there is one factor p for each success, one factor $1-p$ for each failure.
The x factors p and $n-x$ factors $1-p$ are all multiplied together by virtue of the generalized multiplication rule for more than two independent events.

Since this probability applies to any point of the sample space that represents x successes and $n-x$ failures (in any specific order), we have only to count how many points of this kind there are, and then multiply $p^x(1-p)^{n-x}$ by this.

The number of ways in which we can select the x trials on which there is to be a success is *C(n, x)*, the number of combinations of x objects selected from a set of n objects.

Multiplying, we arrive at the following result:
………………

$$b(x; n, p) = \binom{n}{x} p^x(1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n$$

………………

This probability distribution is called the **binomial distribution** because for $x = 0, 1, 2, \ldots,$ and n, the values of the probabilities are the successive terms of the binomial expansion of $[p+(1-p)]^n$.

For the same reason, the combinatorial quantities *C(n, x)* are referred to as **binomial coefficients.**

……………………..

Example. It has been claimed that in 60% of all solar-heat installations the utility bill is reduced by at least one-third. Accordingly, what are the probabilities that the utility bill will be reduced by at least one-third in
(a) four of five installations;
(b) at least four of five installations?

Solution. **(a)** Substituting $x = 4$, $n = 5$, and $p = 0.60$ into the formula for the binomial distribution, we get

$$b(4; 5, 0.60) = C(5,4)\,(0.60)^4(1-0.60)^{5-4} = 0.259;$$

(b) Substituting $x = 5$, $n = 5$, and $p = 0.60$ into the formula for the binomial distribution, we get

$b(5;\ 5,\ 0.60) = C(5, 5)\,(0.60)^5(1 - 0.60)^{5-5}$
$\quad = 0.078$

and the answer is $b(4;\ 5,\ 0.60) + b(5;\ 5,\ 0.60) = 0.259 + 0.078 = 0.337.$ //

……………………………

Example.

Let X have the binomial distribution with probability distribution

$$b(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, \ldots, n$$

Show that

(a) $M(t) = (1 - p + pe^t)^n$ for all t

(b) $E(X) = np$ and $Var(X) = np(1-p)$

Solution.

(a) By definition of the moment generating function

$$\begin{aligned} M(t) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x} \\ &= (pe^t + 1 - p)^n \qquad \text{for all } t \end{aligned}$$

where we have used the binomial formula

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

(b) Differentiating $M(t)$, we find

$$M'(t) = npe^t(pe^t + 1 - p)^{n-1}$$
$$M''(t) = (n-1)np^2e^{2t}(pe^t + 1 - p)^{n-2} + npe^t(pe^t + 1 - p)^{n-1}$$

Evaluating these derivatives at $t = 0$, we obtain the moments

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

Also, the variance is

$$Var(X) = E(X^2) - [E(X)]^2 = np(1-p)$$ ■

Additive property of the binomial distribution

If X is a B (n_1, p) and Y is B (n_2, p) and they are independent then teir sum X + Y also follows B$(n_1 + n_2, p)$.

Proof:

Since,X ~ B(n_1, p), $M_x(t) = (q + pe^t)^{n1}$

Y ~ B(n_2, p), $M_y(t) = (q + pe^t)^{n2}$

We have,

$$M_{x+y}(t) = M_x(t).M_y(t)$$

, since X and Y are independent.

$$= (q+pe^t)^{n1} \times (q+pe^t)^{n2}$$

$$= (q + pe^t)^{n1+n2}$$

$$= mgf of B(n_1 + n_2, p)$$

There fore,

$$X + Y \sim B(n_1 + n_2, p)$$

Recurrence relation for central moments

If $X \sim B(n, p)$, then

$$\mu_{r+1} = pq[nr\mu_{r-1} + \frac{d\mu_r}{dp}]$$

Proof: We have

$$\mu_r = E[X - E(X)]^r$$

$$= E(X - np)^r$$

$$= \sum_0^n (x-np)^r \, {}^nC_x p^x q^{n-x}$$

Therefore,

$$\frac{d\mu_r}{dp} = \sum_0^n [\, {}^nC_x p^x q^{n-x} \, r(x-np)^{r-1}(-n) +$$

$$nCx(x-np)^r \; q^{n-x} x p^{x-1} + {}^n C_x (x-np)^r p^x (n-x) q^{n-x-1}(-1)]$$

$$= -nr\sum_0^n (x-np)^{1-r} \, {}^nC_x p^x q^{n-x} + \sum_o^n (x-np)^r \, {}^nC_x p^x q^{n-x} [\frac{x}{p} - \frac{n-x}{1-p}]$$

$$= -nr\,\mu_{r-1} + \frac{1}{pq}\sum_0^n (x-np)^{r+1} f(x)$$

i.e., $\frac{\mu_r}{dp}$ =-nr μ_r + $\frac{1}{pq}$ μ_{r+1}

There fore,

$$\mu_{r+1} = pq[nr\, \mu_{r-1} + \frac{d\mu_r}{dp}]$$

Using the information , μ_0 = 1 and $\mu_1 = 0$ .Also we can determine the values of μ_2 , μ_3 , μ_4 etc. by this relation.

Recurrence relation for Binomial Distribution

$$B(x + 1; n, p) = \frac{n-x}{x+1}\frac{p}{q}B(x; n, p)$$

Proof:

We have,

$$B(x; n, p) = {}^nC_x\, p^x q^{n-x}$$

$$B(x+1;n,p) =$$

$${}^nC_{x+1}p^{x+1}q^{n-(x+1)}$$

$$\frac{B(x+1;n,p)}{B(x;n,p)} = \frac{{}^nC_{x+1} \quad p^{x+1} \quad q^{n-x-1}}{{}^nC_x \quad p^x \quad q^{n-x}}$$

$$= \frac{n!}{(x+1)!(n-x-1)!} \times \frac{x!(n-x)!}{n} \times \frac{p}{q}$$

$$= \frac{(n-x)}{x+1}\frac{p}{q}$$

There fore,

$$B(x+1;n,p) = \frac{(n-x)}{x+1}\frac{p}{q}B(x;n,p)$$

ANNEXURE-BD
Direct Method:

a) Mean of binomial distribution: $\mu = n \cdot p$

b) Variance of binomial distribution: $\sigma^2 = n \cdot p \cdot (1 - p)$

c)

A Binomial Distribution is positively skewed, symmetric or negatively skewed according as γ_1 >=< 0 which implies q >=< p.

d)

A Binomial Distribution is leptokurtic, mesokurtic or platykurtic according as β_2 >=< 3 which implies pq > = < $\frac{1}{6}$

e) Find the mgf of Binomial distribution.

Solution.

Mean and variance: Mean,

$$E(X)=\sum_x xf(x)$$

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

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

$$np[(p+q)^{n-1}]$$

$$= np[(1)^{n-1}] = np$$

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

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

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

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

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

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

There fore the variance,

$$V(X) = E(X^2) - [E(X)]^2$$

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

$$= n^2p^2 - np_2 + np - n^2p^2$$

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

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

$$E(X^3) = \sum_{x=0}^{n} x^3 x^3[{}^nC_x p^x q^{n-x}]$$

$$=\sum_{x=0}^{n}[x(x-1)(x-2)+3x^2 -2x][\frac{n!}{x!(n-x)!}p^x q^{n-x}$$

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

$$\sum_{x=0}^{n} 3x^2 \frac{n!}{x!(n-x)!} p^x q^{n-x} - \sum_{x=0}^{n} 2x \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= n(n - 1)(n - 2)p^3\sum_{x=0}^{n}[\frac{(n-3)!}{(x-3)!(n-x)!}p^{x-3} + q^{n-x}] + 3E(X^2)-2E(X)$$

$$[n(n - 1)(n-2)p^3 [(p+q)^{n-3}] + 3[n(n-1)p2 +np]- 2np$$

$$= n(n-1)(n-2)p^3 + 3[n(n-1)p^2] + np$$

Similarly,

$$E(X^4) = n(n-1)(n-2)(n-3)p^4 + 3[n(n-1)(n-2)p^3] + 7n(n-1)p^2 + np$$

Beta and Gamma coefficients:

$$\beta_1 = \mu_3^2 = \frac{(q-p)^2}{npq}$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}}$$

A Binomial Distribution is positively skewed, symmetric or negatively skewed according as γ_1 >=< 0 which implies q >=< p.

$$\beta_2 = \frac{\mu 4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

A Binomial Distribution is leptokurtic, mesokurtic or platykurtic according as β_2 >=< 3 which implies pq > = <$\frac{1}{6}$

The mgf,

$$M_x(t) = E(e^{tx})$$

$$=\sum_x e^{tx} P(X=x)$$

$$=\sum_{x=0}^{n} e^{tx} [{}^nC_x p^x q^{n-x}]$$

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

$$(q+pe^t)^n$$

……………………

Assignment- Binomial Distribution:

4.7 Prove that $b(x;\ n,\ p) = b(n - x;\ n,\ 1 - p)$.

4.8 Prove that $B(x;\ n,\ p) = 1 -$on $B(n - x - 1;\ n,\ 1 - p)$.

4.24 Suppose that, next month, the quality control division will inspect 30 units. Among these, 20 will undergo a speed test and 10 will be tested for current flow. If an engineer is randomly assigned 4 units, what are the probabilities that

(a) none of them will need a speed test?
(b) only 2 will need a speed test?
(c) at least 3 will need a speed test?

4.25 A maker of specialized instruments receives shipments of 24 circuit boards. Suppose one shipment contains 4 that are defective. An engineer selects a random sample of size 4. What are the probabilities that the sample will contain

(a) 0 defective circuit boards?
(b) 1 defective circuit board ?
(c) 2 or more defective circuit boards?

4.26 If 6 of 18 new buildings in a city violate the building code, what is the probability that a building inspector, who randomly selects 4 of the new buildings for inspection, will catch

(a) none of the buildings that violate the building code?
(b) 1 of the new buildings that violate the building code?
(c) 2 of the new buildings that violate the building code?
(d) at least 3 of the new buildings that violate the building code?

4.41 Find the mean and the standard deviation of the distribution of each of the following random variables (having binomial distributions):

(a) The number of heads obtained in 676 flips of a balanced coin.
(b) The number of 4's obtained in 720 rolls of a balanced die.
(c) The number of defectives in a sample of 600 parts made by a machine, when the probability is 0.04 that any one of the parts is defective.
(d) The number of students among 800 interviewed who do not like the food served at the university cafeteria, when the probability is 0.65 that any one of them does not like the food.

………………..

POISSON DISTRIBUTION

The **Poisson distribution** often serves as a model for counts which do not have a natural upper bound. The **Poisson distribution**, with mean λ (lambda), has probabilities given by

……………

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{for } x = 0, 1, 2, \ldots \qquad \lambda > 0$$

…………

Moment generating function for Poisson distribution

Let X have the Poisson distribution with probability distribution

$$f(x) = \frac{\lambda^x}{x!}e^{-\lambda} \qquad \text{for } x = 0, 1, \ldots, \infty$$

Show that

(a) $M(t) = e^{\lambda(e^t - 1)}$ for all t
(b) $E(X) = \lambda$ and $Var(X) = \lambda$

The mean and variance of the Poisson distribution are equal.

Solution.

(a) By definition of the moment generating function

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

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)} \qquad \text{for } -\infty < t < \infty$$

where we have used the series $e^y = \sum_{k=0}^{\infty} \frac{y^k}{k!}$

(b) Differentiating $M(t)$, we find

$$M'(t) = \lambda e^t e^{\lambda(e^t - 1)}$$
$$M''(t) = \lambda e^t e^{\lambda(e^t - 1)} + \lambda^2 e^{2t} e^{\lambda(e^t - 1)}$$

Evaluating these derivatives at $t = 0$, we obtain the moments

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

Also, the variance is

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

//

Example. For health reasons, homes need to be inspected for radon gas which decays and produces alpha particles. One device counts the number of alpha particles that hit its detector. To a good pproximation, in one area, the count for the next week follows na Poisson distribution with mean 1.3. Determine

(a) the probability of exactly one particle next week.
(b) the probability of one or more particles next week.
(c) the probability of at least two but no more than four particles next week.
(d) The variance of the Poisson distribution

Solution.
Unlike the binomial case, there is no choice of a fixed Bernoulli trial here because one can always work with smaller intervals.

(a) $P(X = 1) = \lambda^1 e^{-\lambda}/1! = 1.3\, e^{-1.3} = .3543$

(b) $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1.3} = 0.727$
(c) $P(2 \leq X \leq 4) = F(4,\ 1.3) - F(1,\ 1.3) = 0.989 - 0.627 = 0.362$, using Poisson table

....

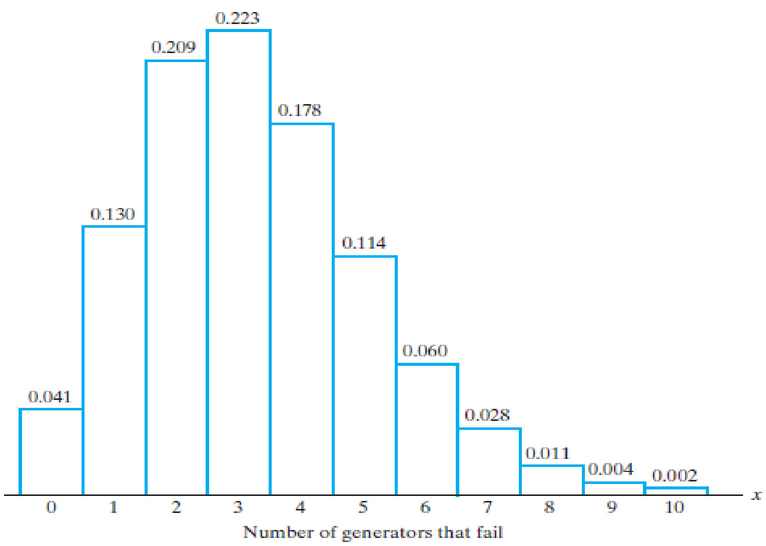
Example. It is known that 5% of the books bound at a certain bindery have defective bindings. Find the probability that 2 of 100 books bound by this bindery will have defective bindings using
(a) the formula for the binomial distribution;
(b) the Poisson approximation to the binomial distribution.

Solution. The binomial distribution could be used when appropriate computer software is available. However, the expected number is small and the number of generators is large so the Poisson approximation is valid.

We take $\lambda = 3,840 \cdot 1/1,200 = 3.2$
Consulting Poisson Table with λ=3.2, and using the identity
$f(x;\ \lambda)=F(x;\ \lambda)-F(x-1;\ \lambda)$,

we obtain the results shown in the probability histogram of Figure:

………………………..



Probability histogram of Poisson distribution with $\lambda = 3.2$

…………………………………

In our justification of the Poisson approximation to the binomial distribution, we let $\lambda = np$.

For the variance we can write $\sigma^2 = np(1-p) = \lambda(1-p)$, which approaches λ as $p \to 0$.

This matches the mean and variance of the Poisson distribution.

The Poisson Approximation to the Binomial Distribution

When n is large and p is small, binomial probabilities are often approximated by means of the Poisson distribution with λ equal to the product np.

Let us now show that when $n \rightarrow \infty$ and $p \rightarrow 0$, while $np = \lambda$ remains constant, the limiting form of the binomial distribution is the Poisson distribution given above.

First let us substitute λ / n for p into the formula for the binomial distribution and simplify the resulting expression; thus, we get

…………..

$$
\begin{aligned}
b(x; n, p) &= \frac{n!}{x!\,(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!\,n^x}(\lambda)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{x!}(\lambda)^x\left(1-\frac{\lambda}{n}\right)^{n-x}
\end{aligned}
$$

Letting $n \rightarrow \infty$, we find that

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right) \rightarrow 1$$

and that

$$\left(1-\frac{\lambda}{n}\right)^{n-x} = \left[\left(1-\frac{\lambda}{n}\right)^{n/\lambda}\right]^{\lambda}\left(1-\frac{\lambda}{n}\right)^{-x} \rightarrow e^{-\lambda}$$

Hence, the binomial distribution $b(x; n, p)$ approaches

$$\frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{for } x = 0, 1, 2, \ldots$$

…………

Theorem. Prove that the Poisson distribution is a positively skewed distribution and leptokurtic

Proof. We know

$$\mu_3 = \mu'_3 - 3\mu_2\mu'_1 + 2(\mu'_1)^3$$

Here $\mu'_1 = \lambda; \mu'_2 = \lambda^2 + \lambda$

Now,

$$\mu'_3 = E(X3) = \sum_0^\infty x^3 f(x)$$

$$= \sum_0^\infty [x(x-1)(x-2) + 3x^2 - 2x] f(x)$$

$$= \sum_0^\infty x(x\text{-}1)(x\text{-}2)\frac{e^{-\lambda}\lambda^x}{x!} + 3\sum_0^\infty x^2 f(x) - 2\sum_0^\infty x f(x)$$

$$= \lambda^3 e^{-\lambda} \sum_0^\infty \frac{x^{x-3}}{(x-3)!} + 3E(X^2) - 2E(X)$$

$$= \lambda^3 e^{-\lambda} e^{-\lambda} + 3(\lambda^2 + \lambda) - 2\lambda$$

$$= \lambda_3 + 3\lambda^2 + \lambda$$

Therefore, $\mu_3 = \lambda^3 + 3\lambda^2 + \lambda - 3(\lambda^2 + \lambda)\lambda + 2\lambda^3$

$$= \lambda$$

In a similar way we can find, $\mu_4 = 3\lambda^2 + \lambda$

Measures of Skewness and Kurtosis:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\mu^2}{\mu^3} = \frac{1}{\lambda}$$

$$\gamma 1 = \sqrt{\beta 1} = \frac{1}{\sqrt{\lambda}}$$

Since $\lambda > 0$, Poisson Distribution is a positively skewed distribution.

Also,

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2+\lambda}{\lambda^2} = 3+\frac{1}{\lambda}$$

$$\gamma 2 = \beta_2 - 3 = \frac{1}{\lambda}$$

Since $\lambda > 0$, Poisson Distribution is leptokurtic.

....

Additive property of Poisson distribution:

Let X_1 and X_2 be two independent Poisson random variables with parameters λ_1 and λ_2 respectively. Then $X = X_1 + X_2$ follows Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Proof:

$$X_1 \sim P(\lambda_1) \Rightarrow M_{X_1}(t) = e^{\lambda 1(e^t-1)}$$

$$X_2 \sim P(\lambda_2) \Rightarrow M_{X_2}(t) = e^{\lambda 2(e^t-1)}$$

$$M_X(t) = M_{X_1+X_2}(t) = M_{X_1}(t).M_{X_2}(t)$$

, Since X_1 and X_2 are independent.

$$\Rightarrow M_X(t) = e^{\lambda 1(e^t-1)} e^{\lambda 2(e^t-1)} = e^{[\lambda 1 + \lambda 2] \quad (e^t-1)}$$

Thus,

$$X = X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$$

Remarks:

In general if $X_i \sim P(\lambda_i)$ for $i = 1, 2, ..., k$; and X_i's are independent, then

$$X = X_1 + X_2 + ... + X_k \sim P(\lambda_1 + \lambda_2 + ... + \lambda_k)$$

Example.

If X and Y are independent poisson variates such that P(X=1) =P(X=2) and P(Y=2)=P(Y=3). Find the variance of X-2Y.

Solution:
Let $X \sim P(\lambda 1)$ and $Y \sim P(\lambda 2)$

Given, $P(X=1)=P(X=2)$

i.e., $$\frac{e^{-\lambda 1}\lambda_1^1}{1!} = \frac{e^{-\lambda 1}\lambda_1^2}{2!}$$

i.e., $$1 = \frac{\lambda 1}{2!}$$

,since $\lambda 1>0$

There fore,

$$\lambda 1 = 2$$

Also

$$P(Y=2)=P(Y=3)$$

$$\frac{e^{-\lambda 2}\lambda_2^2}{2!} = \frac{e^{-\lambda 2}\lambda_2^3}{3!}$$

i.e.,

$$1 = \frac{\lambda 2}{3}$$

There fore, $\lambda_2 = 3$

, since $\lambda_2 > 0$

There fore, $V(X) = \lambda_1 = 2$ and $V(Y) = \lambda_2 = 3$

Then,

$$V(X - 2Y) = V(X) + 4V(Y)$$

Since X and Y are independent.

$$= 2+4 \times 3 = 14$$

ANNEXURE- POISSON DISTRIBUTIO

Direct Method;

Mean and variance of Poisson distribution: $\mu = \lambda$ and $\sigma^2 = \lambda$.

Solution.

Mean:

$$E(X) = \sum_0^\infty x f(x) = \sum_0^\infty x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda\, e^{-\lambda} \sum_0^\infty \frac{\lambda^{x-1}}{(x-1)!}$$

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

Variance,

$$V(X) = E(X2) - [E(X)]2$$

where,

$$E(X^2) = \sum_0^\infty x^2 f(x)$$

$$= \sum_0^\infty [x(x-1) + x] \ f(x)$$

$$= \sum_0^\infty x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_0^\infty x f(x)$$

$$= \lambda^2 e^{-\lambda} \sum_0^\infty \frac{\lambda^{x-2}}{(x-2)!} + E(X)$$

$$= \lambda^2 e^{-\lambda} e^{-\lambda} + \lambda$$

$$= \lambda^2 + \lambda$$

Therefore,

$$V(X) = \lambda^2 + \lambda \ - \lambda^2 = \lambda$$

Also, $SD(X) = \sqrt{\lambda}$

For a Poisson Distribution Mean = Variance

Moment Generating Function

$$Mx(t) = E(e^{tx})$$

$$= \sum_0^\infty e^{tx} f(x) = \sum_0^\infty e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_0^\infty \frac{(\lambda e^t)^x}{x!}$$

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

....

Remark: If λ is an integer $f(\lambda - 1; \lambda) = f(\lambda; \lambda)$ and each is larger than any other probability. Otherwise, when λ is not an integer, the largest probability is assigned to the integer part of λ.

....

ASSIGNMENT – PD

4.50 Prove that for the Poisson distribution

$f(x+1;\ \lambda) / f(x;\ \lambda) = \lambda / (x+1)$ for $x = 0, 1, 2, \ldots$

4.55 In a factory, 8% of all machines break down at least once a year. Use the Poisson approximation to the binomial distribution to determine the probabilities that among 25 machines (randomly chosen in the factory):

(a) 5 will break down at least once a year;
(b) at least 4 will break down once a year;
(c) anywhere from 3 to 8, inclusive, will break down at least once a year.

4.56 During inspection of the continuous process of making large rolls of floor coverings, 0.5 imperfections are spotted per minute on average. Use the Poison distribution to find the probabilities

(a) one imperfection in 4 minutes
(b) at least two in 8 minutes
(c) at most one in 10 minutes.

4.57 The number of gamma rays emitted per second by a certain radioactive substance is a random variable having the Poisson distribution with $\lambda = 5.8$. If a recording instrument becomes inoperative when there are more than 12 rays per second, what is the probability that this instrument becomes inoperative during any given second?

4.58 A consulting engineer receives, on average, 0.7 requests per week. If the number of requests follows a Poisson process, find the probability that

(a) in a given week, there will be at least 1 request;
(b) in a given 4-week period there will be at least 3 requests.

4.59 A conveyor belt conveys finished products to the warehouse at an average of 2 per minute. Find the probabilities that
(a) at most 3 will be conveyed in a given minute;
(b) at least 2 will be conveyed in an interval of 3 minutes;
(c) at most 20 will be conveyed during an interval of 5 minutes.

4.65 During an assembly process, parts arrive just as they are needed. However, at one station, the probability is 0.01 that a defective part will arrive in a one-hour period. Find the probability that

(a) exactly 1 defective part arrives in a 4-hour span;
(b) 1 or more defective parts arrive in a 4-hour span;
(c) exactly 1 defective part arrives in a 4-hour span
and exactly 1 defective part arrives in the next 4-hour span.

4.66 The arrival of trucks at a receiving dock is a Poisson process with a mean arrival rate of 2 per hour.

(a) Find the probability that exactly 5 trucks arrive in a two-hour period.
(b) Find the probability that 8 or more trucks arrive in a two-hour period.
(c) Find the probability that exactly 2 trucks arrive in a one-hour period and exactly 3 trucks arrive in the next one-hour period.

4.67 The number of flaws in a fiber optic cable follows a Poisson process with an average of 0.6 per 100 feet.

(a) Find the probability of exactly 2 flaws in a 200-foot cable.
(b) Find the probability of exactly 1 flaw in the first 100 feet and exactly 1 flaw in the second 100 feet.

Geometric distribution

Suppose that in a sequence of trials we are interested in the number of the trial on which the first success occurs.

The three assumptions for Bernoulli trials are satisfied but the extra assumption underlying the binomial distribution is not. In other words, *n* is not fixed.

Clearly, if the first success is to come on the *x*th trial, it has to be preceded by $x - 1$ failures, and if the probability of a successes is *p*, the probability of $x - 1$ failures in $x - 1$ trials is $(1 - p)^{x-1}$.

Then, if we multiply this expression by the probability *p* of a success on the *x*th trial, we find that the probability of getting the first success on the *x*th trial is given by

$$g(x; p) = p(1 - p)^{x-1} \text{ for } x = 1, 2, 3, 4, \ldots$$

This probability distribution is called the **geometric distribution**

Assignment 2:

Prove that the Mean and variance of geometric distribution are

$\mu = 1/p; \quad \sigma^2 = (1 - p)/p^2$

Solution.

Mean,

$$E(X)=\sum_0^\infty xf(x)$$

$$= \sum_0^\infty xq^x p$$

$$= p[q + 2q^2 + 3q^3 + \ldots]$$

$$=pq[1+2q+3q^2+\ldots]$$

$$= pq(1 - q)^{-2}$$

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

Variance,

$$V(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum_0^\infty x^2 f(x)$$

$$= \sum_0^\infty [x(x - 1) + x]f(x)$$

$$= \sum_0^\infty x(x-1)q^2p+\sum_0^\infty xf(x)$$

$$= p[2.1q^2 + 3.2q^3 + 4.3q^4 + \ldots] + E(X)$$

$$= 2pq^2[1+ 3q + 6q^2 + \ldots] + \frac{q}{p}$$

$$= 2pq^2(1-q)^{-3}+ \frac{q}{p}= \frac{2pq^2}{p^3}+ \frac{q}{p}= \frac{2q^2}{p^2}+\frac{q}{p}$$

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

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

$$= \frac{q^2+pq}{p^2} = \frac{q}{p^2}(p+q) = \frac{q}{p^2}$$

……………..

Example. If the probability is 0.05 that a certain kind of measuring device will show excessive drift, what is the probability that the sixth measuring device tested will be the first to show excessive drift?

Solution. Substituting $x = 6$ and $p = 0.05$ into the formula for the geometric distribution, we get

$$g(6;\ 0.05) = (0.05)(1 - 0.05)^{6-1} = 0.039 \quad //$$

………………..

Moment Generating function:

$$M_X(t)=E(e^{tx}$$

$$= \sum_0^\infty e^{tx} q^x p$$

$$= p\sum_0^\infty (qe^t)^x$$

$$= p[1+qe^t+(qe^t)^2 + \ldots\ldots]$$

$$= p(1-qe^t)^{-1} = \frac{p}{1-qe^t}$$

Lack of memory property:

If X has geometric density with parameter p, then
$P[X \geq s + t/X \geq s] = P(X \geq t)$ for $s,t = 0,1,2,\ldots$

Proof:

$$P[X \geq s + t/X \geq s] = \frac{P(X \geq s + t)}{P(X\geq s)}$$

$$= \frac{\sum_{x=s+t}^\infty q^x p}{\sum_{x=s}^\infty q^x p}$$

$$= \frac{q^{s+t}}{q^s} = qt = P(X \geq t)$$

Thus geometric distribution possesses lack of memory property.

Example.

For a random variable following geometric distribution with parameter p, prove the recurrence formula, $P(x + 1) = q.P(x)$

Solution:

The pmf of X is $p(x) = q^{x}p, x = 0,1,2,\ldots$

Then,

$$P(x+1) = q^{x+1}p$$

$$P(x) = q^{x}p$$

$$\Rightarrow \frac{P(x+1)}{P(x)} = \frac{q^{x+1}p}{q^{x}p} = q$$

Hence

$$P(x+1) = q.P(x)$$

EXAMINATION QUESTIONS :

(2016)

(iii) Show that the moment-generating function of the geometric distribution is given by

$$M_X(t) = \frac{\theta e^t}{1 - e^t(1-\theta)}$$

Use it to show that $\mu = \frac{1}{\theta}$ and $\sigma^2 = \frac{1-\theta}{\theta^2}$.

NEGATIVE BINOMIAL DISTRIBUTION

The negative binomial distribution describes the total number of Bernoulli trials, X, to obtain a specified number r successes.

When r=1, the negative binomial reduces to the geometric distribution.

If the rth success occurs at trial number x, it must be that r–1 successes occurred in the first $x-1$ trials and the last trial is a success.

The probability distribution is then the product of the binomial probability $b(r-1;\ x-1,\ p)$ and p:

$$f(x) = C(x-1, r-1)\ p^{r}(1-p)^{x-r} \quad \text{for } x = r,\ r+1,\ \ldots$$

………. .

Mean and variance of negative binomial distribution are:

$$\mu = r(1-p)/p; \quad \sigma^2 = (1-p)\,r/p^2$$

ASSIGNMENT- UD-GD

4.36 Find the mean and the variance of the uniform probability distribution given by

$f(x) = 1/n$ for $x = 1, 2, 3, \ldots, n$

[*Hint*: The sum of the first n positive integers is $n(n+1)/2$, and the sum of their squares is $n(n+1)(2n+1)/6$.]

4.68 Differentiating with respect to p on both sides of the equation

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

show that the geometric distribution

$$f(x) = p(1-p)^{x-1} \text{ for } x = 1, 2, 3, \ldots$$

has the mean $1/p$.

5.100 Let X have the geometric distribution

$$f(x) = p(1-p)^{x-1} \qquad \text{for } x = 1, 2, \ldots$$

(a) Obtain the moment generating function for

$$t < -\ln(1-p)$$

[*Hint:* Recall that $\sum_{k=0}^{\infty} r^k = \dfrac{1}{1-r}$ for $|r| < 1$.]

(b) Obtain $E(X)$ and $E(X^2)$ by differentiating the moment generating function.

....

Table 5.2(a) Discrete distributions

Distribution	Probability Distribution $f(x)$	Mean	Variance	Moment Generating Function
Binomial $b(x; n, p)$	$\binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$	np	$np(1-p)$	$(pe^t + 1 - p)^n$
Geometric $g(x; p)$	$p(1-p)^{x-1}, \quad x = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric $h(x; n, a, N)$	$\binom{a}{x}\binom{N-a}{n-x} \Big/ \binom{N}{n}$ $x = 0, 1, 2, \ldots, \min(N-a, n)$	$n\frac{a}{N}$ $p = a/N$	$n\frac{a}{N}\left(1-\frac{a}{N}\right)\frac{(N-n)}{(N-1)}$	complicated
Poisson $f(x; \lambda)$	$\frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \ldots$	λ	λ	$e^{\lambda(e^t-1)}$
Negative binomial	$\binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, \ldots$	r/p	$r(1-p)/p^2$	$\left(\frac{pe^t}{1-(1-p)e^t}\right)^r$