Poisson approximation to Binomial

The binomial distribution reduces to Poisson distribution under the following conditions:

(i) If no. of trials 'n' is very large i.e. $n \to \infty$ $(n > 20)$

(ii) If probability of success 'p' is very small i.e. $p \to 0$ $(p \leq 0.05)$

$$\therefore \lambda = np$$

$$\Rightarrow P(X=x) = \frac{e^{-np}(np)^x}{x!}$$

July 20, 2024

Q16) $n = 30$, $p = 0.04$

Using Poisson's approximation to Binomial, we get

$$\lambda = np$$
$$= 30 * 0.04$$
$$= \underline{1.2}$$

$$\therefore P(X=x) = \frac{e^{-np}(np)^x}{x!}$$

(a) $P(X=25) = \dfrac{e^{-1.2}(1.2)^{25}}{25!}$

$= 0$

(b) $P(X=3) = \dfrac{e^{-1.2}(1.2)^3}{3!} = \underline{0.0867}$

ⓒ $P(X=5) = \dfrac{e^{-1.2}(1.2)^5}{5!}$

$= \underline{0.00624}$

<u>Q17.</u> X = number of non-marital disputes

$p = 1 - 0.96$

$= 0.04$

$n = 80$

Here, $n > 20$ & $p \leq 0.05$,

Thus, using Poisson's approximation to binomial, we get:

$\lambda = np$

$= 80 \times 0.04 = 3.2$

$\therefore P(X=x) = \dfrac{e^{-np}(np)^x}{x!}$

ⓐ $P(X=7) = \dfrac{e^{-3.2} \times (3.2)^7}{7!}$

$= \dfrac{0.04076 * 3435.97383}{5040}$

$= \underline{0.02778}$

Q 20.

$p = 1/400 = 0.0025$

$n = 100$

$X =$ no. of defective

Using Poisson's approximation to binomial,

$\lambda = np$

$= 0.0025 \times 100$

$= 0.25$

$\therefore P(X=x) = \dfrac{e^{-np} (np)^x}{x!}$

ⓐ $\therefore P(X=0) = \dfrac{e^{-0.25} (0.25)^0}{0!}$

$= \underline{0.7788}$

ⓑ $P(X<2) = \cancel{1 - P(X \leq 1)}$ $P(X=0) + P(X=1)$

$= 0.7788 + \dfrac{e^{-0.25} (0.25)^1}{1!}$

$= 0.7788 + 0.1947$

$= 0.9735$

(c) $P(X \geq 1) = 1 - P(X=0)$

$= 1 - 0.7788$

$= \underline{0.2212}$

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

$= 1 - \left[0.9735 + \frac{e^{-0.25}(0.25)^2}{2!}\right]$

$= 1 - (0.9735 + 0.0243)$

$= \underline{0.0022}$

Assignment

\# Q 4, 7, 8, 13, 18, 19

Negative Binomial Distribution

$(1-a)^{-n} = \sum_{x=0}^{\infty} {}^{x+n-1}C_x \, a^x$

Introduction: If a trial is repeated independently till fixed number of success occurs with probability 'p' which is constant for each trial, then such an experiment follows Negative Binomial Distribution (NBD).

$P(A \cap B)$ last trial

$= {}^{x+k-1}C_{k-1} \, p^{k-1} q^x \cdot p$

$= {}^{x+k-1}C_{k-1} \, p^k q^x$; $x = 0, 1, 2, \ldots$; $k > 0$

${}^{x+k-1}C_x$

k = fixed number of successes

p = probability of success in each trial

${}^nC_x \, p^x q^{n-x}$

$P(A \cap B) = P(A) \cdot P(B)$ independent

$x+k$ = Total number of trials

x = no. of failures preceding 'k' success

$(x+k-1)$ = trials

$(k-1)$ = successes

Derivation of NBD

Consider a binomial experiment consists of 'n' trials, & the probability of success (p) is constant, for each trial. Let, there be 'x' failures preceding the k^{th} success in $x+k=n$ trials. Here $x+k$ trials are required to produce 'k' successes (which is fixed) and X is a random variable.

Now, we need to find the probability the k^{th} success occurs in $x+k$ trials, which is the probability that exactly 'x' failures preceding the k^{th} success in $x+k$ trials.

Here, the last trial must be a success with probability 'p' and the remaining $(x+k-1)$ trials, there are $(k-1)$ successes and whose probability is given by ${}^{x+k-1}C_{k-1}\, p^{k-1} q^{x}$.

Now, by multiplication theorem of probability, the probability of 'x' failures preceding the k^{th} success in $(x+k)$ trials is given by

$$\left\{ {}^{x+k-1}C_{k-1}\, p^{k-1}\, q^{x} \right\} p$$

$$= {}^{x+k-1}C_{k-1}\, p^{k}\, q^{x}$$

This is the pmf of negative binomial distribution (NBD) with parameters 'k' and 'p'.

If $X \sim NB(k, p)$, then the probability mass function of the distribution is given by

$$P(X=x) = {}^{x+k-1}C_{k-1}\, p^{k} q^{x} \; ; \quad x = 0, 1, 2, \ldots$$
$$k > 0$$
$$p + q = 1$$

NBD as special case of GPSD

In GPSD, $\theta = \dfrac{q/p}{1+q/p}$, $f(\theta) = (1-\theta)^{-K}$

$S = 0, 1, 2, \ldots, \infty$

Now,

$$(1-\theta)^{-k} = \sum_{x=0}^{\infty} {}^{x+k-1}C_x \, \theta^x \quad \text{---} (i)$$

$$f(\theta) = \sum_{x=0}^{\infty} a_x \theta^x \quad \text{---} (ii)$$

From (i) and (ii),

$$a_x = {}^{x+k-1}C_x$$

Since, $P(X=x) = \dfrac{a_x \theta^x}{f(\theta)}$

$$\Rightarrow P(X=x) = \frac{{}^{x+K-1}C_x \left(\frac{q/p}{1+q/p}\right)^x}{\left[1 - \left(\frac{q/p}{1+q/p}\right)\right]^{-K}}$$

$$\Rightarrow P(X=x) = \frac{{}^{x+K-1}C_x \left(\frac{q/p}{1+q/p}\right)^x}{\left(\frac{1}{1+q/p}\right)^{-K}}$$

$$\Rightarrow P(X=x) = \frac{{}^{x+K-1}C_x \cdot \frac{q^x}{(p+q)^x}}{\left(\frac{p}{p+q}\right)^{-K}}$$

$$\Rightarrow P(X=x) = {}^{x+K-1}C_x \, q^x \cdot p^K$$

$$\boxed{\therefore \; P(X=x) = {}^{x+K-1}C_x \, p^K q^x}$$

This is the probability mass function of negative binomial distribution

$P(X=x) = \dfrac{a_x \theta^x}{f(\theta)}$

$f(\theta) = \sum_{x \in S} a_x \theta^x ; \theta > 0$

$\theta = \dfrac{q/p}{1+q/p}$

$f(\theta) = (1-\theta)^{-K}$

$S = \{0, 1, 2, \ldots, \infty\}$

with parameters k and p. Hence, NBD is the special case of GPSD.

July 23, 2024

Mean and Variance of Negative Binomial Distribution.

If $X \sim NB(p,k)$ then, the pmf of the distribution,

$$P(X=x) = {}^{x+k-1}C_x \, p^k . q^x \;;\; x = 0,1,2,\ldots$$
$$p+q=1$$
$$k>0$$

Mean, $E(X) = \sum_{x=0}^{\infty} x \, P(x)$

$$= \sum_{x=0}^{\infty} x \left\{ {}^{x+k-1}C_x \, p^k q^x \right\}$$

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

$$= p^k \sum_{x=0}^{\infty} \cancel{x} \frac{(x+k-1)!}{(\cancel{x}+k-1-\cancel{x})! \, \cancel{x} (x-1)!} q^x$$

$$= k p^k q \sum_{x=1}^{\infty} \frac{(x+k-1)!}{k! \, (x-1)!} q^{x-1}$$

$$= k p^k q \sum_{x=1}^{\infty} {}^{x+k-1}C_{x-1} \, q^{x-1}$$

$$= k p^k q \boxed{\sum_{(x-1)=0}^{\infty} {}^{(x-1)+(k+1)-1}C_{x-1} \, q^{x-1}}$$

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

$$= k p^k q \, p^{-k-1}$$

$$= k \cancel{p^k} q \cdot \frac{1}{\cancel{p^k} \cdot p} = \frac{kq}{p}$$

$$\boxed{\therefore E(X) = \frac{kq}{p}}$$

Variance of NBD

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

$$= E(X^2) - \left(\frac{kq}{p}\right)^2 \text{ ———— (i)}$$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 P(x)$$

$$= \sum_{x=0}^{\infty} [x(x-1) + x] P(x)$$

$$= \sum_{x=0}^{\infty} x(x-1) P(x) + \sum_{x=0}^{\infty} x P(x)$$

$$= \sum_{x=0}^{\infty} x(x-1) \, {}^{x+k-1}C_x \, p^k q^x + \left(\frac{kq}{p}\right)$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{(x+k-1)!}{(\cancel{x}+k-1-\cancel{x})! \, x!} p^k q^x + \left(\frac{kq}{p}\right)$$

$$= p^k q^2 \sum_{x=0}^{\infty} \cancel{x(x-1)} \frac{(x+k-1)!}{(k-1)! \, \cancel{x(x-1)}(x-2)!} q^{x-2} + \left(\frac{kq}{p}\right)$$

$$= p^k q^2 k(k+1) \sum_{x=2}^{\infty} \frac{(x+k-1)!}{k(k+1)(k-1)!(x-2)!} q^{x-2} + (kq/p)$$

$$= p^k q^2 k(k+1) \sum_{x=2}^{\infty} {}^{x+k-1}C_{x-2} \, q^{x-2} + \left(\frac{kq}{p}\right)$$

$$= p^k q^2 k(k+1) \sum_{(x-2)=0}^{\infty} {}^{(x-2)+(k+2)-1}C_{x-2} \, q^{x-2} + \left(\frac{kq}{p}\right)$$

$$= p^k q^2 k(k+1)(1-q)^{-(k+2)} + (kq/p)$$

$$= p^k q^2 k(k+1)(p)^{-k+2} + kq/p = \frac{p^k q^2 k(k+1)}{p^k \cdot p^2} = \frac{q^2}{p^2} k(k+1) + kq/p$$

From ①,

$$Var(X) = \frac{q^2}{p^2} K(K+1) + \frac{kq}{p} - \frac{k^2 q^2}{p^2}$$

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

$$= kq/p \left(\frac{q}{p}+1\right)$$

$$= \frac{kq}{p}\left(\frac{p+q}{p}\right) \qquad [\because p+q=1]$$

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

$$\boxed{\therefore Var(X) = \frac{kq}{p^2}}$$

Moment generating Function (Mgf)

If $X \sim N.B(k, p)$, then $P(X=x) = {}^{x+k-1}C_x \, p^k q^x$; $x = 0, 1, 2, \dots$
$p+q=1$,
$k>0$

Now, mgf of X,

$$M_x(t) = E[e^{tX}]$$

$$= \sum_{x=0}^{\infty} e^{tx} P(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \left\{ {}^{x+k-1}C_x \, p^k q^x \right\}$$

$$\Rightarrow M_x(t) = \left[\sum_{x=0}^{\infty} {}^{x+k-1}C_x \, (qe^t)^x\right] p^K$$

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

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

This is the mgf of negative binomial distribution with parameters K and p.

$$\mu_1' = \text{Mean} = E(X) = \left.\frac{d\, M_x(t)}{dt}\right|_{t=0}$$

$$= \left.\frac{d}{dt}\left[p^K/(1-qe^t)^K\right]\right|_{t=0}$$

$$= \left[p^K \cdot (-K)(1-qe^t)^{-K-1} \cdot \frac{d}{dt}(1-qe^t)\right]_{t=0}$$

$$= -p^K K \cdot (1-qe^t)^{-K-1} \left\{-qe^t\right\}\Big|_{t=0}$$

$$= Kp^K (1-qe^t)^{-K-1} \, qe^t \Big|_{t=0}$$

$$= Kp^K (1-qe^0)^{-K-1} \cdot qe^0$$

$$= Kp^K (1-q)^{-K-1} \cdot q$$

$$= Kp^K q \; p^{-K-1} = \frac{K \cancel{p^K} q}{\cancel{p^K} \cdot p} = Kq/p$$

$$\mu_2' = \frac{d^2 M_x(t)}{dt^2} = kqp^k\left[e^t\left\{(-k-1)(1-qe^t)^{-k-2}(-qe^t)\right\} + (1-qe^t)^{-k-1}e^t\right]$$

$$\mu_2' = kqp^k\left[(-k-1)(1-q)^{-k-2}(-q) + p^{-k-1}\right]$$

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

$$= \frac{kq^2(k+1)}{p^2} + \frac{kq}{p}$$

$$\therefore \mu_2 = \mu_2' - (\mu_1')^2$$

$$= \frac{kq^2(k+1)}{p^2} + \frac{kq}{p} - \frac{k^2q^2}{p^2}$$

$$= \cancel{\frac{k^2q^2}{p^2}} + \frac{kq^2}{p^2} + \frac{kq}{p} - \cancel{\frac{k^2q^2}{p^2}}$$

$$= \frac{kq}{p}\left(q/p + 1\right)$$

$$= \frac{kq}{p}\left(\frac{p+q}{p}\right) = \frac{kq}{p^2}$$

$$\boxed{\therefore Var(x) = kq/p^2}$$

Q. The probability of hitting the target at any trial 0.2. If a shooter aims at a target, find the probability that fifth fire is second hit.

$p = 0.2$

$x + k = 5$

$k = 2$

$\Rightarrow x + 2 = 5$

$\therefore x = 3$

$P(X = 3)$

Given, p = probability of hitting a target in any trial $= 0.2$

x = number of mis hits

$k = 2$

$x + k = 5$

$\Rightarrow x = 3$

Here, $X \sim NBD(p, k)$ so $P(X = x) = {}^{x+k-1}C_x \, p^k q^x$; $x = 0, 1, 2, \ldots$

$p + q = 1$

$k > 0$

Now, probability that the fifth fire is second hit, $P(X = 3) = {}^{3+2-1}C_3 (0.2)^2 (0.8)^3$

$= {}^4C_3 (0.2)^2 (0.8)^3$

$= 0.08192$

Assignment

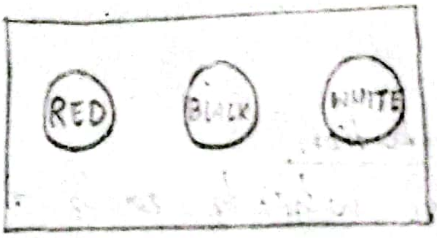
Q1. A boy is throwing stones at a target. If the probability of hitting the target at any trial is 0.5, what is the prob. that

10^{th} throw is 5^{th} hit?

Q2. The prob. that a striker will make a goal in a football match is 0.3. What is the probability that he'll score fourth goal in 5^{th} shot?

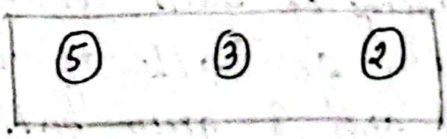
July 25, 2024

— MULTINOMIAL DISTRIBUTION —

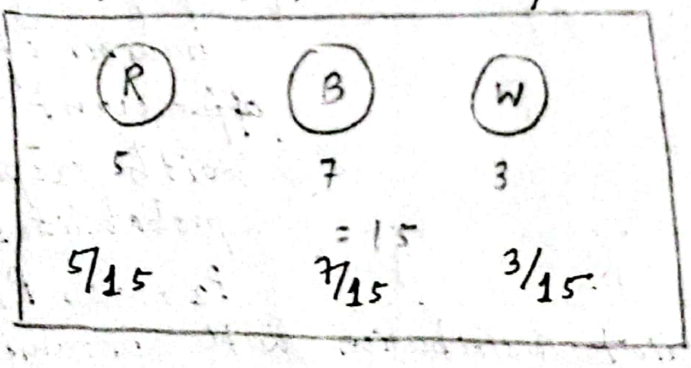


$n = 10$

$P(X_1 = 5, X_2 = 3, X_3 = 2)$



No. of trials = 5 using w/ replacement



$$P(X_1 = 2, X_2 = 2, X_3 = 1) = \frac{5!}{(2!)(2!)(1!)}\left(\frac{5}{15}\right)^2\left(\frac{7}{15}\right)^2\left(\frac{3}{15}\right)^1$$

At least 3 possible outcomes

→ with replacement

$$P(x_1, x_2, \ldots, x_K) = \frac{n!}{x_1!\, x_2! \ldots x_K!} p_1^{x_1} p_2^{x_2} \ldots p_K^{x_K}$$

$$\sum_{i=1}^{K} x_i = n$$

$$p_1 + p_2 + \ldots + p_K = \sum_{i=1}^{K} p_i = 1$$

Introduction

An experiment is said to be a multinomial ~~distribution~~ experiment if

i) the experiment consists of a fixed number of trials 'n'.

ii) for each trial there are $K (\geq 3)$ possible outcomes.

iii) Trials are statistically independent.

iv) The probability of each outcome remains the same from trial to trial.

Examples of the multinomial distribution:

(i) No. of throws of a fair dice in which each throw can result six different outcomes.

(ii) Number of selection or drawings of balls at random with replacement from a box containing 20 balls of which 2 are white, 4 are black, 6 Red and 8 green balls.

Definition: The random variable $X = (x_1, x_2, \ldots, x_K)$ denoting the outcome of 'n' trials where X_i = frequency or number of outcomes of event E_i with respective probabilities (P_i) $P_1, P_2, \ldots, P_K$.

is said to have multinomial distribution with parameters $(n, P_1, P_2, \ldots, P_K)$, ~~then the~~ if its pmf is given by

$$P(X = x) = P(X = x_1, x_2, \ldots, x_K)$$

$$\boxed{= \frac{n!}{x_1!\, x_2! \ldots x_K!} P_1^{x_1} P_2^{x_2} \ldots P_K^{x_K}}$$

where, $n = x_1 + x_2 + \ldots + x_K = \sum_{i=1}^{K} x_i$

and $p = p_1 + p_2 + \ldots + p_K = \sum_{i=1}^{K} P_i$

$$\left[\sum_x P(x_1, x_2, \dots, x_K) = \sum_x \left[\frac{n!}{x_1! x_2! \dots x_K!} p_1^{x_1} p_2^{x_2} \dots p_K^{x_K}\right] = (p_1+p_2+\dots+p_K)^n = 1\right.$$

Reduction of multinomial distribution when $k=2$

when $k=2$, then the pmf of multinomial distribution becomes

$$P(x_1, x_2) = \frac{n!}{x_1! x_2!} p_1^{x_1} p_2^{x_2}$$

where, $n = x_1 + x_2$ and $p_1 + p_2 = 1$

$$\Rightarrow x_2 = (n - x_1) \qquad \Rightarrow p_2 = (1 - p_1)$$

$$P(x_1, n - x_1) = \frac{n!}{x_1! (n-x_1)!} p_1^{x_1} (1-p_1)^{(n-x_1)}.$$

Consider $x_1 = x$ and $p_1 = p$

$$P(x, n-x) = \frac{n!}{x! (n-x)!} p^x (1-p_1)^{n-x}$$

$$\Rightarrow \boxed{P(x) = {}^nC_x \, p^x (1-p)^{n-x}}$$

This is the pmf of binomial distribution with parameters 'n' and 'p'.

Hence, when $k=2$, the multinomial distribution reduces to binomial distribution with parameters 'n' and 'p'.

Moment generating function of Multinomial distribution

If $X \sim M_K(n, p_1, p_2, \dots, p_K)$ then pmf

$$P(x_1, x_2, \dots, x_K) = \frac{n!}{x_1! x_2! \dots x_K!} p_1^{x_1} p_2^{x_2} \dots p_K^{x_K}$$

Now, the moment generating function,

$$M_X(t) = M_{(X_1, X_2, \dots, X_K)}(t_1, t_2, \dots, t_K)$$

CamScanner

$$= E\left[e^{t_1 x_1} e^{t_2 x_2} \cdots e^{t_k x_k}\right]$$

$$= \sum_{x} \left(e^{t_1 x_1}\right) \cdots \left(e^{t_k x_k}\right) \frac{n!}{x_1!\, x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

$$= \sum_{x} \left(p_1 e^{t_1}\right)^{x_1} \cdots \left(p_k e^{t_k}\right)^{x_k} \frac{n!}{x_1!\, x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

$$\cancel{= p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t}}$$

$$= \left(p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t_k}\right)^n$$

$$\because \sum_{x} \frac{n!}{x_1!\, x_2! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} = (p_1 + p_2 + \ldots + p_k)^n$$

If $t_1 \neq 0$ and $t_i = 0$ for $i = 2, 3, \ldots, k$, we get:

$$M_x(t) = M_{x_1, x_2, \ldots x_k}(t_1, 0, \ldots, 0)$$

$$= \left(p_1 e^{t_1} + \underline{p_2 + \ldots + p_k}\right)^n$$

$$= \left(p_1 e^{t_1} + (1 - p_1)\right)^n$$

$$= \left[q_1 + p_1 e^{t_1}\right]^n \qquad \left(\because p_1 + q_1 = 1 \Rightarrow q_1 = 1 - p_1\right)$$

This is the moment generating function of Binomial distribution with parameters n and p_1.

Hence by uniqueness property of moment generating function, $X_1 \sim B(n, p_1)$.

Since each X_i is a binomial variate with parameters n and p_i so we have

$E(X_i) = np_i$

and $V(X_i) = np_i q_i \quad \forall\, i = 1, 2, \ldots, k$