Geometric and Negative Binomial Distribution

Negative binomial distribution is the generalization of geometric distribution. Probability of getting 1st success from the independent trials is based on geometric distribution. and probability of getting kth success ~~from~~ in nth position is negative binomial distribution.

Geometric distribution

Let X counts no. of trials of getting 1st success. Let $P(S) = P(\text{success}) = p$ and $P(F) = P(\text{failure}) = q$ then $p+q=1$ i.e. $q = 1-p$, the pmf of $X = p(x) = q^{x-1}p,\ x = 1, 2 \ldots$

sample point (S)	S	FS	FFS	FFFS	—
X	1	2	3	4	---
$p(x) = P[X=x]$	p	qp	q^2p	q^3p	---

$$\sum_{x=1}^{\infty} p(x) = p(1) + p(2) + p(3) + p(4) + \ldots$$

$$= p + qp + q^2p + q^3p + \ldots$$

$$= p(1 + q + q^2 + q^3 + \ldots)$$

$$= p \cdot \frac{1}{1-q} = p \cdot \frac{1}{p} = 1$$

Since the probability summation is depending on geometric series, it is known as geometric distribution.

Mean & Variance of geometric distribution

We know that

$$\frac{1}{1-q} = 1+q+q^2+q^3+q^4+\cdots \quad (-1<q<1)$$

$$\Rightarrow \frac{d}{dq}\left(\frac{1}{1-q}\right) = 1+2q+3q^2+4q^3+\cdots$$

$$\Rightarrow \frac{1}{(1-q)^2} = 1+2q+3q^2+4q^3+\cdots \quad \text{—— (1)}$$

$$\Rightarrow \frac{d}{dq}\left(\frac{1}{(1-q)^2}\right) = 2+6q+12q^2+\cdots$$

$$\Rightarrow \frac{2}{(1-q)^3} = 2+6q+12q^2+\cdots \quad \text{—— (2)}$$

$$\text{mean}(X) = \sum_{x=1}^{\infty} x\, p(x) \qquad \text{where } p(x) = \begin{cases} p q^{x-1}, & x=1,2,\cdots \\ 0 & \text{otherwise} \end{cases}$$

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

$$\Rightarrow E(X) = p(1+2q+3q^2+\cdots)$$

$$= p \cdot \frac{1}{(1-q)^2}$$

$$= p \cdot \frac{1}{p^2} = \frac{1}{p}$$

Variance(X)

$$V(X) = \sigma^2 = E(X^2) - (E(X))^2$$

$$= E(X^2) - \frac{1}{p^2}$$

Where

$$E(X^2) = \sum_{x=1}^{\infty} x^2 p(x) = \sum_{x=1}^{\infty} x^2 p q^{x-1}$$

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

$$E(X^2) = E(X(X-1)+X) = \sum_{x=1}^{\infty} (x(x-1)+x) p q^{x-1}$$

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

$$= p\left[2q + 6q^2 + 12q^3 + \cdots\right] + \frac{1}{p}$$

$$= pq(2 + 6q + 12q^2 + \cdots] + \frac{1}{p}$$

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

$$= pq \cdot \frac{2}{p^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}$$

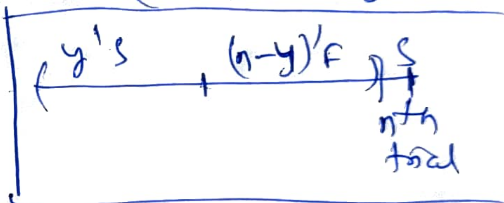
$$V(X) = \sigma^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$\boxed{\begin{aligned} &2q+p-1 \\ &= q+(q+p)-1 \\ &= q+1-1 \\ &= q \end{aligned}}$$

$$= \frac{2q+p-1}{p^2} = \frac{q}{p^2}$$

Negative Binomial distribution

Let Y be the r.v. that counts the trials of getting yth success. If $P(\text{success}) = P(S) = p$ and $P(\text{failure}) = P(F) = q$, then $p + q = 1$, implying $q = 1 - p$. Let yth success is obtained in nth trial, then before it, there are $(y-1)$ successes in $(n-1)$ trials which can be obtained in $\binom{n-1}{y-1}$ ways. No. of failures obtained is $(n-1)-(y-1) = n-y$.



Prob. of getting $(y-1)$ successes is p^{y-1}

Prob. of getting $(n-y)$ failures is q^{n-y}

Hence probability of getting yth success in nth trial is

$$P(y) = \binom{n-1}{y-1} p^{y-1} q^{n-y} \cdot p$$

$$= \binom{n-1}{y-1} p^{y} q^{n-y}, \quad n = y, y+1, \ldots$$

* Negative binomial r.v. with y successes can be view as sum of y independent geometric r.v., i.e. $Y = \sum_{i=1}^{y} Y_i$, Y_i = g.r.v.

If the r.v. X counts no. of failure before getting yth success from nth trial, then we have $n = y + x = x + y \Rightarrow x = n - y$.

Thus the pmf of X is

$$p(x) = \binom{x+y-1}{y-1} p^{y} q^{x}, \quad x = 0, 1, 2, \ldots$$

* For $y=1$, we have $n = x+1 \Rightarrow x = n-1$

$$\therefore\ P(x) = \binom{n-1+1-1}{0} p^1 q^{n-1}$$

$$= \binom{n-1}{0} p q^{n-1} = p q^{n-1}, \quad n = 1, 2, \ldots$$

which is geometric distribution.

Ex A fair die is rolled. Find the prob. of getting 1st six in 4th trial.

Sol^n Let X = counts the position of getting 1st six (success or S)

$p = P(S)$ = probability of getting a six
$= \frac{1}{6}$

$q = P(F)$ = Prob. of not getting a six
$= 1 - \frac{1}{6} = 5/6$

$\therefore P\{$getting 1st six in 4th trial$\}$

$$= P[X=4] = p(x) = p q^{4-1} = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 = 0.0955$$

Ex A fair die is rolled. Find the prob. of getting 2nd six in 4th trial.

Sol^n Let Y = counts the position at which 2nd six is obtained.

To find $P[Y=4] = \binom{4-1}{2-1} p^2 q^{4-2} = \binom{3}{1}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
$= 0.0579$

Alternative

Let X = no. of failures before getting 2nd six

We have $y = 2$ (no. of success), $x = 2$ no. of failures

$\boxed{n=4}$ $\therefore P[X=2] = \binom{x+y-1}{y-1} p^y q^{n-y} = \binom{3}{1} p^2 q^2 = 0.0579$

The negative binomial distribution is ~~defined~~ denoted by, for $y = 1, 2, \ldots$

$$nb(x; y, p) = {}^{x+y-1}C_{y-1}\, p^y q^{x-y},$$

where $x = 0, 1, 2, \ldots$

X = no. of failures

Y = no. of successes

p = prob. of success

Ex $$nb(10; 5, 0.2) = {}^{10+5-1}C_{5-1} (0.2)^5 (0.8)^5$$

$$= {}^{14}C_4 (0.2)^5 (0.8)^5$$

$$= 0.034$$

The PDF $X \sim NB(y, p)$

$$F(x) = P[X \le x] = \sum_{k=0}^{x} nb(k; y, p)$$

Ex If $X \sim NB(3, 0.3)$, then

$$F(2) = P[X \le 2]$$

$$= \sum_{x=0}^{2} nb(x; y, p)$$

$$= \sum_{x=0}^{2} {}^{x+y-1}C_{y-1}\, p^y q^x \qquad \boxed{\because y = 3,\ p = 0.3}$$

$$= \sum_{x=0}^{2} {}^{x+2}C_2 (0.3)^2 (0.7)^x$$

$$= \underline{0.5436}$$

Mean & variance of NB dist.

$$\mu = E(X) = \frac{yq}{p}, \quad \sigma^2 = V(X) = \frac{yq}{p^2}, \quad q = 1 - p$$