

Geometric & Negative Binomial Distributions

(5 pages; 11/2/25)

Geometric Distribution

(1) Let X be the number of attempts needed to achieve a success, where the probability of success at each attempt is p (a constant).

(The attempts are assumed to be independent.)

Then $X \sim \text{Geo}(p)$ (a discrete distribution)

and $P(X = r) = (1 - p)^{r-1}p$, for $r = 1, 2, \dots$

[ie $r - 1$ failures, followed by a success]

(2) To confirm that $(1 - p)^{r-1}p$ is a pdf:

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

[being the sum of a geometric series, with 1st term p and common ratio $1 - p$]

$$\begin{aligned} (3) E(X) &= \sum_{r=1}^{\infty} r(1 - p)^{r-1}p = -p \frac{d}{dp} \sum_{r=1}^{\infty} (1 - p)^r \\ &= -p \frac{d}{dp} \left\{ \frac{1 - p}{1 - (1 - p)} \right\} = -p \frac{d}{dp} \left\{ \frac{1}{p} - 1 \right\} = -p \left(-\frac{1}{p^2} \right) = \frac{1}{p} \end{aligned}$$

$$(4) E(X(X - 1)) = \sum_{r=1}^{\infty} r(r - 1)(1 - p)^{r-1}p$$

$$\text{Now } \frac{d}{dp} \{(1 - p)^r\} = -r(1 - p)^{r-1}$$

$$\text{and } \frac{d}{dp} \{-r(1 - p)^{r-1}\} = r(r - 1)(1 - p)^{r-2},$$

$$\text{so that } r(r-1)(1-p)^{r-1}p = p(1-p)\{r(r-1)(1-p)^{r-2}\}$$

$$= p(1-p)\frac{d^2}{dp^2}\{(1-p)^r\}$$

$$\text{and } E(X(X-1)) = p(1-p)\sum_{r=1}^{\infty}\frac{d^2}{dp^2}\{(1-p)^r\}$$

$$= p(1-p)\frac{d^2}{dp^2}\{\sum_{r=1}^{\infty}(1-p)^r\}$$

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

$$= p(1-p)\frac{d}{dp}\left\{-\frac{1}{p^2}\right\} = p(1-p)\left(\frac{2}{p^3}\right) = \frac{2(1-p)}{p^2}$$

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

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

$$= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2}(2 - 2p + p - 1) = \frac{1-p}{p^2} \text{ or } \frac{q}{p^2}$$

(5) Cumulative probabilities

Approach A

$$P(X \leq k) = 1 - P(X > k) = 1 - P(k \text{ failures})$$

$$= 1 - (1-p)^k$$

Approach B

$$P(X \leq k) = \sum_{r=1}^k (1-p)^{r-1}p = \frac{p\{1-(1-p)^k\}}{1-(1-p)}$$

[the sum of k terms of a geometric series with 1st term p and common ratio $1-p$]

$$= 1 - (1-p)^k$$

(6) Probability generating function [see "PGFs"]

$$G_X(s) = E(s^X) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=1}^{\infty} q^{k-1} p s^k$$

$$= ps \sum_{k=1}^{\infty} (qs)^{k-1} = \frac{ps}{1-qs} \text{ if } |qs| < 1; \text{ ie } |s| < \frac{1}{q}$$

$$G'_X(s) = \frac{(1-qs)p - ps(-q)}{(1-qs)^2} = \frac{p}{(1-qs)^2}$$

$$\text{and } G''_X(s) = \frac{-2(-q)p}{(1-qs)^3} = \frac{2qp}{(1-qs)^3}$$

$$\text{Then } E[X] = G'_X(1) = \frac{1}{p}$$

$$\text{and } Var(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2$$

$$= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2(1-p)+p-1}{p^2} = \frac{1-p}{p^2} \text{ or } \frac{q}{p^2}$$

Negative Binomial Distribution

(1) Let Y be the number of attempts needed to achieve n successes, where the probability of success at each attempt is p (a constant), and the attempts are assumed to be independent.

Then the following notation is sometimes used:

$Y \sim nb(n, p)$ (a discrete distribution)

$$\text{and } P(Y = r) = \binom{r-1}{n-1} p^{n-1} q^{(r-1)-(n-1)} p$$

[The probability of achieving $n - 1$ successes, followed by $(r - 1) - (n - 1)$ failures (making a total of $r - 1$ attempts), and then one success, is $p^{n-1} q^{(r-1)-(n-1)} p$,

and $\binom{r-1}{n-1}$ is the number of ways of having $n - 1$ successes

amongst $r - 1$ attempts.]

$$= \binom{r-1}{n-1} p^n q^{r-n}, \text{ for } r = n, n+1, \dots$$

Notes

(i) The Geometric distribution is the special case of the Negative Binomial distribution when $n = 1$.

(ii) The Negative Binomial distribution is (sometimes) said to get its name from the fact that it is effectively measuring the number of failures occurring before n successes are achieved.

(2) Probability generating function of Y

(i) $Y = X_1 + \dots + X_n$, where $X_i \sim \text{Geo}(p)$

(ii) $G_{X_1+X_2}(s) = G_{X_1}(s)G_{X_2}(s)$, where X_1 & X_2 are independent random variables [see "PGFs"]

(iii) $G_{X_i}(s) = \frac{ps}{1-qs}$, so that (extending (i) to a sum of n variables)

$$G_Y(s) = \left(\frac{ps}{1-qs} \right)^n$$

(3) Derivation of $E(Y)$

$$G_Y(s) = \left(\frac{ps}{1-qs} \right)^n$$

So $G'_Y(s) = n \left(\frac{ps}{1-qs} \right)^{n-1} \cdot \frac{(1-qs)p - ps(-q)}{(1-qs)^2}$ (by the Quotient rule)

$$= n \left(\frac{ps}{1-qs} \right)^{n-1} \cdot \frac{p}{(1-qs)^2} = \frac{np^n s^{n-1}}{(1-qs)^{n+1}}$$

$$\text{and } E(Y) = G'_Y(1) = \frac{np^n}{p^{n+1}} = \frac{n}{p}$$

(4) Derivation of $Var(Y)$

$$Var(Y) = G''_Y(1) + G'_Y(1) - [G'_Y(1)]^2$$

$$\text{From (3), } G'_Y(s) = \frac{np^n s^{n-1}}{(1-qs)^{n+1}},$$

so that, by the Quotient rule,

$$G''_Y(s) = \frac{np^n[(1-qs)^{n+1}(n-1)s^{n-2} - s^{n-1}(n+1)(1-qs)^n(-q)]}{(1-qs)^{2(n+1)}}$$

$$\text{and hence } G''_Y(1) = \frac{np^n[p^{n+1}(n-1) - (n+1)p^n(-q)]}{p^{2(n+1)}}$$

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

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

$$= \frac{n[n+q-p]}{p^2}$$

$$\text{Then } Var(Y) = G''_Y(1) + G'_Y(1) - [G'_Y(1)]^2$$

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

$$= \frac{n[n+q-p+p-n]}{p^2}$$

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