Geometric and Negative Binomial Distributions

Definition 6 Geometric distribution A random variable X is defined to have *geometric* (or *Pascal*) *distribution* if the density of X is given by

$$f_X(x) = f_X(x; p)$$

$$= \begin{cases} p(1-p)^x & \text{for } x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases} = p(1-p)^x I_{\{0, 1, \ldots\}}(x), \qquad (11)$$

where the parameter p satisfies $0 < p \leq 1$. (Define $q = 1 - p$.) ////

Definition 7 Negative binomial distribution A random variable X with density

$$f_X(x) = f_X(x; r, p)$$

$$= \begin{cases} \binom{r+x-1}{x} p^r q^x = \binom{-r}{x} p^r (-q)^x & \text{for } x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \qquad (12)$$

$$= \binom{r+x-1}{x} p^r q^x I_{\{0, 1, \ldots\}}(x),$$

where the parameters r and p satisfy $r = 1, 2, 3, \ldots$ and $0 < p \leq 1$ ($q = 1 - p$), is defined to have a *negative binomial distribution*. The density given by Eq. (12) is called a *negative binomial density*.

Remark If in the negative binomial distribution $r = 1$, then the negative binomial density specializes to the geometric density. ////

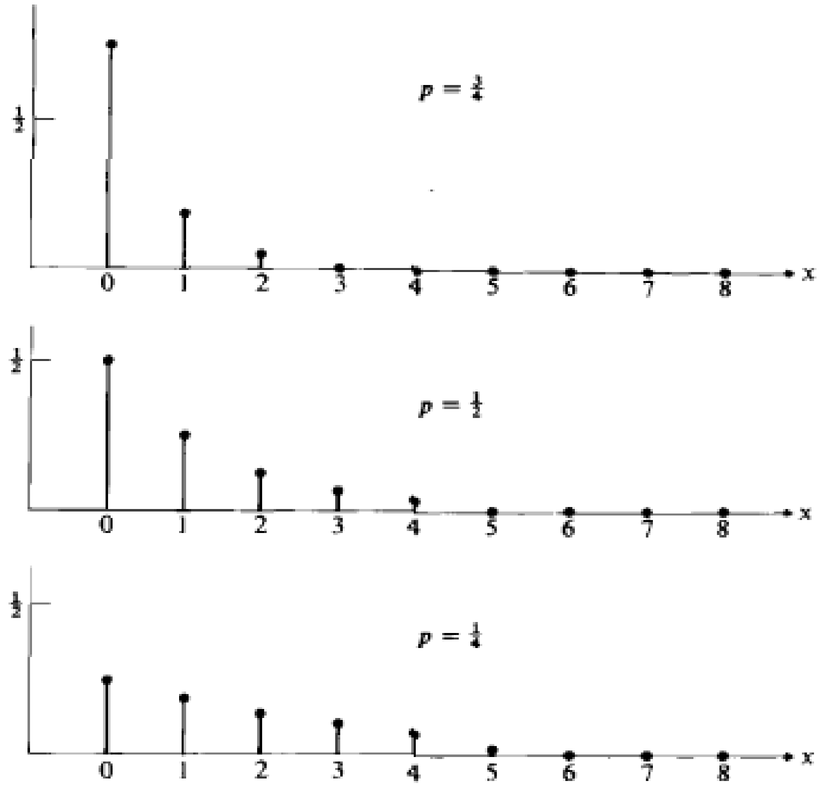

FIGURE 7
Geometric densities

Theorem 9 If the random variable X has a geometric distribution, then

$$\mathscr{E}[X] = \frac{q}{p}, \qquad \text{var}\,[X] = \frac{q}{p^2}, \qquad \text{and} \qquad m_X(t) = \frac{p}{1 - qe^t}. \qquad (13)$$

Before leaving the geometric distribution, we note that some authors define the geometric distribution by assuming 1 (instead of 0) is the smallest mass point. The density then has the form

$$f(x; p) = p(1 - p)^{x-1} I_{\{1, 2, \dots\}}(x), \qquad (14)$$

and the mean is $1/p$, the variance is q/p^2, and the moment generating function is $pe^t/(1 - qe^t)$.

Theorem 11 Let X have a negative binomial distribution; then

$$\mathscr{E}[X] = \frac{rq}{p}, \qquad \text{var}[X] = \frac{rq}{p^2}, \qquad \text{and} \qquad m_X(t) = \left[\frac{p}{1 - qe^t}\right]^r. \tag{15}$$

PROOF

$$m_X(t) = \mathscr{E}[e^{tX}] = \sum_{x=0}^{\infty} e^{tx}\binom{-r}{x}p^r(-q)^x$$

$$= \sum_{x=0}^{\infty}\binom{-r}{x}p^r(-qe^t)^x = \left[\frac{p}{1 - qe^t}\right]^r$$

[see Eq. (33) in Appendix A].

$$m'_X(t) = p^r(-r)(1 - qe^t)^{-r-1}(-qe^t)$$

and

$$m''_X(t) = rqp^r[q(r+1)e^{2t}(1 - qe^t)^{-r-2} + e^t(1 - qe^t)^{-r-1}];$$

hence

$$\mathscr{E}[X] = m'_X(t)\Big|_{t=0} = \frac{rq}{p}$$

and

$$\text{var}[X] = m''_X(t)\Big|_{t=0} - (\mathscr{E}[X])^2 = rqp^r[qp^{-r-2}(r+1) + p^{-r-1}] - \left(\frac{rq}{p}\right)^2$$

$$= \frac{rq^2}{p^2} + \frac{rq}{p} = \frac{rq}{p^2}. \qquad ////$$

EXAMPLE 10 Consider a sequence of independent, repeated Bernoulli trials with p equal to the probability of success on an individual trial. Let the random variable X represent the number of failures prior to the rth success; then X has the negative binomial density given by Eq. (12), as the following argument shows: The last trial must result in a success, having probability p; among the first $x + r - 1$ trials there must be $r - 1$ successes and x failures, and the probability of this is

$$\binom{x+r-1}{r-1}p^{r-1}q^x = \binom{r+x-1}{x}p^{r-1}q^x,$$

which when multiplied by p gives the desired result. ////

A random variable X having a negative binomial distribution is often referred to as a discrete *waiting-time* random variable. It represents how long (in terms of the number of failures) one waits for the rth success.