Geometric Distribution

Geometric random variables arise in experiments with the following properties:

- The experiment consists of a series of trials. The outcome of each trial can be either a “success (s)” or a “failure (f)” – “Bernoulli trial” with success probability of p.
- The trials are identical and independent. The probability of success p, remains the same from trial to trial.
- The random variable X denotes ***the number of trials needed to obtain the first success***.

The sample space of the outcome of the experiment: $S = \{s, fs, ffs, fffs, \ldots\}$

X takes values from: $\{1,2,3, \ldots\}$

Can X take the value of 0?

Density Function of Geometric Distribution

Let's find the density of a geometric distribution:

$f(1) = P[X = 1] = P[success\ on\ the\ first\ trial] = p$

$f(2) = P[X = 2] = P[fail\ on\ the\ first\ trial\ and\ success\ on\ the\ second\ trial]$

$\qquad = P[fail\ on\ the\ first\ trial]P[succeed\ on\ the\ second\ trial]$

$\qquad = (1 - p)p$

$f(3) = P[X = 3] = P[fail\ on\ the\ first\ trial\ and\ fail\ on\ the\ second\ trial\ and\ success\ on\ the\ third\ trial]$

$\qquad = P[fail\ on\ the\ first\ trial]P[fail\ on\ the\ second\ trial]P[success\ on\ the\ third\ trial]$

$\qquad = (1 - p)^2 p$

$\vdots$

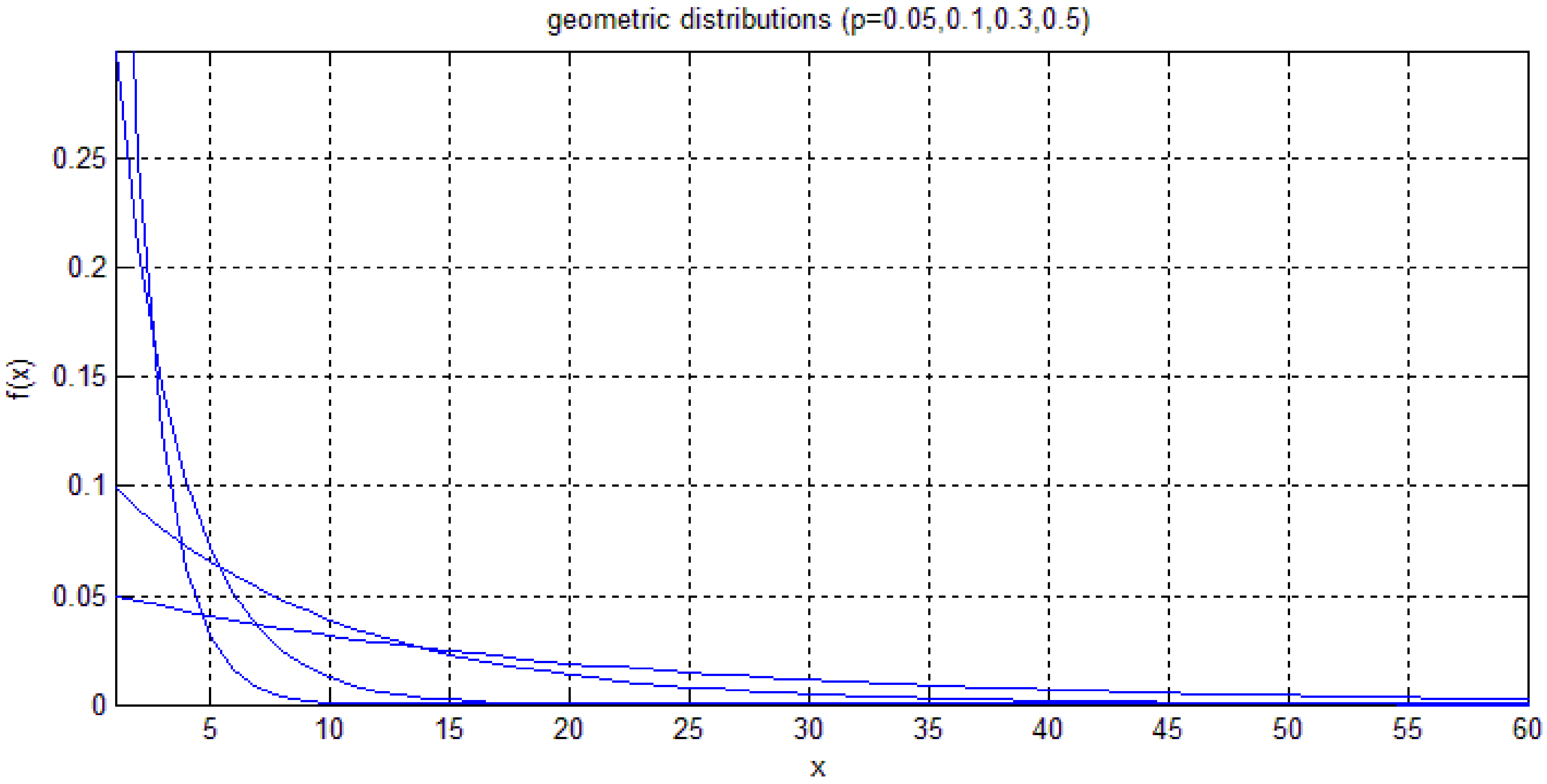
In general, we have, $f(x) = P[X = x] = (1 - p)^{x-1}p,\ \ x = 1,2,3, \ldots$

We define a random variable X as having a geometric distribution with parameter p, if its density function is given by

$$f(x) = (1 - p)^{x-1}p, \qquad 0 < p < 1, \qquad x = 1,2,3, \ldots$$

Geometric Distribution

$$f(x) = P[X = x] = (1-p)^{x-1}p, \qquad 0 < p < 1, \qquad x = 1,2,3,\ldots$$



Verify that $f(x) = P[X = x] = (1-p)^{x-1}p,\ 0 < p < 1,\ x = 1,2,3,\ldots$
is a probability density function.

Geometric Distribution

Example: In a simulation experiment, one want to generate random digits by randomly selecting one digit from $\{0,1,2,3,4,5,6,7,8,9\}$ at each stage. Let X denote the number of trials needed to obtain the first zero. What is the density function of X?

X is a geometric random variable with $p = \frac{1}{10}$
$f(x) = (1-p)^{x-1}p = 0.1(0.9)^{x-1}, x = 1,2,3, \ldots$

Guess the value of $E[X]$

Geometric Distribution

Example: Given a geometric random variable X with parameter p. Find $E[X]$

Solution:

The probability density function of X is $f(x) = (1-p)^{x-1}p, x = 1,2,3, \dots, where, 0 < p < 1, is\ the\ success\ rate.$

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

Where,

$$\begin{aligned}
\textstyle\sum_{x=1}^{\infty} x(1-p)^{x-1} &= 1 + 2(1-p) + 3(1-p)^2 + 4(1-p)^3 + 5(1-p)^4 + \cdots \\
&= 1 + (1-p) + (1-p)^2 + (1-p)^3 + (1-p)^4 + \cdots \\
&\quad + (1-p) + (1-p)^2 + (1-p)^3 + (1-\mathrm{p})^4 + \cdots \\
&\quad + (1-p)^2 + (1-p)^3 + (1-p)^4 + \cdots \\
&\quad + (1-p)^3 + (1-p)^4 + \cdots \\
&\quad \vdots \\
&= \textstyle\sum_{x=1}^{\infty}(1-p)^{x-1} + \sum_{x=2}^{\infty}(1-p)^{x-1} + \sum_{x=3}^{\infty}(1-p)^{x-1} + \cdots \\
&= \tfrac{1}{p} + \left[\tfrac{1}{p} - 1\right] + \left[\tfrac{1}{p} - 1 - (1-p)\right] + \left[\tfrac{1}{p} - 1 - (1-p) - (1-p)^2\right] + \cdots \\
&= \tfrac{1}{p} + \tfrac{1-p}{p} + \tfrac{(1-p)^2}{p} + \tfrac{(1-p)^3}{p} + \cdots
\end{aligned}$$

Recall geometric series result:

$$\sum_{k=1}^{\infty} ar^{k-1} = \frac{a}{1-r}, |r| < 1$$

Therefore,

$$E[X] = p\left[\sum_{x=1}^{\infty} x(1-p)^{x-1}\right] = 1 + (1-p) + (1-p)^2 + (1-p)^3 + \cdots = \sum_{x=1}^{\infty}(1-p)^{x-1} = \frac{1}{p}$$

Ordinary moments

Definition: Let X be a random variable. The kth ordinary moment for X is defined as $E[X^k]$
$E[X] = \mu$ is the first ordinary moment of X
$E[X^2]$ is the second ordinary moment of X

The moment generating function is a way to find these moments for a RVs

Moment generating function (m.g.f)

Definition: Let X be a random variable with density f. The moment generating function for X, denoted by $m_x(t)$ is given by

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

Provided this expectation is finite for all real number t in some open interval.

The m.g.f of a distribution is unique.

Moment generating function (mgf)

Example: Let X be a geometric random variable. Find the m.g.f of X

Solution:
The probability density function of a geometric random variable X is
$f(x) = (1-p)^{x-1}p = q^{x-1}p, q = 1-p, x = 1,2,3, \ldots, where, 0 < p < 1, is\ the\ success\ rate.$
Then, by definition,

$$m_X(t) = E[e^{tX}] = \sum_{all\ x} e^{tx} f(x) = \sum_{all\ x} e^{tx} q^{x-1} p = pq^{-1} \sum_{all\ x} e^{tx} q^x = pq^{-1} \sum_{all\ x} (e^t q)^x$$

Recall the result from geometric series, that,

$$\sum_{k=1}^{\infty} ar^{k-1} = \frac{a}{1-r}, |r| < 1$$

Then,

$$m_X(t) = E[e^{tX}] = pq^{-1}\left(\frac{qe^t}{1-qe^t}\right) = \frac{pe^t}{1-qe^t}$$

Provided that $|r| = |qe^t| < 1$. This requires $qe^t < 1 \Rightarrow t < -\ln q$

Moment generating function (m.g.f)

Theorem 3.1: Let $m_x(t)$ be the m.g.f of a random variable X. Then $\frac{d^k m_x(t)}{dt^k}|_{t=0} = E[X^k]$

Proof:

Recall that, $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$. Then, we have, $e^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots$

Therefore,

$$m_X(t) = E[e^{tX}] = E\left[1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right] = 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \cdots$$

$$\frac{dm_X(t)}{dt} = E[X] + tE[X^2] + \frac{t^2}{2!}E[X^3] + \cdots$$

Hence,

$$\frac{dm_X(t)}{dt}|_{t=0} = E[X]$$

Also,

$$\frac{d^2 m_X(t)}{dt^2} = E[X^2] + tE[X^3] + \frac{t^2}{2!}E[X^4] + \cdots$$

And,

$$\frac{d^2 m_X(t)}{dt^3}|_{t=0} = E[X^2]$$

In general,

$$\frac{d^k m_x(t)}{dt^k}|_{t=0} = E[X^k]$$

Moment generating function (m.g.f)

Practice Example: Let X be a geometric random variable with parameter p. The moment generating function of X is $m_X(t) = \frac{pe^t}{1-qe^t}, q = 1-p$. Find $E[X], E[X^2]$ and $VarX$.

The moment generating function of a geometric random variable with parameter p is:

$$m_X(t) = E[e^{tX}] = pq^{-1}\left(\frac{qe^t}{1-qe^t}\right) = \frac{pe^t}{1-qe^t}, q = 1-p$$

Then,

$$\frac{dm_X(t)}{dt} = \frac{pe^t(1-qe^t) + pe^tqe^t}{(1-qe^t)^2} = \frac{pe^t}{(1-qe^t)^2}$$

$$E[X] = \frac{dm_X(t)}{dt}|_{t=0} = \frac{p}{(1-q)^2} = \frac{1}{p}$$

Also,

$$\frac{d^2m_X(t)}{dt^2} = \frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

$$E[X^2] = \frac{d^2m_X(t)}{dt^3}|_{t=0} = \frac{p(1+q)}{(1-q)^3} = \frac{1+q}{p^2}$$

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

Binomial Distribution

Binomial random variables arise in experiments with the following properties:

- The experiment consists of a fixed number, n, of Bernoulli trials each with probability p of success.
- The trials are identical and independent. The probability of success p, remains the same from trial to trial.
- The random variable X denotes ***the number of success obtained in the*** $\boldsymbol{n}$ ***trials***.

Let's take a look at a special case with $n = 3$:

The sample space is: $S = \{fff, sff, fsf, ffs, ssf, sfs, fss, sss\}$

X can take values from: {0,1,2,3}

$$f(0) = P[X = 0] = (1 - p)^3$$
$$f(1) = P[X = 1] = 3 \times (1 - p)^2 p$$
$$f(2) = P[X = 2] = 3 \times (1 - p) p^2$$
$$f(3) = P[X = 3] = p^3$$

i.e., $f(x) = \binom{3}{x} p^x (1 - p)^{3-x}, x = 0,1,2,3$