Chapter 10

Geometric population distribution $X \sim Ge(p)$

10.1 Definition of a Geometric distribution, $X \sim Ge(p)$

If the random variable X has a Geometric population distribution,

i.e. $X \sim Ge(p)$, then its probability function is given by

$$\boxed{p(X = x) = p(1 - p)^{x-1}} \qquad \text{for } x = 1, 2, 3, \ldots\ldots$$

A Geometric distribution is often a good model for discrete random variables which are

***X* = number of trials until the first time event A occurs**

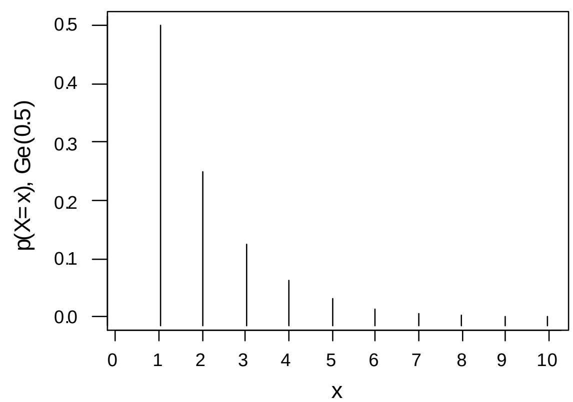
e.g.

***X* = number of trials to the first 'success'**
X = number of tosses of a coin to the first heads
X = number of tosses of a dice to the first six

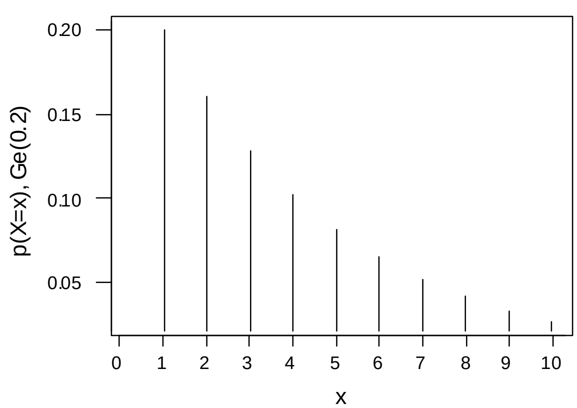
***X* = number of people interviewed to the first person of type A**
X = number of people interviewed to the first Labour supporter
X = number of babies to the first girl
X = number of people tested to the first person cured

NOTE the range of possible values of the variable *X* is $x = 1, 2, 3, \ldots\ldots,$
i.e. *X* is a discrete variable which takes positive integer values.

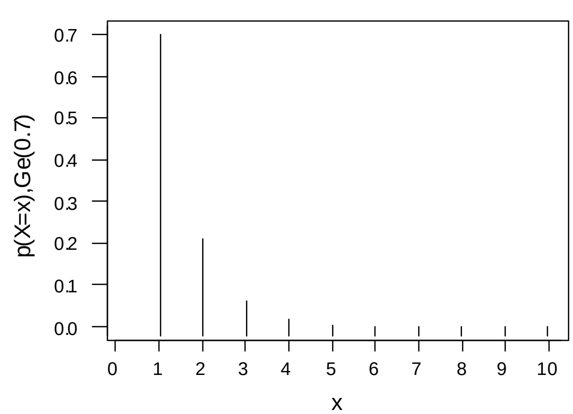
probability function for $X \sim Ge(0.5)$



probability function for $X \sim Ge(0.2)$



probability function for $X \sim Ge(0.7)$



10.2 Theoretical conditions leading to a Geometric variable $X \sim Ge(p)$

Theorem 1

If X counts the number of times event trials (i.e. repetitions of an experiment) until the first time event A occurs, then X has a Geometric distribution,

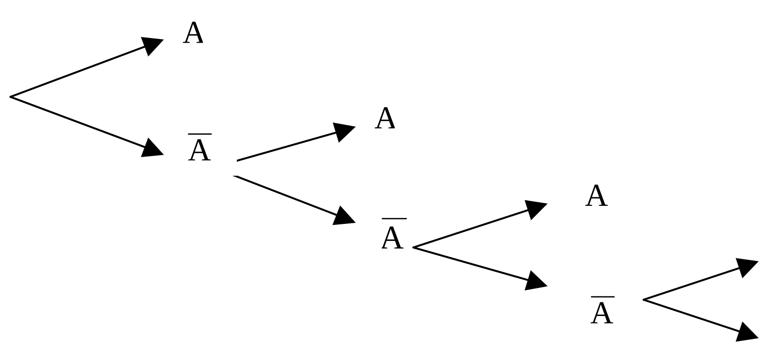
$$X \sim Ge(p)$$

with $$p(X = x) = p(1-p)^{x-1} \qquad \text{for } x = 1, 2, 3, \ldots$$

provided

i) the outcomes of the trials are independent of each other
ii) the probability of event A occurring is the same value p for each of the n trials

Proof 1



X = number of trials until the first time event A occurs e.g. X = number of tosses until the first heads occurs

where p = probability of heads on each toss of the coin

and $q = (1 - p)$ = probability of tails on each toss of the coin

sample outcomes	x	p($X = x$)
H	1	p
TH	2	qp
TTH	3	$q^2 p$
TTTH	4	$q^3 p$
TTTTH	5	$q^4 p$
….	…	…

$$\mathrm{p}(X = x) = \mathrm{p}(\underbrace{\mathrm{TT\ldots\ldots T}}_{x-1}\underbrace{\mathrm{H}}_{1}) = qq\ldots q\,p = q^{x-1}p = p(1-p)^{x-1}$$

Hence $$\boxed{p(X = x) = p(1-p)^{x-1}}$$ for $x = 1, 2, 3, \ldots\ldots$

10.3 Population summary measures for a Geometric variable $X \sim Ge(p)$

mean $\mu_X = \dfrac{1}{p}$

variance $\sigma_X^2 = \dfrac{q}{p^2}$

standard deviation $\sigma_X = \sqrt{\dfrac{q}{p^2}}$

10.4 Cumulative distribution function for $X \sim Ge(p)$

$$F_X(x) = p(X \le x) = \sum_{t=1}^{x} p(X = t) = \sum_{t=1}^{x} pq^{t-1} = p + pq + pq^2 + pq^3 + ... + pq^{x-1} =$$

$$\frac{p(1-q^x)}{(1-q)}$$

$$\boxed{F_X(x) = (1-q^x)}$$

10.5 Procedure for finding a Geometric probability

1) identify the event A, 'success'

2) find the probability p of event A occurring in each of the trials

Calculate $p(X = x) = p(1-p)^{x-1}$ for $x = 1, 2, 3, \ldots.$

10.6 Examples

Example 1

Toss a fair coin repeatedly and count the number of tosses until the first heads occurs

What is the distribution of *X*?

$$X \sim Ge(p) = Ge(0.5)$$

What is the probability of getting the first heads on the 5th toss?

$$p(X = 5) = p(1-p)^4 = 0.5 * 0.5^4 = 0.5^5 = 0.03125$$

Example 2

Toss a fair dice repeatedly and count the number of tosses until the first six occurs

What is the distribution of *X*?

$$X \sim Ge(p) = Ge(\frac{1}{6})$$

What is the probability of getting the first six on the 10th toss?

$$p(X = 10) = p(1-p)^9 = \frac{1}{6} * \frac{5}{6}^9 = 0.0323$$

What is the probability of taking more than 6 tosses to get the first six?

$$p(X > 6) = 1 - p(X \le 6) = 1 - F_X(6) = \left(\frac{5}{6}\right)^6 0.335$$

10.7 Comparison between a Binomial and Geometric distribution

	number of times A occurs	number of trials
Binomial	X	n
Geometric	1	X
Negative Binomial	k	X