10 GEOMETRIC DISTRIBUTION

EXAMPLES:

1. Terminals on an on-line computer system are attached to a communication line to the central computer system. The probability that any terminal is ready to transmit is 0.95.
   Let X = number of terminals polled until the first ready terminal is located.

2. Toss a coin repeatedly.
   Let X = number of tosses to first head

3. It is known that 20% of products on a production line are defective. Products are inspected until first defective is encountered.
   Let X = number of inspections to obtain first defective

4. One percent of bits transmitted through a digital transmission are received in error. Bits are transmitted until the first error.
   Let X denote the number of bits transmitted until the first error.

GEOMETRIC DISTRIBUTION
Conditions:

1. An experiment consists of repeating trials until first success.
2. Each trial has two possible outcomes;
   (a) A success with probability p
   (b) A failure with probability $q = 1 - p$.
3. Repeated trials are independent.

X = number of trials to first success

X is a **GEOMETRIC RANDOM VARIABLE**.

PDF:

$$P(X = x) = q^{x-1}p; \quad x = 1, 2, 3, \cdots$$

CDF:

$$\begin{aligned} P(X \leq x) &= P(X = 1) + P(X = 2) \cdots P(X = x) \\ &= p + qp + q^2 p \cdots + q^{x-1}p \\ &= p[1 - q^x]/(1 - q) \\ &= 1 - q^x \end{aligned}$$

Example:
Products produced by a machine has a 3% defective rate.

- What is the probability that the first defective occurs in the fifth item inspected?

$$
\begin{aligned}
P(X=5) &= P(\text{1st 4 non-defective })P(\text{ 5th defective}) \\
&= \left(0.97^4\right)(0.03)
\end{aligned}
$$

In R

```
>dgeom (x= 4, prob = .03)
[1] 0.02655878
```

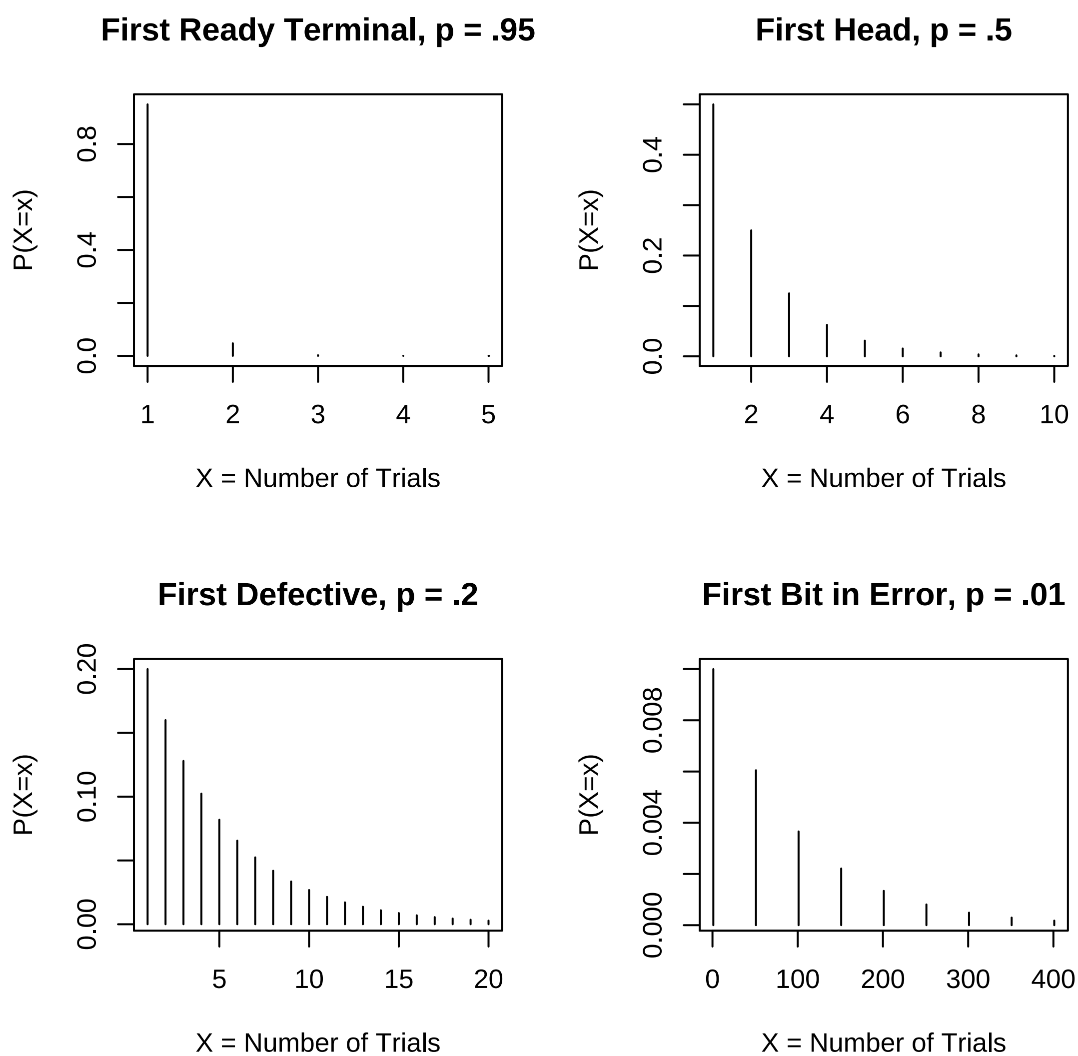
The convention in R is to record X as the number of failures that occur **before** the first success.

- What is the probability that the first defective occurs in the first five inspections?

$$
\begin{aligned}
P(X \leq 5) &= 1 - P(\text{First 5 non-defective}) \\
&= 1 - 0.97^5
\end{aligned}
$$

```
> pgeom(4, .03)
[1] 0.1412660
```

Geometric *pdf s*



Calculating pdfs in *R*

```
par (mfrow = c(2,2))

x<-0:4
plot(x+1, dgeom(x, prob = .95),
   xlab = "X = Number of Trials", ylab = "P(X=x)",
   type = "h", main = "First Ready Terminal, p = .95")

x<-0:9
plot(x+1, dgeom(x, prob = .5),
  xlab = "X = Number of Trials", ylab = "P(X=x)",
   type = "h", main = "First Head, p = .5")

x<- 0:19
plot(x+1, dgeom(x, prob = .2),
  xlab = "X = Number of Trials", ylab = "P(X=x)",
  type = "h", main = "First Defective, p = .2")


x<- seq(0, 400, 50)
plot(x+1, dgeom(x, prob = .01),
  xlab = "X = Number of Trials", ylab = "P(X=x)",
  type = "h", main = "First Bit in Error, p = .01")
```

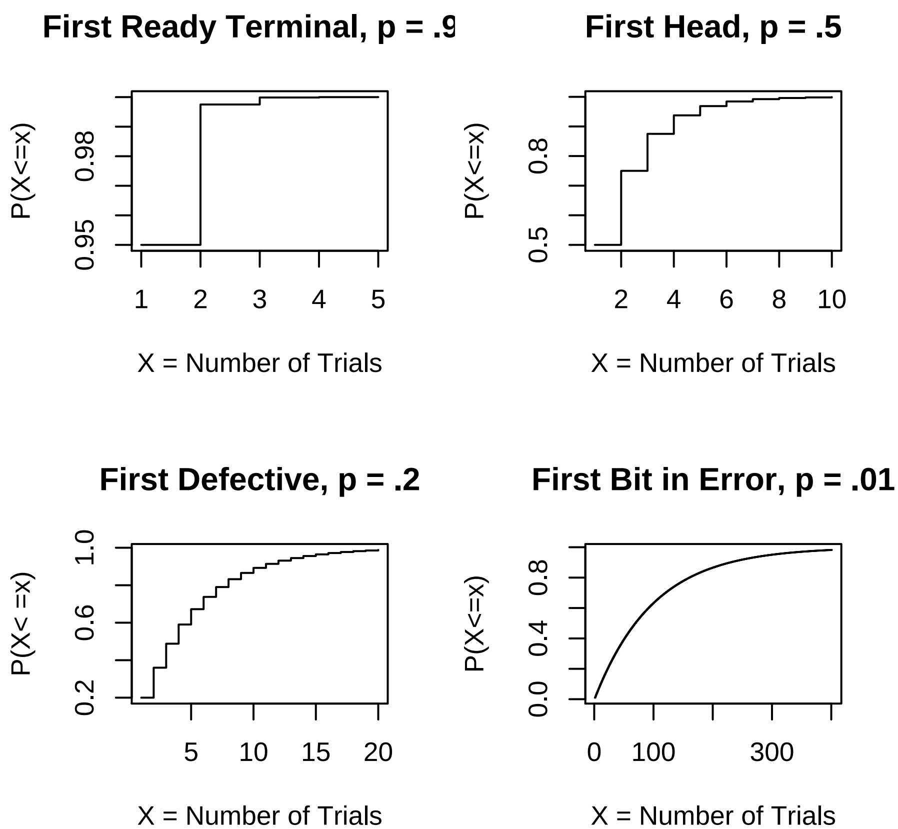


Figure 1: Geometric cdfs

```
par (mfrow = c(2,2))

x<-0:4
plot(x+1, pgeom(x, prob = .95),
   xlab = "X = Number of Trials", ylab = "P(X<=x)",
   type = "s", main = "First Ready Terminal, p = .95")

x<-0:9
plot(x+1, pgeom(x, prob = .5),
   xlab = "X = Number of Trials", ylab = "P(X<=x)",
   type = "s", main = "First Head, p = .5")

x<-0:19
plot(x+1, pgeom(x, prob = .2),
   xlab = "X = Number of Trials", ylab = "P(X< =x)",
   type = "s", main = "First Defective, p = .2")

x<- seq(0, 399)
plot(x+1, pgeom(x, prob = .01),
   xlab = "X = Number of Trials", ylab = "P(X<=x)",
   type = "s", main = "First Bit in Error, p = .01")
```

The Quantile Function

In Example 3, a production line which has a 20% defective rate, what is the minimum number of inspections, that would be necessary so that the probability of observing a defective is more that 75%?

Choose k so that

$$P(X \leq k) \geq .75.$$

In *R*

```
qgeom(.75, .2)
[1] 6
```

i.e. 6 failures before first success.

or with 7 inspections, there is at least a 75% chance of obtaining the first defective.

Mean of geometric distribution:

Example:

If a production line has a 20% defective rate. What is the average number of inspections to obtain the first defective?

$$
\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x q^{x-1} p \\
&= p \sum_{x=1}^{\infty} x q^{x-1} \\
&= p \sum_{x=1}^{\infty} \frac{dq^x}{dq} \\
&= p \frac{d \sum_{x=1}^{\infty} q^x}{dq} \\
&= p \frac{d(q/(1-q)}{dq} \\
&= p \frac{[(1-q)+q]}{(1-q)^2} \\
&= \frac{p}{p^2} = \frac{1}{p}
\end{aligned}
$$

Average number of inspections to obtain the first defective:

$$E(X) = \frac{1}{.2} = 5$$

The Markov Property:

If the probability of events happening in the future is independent of what went before, then the random variable is said to have the **Markov property**.

MARKOV PROPERTY
$\Longrightarrow$ MEMORYLESS PROPERTY

Example:

Products are inspected until first defective is found. X is a geometric random variable with parameter p. The first 10 trials have been found to be free of defectives. What is the probability that the first defective will occur in the 15th trial?

Let E_1 be the event that first ten trials are free of defectives.
Let E_2 be the event that that first defective will occur on the 15th trial.

$$\begin{aligned} P(X = 15|X > 10) &= P(E_2|E_1) \\ &= \frac{P(E_1 \cap E_2)}{P(E_1)} \\ &= \frac{P(X = 15 \cap X > 10)}{P(X > 10)} \\ &= \frac{P(X = 15)}{P(X > 10)} \\ &= \frac{q^{14}p}{q^{10}} = q^4 p = P(X = 5) \end{aligned}$$

MARKOV PROPERTY

Generally, the Markov property states:

$$P(X = x + n|X > n) = P(X = x)$$

Proof:
Let

$$\begin{aligned} E_1 &= \{X > n\} \\ E_2 &= \{X = x + n\} \end{aligned}$$

Then we may write

$$P(X = x + n|X > n) = P(E_2|E_1)$$

But

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

Now

$$P(E_1 \cap E_2) = P(X = x + n) = q^{x+n-1}p$$

And

$$P(E_1) = P(X > n) = q^n$$

Thus

$$\begin{aligned} P(E_2|E_1) &= \frac{q^{x+n-1}p}{q^n} \\ &= q^{x-1}p \end{aligned}$$

But

$$P(X = x) = q^{x-1}p$$

Hence

$$P(X = x + n|(X > n) = P(X = x)$$

R Functions for the Geometric Distribution

- dgeom

  `dgeom (x= 4, prob = .03)`

  the probability of

  exactly 4 trials before first defective or

  exactly 5 trials to first defective

- pgeom

  `pgeom (x= 4, prob = .03)`

  the probability of

  up to 4 trials before first defective or

  up to 5 trials to first defective

- qgeom

  `qgeom(.75, .2)`

  returns the number of trials before first defective that has a probabilty of .75.