Practical Sheet in R

Distributions

Binomial Distribution

R has four in-built functions to generate binomial distribution. They are described below.

dbinom(*x*, size, prob)
pbinom(*x,* size, prob)
qbinom(*p*, size, prob)
rbinom(*n*, size, prob)

Following is the description of the parameters used −

- ***x*** is a vector of numbers.
- ***p*** is a vector of probabilities.
- ***n*** is number of observations.
- **size** is the number of trials.
- **prob** is the probability of success of each trial.

dbinom()

This function gives the probability density distribution at each point.

```
x <- seq(0,50,by = 1)
y <- dbinom(x,50,0.5)   (# Creates a binomial distribution for 50 numbers with p = 0.5)
png(file = "dbinom.png")
plot(x,y)
dev.off()
```

pbinom()

This function gives the cumulative probability of an event. It is a single value representing the probability.

 # Probability of getting 26 or less heads from a 51 tosses of a coin.

```
print(x <- pbinom(26,51,0.5))
```

qbinom()

This function takes the probability value and gives a number whose cumulative value matches the probability value.

 # How many heads will come out with a probability of 0.25 when a coin is tossed 51 times

```
print(x <- qbinom(0.25,51,1/2))
```

rbinom()

This function generates required number of random values of given probability from a given sample.

 # Find 8 random values from a sample of 150 with probability of 0.4.

```
print(x <- rbinom(8,150,.4))
```

Poisson Distribution

R has four in-built functions to generate poisson distribution. They are described below.

dpois(*x*, lambda)
ppois(*q*, lambda)
qpois(*p*, lambda)
rpois(*n*, lambda)

- ***x*** is a vector of numbers.
- ***p*** is a vector of probabilities.
- ***q*** is the vector of quantiles
- ***n*** is number of observations.

dpois()

This function gives the probability density distribution at each point.

```
x <- seq(0,20,by = 1)
y<-dpois(x,1)
png(file = "dpois.png")
plot(x,y)
dev.off()
```

ppois()

 # If there are twelve cars crossing a bridge per minute on average, find the probability of having sixteen or less cars crossing the bridge in a particular minute.

```
ppois(16, lambda=12)
```

 # If there are twelve cars crossing a bridge per minute on average, find the probability of having seventeen or more cars crossing the bridge in a particular minute.

```
ppois(16, lambda=12, lower=FALSE)
```

qpois()

qpois(*p*, lambda) is the smallest integer x such that $P(X \le x) \ge p$.

Find the smallest integer x such that $P(X \leq x) \geq 0.75$.

```
qpois(0.75, lambda=12)
```

rpois()

Find 16 values of a random value that is Poisson distributed with mean = 16

```
rpois(16, lambda=12)
```

Geometric Distribution

R has four in-built functions to generate geometric distribution. They are described below.

```
dgeom(x, prob, log = FALSE )
pgeom(q, prob, lower.tail = TRUE, log.p = FALSE)
qgeom(p, prob, lower.tail = TRUE, log.p = FALSE)
rgeom(n, prob)
```

- ***x*** is a vector of numbers of failures in a sequence of Bernoulli trials before success occurs
- **prob** is the success of each trial.
- ***p*** is a vector of probabilities.
- ***q*** is the vector of quantiles
- ***n*** is number of observations.
- Lower.tail = TRUE represents $P[X \leq x]$ else $P[X > x]$

dgeom()

This function gives the probability density distribution at each point.

```
x <- seq(0,20,by = 1)
y<-dgeom(x,0.6,log = FALSE)
png(file = "dgeom.png")
plot(x,y)
dev.off()
```

pgeom()

The probability of having at most three failures before a success with probability of success 0.1

```
pgeom(3, 0.1, lower.tail = TRUE, log.p = FALSE)
```

qgeom()

 # What is the smallest value of *x* such that *P*(*X* ≤ *x*) ≥ 0.3 for a geometric distribution with probability of success 0.1?

qgeom(0.3, 0.1, lower.tail = TRUE, log.p = FALSE)

Gamma Distribution

R has four in-built functions to generate geometric distribution. They are described below.

dgamma(*x*, *k*, theta, log = FALSE)
pgamma(*q*, *k*, theta, lower.tail = TRUE, log.p = FALSE)
qgeom(*p*, *k*, theta, lower.tail = TRUE, log.p = FALSE)
rgeom(*n*, *k*, theta)

- ***x*** is a vector of numbers of failures in a sequence of Bernoulli trials before success occurs
- ***k*** is the shape parameter
- **theta** is the scale parameter.
- ***p*** is a vector of probabilities.
- ***q*** is the vector of quantiles
- ***n*** is number of observations.
- Lower.tail = TRUE represents $P\ [X \le x]$ else $P\ [X > x]$

dgamma()

This function gives the probability density distribution at each point.

```
x <- seq(0,20,by = 0.01)
y<-dgamma(x,2,2,log = FALSE)
png(file = "dgamma.png")
plot(x,y)
dev.off()
```

Exercise

Q1) Plot the curve for a Binomial distribution with n = 10, p = 0.4. Increase the value of n in steps of 5 till 30. What do you observe?
Q2) Plot the curve for a Poisson distribution with lambda=0.1, 1, 10, 20 . What do you observe?
Q3) Plot the curve for a Binomial distribution with n = 50, p = 0.1. Plot the curve for Poisson distribution with lambda = 5. Give your observation.
Q4) Try to do a similar exercise by changing the value of *prob* in the geometric distribution. Can you explin the trend with mathematical explanation
Q5) Plot the curves for gamma distribution for theta = 2, 1, 0.5 and k = 1, 2, 3, 5, 9