Statistical Distribution in R

R functions produce 4 important values for commonly statistical distributions. The four important functions are:

- The density function (d)
- The probability function (p) – or cumulative density function $P(X \leq x)=F(x)$
- The quantile function (q) – inverse of the probability function $p(q(x))=x$ & $q(p(x))=x$.
- The random number generation function (r) – generate random numbers from specified distribution.

For a normal distribution

To compute density at x

dnorm(x, mean=0, sd=1) , x is a vector of quantiles.

To compute the cumulative

pnorm (q, mean=0, sd=1) , q vector of quantiles.

To compute the pth quantile – inverse of the prob. function

qnorm(p, mean=0, sd=1) , p vector of probabilities.

To generate a random sample of size n from normal distribution

rnorm (n, mean=0, sd=1) , n sample size

A powerful feature of the R language is the ability to generate random numbers from a specified distribution. For example, rnorm (10) generates 10 random numbers from the standard normal distribution; rnorm (10, 12, 3) generates them from N (12, 9)

The quantile x is the value such that

p (random variable $\leq$ x) = F(x) = p(probability).

pnorm(1) ; pnorm(1.96) ; pnorm(1.64) ; pnorm(-.5 : .5)

pnorm(seq(-2, 2, 1);pnorm (2, 0, 2); pnorm (1:5, 3,1.5)

Geometrically pnorm is the area of pdf to the left of the value x under the curve in standard normal distribution.

qnorm (p) is the inverse function of (p) which gives you the quantile x.

Geometrically qnorm is the value associated with the area size p from standard normal distribution.

Examples:

rnorm(10) # generate 10 numbers from normal(0,1)
rnorm(10,12,3) # generate 10 numbers from normal(12,9)
pnorm(1);pnorm(1.96);pnorm(-.5:.5); pnorm(seq(-2,2,1))
pnorm(2,0,2); pnorm(1:5,3,1.5)

qnorm(0.8);qnorm(0.975);qnorm(0.95); qnorm(0.05);qnorm(0.025)

Note that

f (f (p)) = p
check this by
pnorm(qnorm(0.9,0.95,0.975,0.99)) # give prob back.

The same way we can use the four functions (d,p,q,r) for the other distributions in R. A command are invoked by placing one of four symbols (p, q, d, r) on the front of R root name associated with a specific distribution.

Distribution	R root name	Parameters	
Normal	norm	Mean = 0	sd =1
Student's t	t	df	
Chi-square	chisq	df	
F	F	df1	df2
Gamma	gamma	shape	
Beta	beta	shape1	shape2
Uniform	unif	min=0	max=1
Lognormal	lnorm	meanlog=0	sdlog=1
Logistic	logis	location=0	scale=1
Cauchy	cauchy	location=0	scale=1
Exponential	exp	rate=1	
Binomial	binom	size	probability
Poisson	pois	lambda	
Weibull	weibull	shape	

- A special distributions the uniform distribution.

Ex:
runif (12) #generates 12 random numbers from U(0, 1)

- To generate 12 random numbers equally likely between 0 and 10, then
  10*runif (12) # or runif (12, 0, 10)

- To generate random integers between 1, 10
  ceiling (runif (12, 0, 10))

- To fix the random seed to regenerate same sequence of random numbers, you can use
  set.seed (111)

More examples:
qt(p=0.975,df=9) #the 5% critical value for a two sided t-test on 9 d.f.
dpois(x=3,lambda=5) # the value of density of poisson dist. With rate =5
dnorm(-2:2,2,2) # find density at ……. From Normal(,)
dnorm(qnorm(c(0.05,0.1,0.9,0.95))) #to check symmetry of distribution

Questions:

1. find probability that x=6 where x is poisson(lambda=4)
2. find probability that x less than or equal to 6 where x is Normal(15,16)
3. find probability that P(X≤x)=0.5 from uniform(2,8)
4. find probability that P(X≤x)=0.5,0.95,0.65 from t(2,8)
5. Explore the following:
   xnorm=rnorm(1000)
   stem(xnorm)
   hist(xnorm)
   boxplot(xnorm)
   mean(xnorm)
   median(xnorm)
   range(xnorm)
   quantile(xnorm)
   sqrt(var(xnorm))