**In central tendency, if sd of population is not given then infer sd of sample as sd of population.**

**Pnorm(x, mean, sd)**

**Solved Examples**

**Question 1:**The record of weights of male population follows normal distribution. Its mean and standard deviation are 70 kg and 15 kg respectively. If a researcher considers the records of 50 males, then what would be the mean and standard deviation of the chosen sample?  
  
**Solution:**

Mean of the population μ = 70 kg  
Standard deviation of the population = 15 kg  
sample size n = 50  
Mean of the sample is given by:  
μx¯=μμx¯=μ  
μx¯μx¯ = 70 kg  
Standard deviation of the sample is given by:  
σx¯σx¯ = σn√σn  
σx¯σx¯ = 1550√1550  
σx¯σx¯ = 2.121 = 2.1 kg (approx)

**Question 2:**At a coastal area, the number of crabs caught per day are recorded. The average of which is 10 and standard deviation is 3. If the record of 60 days is chosen randomly, estimate the mean and standard deviation of the chosen sample.  
  
**Solution:**

Mean of the population μ = 10  
Standard deviation of the population = 3  
sample size n = 60  
Mean of the sample is given by:  
μx¯=μμx¯=μ  
μx¯μx¯ = 10  
Standard deviation of the sample is given by:  
σx¯σx¯ = σn√σn  
σx¯σx¯ = 360√360  
σx¯σx¯ = 0.387

Example 2:  
In a survey of a company, mean salary of employees is 29321 dollars with SD of 2120 dollars. Consider the sample of 100 employees and find the probability their mean salary will be less than 29000 dollars?  
  
Solution:   
Total number of employees (n)   
4= 100  
Mean (μ)  
= 29321  
standard deviation (σ)  
= 2120  
z  
=-1.51  
Step 1: Insert the values into the z-formula:  
= (29,000 – 29,321) / (2,120/√100) = -321/212 = -1.51.  
Step 2: Look up the [z-score](http://www.statisticshowto.com/probability-and-statistics/z-score/) in the [left-hand z-table](http://www.statisticshowto.com/tables/z-table/#left) (or use technology). -1.51 has an area of 93.45%.  
However, this is not the answer, as the question is asking for LESS THAN, and 93.45% is the area “greater than” so you need to subtract from 100%.  
100% – 93.45% = **6.55% or about 0.07.**Conclusion:  
The probability of employees having mean salary less than 29000 dollars is 0.07%.  
pnorm(29000,29321,212)

pnorm(29000,29321,212)

[1] 0.06499378

Example 3  
The average GPA at a particular school is m=2.89 with a standard deviation s=0.63. A random sample of 25 students is collected. Find the probability that the average GPA for this sample is greater than 3.0.  
  
mean = 2.89  
standard error = .126  
z-score = .87  
  
Looking up this z-score in the normal curve table yields a probability of .8078. The final answer is 1-.8078=.1922.

1-pnorm(3,2.89,0.63/5)

Q. A certain group of welfare recipients receives SNAP benefits of $110 per week with a standard deviation of $20. If a [random sample](http://www.statisticshowto.com/simple-random-sample/) of 25 people is taken, what is the probability their mean benefit will be greater than $120 per week? (didn’t get answer)

Step 1: Insert the information into the z-formula:  
= (120-110)/20 √25 = 10/ (20/5) = 10/4 = 2.5.  
Step 2: Look up the z-score in a table (or calculate it using technology). A z-score of 2.5 has an area of roughly 49.38%. Adding 50% (for the left half of the curve), we get 99.38%.

Tried in this way :

Answer incorrect

1-pnorm(120,110,4)

[1] 0.006209665

Correct answer:

pnorm(120,110,4)

[1] 0.9937903

Try with z score :

Z = (120 – 110)/ 4 = 2.5

Sample problem: There are 250 dogs at a dog show who weigh an average of 12 pounds, with a [standard deviation](http://www.statisticshowto.com/probability-and-statistics/standard-deviation/)of 8 pounds. If 4 dogs are chosen at random, what is the [probability](http://www.statisticshowto.com/probability-and-statistics/probability-main-index/) they have an average weight of greater than 8 pounds and less than 25 pounds?

pnorm(25,12,8/sqrt(4)) - pnorm(8,12,8/sqrt(4))

**Sample problem:** A population of community college students includes inner city students (p = .33). What is the **probability** that a random sample of 45 students from the [population](http://www.statisticshowto.com/what-is-a-population/) will have from 20% to 40% inner city students? (didn’t get answer)

Random sample is mentioned and we cannot apply directly pnorm, go with binomial distribution with continuity correction

P =.33

Q= 1-p = 0.67

Sample :

N= 45

What is the **probability** that a random sample of 45 students from the [population](http://www.statisticshowto.com/what-is-a-population/) will have from 20% to 40% inner city students?

Probability of 20% to 40% inner city students : 20% of 45 = 45 \* 20/100 = **9**

40% of 45 : 45 \* 40/100 = 18

So range between 9 to 18 :

9 starts from 8.5 and 18 comes till 18.5

pbinom(18.5,45,0.33) - pbinom(8.5, 45, 0.33)

[1] 0.8571422

**Step 1:** Press APPS. Highlight the **Stats/List Editor** by using the scroll keys. Press ENTER.  
If you don’t see the Stats/List editor you need to load the app. See instructions [here](http://education.ti.com/en/us/software/details/en/31FC737C43CF43B0ADA1CF67420C3AA8/89statisticswithlisteditor).

**Step 2:** Press F5 and scroll down to C: **BinomialCdf**.

**Step 3:** Enter 45 in the **Num Trials** box.

**Step 4:** Scroll down and enter .33 in the **Prob Success** box.

**Step 5:** Scroll down and enter 9 in the **Lower Value** box (because 20% of 45 = 9).

**Step 6:** Scroll down and enter 18 in the **Upper Value** box (because 40% of 45 = 18). Press ENTER.

**Step 7:** Read the Result: **Cdf = .857142**. This means that the probability your random sample will have 20 – 40% inner city students is **85.71%**.

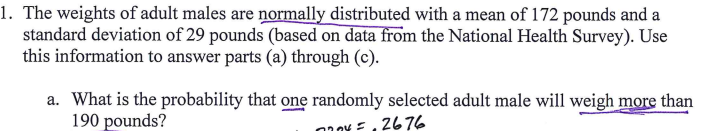
**Sample problem**: A fertilizer company manufactures organic fertilizer in 10 pound bags with a standard deviation of 1.25 pounds per bag. What is the probability that a random sample of 15 bags will have a mean between 9 and 9.5 pounds?

**Step 1**: 2nd VARS 2.

**Step 2**: Enter your variables (lower bound, upper bound, mean, and standard deviation). Separate each variable by a comma: 9,9.5, 10,(1.25/√15)).

**Step 3**: Press ENTER. This returns the probability of .05969, or **.05969%**.

pnorm(9.5, 10, 1.25/sqrt(15)) - pnorm(9, 10, 1.25/sqrt(15)) = 0.05969474



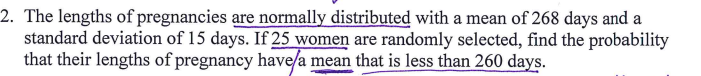
1-pnorm(190,172,29)

[1] 0.2674019



1-pnorm(190,172,29/sqrt(25))

[1] 0.0009563983



Mean(population) = sample(population) = 268 days

Sd = 15

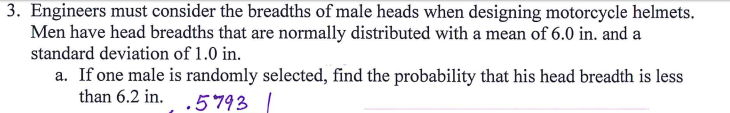
N = 25

Sd(sample) = 15/ sqrt(25)

Less than is pnorm : pnorm(260,268,3)

pnorm(260,268,3)

[1] 0.003830381



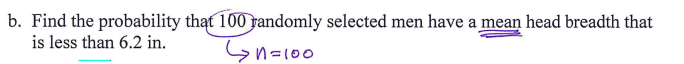
Mean = 6.0

Sd = 1.0

X = 6.2 , less than

pnorm(6.2, 6.0,1)

[1] 0.5792597



pnorm(6.2, 6.0,1/sqrt(100))

[1] 0.9772499

Sample questions : http://www.probabilityformula.org/normal-distribution-examples.html#

A test is normally distributed with a mean of 70 and a standard deviation of 8. (a) What score would be needed to be in the 85th percentile? (b) What score would be needed to be in the 22nd percentile? ([relevant section](http://onlinestatbook.com/2/normal_distribution/areas_normal.html))

Ans (a) 78.3

> qnorm(0.85,70,8)

[1] 78.29147

General Hospital's patient account division has compiled data on the age of accounts receivables. The data collected indicate that the age of the accounts follows a normal distribution with mean 28 days and standard deviation 8 days.

1. What proportion of the accounts are between 20 and 40 days old?
2. What proportion of the accounts are less than 30 days old?
3. What is the number of days in which 75% of all accounts are above?

|  |
| --- |
| pnorm(40,28,8) - pnorm(20,28,8)  [1] 0.7745375  > pnorm(30,28,8)  [1] 0.5987063  > qnorm(.25,28,8)  [1] 22.60408 |
|  |
| |  | | --- | | > The number of days in which 75% of all accounts are above is the 25*th* percentile.  normal curve  normal curve | |

Based on past experience, the main printer in a university computer center is operating properly 90% of the time. Suppose inspections are made at 10 randomly selected times.

1. What is the probability that the main printer is operating properly for exactly 9 of the inspections?
2. What is the probability that the main printer is *not* operating properly no more than 1 inspection?
3. What is the expected number of inspections in which the main printer is operating properly?

Sol :

Prob. Of working properly = 0.9

Not working = 0.1

N =10

whatever is asking in question that has to be taken as success/ p.

1. the main printer is operating properly

p = 0.9, q = 0.1

10C9 \* (0.9)^9 \* (0.1) ^(10-9) = nCr \* p ^r \* q ^(n-r)

dbinom(9, size=10, prob=0.9)

[1] 0.3874205

b.probability that the main printer **is *not* operating properly** no more than 1 inspection?

p = 0.1 , q = 0.9

r = 0,1

dbinom(0, size=12, prob=0.2) +   
+ dbinom(1, size=12, prob=0.2) +   
+ dbinom(2, size=12, prob=0.2)

can be written as

 pbinom(2, size=12, prob=0.2)

so in this case

pbinom(1, size=10, prob=0.1)

[1] 0.7360989

c.What is the expected number of inspections in which **the main printer is operating properly?**

**P =0.9, q =0.1**

E(x) = np = 10 \* 0.9

The time required to complete a final examination in a particular college course is normally distributed, with mean of 80 minutes and a standard deviation of 10 minutes. Answer the following questions.

1. What is the probability of completing the exam in one hour or less? : (60 minutes or less)
2. What is the probability a student will complete the exam in a time between 60 and 75 minutes?
3. What is the interquartile range for completion times?

Ans : for c.

*IQR* = *Q*3 - *Q*1 , *Q*1, the 25*th* percentile, and *Q*3, the 75*th* percentile.

qnorm(.75,80,10) - qnorm(.25,80,10) = 13.4898

Assume that the dividends of electric utility stocks as of a given date have a symmetric distribution with mean of 8.5 percent and standard deviation of 2.5 percent. Find the probability that the average dividend of 25 such stocks will exceed 10 percent.

Ans : 1- pnorm(10, 8.5 , 2.5/sqrt(25))

The probability that a patient fails to recover from a particular operation is 0.1. Suppose that eight patients having this operation are selected at random. Answer the following questions.

1. What is the probability that at most one patient will not recover?
2. What is the probability that at least 2 but no more than 3 patients will not recover?
3. What is the probability that all patients will not recover?
4. What is the expected number of patients that will not recover?

Ans :

Let the Bernoulli trials is patient not recover.

Success , p = 0.1

Q = 1-p = 0.9

1. Atmost one : r = 0,1

Pbinom(1, size = 8, prob = 0.1) i.e P(x=0) + P(x=1)

8C0 \* (0.1) ^0 \* (0.9) ^ (8-0) + 8C1 \* ……………………………………………

pbinom(1, size = 8, prob = 0.1)

[1] 0.8131047

1. What is the probability that at least 2 but no more than 3 patients will not recover?

dbinom(2, size = 8, prob = 0.1) + dbinom(3, size = 8, prob = 0.1)

[1] 0.1818709

1. What is the probability that all patients will not recover?

dbinom(8, size = 8, prob = 0.1)

[1] 1e-08

d.What is the expected number of patients that will not recover?

Ans : E(x)= np = 8(.1) = 0.8

An insurance company states that 10% of all fire insurance claims are fraudulent. Suppose the company is correct, and that it receives 125 claims.

1. What's the probability that at least 15 claims are fraudulent?
2. What's the probability that less than 10 claims are fraudulent?

Ans : binomial

N = 125, p = 0.1 , q = 0.9

**Atleast 15 claims** : r =15,16,17……125

Calc. pbinom(14 , 125, 0.1)= 0.732 i.e. till r =0 to 14 and then subtract by 1

1 – [p(x=0) +……….p(x=14)] = 1- 0.7329886

[1] 0.2670114

**less than 10 claims :** r = 0 to 9

pbinom(9 , 125, 0.1) : getting wrong answer

What type of data would be collected by the following survey questions?

1. "How many pairs of shoes do you own?"
2. "What color are your eyes?"
3. "Which brand of soft drink do you prefer?"
4. "What is the circumference of that tree?"
5. "What is your GPA?"
6. "Are you planning to vote this year?"

(a)

Numerical, discrete

(b)

Categorical

(c)

Categorical

(d)

Numerical, continuous

(e)

Numerical, continuous

(f)

Categorical ("Yes" or "No")

Seventy percent of small businesses experience cash flow problems during their first year of operation. A consultant takes a random sample of 50 small businesses that have been in business for one year.

1. What is **the probability** that more than 80% of the sample have experienced cash flow problems?
2. What is the probability that more than half of the sample has had cash flow problems?

Ans : n = 50

P = 0.7

Q = 1- 0.7 = 0.3

Mean (population)= np = 0.7

Variance = npq = 0.7 \* 0.3 = 0.21

Sd = sqrt(0.21) = 0.4582576

1-pnorm(0.8 , 0.7 , 0.458/sqrt(50)) = 0.06130613

1-pnorm(0.5 , 0.7 , 0.458/sqrt(50))

[1] 0.9989918

A company estimates that there is an 80% chance of an order arriving on time from a supplier. Suppose 5 orders are placed this week.

1. What is the probability that at least 4 orders arrive on time?
2. What is the probability that none of the orders arrive on time?
3. How many orders would you expect to arrive on time?

BERNOULlis trail is orders arrive on time.

N = 5 , p = 0.8 , q= 0.2

1. Pbinom( 3, 5, 0.8)

pbinom( 3, 5, 0.8)

[1] 0.26272

1 - 0.262= 0.738

1. R = 0 i.e P(x=0)

pbinom( 0, 5, 0.8)

0.00032

(c)

np = 5(.8) = 4

The reaction time to a certain psychological experiment is considered to be normally distributed with a mean of 20 seconds and a standard deviation of 4 seconds.

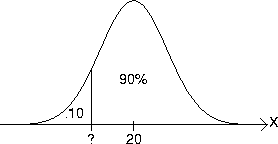
1. What proportion of subjects take between 15 and 30 seconds to react?
2. What proportion of subjects take longer than 30 seconds to react?
3. What is the reaction time such that only 10% of subjects are *faster*?

Ans :

Mean = 20, sd = 4

© qnorm(0.1, 20,4) = 14.87379

"Faster" means less time. 10% have quicker (smaller) times, and 90% have slower (larger) times. This is the 10*th* percentile. 



A natural gas exploration company averages four *strikes* (that is, natural gas is found) per 100 holes drilled.

1. If 20 holes are to be drilled, what is the probability that **no strikes** will be made?
2. What is the probability that **at least** one strike will be made?

Ans :binomial

strike will be made be Bernoulli’s trial

mean = 4

for n = 100

p , q =??

Mean = np i.e p = mean/n

P = 4/100 = 0.04

1. N = 20 , r = 0 (no strikes)

so p(x=0) = ?

dbinom(0,20,0.04) or pbinom

[1] 0.4420024

b.atleast one is r = 1,2….20

p(x>=1) = 1- P(x=0) = 1- 0.4420024

[1] 0.5579976

Wages for workers in a particular industry average $11.90 per hour with a standard deviation of 40 cents. The wages are considered to be normally distributed.

1. Suppose you are employed in this industry. What would your wage have to be if 75% of all workers earn more than you?

Ans: mean = 11.90, sd = 40 cents

75% workers earn more than me = to the left it will be 100-75% = 25%

qnorm(0.25, 11.90,0.40)

[1] 11.6302

**Questions on poisson and binomial :** <https://www.intmath.com/counting-probability/13-poisson-probability-distribution.php>

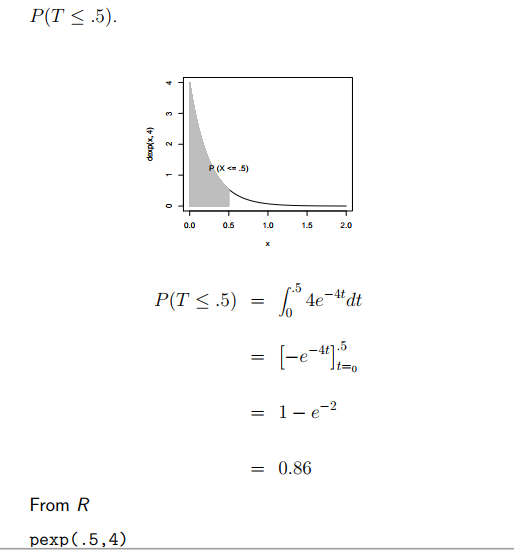
[http://www.probabilityformula.org/poisson-distribution-examples.html#](http://www.probabilityformula.org/poisson-distribution-examples.html)

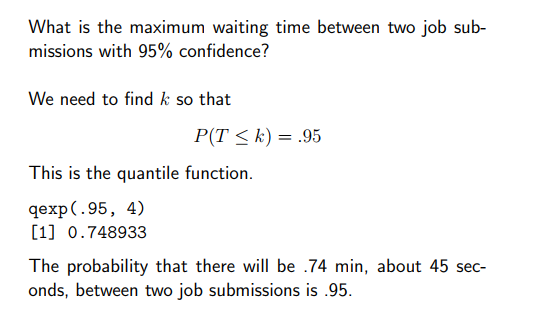
https://www.examsolutions.net/tutorials/exam-questions-poisson-distribution/

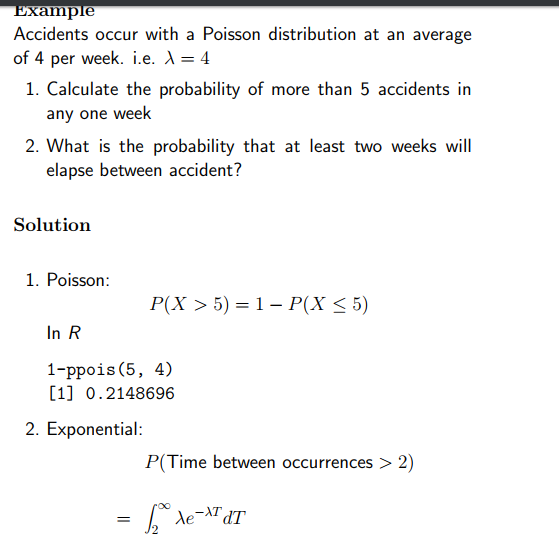
xample 1:  
A bank is interested in studying the number of people who use the ATM located outside its office late at night.  
On average, 1.6 customers walk up to the ATM during any 10 minute interval between 9pm and midnight.  
What is lambda λ for this problem?  
What is the probability of exactly 3 customers using th ATM during any 10 minute interval?  
What is the probability of 3 or fewer people?

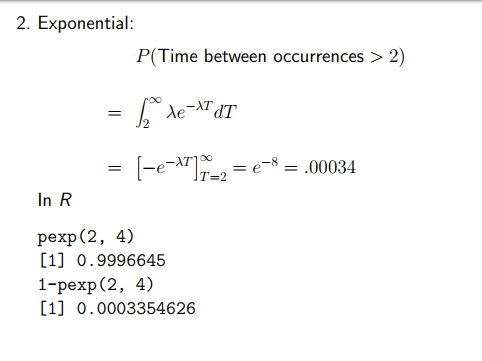
Example 2:  
The Indiana Department of Transportation is concerned about the number of deer being struck by cars between Martinsville and Bloomington. They note the number of deer carcasses and other deer-related accidents over a 1-month period in a 2-mile intervals. What is the probability of zero deer strike incidents during any 2-mile interval between Martinsville and Bloomington?

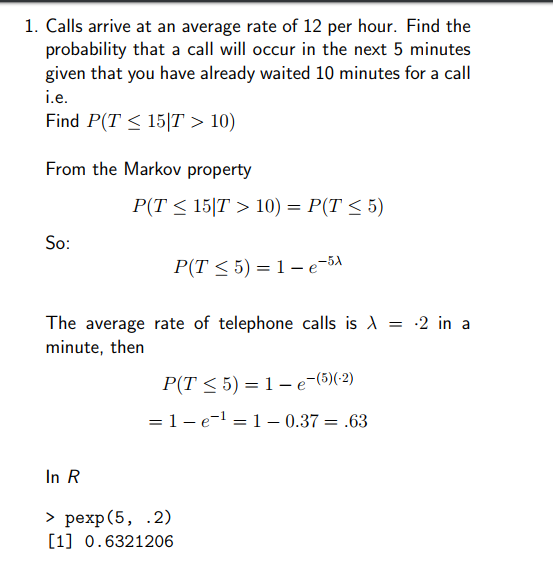
Example: If jobs arrive every 15 seconds on average, λ = 4 per minute, what is the probability of waiting less than or equal to 30 seconds, i.e .5 min?

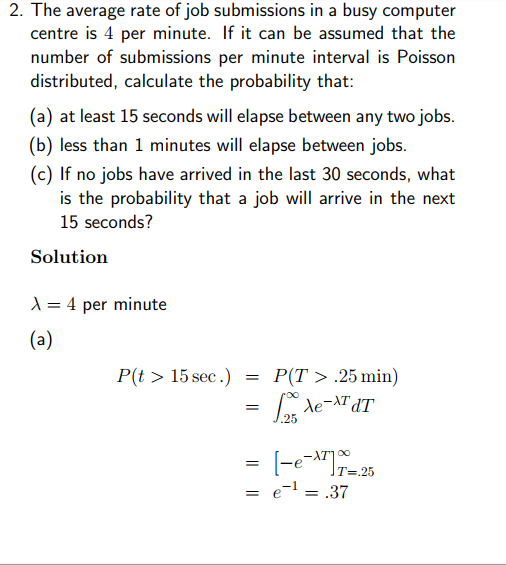


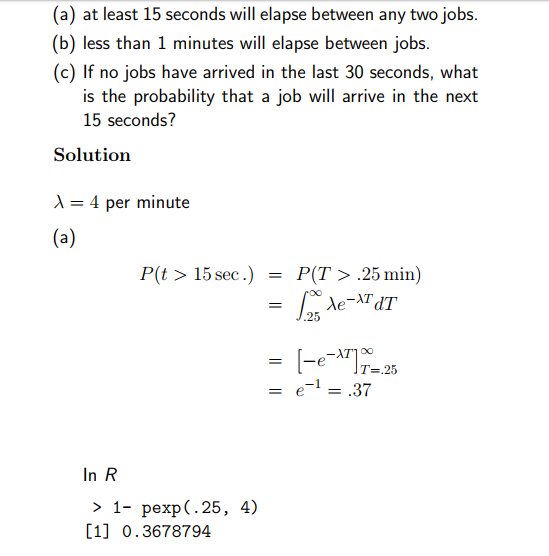












**R Lab Session : Part 2**   
(Solutions)

To see a review of how to start R, look at the beginning of [Lab1](http://www-stat.stanford.edu/~epurdom/RLab.htm)   
Lab1 http://www-stat.stanford.edu/ epurdom/RLab.htm

**Probability Calculations**

The following examples demonstrate how to calculate the value of the cumulative distribution function at (or the probability to the left of) a given number.

* **Normal(0,1) Distribution :**
* > x <- c(-2,-1,0,1,2)
* > x
* [1] -2 -1 0 1 2
* > pnorm(x)
* [1] 0.02275013 0.15865525 0.50000000 0.84134475 0.97724987
* **Binomial($n$,$p$) Distribution :**
* > x <- c(0,1,2,5,8,10,15,20)
* > pbinom(x,size=20,prob=.2)
* [1] 0.01152922 0.06917529 0.20608472 0.80420779 0.99001821 0.99943659 0.99999999
* [8] 1.00000000
* **Poisson($\lambda$) Distribution :**
* > x <- c(0,1,2,5,8,10,15,20)
* > ppois(x,6)
* [1] 0.002478752 0.017351265 0.061968804 0.445679641 0.847237494 0.957379076
* [7] 0.999490902 0.999998545

**Exercise :** Calculate the following probabilities :

1.

Probability that a normal random variable with mean 22 and variance 25

(i)

lies between 16.2 and 27.5   
pnorm(27.5,22,sd=5)-pnorm(16.2,22,sd=5)   
[1] 0.7413095

(ii)

is greater than 29 1-pnorm(29,22,sd=5)   
[1] 0.08075666

(iii)

is less than 17 pnorm(17,22,sd=5)   
[1] 0.1586553

(iv)

is less than 15 or greater than 25 pnorm(15,22,sd=5)+1-pnorm(25,22,sd=5)   
[1] 0.3550098

2.

Probability that in 60 tosses of a fair coin the head comes up

(i)

20,25 or 30 times   
sum(dbinom(c(20,25,30),60,prob=0.5))   
[1] 0.1512435

(ii)

less than 20 times   
pbinom(19,60,prob=0.5)   
[1] 0.0031088

(iii)

between 20 and 30 times pbinom(30,60,prob=0.5)-pbinom(20,60,prob=0.5)   
[1] 0.5445444

3.

A random variable X has Poisson distribution with mean 7. Find the probability that

(i)

X is less than 5 less or equal is:   
> ppois(5,7)   
[1]0.3007083   
less than is   
> ppois(4,7)   
[1]0.1729916

(ii)

X is greater than 10 (strictly)   
> 1-ppois(10,7) [1] 0.0985208

(iii)

X is between 4 and 16 > ppois(16,7)-ppois(3,7) [1] 0.9172764

**Quantiles**

The following examples show how to common the quantiles of some common distributions for a given probability (or a number between 0 and 1).

* **Normal(0,1) Distribution :**
* > y <- c(.01,.05,.1,.2,.5,.8,.95,.99)
* > qnorm(y,mean=0,sd=1)
* [1] -2.3263479 -1.6448536 -1.2815516 -0.8416212 0.0000000 0.8416212 1.6448536
* [8] 2.3263479
* **Binomial($n$,$p$) Distribution :**
* > y <- c(.01,.05,.1,.2,.5,.8,.95,.99)
* > qbinom(y,size=30,prob=.2)
* [1] 1 3 3 4 6 8 10 11
* **Poisson($\lambda$) Distribution :**
* > y <- c(.01,.05,.1,.2,.5,.8,.95,.99)
* > qpois(y,6)
* [1] 1 2 3 4 6 8 10 12

**Random Variable generation**

The following examples illustrate how to generate random samples from some of the well-known probability distributions.

* **Normal($\mu$,$\sigma^2$) Distribution :**

The first sample is from $N(0,1)$ distribution and the next one from $N(5,1)$ distribution.

> z <- rnorm(10)

> z

[1] -0.90361592 -1.96522764 -1.35107949 -0.10846423 0.29756634 1.40831606

[7] -0.07844737 1.40575257 -0.97511415 -0.33418299

If you would like to see how the distribution of the sample points looks like ....

> w <- rnorm(1000,mean=5,sd=1)

> hist(w)

* **Binomial($n$,$p$) Distribution :**
* > k <- rbinom(20,size=5,prob=.2)
* > k
* [1] 1 2 0 1 0 0 0 2 0 1 0 0 0 0 0 2 4 1 1 1
* **Poisson($\lambda$) Distribution :**
* > x <- rpois(20,6)
* > x

[1] 2 8 7 5 5 5 3 8 5 5 1 8 5 5 5 4 10 7 3 4