HOMEWORK 6: MATH 183 WINTER 2013 (UPDATED VERSION)

DUE IN CLASS ON FRIDAY FEB 15TH

Useful R commands for this homework are

```
mean - computes mean of data
var - computes variance of data
rchisq -  generates random numbers from chi squared distribution
rnorm -  generates random numbers from normal distribution
rexp -  generates random numbers from exponential distribution
rpois -  generates random numbers from poisson distribution
rbinom -  generates random numbers from binomial distribution
pbinom - gives values of probability mass function of binomial values
                (similar for all other distribution)
dorm - gives values of density function of normal random variable
                (similar for all other distribution)
qqnorm - produces qq plot between sample and standard normal distribution
qqline - produces straight line on qqplot
               (ideal situation if two distributions were the same)
hist - produces histogram of points
```

(1) Repeat all the exercises from the precept for Chi-squared distribution with 15 degrees of freedom. Visualize the probability that this random variable X taking values between 20 and 40? (do it for the histogram as well as for the density plots)

(2) Lets have a look what is the difference between normal and binomial probability density functions. How would you approach solving this problem in R? Do it. Would you expect them to be different or not and why ? (hint: what are the types of random variables)

(3) Suppose that for the certain disease it is known to have a postoperative complication frequency of 20%.
   (a) What is the probability of operating on 4 patients successfully if the total number of patients that were operated was 20?
   (b) What is the probability of operating on more than 4 patient successfully ?
   (c) Formulate and solve this problem both mathematically and using R. Give graphical/visual/colorful interpretation of the solution.
   (d) Which of the following three lines is giving you the correct answer to the first question and which to the second question?

```
> pbinom(4,size=20,prob=0.20)
> 1-pbinom(4,size=20,prob=0.20)
```

```
> pbinom(4,size=20,prob=0.20)-pbinom(3,size=20,prob=0.20)
> dbinom(4,size=20,prob=0.20)
```

(4) Generate a random sample of size 100 from standard normal random variable and one of size 1000 from exponentially distributed random variable exp(1).
  (a) Compare the distributions of the two samples ? Draw two histograms and compare the probability density functions of the two. [Hint: Density function can be taken to be the line describing upper border of heights of bins]
  (b) Use qq plot to conclude which one has lighter/heavier tails.
  (c) If X is standard normal random variable compute the probability $P(2 < X < 5)$ and show it graphically! [Hint: Use equivalent of Problem 1 but now for normal distribution]

(5) A machine produces memory sticks of varying lengths, distributed uniformly between 2 and 12 mm. Memory sticks longer then 10mm do not meet the design criterion and much be scraped.
  (a) Calculate the proportion of memory sticks that will be scrapped.
  (b) Simulate 50 memory sticks lengths and obtain a histogram of the simulated values. Calculate the simulated mean and variance. Is it expected?

(6) Records show that the job submissions to a busy computer center have a Poisson distribution with an average of 4 per minute. Let T be the time between submissions.
  (a) Simulate 1000 values that random variable T can take. Show the histogram.
  (b) Compute simulated mean. Is it close to what you expected ?
  (c) Use R to compute $P(T \leq 0.25)$.

(7) An analogue signal received at a detector (measured in microvolts) is normally distributed with a mean of 100 and variance 256.
  (a) What is the probability that the signal is less then 120 microvolts?
  (b) What is the probability that the signal is less then 120 microvolts given that it is larger then 110 microvolts? [Hint: Remember the definitions of conditional probabilities] How would you answer both questions using simulations in R (and do it)?
  (c) Estimate interquartile range of the signal?

(8) A rule of thumb is that 5% of the normal distribution lies outside an interval approximately $\pm 2\sigma$ about the mean.
  (a) To what extend is this true?
  (b) What are positions of the quartiles measured in standard deviation units.]

Challenge Problem Another way of simulating Poisson Random Variable

Note that among poisson probabilities there is a recursive relationship:

$$p_{k+1} = \frac{\lambda}{k+1} p_k,$$

$$p_k = \frac{\lambda^k e^{-\lambda}}{k!}$$

Use it to make your own random number generator from Poisson distribution. The outline of the algorithm is :

Step 1 Generate a uniform random number $U(0, 1)$
Step 2 Set $i = 0, p = \exp(-\lambda), F = p$
Step 3 If $U < F$, set $X = i$ and STOP
Step 4 Set $p = \frac{\lambda}{i+1}p$, $F = F + p$ and $i = i + 1$
Step 5 return to Step 3

Write down the algorithm in R. Explain what each step of the algorithm does and explain why is it necessary.