Statistics - Homework 7 (Due Dec. 27, 2017)

1. A random variable *X* with normal density (μ=16, σ=5).
   (1) Generate 10,000 random samples (size *n*=30) from the above density. From the 10,000 samples, plot the corresponding 95% confidence intervals of the expected value of *X*. [Assuming the variance of *X* is unknown.]
   (2) What percentages of the 10,000 confidence intervals actually cover the expected value of *X* ?
   (3) What is the average length of the 10,000 confidence intervals?
   (4) What is the expected value of the length of confidence interval?

   The following R-code is provided for your reference.

```
# Interval estimation -- Normal Density
# Demonstration of confidence interval estimation for mean of a normal density
# -----------------------------------------------------------
ns=10
nr=100
siglevel=0.05
mx=c();stdx=c();lb=c();ub=c()
a=16 #location parameter
s=5  #scale parameter
exp.x=a
d=c() # 2d = length of the confidence interval
count=0
windows()
plot(0,0,type="n",xlim=c(a-4*s/sqrt(ns),a+4*s/sqrt(ns)),ylim=c(0,2*nr))
for (i in 1:nr)
{
      x=rnorm(ns,a,s)
      mx[i]=mean(x);stdx[i]=sqrt(var(x))
      d[i]=qt((1-siglevel/2),ns-1)*stdx[i]/sqrt(ns)
      lb[i]=mx[i]-d
      ub[i]=mx[i]+d
      if (lb[i]>exp.x |ub[i]<exp.x) count=count+1
      u=c(lb[i],ub[i])
      v=c(2*i,2*i)
      lines(u,v,type="l",lwd=2)
}
lines(c(a,a),c(0,2*nr),type="l",col="red",lwd=2)
print(c(count,count/nr,mean(2*d)))
```

   Note: The sample standard deviation (s) is a biased estimator of population standard deviation σ, i.e., $E(s) = CF \cdot \sigma$, where CF is the correction factor and can be approximated by the following equation:

$$CF \approx 1 - \frac{1}{4n} - \frac{7}{32n^2} - \frac{9}{128n^3} \quad (n\text{: sample size})$$

2. The data in the following table is a random sample from an exponential density $f_X(x;\ \lambda) = \lambda e^{-\lambda x}$.

62.48	10.37	31.19	9.57	85.27	162.96	16.11	133.86	47.28	34.79
70.15	75.26	25.90	9.34	47.28	4.33	35.09	48.69	83.99	158.74
60.66	56.81	64.01	89.74	7.02	64.72	100.14	16.64	147.06	28.00
116.40	11.01	29.21	12.54	43.04	18.69	50.47	17.47	56.22	25.18

   (1) Prove that $n\lambda\overline{X}_n$ is a pivotal quantity.
   (2) Find a 90% confidence interval of λ.

(3) Find the lower bound of the one-sided 95% confidence interval of λ.

3. Assuming that the data in Problem 2 were generated from an exponential distribution with $\lambda = 0.02$, write your own R code to generate 1000 random samples of the exponential distribution, each of sample size n=40.
   (1) Find the 90% confidence interval of λ for each random sample and then calculate the proportion of the random samples whose 90% confidence intervals do not cover the true value of λ.
   (2) Plot the histogram of the lengths of the 90% confidence intervals.
   (3) Determine the 90% acceptance interval of the sample mean ($\overline{X}_n$) and calculate the proportion of sample means that fall within the acceptance interval.
   (4) Plot the scatter plot of sample means and lengths of confidence intervals. Explain the relationship which can be observed in the plot.

4. A random sample of size 20 ($(x_1, x_2, \cdots, x_n)$, $n = 20$) drawn from a uniform density $U[0, \theta]$ is shown in the following table.

11.90	6.94	14.90	4.36	14.00
15.21	18.87	4.75	7.67	6.57
5.15	14.83	7.79	13.88	18.55
2.49	9.90	9.68	11.10	9.92

   (1) Show that Y_n / θ ($Y_n = \max(X_1, \ldots, X_n)$) is a pivotal quantity.

   (2) Find a 90% confidence interval of θ.

5. The radius (r) of a circle is measured with an error of measurement which is normally distributed with zero mean and unknown variance σ^2. Given the following independent measurements of the radius:

51.07	52.49	48.07	48.94	49.33
50.34	51.30	40.33	41.38	44.05
51.45	48.89	56.42	49.21	49.41
46.03	53.64	41.13	46.17	57.07

   Find the 95% confidence interval of the area of the circle.