

# STATISTICS

Due June 4, 2021

## Homework 7

1. Let  $X$  be a normally distributed random variable with expected value and standard deviation being 60 and 20, respectively. Let  $\bar{X}_n$  be the sample mean of a random sample of size  $n$  from  $X$ . A random sample of size 25 from  $X$  is given in the following table:

37.3332	84.75534	56.2749	27.09361	63.11717
44.61746	46.38288	73.65585	91.7605	50.46811
86.2026	51.86157	78.05359	33.82873	75.01817
52.57203	19.59978	80.21883	72.44076	42.92938
61.5187	68.02203	68.10625	81.53383	60.46798

Determine a 95% confidence interval of the population mean of  $X$  using the above data. (Assuming the standard deviation of  $X$  is unknown.)

2. A random variable  $X$  has a gamma density with expected value 180 and standard deviation 104.
- (1) Generate 10,000 random samples (size  $n=100$ ) 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) Among the 10,000 confidence intervals, how many of them actually contain the expected value of  $X$ ?
  - (3) What is the average length of the 10,000 confidence intervals?

[An R code at the end of this file is provided for your reference.](#)

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

37.01	52.20	7.55	2.31	29.48	111.70	39.63	3.89	201.60	20.60
41.23	18.69	241.49	4.12	15.67	6.74	16.72	24.86	9.91	115.26
16.00	27.65	66.90	21.81	4.92	46.05	58.79	65.23	7.79	51.24
163.89	9.49	61.21	181.03	71.65	74.64	43.06	13.57	35.89	136.47

- (1) Find a two-sided 90% confidence interval of  $\lambda$ .
  - (2) Find the lower bound ( $a$ ) of the one-sided 95% confidence interval of  $\lambda$ , i.e.  $P[a \leq \lambda] = 0.95$ .
4. Assuming that the data in Problem 3 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  $\lambda$  for each random sample and then calculate the proportion of the random samples whose 90% confidence intervals do not cover the true value of  $\lambda$ .
  - (2) Plot the histogram of the lengths of the 90% confidence intervals.
  - (3) Determine the 90% acceptance interval of the sample mean ( $\bar{X}_n$ ) 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.

5. A random sample of size 20  $((x_1, x_2, \dots, x_n), n = 20)$  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  $Y_n / \theta$  ( $Y_n = \max(X_1, \dots, X_n)$ ) is a pivotal quantity.

- (2) Find a 90% confidence interval of  $\theta$ .

6. The radius ( $r$ ) of a circle is measured with an error of measurement which is normally distributed with zero mean and unknown variance  $\sigma^2$ . 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.

### Reference R code for problem 2

```
# 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)))
```