MAT 2377, Probability and statistics for engineers

Assignment 4 - solutions

[4] **Exercise 7.27**

(a) Since

$$E(\hat{\Theta}_1) = \frac{E(X_1) + \cdots + E(X_7)}{7} = \frac{7\mu}{7} = \mu$$

and

$$E(\hat{\Theta}_2) = \frac{2E(X_1) - E(X_6) + E(X_4)}{2} = \frac{2\mu - \mu + \mu}{2} = \mu,$$

both $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are unbiased estimators.

(b) We have

$$Var(\hat{\Theta}_1) = \frac{\sigma^2}{7}$$

and

$$Var(\hat{\Theta}_2) = \frac{4\sigma^2 + \sigma^2 + \sigma^2}{4} = 1.5\sigma^2.$$

and the relative efficiency of $\hat{\Theta}_2$ to $\hat{\Theta}_1$ is $1/1.5 = 2/3$. Therefore $\hat{\Theta}_1$ is the better estimator.

[4] **Exercise 8.6 :**

(a) We have

$$\bar{x} = (37.53 + 49.87)/2 = (35.59 + 51.81)/2 = 43.7.$$

(b) Since the the shorter confidence interval has lower confidence level, the confidence interval $(37.53, 49.87)$ is the 95% confidence interval and $(35.59, 51.81)$ is the 99% confidence interval.

[4] **Exercise 8.16 :**

(a) the 95% confidence interval is

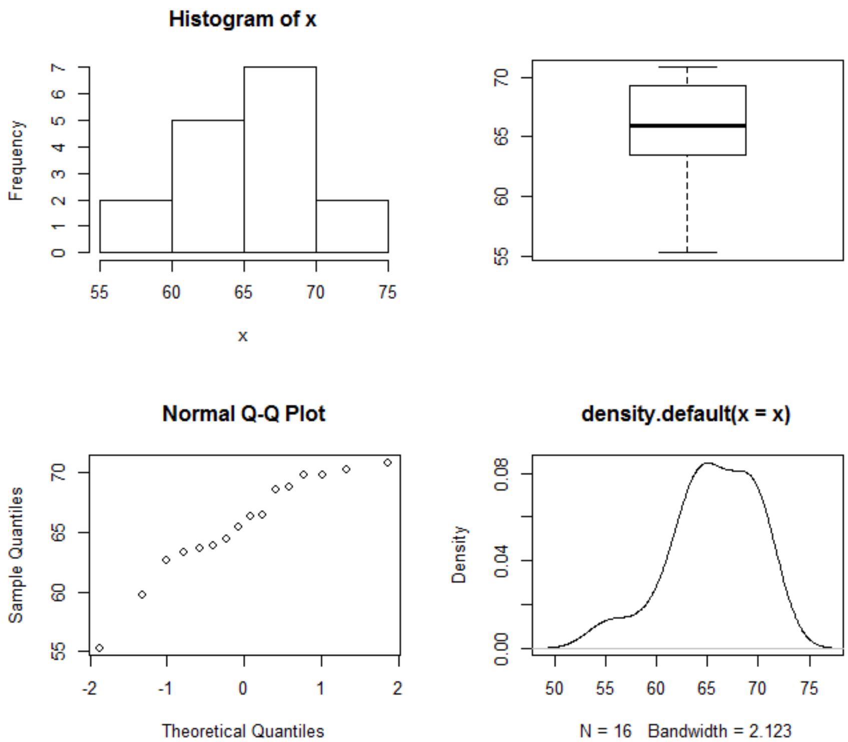
$$\bar{x} \pm z_{0.025}\sigma/\sqrt{n} = 1014 + (1.96)(25)/\sqrt{20} = (1003.043, 1024.957).$$

(b)

$$\bar{x} - z_{0.05}\sigma/\sqrt{n} = 1014 - (1.645)(25)/\sqrt{20} = (1004.804, \infty).$$

[4] **Exercise 8.36 :**

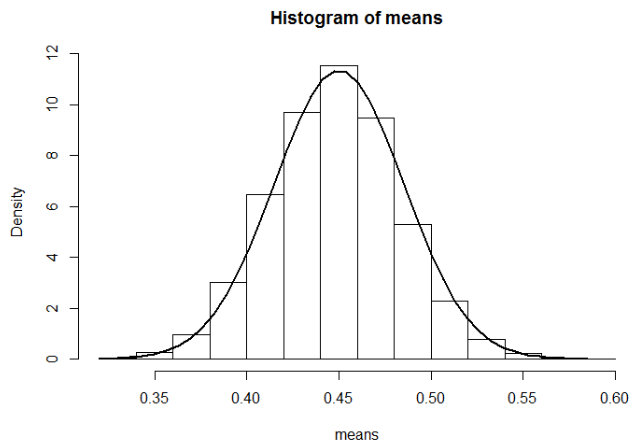
```
> x=c(55.291,66.458,70.237,65.454,59.718,68.548,69.787,
+ 64.391,62.688,69.857,68.793,63.62,63.886,70.833,66.388,
+ 63.287)
> par(mfrow=c(2,2))
> hist(x)
> boxplot(x)
> qqnorm(x)
> plot(density(x))
> t.test(x)

        One Sample t-test

data:  x
t = 62.0782, df = 15, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 63.32566 67.82884
sample estimates:
mean of x
 65.57725
```



[4] **[R Problem] : Bernoulli Random Variable :** With the following program, we produced the following diagram.

```
#Bernoulli trials
x=rbinom(5000*200,1,0.45)
# 5000 samples of size 200 stored in a 5000 by 200 matrix
A=matrix(x,ncol=200)
# calculates the proportion of 1's for each row
means=apply(A,1,mean)
#draws density histogram for sample means to check normality
hist(means,prob=TRUE)
#overlay a normal density
```

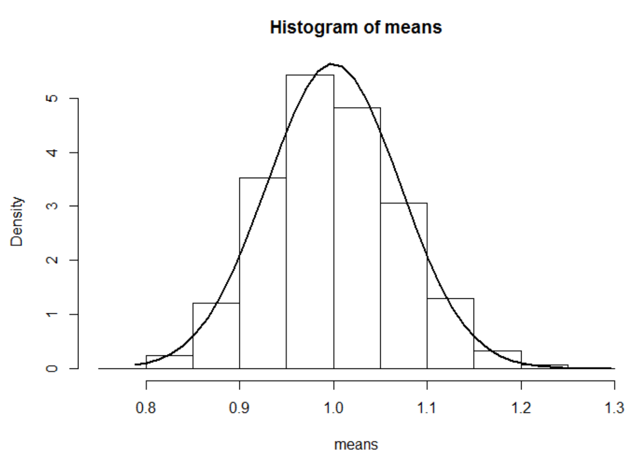
```
xnorm <- seq(min(means),max(means),length=40)
ynorm <-dnorm(xnorm,mean=0.45,sd=sqrt((0.45)*(1-0.45)/200))
lines(xnorm, ynorm, lwd=2)
```



It is the density histogram of 5000 means, each for a sample of size $n = 200$. We overlayed the density for a normal distribution with mean $\mu_{\overline{X}} = p = 0.45$ and standard deviation $\sigma_{\overline{X}} = \sqrt{p(1-p)/n} = \sqrt{0.45(1-p)/n}$. This is an illustration of the **Central Limit Theorem.**

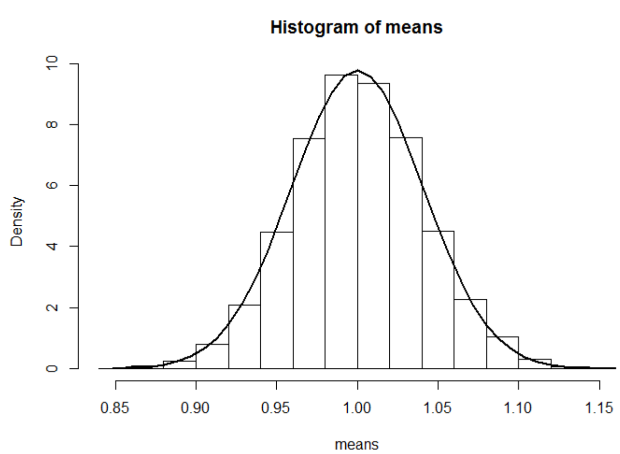
Exponential Distribution : Suppose that X has an exponential distribution with mean $1 = 1/\lambda$. So $\lambda = 1$ and furthermore its variance is $1/\lambda^2 = 1$. So by the Central Limit Theorem, for large n, $\overline{X}$ has approximately a normal distribution with mean $\mu_{\overline{X}} = \mu = 1$ and standard deviation $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 1/\sqrt{200}$. Here, is the program to generate 5000 samples, each of size $n = 200$, from an exponential distribution with mean 1. We see that there is a good fit between the histogram of the means and the approximate sampling distribution from the Central Limit Theorem.

```
#Exponential population
x=rexp(5000*200,1)
# 5000 samples of size 200 stored in a 5000 by 200 matrix
A=matrix(x,ncol=200)
# calculates the proportion of 1's for each row
means=apply(A,1,mean)
#draws density histogram for sample means to check normality
hist(means,prob=TRUE)
#overlay a normal density
xnorm <- seq(min(means),max(means),length=40)
ynorm <-dnorm(xnorm,mean=1,sd=1/sqrt(200))
lines(xnorm, ynorm, lwd=2)
```



Uniform Distribution : Suppose that X has a $U(0,2)$ distribution. Its mean $\mu = (0+2)/2 = 1$ and its standard deviation is $\sigma = (2-0)/\sqrt{12}$. So by the Central Limit Theorem, for large n, $\overline{X}$ has approximately a normal distribution with mean $\mu_{\overline{X}} = \mu = 1$ and standard deviation $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 2/\sqrt{12(200)}$. Here, is the program to generate 5000 samples, each of size $n = 200$, from an exponential distribution with mean 1. We see that there is a good fit between the histogram of the means and the approximate sampling distribution from the Central Limit Theorem.

```
#uniform population
x=runif(5000*200,0,2)
# 5000 samples of size 200 stored in a 5000 by 200 matrix
A=matrix(x,ncol=200)
# calculates the proportion of 1's for each row
means=apply(A,1,mean)
#draws density histogram for sample means to check normality
hist(means,prob=TRUE)
#overlay a normal density
xnorm <- seq(min(means),max(means),length=40)
ynorm <-dnorm(xnorm,mean=1,sd=2/sqrt(12*200))
lines(xnorm, ynorm, lwd=2)
```



[/**16**]