MAT 2377
Probability and Statistics for Engineers

Practice Set

P. Boily (uOttawa)

Winter 2021

Based on course notes by Rafał Kulik

Q78. Consider the following R output:

```
> pbinom(15,100,0.25)
[1] 0.01108327
> pbinom(17,100,0.25)
[1] 0.03762626
> pbinom(31,100,0.25)
[1] 0.9306511
```

```
> pbinom(16,100,0.25)
[1] 0.02111062
> pbinom(30,100,0.25)
[1] 0.8962128
> pbinom(32,100,0.25)
[1] 0.9554037
```

Let X be a binomial random variable with $n = 100$ and $p = 0.25$. Using the R output above, calculate $P(16 \leq X \leq 31)$.

a) 0.9196 b) 0.9095 c) 0.9348 d) 0.9443 e) none of the preceding

Solution: We have

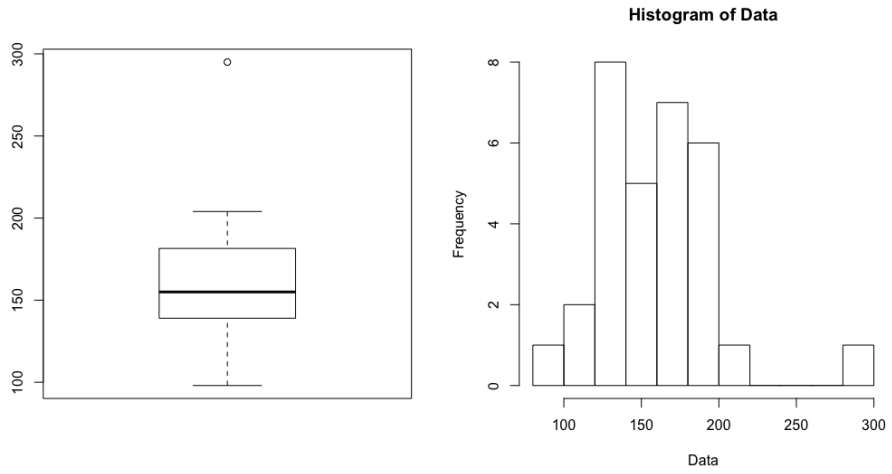
$$\begin{aligned} P(16 \leq X \leq 31) &= P(X \leq 31) - P(X < 16) \\ &= P(X \leq 31) - P(X \leq 15) \\ &= \texttt{pbinom}(31, 100, 0.25) - \texttt{pbinom}(15, 100, 0.25) \\ &\approx 0.9307 - 0.0111 \approx 0.9196 \end{aligned}$$

Q84. Discuss the normality of the following dataset:

170,295,200,165,140,190,195,142,138,148,110,140,103,176,125,
126,204,196,98,123,124,152,177,168,175,186,140,147,174,155,195

Solution: the following piece of R code will produce a boxplot and a histogram for the data:

```
Data=c(170,295,200,165,140,190,195,142,138,148,110,140,103,
176,125,126,204,196,98,123,124,152,177,168,175,186,140,147,174,155,195);
par(mfrow=c(1,2)); boxplot(Data);hist(Data,breaks=10)
```



The observations do not seem to be symmetric; there is an outlier; the data does not seem to be normal.

Q85. Using R, illustrate the central limit theorem by generating $M = 300$ samples of size $n = 30$ from:

- a normal random variable with mean 10 and variance 0.75;
- a binomial random variable with 3 trials and probability of success 0.3.

Repeat the same procedure for samples of size $n = 200$. What do you observe?

Hint: In each case, assess the normality using a histogram and a QQ plot.

Solution: let's try the following code

```
set.seed(1234)
n=30
M=300
x <- rnorm(n, mean=10, sd=sqrt(0.75))
hist(x)

mean(x)
sd(x)
means.x=c()

for(m in 1:M){
  x <- rnorm(n, mean=10, sd=sqrt(0.75))
  means.x[m] = mean(x)
}
```

```
hist(means.x)
mean(means.x)
sd(means.x)
qqnorm(means.x, pch = 1, frame = FALSE)
qqline(means.x, col = "steelblue", lwd = 2)
```

```
sqrt(0.75)/sqrt(n)
```

What do we need to modify to do the rest? `rbinom(n,3,0.3)`