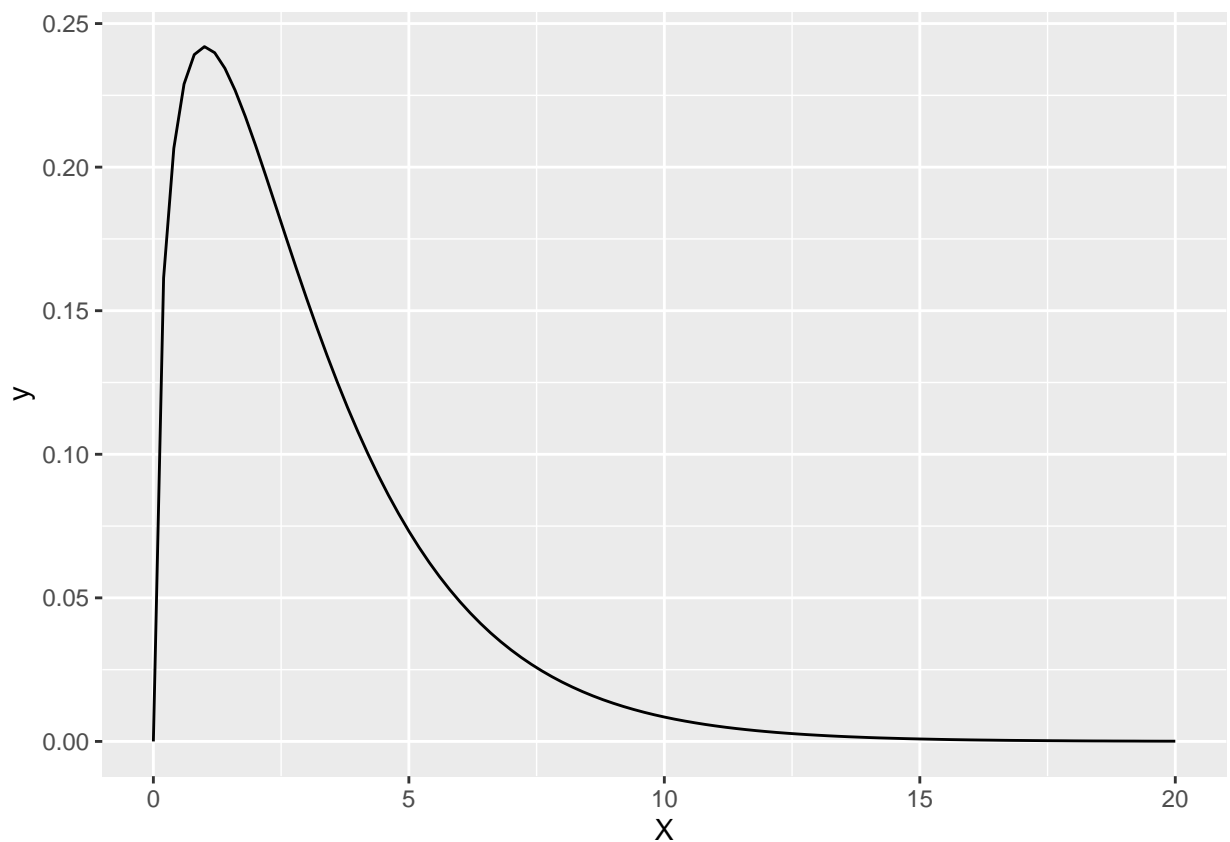# Homework 7

Noah McIntire

Problem 1

```
library(ggplot2)
set.seed(07312001)
Xdata2 <- data.frame(X=c(0,20))
ggplot(Xdata2, aes(x=X)) + stat_function(fun=dchisq, args=list(df=3))
```


The shape of the curve is right skewed with one peak, with a center around 2.

Problem 2

a)

```r
ztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  zstat <- (xbar - 3 )/(2.45/sqrt(x))
  pvalue <- pnorm(-abs(zstat))
  reject_null <- (pvalue < 0.025)
  sum(reject_null)/10000
}
sizes<-c(9,27,51)
samp.means <- sapply(sizes, ztest)
samp.means
```

```
## [1] 0.0460 0.0446 0.0502
```

b)

```r
lztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  zstat <- (xbar - 3 )/(2.45/sqrt(x))
  pvalue <- pnorm(zstat)
  reject_null <- pvalue < 0.05
  sum(reject_null)/10000
}

samp.means1 <- sapply(sizes, lztest)
samp.means1
```

```
## [1] 0.0278 0.0376 0.0387
```

2

c)

```
rztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  zstat <- (xbar - 3 )/(2.45/sqrt(x))
  pvalue <- pnorm(zstat, lower.tail=F)
  reject_null <- (pvalue < 0.05)
  sum(reject_null)/10000
}
samp.means2 <- sapply(sizes, rztest)
samp.means2
```

```
## [1] 0.0601 0.0604 0.0553
```

Problem 3

```
df<-data.frame(sizes, samp.means, samp.means1, samp.means2)
colnames(df) <- c("Sample Size", "Two-sided", "Left-sided", "Right-sided")
df
```

```
##   Sample Size Two-sided Left-sided Right-sided
## 1           9    0.0460     0.0278      0.0601
## 2          27    0.0446     0.0376      0.0604
## 3          51    0.0502     0.0387      0.0553
```

Using the empirical time one error from each of our z-test, we can see that as the sample sizes increases within each simulation, the probability of reject of the null hypothesis when it is true comes closer and closer to our assumed significance level of 0.05. The trend also demonstrates an important assumption made when performing a z-test, that we are using an adequate sample size.

Problem 4

a)

```
ttest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  tstat <- (xbar - 3 )/(sqrt((std^2)/x))
  pvalue <- pnorm(-abs(tstat))
  reject_null <- (pvalue < 0.025)
  sum(reject_null)/10000
}
tsamp.means <- sapply(sizes, ttest)
tsamp.means
```

```
## [1] 0.1228 0.0748 0.0640
```

b)

```
lttest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  tstat <- (xbar - 3 )/(sqrt((std^2)/x))
  pvalue <- pnorm(tstat)
  reject_null <- (pvalue < 0.05)
  sum(reject_null)/10000
}
tsamp.means1 <- sapply(sizes, lttest)
tsamp.means1
```

```
## [1] 0.1429 0.0934 0.0814
```

c)

```
rttest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  tstat <- (xbar - 3 )/(sqrt((std)^2/x))
  pvalue <- pnorm(tstat, lower.tail = F)
  reject_null <- (pvalue < 0.05)
  sum(reject_null)/10000
}

tsamp.means2 <- sapply(sizes, rttest)
tsamp.means2
```

```
## [1] 0.0270 0.0304 0.0325
```

Problem 5

```
df<-data.frame(sizes, tsamp.means, tsamp.means1, tsamp.means2)
colnames(df) <- c("Sample Size", "Two-sided", "Left-sided", "Right-sided")
df
```

```
##   Sample Size Two-sided Left-sided Right-sided
## 1           9    0.1228     0.1429      0.0270
## 2          27    0.0748     0.0934      0.0304
## 3          51    0.0640     0.0814      0.0325
```

The same trend that was seen from our z tests can be seen here: as the sample sizes increases (as well as degrees of freedom) within each simulation, the probability of reject of the null hypothesis when it is true comes closer and closer to our assumed significance level of 0.05. However, these are less accurate then the results shown in the table from problem 3. The trend above can also demonstrates an important assumption made when performing a t-test, that we are using an adequate sample size.

Problem 6

a)

```r
iztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  zstat <- (xbar - 3 )/(std/sqrt(x))
  pvalue <- pnorm(-abs(zstat))
  reject_null <- (pvalue < 0.025)
  sum(reject_null)/10000
}

isamp.means <- sapply(sizes, iztest)
isamp.means
```

```
## [1] 0.1187 0.0735 0.0651
```

b) Left is less than or equal to

```r
ilztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  zstat <- (xbar - 3 )/(std/sqrt(x))
  pvalue <- pnorm(zstat)
  reject_null <- (pvalue < 0.05)
  sum(reject_null)/10000
}
isamp.means1 <- sapply(sizes, ilztest)
isamp.means1
```

```
## [1] 0.1429 0.0964 0.0829
```

c)

```
irztest <- function(x){
  samples <- replicate(10000,rchisq(x,3))
  xbar<- apply(samples,2, mean)
  std<- apply(samples,2, sd)
  zstat <- (xbar - 3 )/(std/sqrt(x))
  pvalue <- pnorm(zstat, lower.tail = F)
  reject_null <- (pvalue  < 0.05)
  sum(reject_null)/10000
}
isamp.means2 <- sapply(sizes, irztest)
isamp.means2
```

```
## [1] 0.0283 0.0263 0.0323
```

Problem 7

```
df<-data.frame(sizes, isamp.means, isamp.means1, isamp.means2)
colnames(df) <- c("Sample Size", "Two-sided", "Left-sided", "Right-sided")
df
```

```
##   Sample Size Two-sided Left-sided Right-sided
## 1           9    0.1187     0.1429      0.0283
## 2          27    0.0735     0.0964      0.0263
## 3          51    0.0651     0.0829      0.0323
```

Because we are working with incorrectly performed z-tests, it would make sense that the results are less accurate then the correct z-tests performed in problem 3. However it still follows the trend that as the sample sizes increases within each simulation, the probability of reject of the null hypothesis when it is true comes closer and closer to our assumed significance level of 0.05.