Homework7

Akhil Havaldar

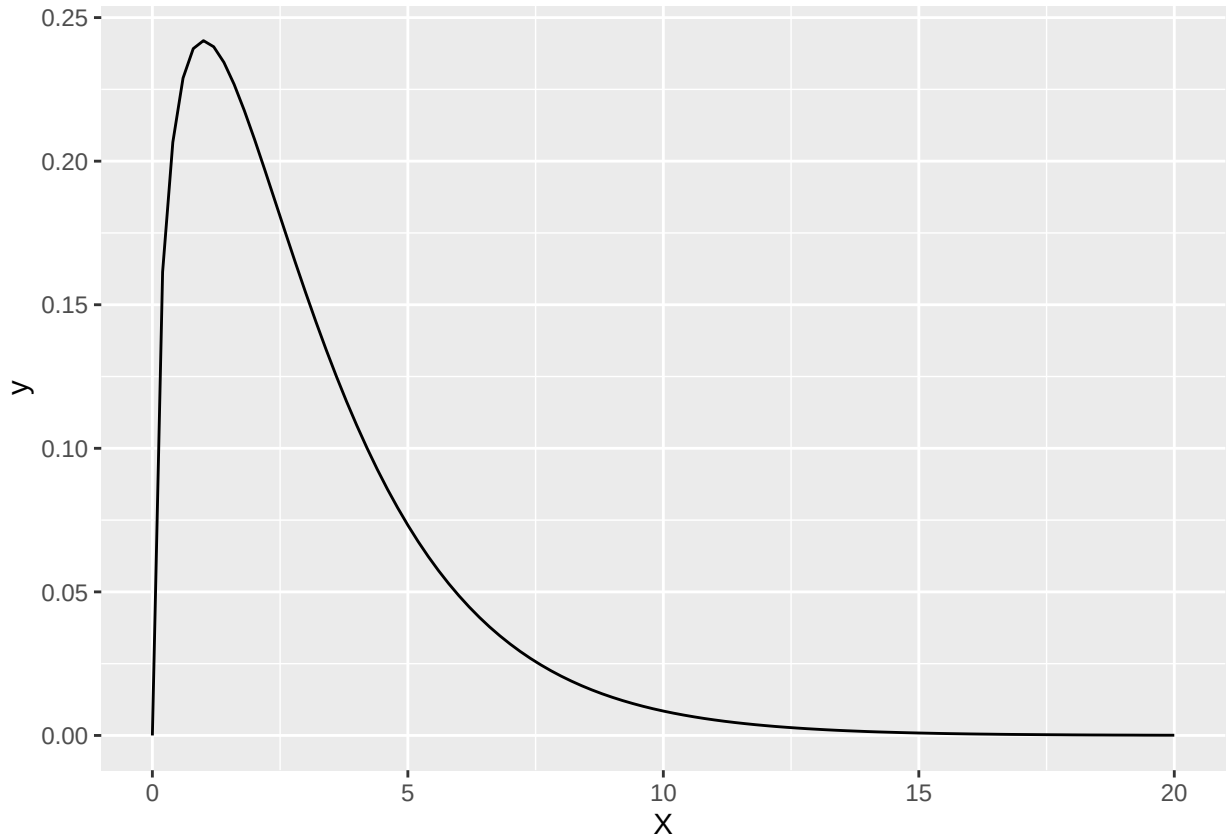
10/24/2022

Question 1

```
set.seed(05202001)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

```
dat <- data.frame(X=c(0,20))
ggplot(dat, aes(x=X)) + stat_function(fun=dchisq, args=list(df=3))
```



- The density curve is heavily right skewed because of its low degrees of freedom.

Question 2

2a)

```
mu <- 3
sig <- 2.45
alpha <- .05

ztest2 <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sig/sqrt(size))
  pvalue <- 2*pnorm(-(abs(teststat)))
  if(pvalue < alpha){TRUE}
  else{FALSE}
}

K <- 10000
t1two <- function(size){
  samps <- replicate(K, ztest2(size))
  length(which(samps == TRUE)) / 10000
}

q2a <- sapply(c(8,23,52), t1two)
q2a
```

```
## [1] 0.0455 0.0513 0.0526
```

2b)

```
ztleft <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sig/sqrt(size))
  pvalue <- pnorm(teststat, lower.tail = TRUE)
  if(pvalue < alpha){TRUE}
  else{FALSE}
}

K <- 10000
t1lft <- function(size){
  samps <- replicate(K, ztleft(size))
  length(which(samps == TRUE)) / 10000
}

q2b <- sapply(c(8,23,52), t1lft)
q2b
```

```
## [1] 0.0264 0.0375 0.0424
```

2c)

```
ztright <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sig/sqrt(size))
  pvalue <- pnorm(teststat, lower.tail = FALSE)
  if(pvalue < alpha){TRUE}
  else{FALSE}
}

K <- 10000
type1right <- function(size){
samps <- replicate(K, ztright(size))
length(which(samps == TRUE)) / 10000
}

q2c <- sapply(c(8,23,52), type1right)
q2c
```

```
## [1] 0.0593 0.0588 0.0607
```

Question 3

```
out <- data.frame(q2a, q2b, q2c)
rownames(out) <- c("n=8", "n=23", "n=52")
colnames(out) <- c("two-sided","left-sided", "right-sided")
out
```

```
##      two-sided left-sided right-sided
## n=8     0.0455     0.0264      0.0593
## n=23    0.0513     0.0375      0.0588
## n=52    0.0526     0.0424      0.0607
```

- As the sample size increases for each test, the type 1 error gets closer and closer to the significance level of 0.05.

Question 4

4a)

```
t2 <- function(size){
  sample <- rchisq(size, df=3)
  teststat <- (mean(sample) - mu) / (sd(sample)/sqrt(size))
  pvalue <- 2*pt(-abs(teststat), df=(size-1))
  pvalue < alpha
}
```

```
K <- 10000
t1twoT <- function(size){
  samps <- replicate(K, t2(size))
  length(which(samps == TRUE)) / 10000
}

q4a <- sapply(c(8,23,52), t1twoT)
q4a
```

```
## [1] 0.0882 0.0684 0.0611
```

4b)

```
ttleft <- function(size){
  sample <- rchisq(size, df=3)
  teststat <- (mean(sample) - mu) / (sd(sample)/sqrt(size))
  pvalue <- pt(teststat, df=(size-1), lower.tail = TRUE)
  pvalue < alpha
}

K <- 10000
t1lftT <- function(size){
  samps <- replicate(K, ttleft(size))
  length(which(samps == TRUE)) / 10000
}

q4b <- sapply(c(8,23,52), t1lftT)
q4b
```

```
## [1] 0.1141 0.0943 0.0815
```

4c)

```
ttright <- function(size){
  sample <- rchisq(size, df=3)
  teststat <- (mean(sample) - mu) / (sd(sample)/sqrt(size))
  pvalue <- pt(teststat, df=(size-1), lower.tail = FALSE)
  pvalue < alpha
}

K <- 10000
type1rightT <- function(size){
  samps <- replicate(K, ttright(size))
  length(which(samps == TRUE)) / 10000
}

q4c <- sapply(c(8,23,52), type1rightT)
q4c
```

```
## [1] 0.0147 0.0222 0.0305
```

Question 5

```
out2 <- data.frame(q4a, q4b, q4c)
rownames(out2) <- c("n=8", "n=23", "n=52")
colnames(out2) <- c("two-sided","left-sided", "right-sided")
out2
```

```
##      two-sided left-sided right-sided
## n=8     0.0882     0.1141      0.0147
## n=23    0.0684     0.0943      0.0222
## n=52    0.0611     0.0815      0.0305
```

- As the sample size increases for each test, the type 1 error gets closer and closer to the significance level of 0.05.

Question 6

6a)

```
zbad <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sd(sample)/sqrt(size))
  pvalue <- 2*pnorm(-(abs(teststat)))
  if(pvalue < alpha){TRUE}
  else{FALSE}
}

K <- 10000
t1twoZbad <- function(size){
  samps <- replicate(K, zbad(size))
  length(which(samps == TRUE)) / 10000
}

q6a <- sapply(c(8,23,52), t1twoZbad)
q6a
```

```
## [1] 0.1305 0.0844 0.0653
```

6b)

```
ztleftbad <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sd(sample)/sqrt(size))
  pvalue <- pnorm(teststat, lower.tail = TRUE)
  if(pvalue < alpha){TRUE}
  else{FALSE}
```

```
}

K <- 10000
t1lftZbad <- function(size){
  samps <- replicate(K, ztleftbad(size))
  length(which(samps == TRUE)) / 10000
}

q6b <- sapply(c(8,23,52), t1lftZbad)
q6b
```

```
## [1] 0.1505 0.1071 0.0819
```

6c)

```
ztrightbad <- function(size) {
  sample <- rchisq(size, df=3)
  samplemean <- mean(sample)
  teststat <- (samplemean - mu) / (sd(sample)/sqrt(size))
  pvalue <- pnorm(teststat, lower.tail = FALSE)
  if(pvalue < alpha){TRUE}
  else{FALSE}
}

K <- 10000
type1rightZbad <- function(size){
  samps <- replicate(K, ztrightbad(size))
  length(which(samps == TRUE)) / 10000
}

q6c <- sapply(c(8,23,52), type1rightZbad)
q6c
```

```
## [1] 0.0332 0.0302 0.0353
```

Question 7

```
out3 <- data.frame(q6a, q6b, q6c)
rownames(out3) <- c("n=8", "n=23", "n=52")
colnames(out3) <- c("two-sided","left-sided", "right-sided")
out3
```

```
##      two-sided left-sided right-sided
## n=8     0.1305     0.1505      0.0332
## n=23    0.0844     0.1071      0.0302
## n=52    0.0653     0.0819      0.0353
```

- As the sample size increases for each test, the type 1 error gets closer and closer to the significance level of 0.05.