# Homework 7

Santhosh Rajendran

## Seed Setting

**Multiple parts can be separated by secondary headers.**

```
#Seed Setting
set.seed(06072001)
```

## Problem 1

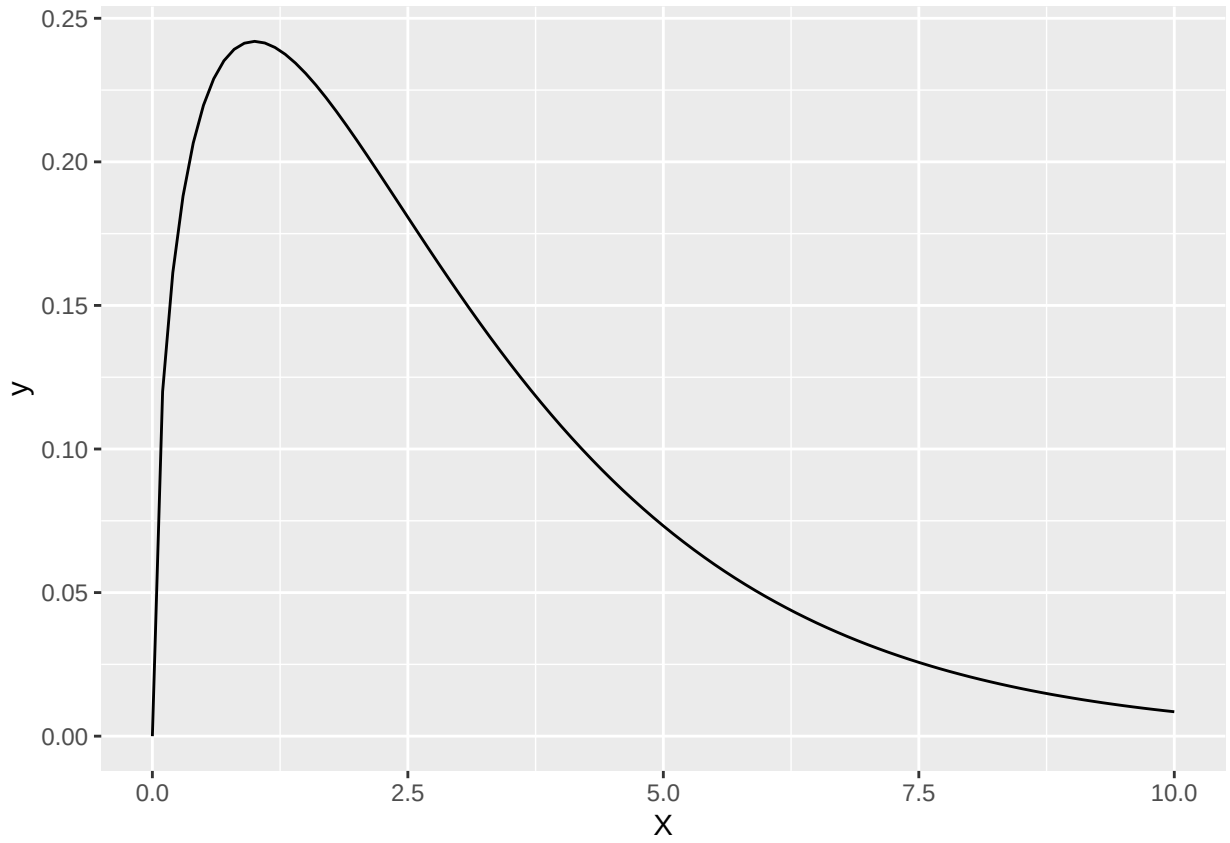
```
library('tidyverse')
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
chidist = data.frame(X=c(0,10))
ggplot(chidist,aes(x=X)) + stat_function(fun=dchisq, args=list(df=3))
```




```
# The distribution is not symmetric nor does it have only one direction;
# it starts out trending positively very quickly then trending towards
# decreasing to 0 when x=10.
```

## Problem 2

```
#a
mu=3
stdev=2.45
rep=10000
sig=.05
sizes=c(9,27,51)

twotailp2 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  z = ((xbar-mu) / (stdev/sqrt(size)))
  p_val = 2*(1-pnorm(abs(z)))
  p_val < sig
}
```

```
twotailp2results =
  lapply(sizes, function(size) replicate(rep, twotailp2(size)))
type1twop2 =
  sapply(twotailp2results, function(tests) sum(tests == TRUE) / rep)
type1twop2
```

```
## [1] 0.0473 0.0472 0.0500
```

```
#b
lefttailp2 <- function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  z = ((xbar-mu) / (stdev/sqrt(size)))
  p_val = pnorm(z)
  p_val < sig
}
lefttailp2results =
  lapply(sizes, function(size) replicate(rep, lefttailp2(size)))
type1leftp2 =
  sapply(lefttailp2results, function(tests) sum(tests == TRUE) / rep)
type1leftp2
```

```
## [1] 0.0276 0.0383 0.0447
```

```
#c
righttailp2 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  z = ((xbar-mu) / (stdev/sqrt(size)))
  p_val = 1-pnorm(z)
  p_val < sig
}
righttailp2results =
  lapply(sizes, function(size) replicate(rep, righttailp2(size)))
type1rightp2 =
  sapply(righttailp2results, function(tests) sum(tests == TRUE) / rep)
type1rightp2
```

```
## [1] 0.0629 0.0615 0.0578
```

## Problem 3

```
prob2df <- data.frame("Sample.size" = c(9, 27, 51), "Two.sided" = type1twop2,
"Left.sided" = type1leftp2, "Right.sided" = type1rightp2)
prob2df
```

```
##   Sample.size Two.sided Left.sided Right.sided
## 1           9    0.0473     0.0276      0.0629
## 2          27    0.0472     0.0383      0.0615
## 3          51    0.0500     0.0447      0.0578

# As the sample size increases, probability of a type 1 error, incorrectly rejecting
# the null hypothesis approaches .05, the significance level of the test.
# The two sided and left sided approach .05 from the lower end, while
# the right sided test is consistently decreasing towards .05.
```

## Problem 4

```
#a
twotailp4 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
  t = ((xbar-mu) / (s/sqrt(size)))
  p_val = 2*(1-pt(abs(t), df=(size-1)))
  p_val < sig
}
twotailp4results =
  lapply(sizes, function(size) replicate(rep, twotailp4(size)))
type1twop4 =
  sapply(twotailp4results, function(tests) sum(tests == TRUE) / rep)
type1twop4

## [1] 0.0856 0.0684 0.0571

#b
lefttailp4 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
  t = ((xbar-mu) / (s/sqrt(size)))
  p_val = pt(t, df=(size-1))
  p_val < sig
}
lefttailp4results =
  lapply(sizes, function(size) replicate(rep, lefttailp4(size)))
type1leftp4 =
  sapply(lefttailp4results, function(tests) sum(tests == TRUE) / rep)
type1leftp4
```

```
## [1] 0.1142 0.0897 0.0824
```

```
#c
righttailp4 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
  t = ((xbar-mu) / (s/sqrt(size)))
  p_val = 1-pt(t, df=(size-1))
  p_val < sig
}
righttailp4results =
  lapply(sizes, function(size) replicate(rep, righttailp4(size)))
type1rightp4 =
  sapply(righttailp4results, function(tests) sum(tests == TRUE) / rep)
type1rightp4
```

```
## [1] 0.0187 0.0250 0.0320
```

## Problem 5

```
prob4df = data.frame("Sample.size" = c(9, 27, 51), "Two.sided" = type1twop4,
                     "Left.sided"=type1leftp4, "Right.sided"= type1rightp4)
prob4df
```

```
##   Sample.size Two.sided Left.sided Right.sided
## 1           9    0.0856     0.1142      0.0187
## 2          27    0.0684     0.0897      0.0250
## 3          51    0.0571     0.0824      0.0320
```

```
# The values here are further from the significance level than the
# values are in the z-test. Even at sample size 51, the only value
# close to .05 is the two-sided test. The left sided and right sided
# test values are trending towards .05 but still very far from it.
```

## Problem 6

```
#a
twotailp6 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
```

```
  z = ((xbar-mu) / (s/sqrt(size)))
  p_val = 2*(1-pnorm(abs(z)))
  p_val < sig
}
twotailp6results =
  lapply(sizes, function(size) replicate(rep, twotailp6(size)))
twotailp6_in =
  sapply(twotailp6results, function(tests) sum(tests == TRUE) / rep)
twotailp6_in
```

```
## [1] 0.1302 0.0780 0.0655
```

```
#b
lefttailp6 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
  z = ((xbar-mu) / (s/sqrt(size)))
  p_val = pnorm(z)
  p_val < sig
}
lefttailp6results =
  lapply(sizes, function(size) replicate(rep, lefttailp6(size)))
lefttailp6_in =
  sapply(lefttailp6results, function(tests) sum(tests == TRUE) / rep)
lefttailp6_in
```

```
## [1] 0.1438 0.0990 0.0820
```

```
#c
righttailp6 = function(size){
  samp = rchisq(size, df=3)
  xbar = mean(samp)
  s = sd(samp)
  z = ((xbar-mu) / (s/sqrt(size)))
  p_val = pnorm(z)
  p_val < sig
}
righttailp6results =
  lapply(sizes, function(size) replicate(rep, righttailp6(size)))
righttailp6_in =
  sapply(righttailp6results, function(tests) sum(tests == TRUE) / rep)
righttailp6_in
```

```
## [1] 0.1414 0.0956 0.0842
```

## Problem 7

```
prob6df=data.frame("Sample.size"=c(9,27,51),"Two-sided"=twotailp6_in,
                   "Left-sided"=lefttailp6_in,"Right-sided"=righttailp6_in)
prob6df
```

```
##   Sample.size Two.sided Left.sided Right.sided
## 1           9    0.1302     0.1438      0.1414
## 2          27    0.0780     0.0990      0.0956
## 3          51    0.0655     0.0820      0.0842
```

```
# As the sample size increases from 7 to 51, the two-sided test continues to get
# significantly large. The left sided test values decrease steadily from .14 to .08
# while the right sided test values are also decreasing steadily at a similar rate.
# The probability of a type 1 error under an incorrectly conducted z-test is larger
# than the values found when correctly performing a z or t test.
```

Your explanation and comments go here, when needed.

Linebreaks are created using two spaces at the end of the previous line of text, or using the command

.

Pagebreaks are created using the command

.
Note: These commands execute when knitting, moving the end-of-sentence period in the resulting PDF.

## References

1. Your resources go here. Links should be surrounded by <>.