Question 7.1

Estimate the MSE of the level k trimmed means for random samples of size 20 generated from a standard Cauchy distribution. (The target parameter θ is the center or median; the expected value does not exist.) Summarize the estimates of MSE in a table for k= 1,2,...,9

R Code:

```
set.seed(5)
n <- 20
K <- n/2 - 1
m <- 1000
mse.se <- matrix(0, K, 2)
trimmed.mse <- function(n, m, k)
{
    tmean <- numeric(m)
    for (i in 1:m)
    {
        x <- sort(rcauchy(n))
        tmean[i] <- sum(x[(k+1):(n-k)]) / (n-2*k)
    }
    mse.est <- mean((tmean-0)^2)
    se.mse <- sqrt(var((tmean-0)^2)/m)
    return(c(mse.est, se.mse))
}
for (k in 1:K)
{
    mse.se[k, 1:2] <- trimmed.mse(n=n, m=m, k=k)
}
cbind(mse.se, seq(9))
```

Results:

```
> cbind(mse.se, seq(9))
           [,1]        [,2] [,3]
[1,] 1.7070424 0.288101467    1
[2,] 0.4117096 0.034415918    2
[3,] 0.2428884 0.016296178    3
[4,] 0.1842113 0.010088513    4
[5,] 0.1512848 0.009169079    5
[6,] 0.1344769 0.006704330    6
[7,] 0.1368011 0.007108228    7
[8,] 0.1373021 0.007471793    8
[9,] 0.1490710 0.011420845    9
```

Column 1: MSE
Column 2: Standard Error of MSE
Column 3: k # of data points trimmed from the mean calculation

Approach and Conclusion:
- 1000 trials of n-k*2 samples from the cauchy distribution were generated and evaluated for their Mean Square Error and standard deviation of mean square error
- The lowest MSE was found to occur when 6 data points were trimmed from the mean calculation on both the high and low side.
- In general, the MSE was higher and less precise with very little trimming of the data. This is because outliers/extreme points were kept in the data set.

Question 7.2

Plot the empirical power curve for the t-test in Example 7.9, changing the alternative hypothesis to H1 : $\mu \neq$ 500, and keeping the significance level α= 0.05.

R Code:

```
n <- 20
m <- 1000
mu0 <- 500
sigma <- 100
mu <- c(seq(450, 650, 10)) #alternatives
M <- length(mu)
power <- numeric(M)
for (i in 1:M) {
   mu1 <- mu[i]
   pvalues <- replicate(m, expr = {
      x <- rnorm(n, mean = mu1, sd = sigma)  #simulate under alternative mu1
      ttest <- t.test(x, alternative = "two.sided", mu = mu0)
      ttest$p.value } )
   power[i] <- mean(pvalues <= .05)
}
se <- sqrt(power * (1-power) / m)
library(ggplot2)
df <- data.frame(mean=mu, power=power,
            upper=power+2*se, lower=power-2*se)
ggplot(df, aes(x=mean, y=power)) +
   geom_line() +
   geom_vline(xintercept=500, lty=2) +
   geom_hline(yintercept=c(0,.05), lty=1:2) +
   geom_errorbar(aes(ymin=lower, ymax=upper), width = 0.2, lwd=1.5)
```

Results:
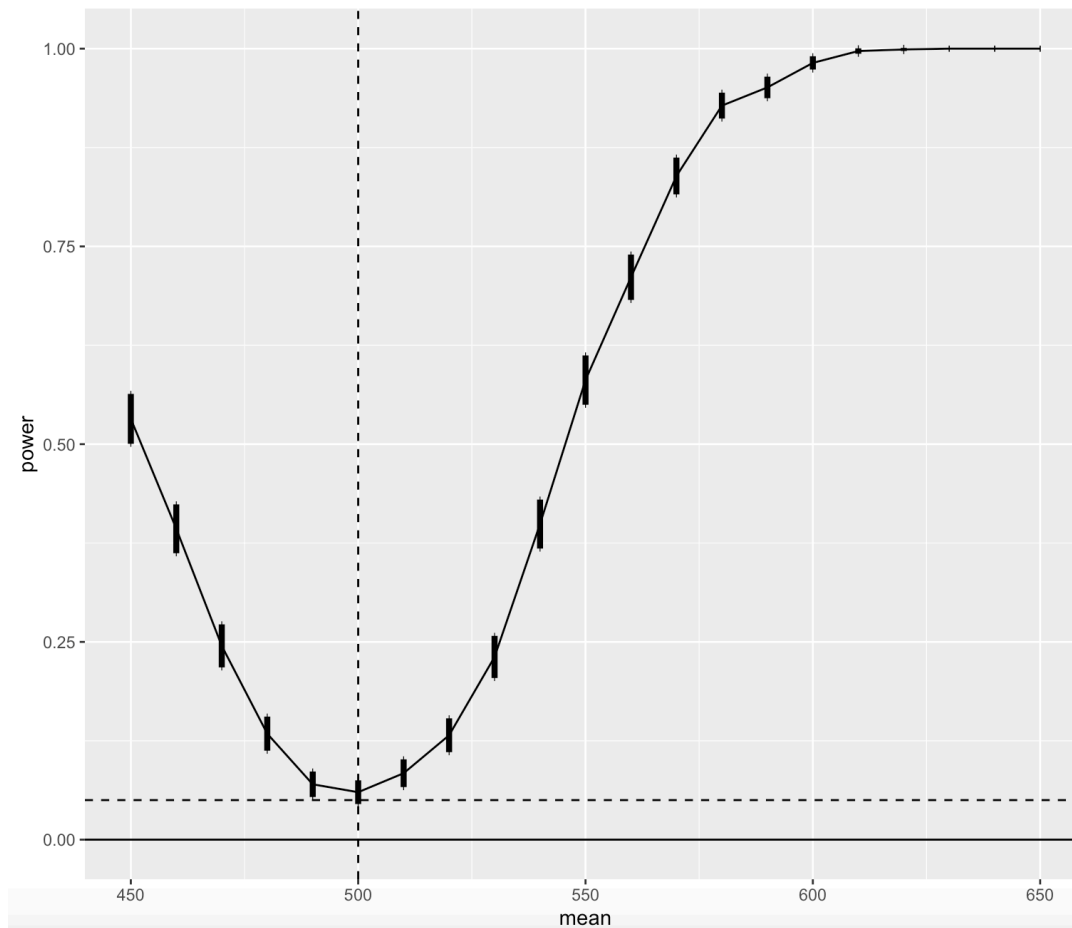
Approach and Conclusion:
- Random normal distributions with standard deviation of 100 and mean moving from 450 to 650 by 10 were generated via 1000 trials containing 20 samples each
- Each of these samples of randomly generated normal distributions were put through a t test to see if their means were significantly different from 500.
- If the means were different at a 0.05 significance level, that sample took on a value of 1, otherwise that sample took on a value of 0.
- The proportion of 1000 trials that were significantly different from 500 were then plotted in the plot above
- As the mean grew further and further from 500 in either direction (it was a 2 sided t test) then it was found that the proportion of samples that were correctly rejected increased

4

Question 7.3

Plot the power curves for the t-test in Example 7.9 for sample sizes 10, 20, 30, 40, and 50, but omit the standard error bars. Plot the curves on the same graph, each in a different color or different line type, and include a legend. Comment on the relation between power and sample size.

R Code:

```
n <- c(10, 20, 30, 40, 50)
m <- 1000
mu0 <- 500
sigma <- 100
mu <- c(seq(450, 650, 10)) #alternatives
M <- length(mu)
power <- matrix(0, M, length(n))
for (j in 1:length(n)){
    for (i in 1:M) {
        mu1 <- mu[i]
        pvalues <- replicate(m, expr = {
            x <- rnorm(n[j], mean = mu1, sd = sigma)  #simulate under alternative mu1
            ttest <- t.test(x, alternative = "two.sided", mu = mu0)
            ttest$p.value } )
        power[i, j] <- mean(pvalues <= .05)
    }
}

library(ggplot2)
df <- data.frame(mean=mu, power=power)
ggplot(df, aes(x=mean, y=power)) +
    geom_line(aes(y = power[,1], colour = "n = 10")) +
    geom_line(aes(y = power[,2], colour = "n = 20")) +
    geom_line(aes(y = power[,3], colour = "n = 30")) +
    geom_line(aes(y = power[,4], colour = "n = 40")) +
    geom_line(aes(y = power[,5], colour = "n = 50")) +
    geom_vline(xintercept=500, lty=2) +
    geom_hline(yintercept=c(0,.05), lty=1:2)
```

Results:
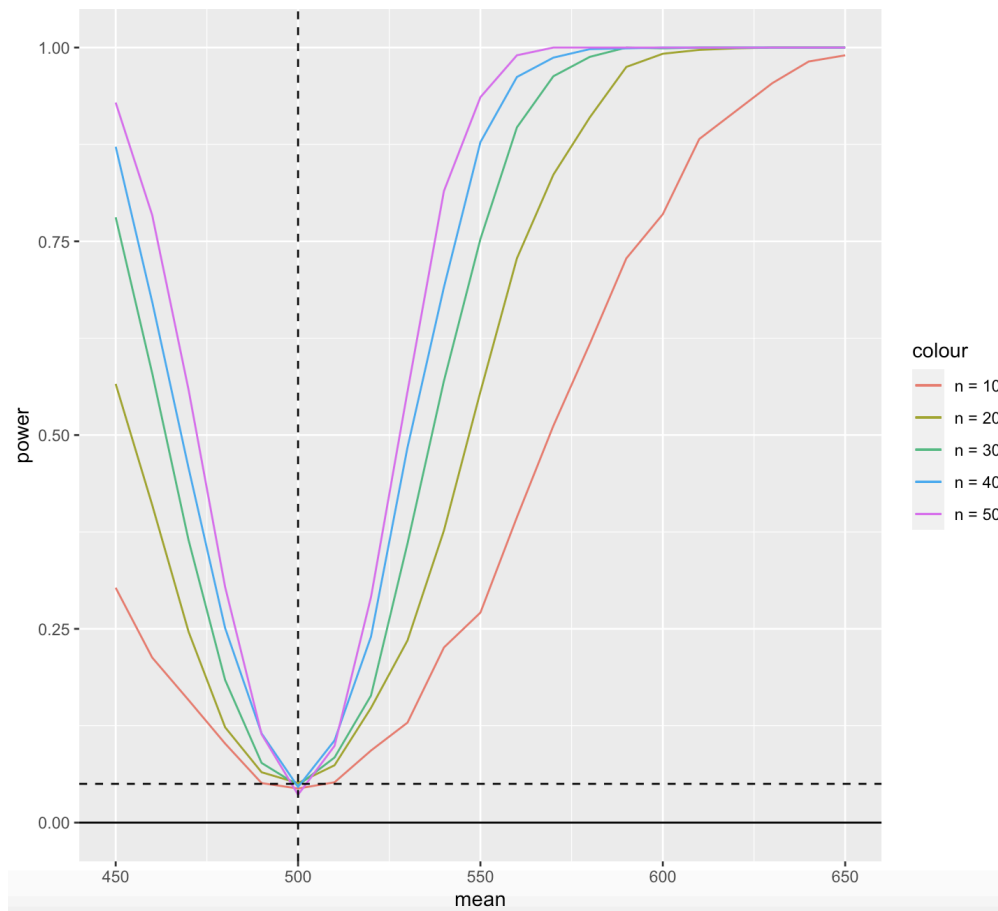
Approach and Conclusion:
- The approach for this question was very similar to problem 7.2.
- The difference between this question and problem 7.2 is that multiple sample levels were taken at each mean for the randomly generated normal distributions
- A similar trend was found as problem 7.2 where as the mean grew further and further from 500 in either direction the proportion of samples that were correctly rejected increased
- One additional piece of insight that this problem provided was that the power of the t test increases faster when the sample size is higher. Rejections of a mean of 500 were more probable when the sample size was larger.

Question 7.6

Suppose a 95% symmetric t-interval is applied to estimate a mean, but the sample data are non-normal. Then the probability that the confidence interval covers the mean is not necessarily equal to 0.95. Use a Monte Carlo experiment to estimate the coverage probability of the t-interval for random samples of $\chi^2(2)$ data with sample size n = 20. Compare your t-interval results with the simulation results in Example 7.4. (The t-interval should be more robust to departures from normality than the interval for variance.)

R Code:

```
set.seed(5)
m = 1000
n = 20
pvalues <- replicate(m, expr = {
    x <- rchisq(n, df = 2)  #simulate 20 samples from χ2(2)
    ttest <- t.test(x, alternative = "two.sided", mu = 2)  #mean of χ2(2) is 2, see if t test captures
    ttest$p.value })
mean(pvalues >= .05)
```

Results:

```
> print(mean(pvalues >= .05))
[1] 0.917
```

Approach and Conclusion:
- 1000 trials of 20 samples were generated from a Chi square distribution with 2 degrees of freedom.
- The mean or expected value of a chi square distribution is commonly known to equal to its degrees of freedom
- A t test was conducted to compare the mean of the randomly generated samples with the expected value of what the mean should be (the degrees of freedom)
- Approximately 91.7% of the t tests captured the correct expected value at the 95% confidence level. This indicates that a t test has a harder time capturing the mean of a non-normal distribution. This also indicates that the actual capture rate of the mean may differ from the confidence level anticipated from the t test if used for non normal distributions.
- Ultimately, the t-test captured the correct expected value a higher number of times than it captured the correct variance. This indicates more robustness to the normality assumption when attempting to capture the mean than the variance as the problem statement alludes to.

Question 7.8

Estimate the power of the skewness test of normality against symmetric Beta(α,α) distributions and comment on the results. Are the results different for heavy-tailed symmetric alternatives such as t(v)?

R Code:

```
a <- c(1, 2, 3, 4, 5, 10, 50, 100) #holds different values of alpha for beta distribution
pwr.beta <- numeric(length(a)) #holds the power of rejecting normality at that particular alpha
alpha <- 0.1 #confidence in normality
n <- 30
m <- 2500
cv <- qnorm(1-alpha/2, 0, sqrt(6*(n-2) / ((n+1)*(n+3)))) # critical value for the skewness test
sk <- function(x) #computes the sample skewness coeff.
{
   xbar <- mean(x)
   m3 <- mean((x - xbar)^3)
   m2 <- mean((x - xbar)^2)
   return( m3 / m2^1.5 )
}
for (j in 1:length(a))  #for each value of alpha
{
   sktests <- numeric(m)
   for (i in 1:m)       #for each replicate
   {
      x <- rt(n, df)
      sktests[i] <- as.integer(abs(sk(x)) >= cv)
   }
   pwr.beta[j] <- mean(sktests)
}
cbind(a, pwr.beta)
df <- c(1, 2, 3, 4, 5, 10, 50, 100) #holds different degrees of freedom for t distribution
pwr.t <- numeric(length(a)) #holds the power of rejecting normality at that particular df
for (j in 1:length(df))  #for each value of alpha
{
   sktests <- numeric(m)
   for (i in 1:m)       #for each replicate
   {
      x <- rt(n, df[j])
      sktests[i] <- as.integer(abs(sk(x)) >= cv)
   }
   pwr.t[j] <- mean(sktests)
}
cbind(df, pwr.t)
```

Results:

```
         a pwr.beta
[1,]     1    0.6512
[2,]     2    0.6604
[3,]     3    0.6612
[4,]     4    0.6408
[5,]     5    0.6488
[6,]    10    0.6596
[7,]    50    0.6496
[8,]   100    0.6504
```

```
        df   pwr.t
[1,]     1 0.8716
[2,]     2 0.6564
[3,]     3 0.4884
[4,]     4 0.4068
[5,]     5 0.3292
[6,]    10 0.2140
[7,]    50 0.1252
[8,]   100 0.1128
```

Approach and Conclusion:
- Different values of a were taken to generate random beta(a,a) samples
- 1000 trials of 30 samples were taken from the various beta(a,a) distributions, the skewness value was calculated and compared against the critical value of the normal distribution

- If the skewness value was greater than the critical value it was correctly indicated as non normal. The proportion of skewness values that were found to have a magnitude greater than the critical value were noted as the power.
- The power of the skewness test remained consistent regardless of the a chosen to generate the beta(a,a) distributions.
- Different values of degrees of freedom were taken to generate random t(df) distributions
- 1000 trials of 30 samples were taken from the various t(df) distributions, the skewness value was calculated and compared against the critical value of the normal distribution
- If the skewness value was greater than the critical value it was correctly indicated as non normal. The proportion of skewness values that were found to have a magnitude greater than the critical value were noted as the power.
- The power of the skewness test decreased as the degrees of freedom of the t distribution was increased. This is likely because it became closer and closer to resembling a normal distribution.

Question 7.9

Refer to Example 7.16. Repeat the simulation, but also compute the F test of equal variance, at significance level α. = 0.055. Compare the power of the Count Five test and F test for small, medium, and large sample sizes. (Recall that the F test is not applicable for non-normal distributions.)

R Code:

```
count5test <- function(x, y) {
   X <- x - mean(x)
   Y <- y - mean(y)
   outx <- sum(X > max(Y)) + sum(X < min(Y))
   outy <- sum(Y > max(X)) + sum(Y < min(X))
   # return 1 (reject) or 0 (do not reject H0)
   return(as.integer(max(c(outx, outy)) > 5))
}
n <- c(10, 20, 30, 40, 50, 100)
m <- 1000
sigma1 <- 1
sigma2 <- 1.5
count5 <- numeric(m)
f <- numeric(m)
power.count5 <- numeric(length(n))
power.f <- numeric(length(n))
for (j in 1:length(n))  #for each value of n
{
   for (i in 1:m)  #for each value of n
   {
      x <- rnorm(n[j], 0, sigma1)
      y <- rnorm(n[j], 0, sigma2)
      count5[i] <- count5test(x, y)
      f[i] <- as.integer(var.test(x, y, alternative = "two.sided")$p.value < 0.055)
   }
   power.count5[j] <- mean(count5)
   power.f[j] <- mean(f)
}

cbind(power.count5, n)
cbind(power.f, n)
```

Results:

```
        power.count5    n
[1,]            0.101   10
[2,]            0.312   20
[3,]            0.463   30
[4,]            0.601   40
[5,]            0.672   50
[6,]            0.858  100


        power.f    n
[1,]    0.202   10
[2,]    0.407   20
[3,]    0.566   30
[4,]    0.733   40
[5,]    0.816   50
[6,]    0.975  100
```

Approach and Conclusion:
- 1000 trials of different sample sizes were randomly generated from N(0, 1) and N(0, 1.5) distributions
- Both the count five test and the F test were used to see if the variance differed between these randomly generated distributions
- The proportion of tests that correctly identified that the variances between the two distributions were different were noted.
- For both tests, it was found that if the sample size was increased, the power (or ability to correctly tell that the variances were different) was increased.
- The F test did a better job of finding that the variances were different than the count five test at every level of sample size n.