MATH-472: Homework 4

Andrew Moore

3/19/23

Question 1

Do 7.1, 7.4, and 7.6 from the exercises in Chapter 7.

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 $k = 1, 2, \dots, 9$.

```r
K <- numeric(9)
for (k in 1:9) {
  hat <- replicate(n = 1000, expr = {
    x <- rcauchy(n = 20, location = 0, scale = 1)
    x <- sort(x)
    mean(x[k:(length(x) - k)])
  })

  K[k] <- mean(hat^2)
}

out <- data.frame(matrix(nrow = 1, ncol = 9))
colnames(out) <- c("k = 1", 2:9)
rownames(out) <- "MSE"
out[1, ] <- K

knitr::kable(out, digits = 4)
```

	k = 1	2	3	4	5	6	7	8	9
MSE	67.4558	0.8918	0.3615	0.2506	0.193	0.1601	0.1423	0.1513	0.1406

7.4

Suppose that $X_1, \dots, X_n$ are a random sample from a lognormal distribution. Construct a 95% confidence interval for the parameter μ. Use a Monte Carlo method to obtain an empirical estimate of the confidence level when data is generated from standard lognormal.

```r
m <- 1000
n <- 100

calc_ci <- function(x, z = 1.96) {
  m <- mean(x)
  se <- sd(x) / sqrt(length(x))
  m + c(-1, 1) * (z * se)
}

hat <- replicate(n = m, expr = calc_ci(log(rlnorm(n, 0, 1))))
mean(map2_lgl(hat[1, ], hat[2, ], ~between(0, .x, .y)))
```

```
[1] 0.948
```

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 saples 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.)

Question 2

Suppose that $x_1, x_2, \dots, x_{15}$ are a random sample of size 15 from $N(\mu, \sigma^2 = 4)$. Calculate an empirical Type I error for $H_0 : \mu \leq 5$ v $H_1 : \mu > 5$ with $\alpha = 0.01$ and the number of replicates 1000.

```r
m <- 1000
n <- 15
sigma2 <- 4
alpha <- 0.01

calc_t_stat <- function(n, mu0, sigma2) {
  x <- rnorm(n, mean = mu0 + 0.05, sd = sqrt(sigma2))
  t.test(x, alternative = "greater", mu = m0)$statistic
}
```