Exercise 2

EMAIL: n.c.deschipper

Question 1 part 1 1.1 population standard error of the mean difference, you can estimate this with SE of student t and SE of welch

1.2

1.3 Input: S: integer deining the number of independent data sets to generate n: the sample size of the data sets $$, and vlaues fo the population parameters ofthe data generating distribution Output: S estimates o the statistics of interest, namely bias, variance, MSE and RE

1. initialize out students\_bias, students\_var, students\_mse, welch\_bias, welch\_var, welch\_mse and relative\_efficiency all as vector of length S
2. For s in 1:S: A. Generate data with sample size n1 from N() and sample size n2 from N() B. Obtain sample variance, bias and mse for both
3. intialize bias, variance, mse , bias, , mse all as vector of lentgh S and RE for relative efficency

set.seed(200)  
n1 <- c(10, 100, 200)  
n2 <- 200  
var2 <- c(1, 2 , 10)  
var1 <- 1  
mean1 <- 0  
mean2 <- 1  
S <- 10000  
conditions <- expand.grid(n1, var2)  
  
MonteCarlo\_SE <- function(n1, var2) {  
 data\_1 <- rnorm(n1, mean1, sqrt(var1))  
 data\_2 <- rnorm(n2, mean2, sqrt(var2))  
 SE\_student <- rep(NA, S)  
 SE\_welch <- rep(NA, S)  
 for (i in 1:S) {  
 data\_1 <- rnorm(n1, mean1, sqrt(var1))  
 data\_2 <- rnorm(n2, mean2, sqrt(var2))  
 pooled\_sd <- sqrt(((n1-1)\*var(data\_1)+(n2-1)\*var(data\_2))/(n1+n2-2))  
 SE\_student[i] <- pooled\_sd\*sqrt(((1/n1)+(1/n2)))  
 SE\_welch[i] <- sqrt((var(data\_1)/n1)+(var(data\_2)/n2))  
 }  
 bias\_Student <- sqrt((var1/n1)+(var2/n2)) - mean(SE\_student)  
 bias\_Welch <- sqrt((var1/n1)+(var2/n2)) - mean(SE\_welch)  
 Out\_variance\_student <- var(SE\_student)  
 Out\_variance\_welch <- var(SE\_welch)  
 Out\_bias\_student <- bias\_Student  
 Out\_bias\_welch <- bias\_Welch  
 Out\_MSE\_student <- (bias\_Student)^2 + Out\_variance\_student  
 Out\_MSE\_welch <- (bias\_Welch)^2 + Out\_variance\_welch  
 RE <- (sum(SE\_welch)/sum(SE\_student))  
 return(c(Out\_variance\_student, Out\_variance\_welch, Out\_bias\_student, Out\_bias\_welch, Out\_MSE\_student, Out\_MSE\_welch, RE))  
}  
  
results <- matrix(NA, 7, 9)  
for (i in 1:nrow(conditions)) {  
 results[, i] <- MonteCarlo\_SE(n1 = conditions[i, 1], var2 = conditions[i,2])  
}  
  
out\_results <- cbind(conditions, t(round(results, digits = 3)))  
colnames(out\_results) <- c("n1", "var2", "Variance\_S", "variance\_W", "Bias\_S", "Bias\_W", "MSE\_S", "MSE\_W", "Relative Efficiency")  
out\_results

## n1 var2 Variance\_S variance\_W Bias\_S Bias\_W MSE\_S MSE\_W  
## 1 10 1 0.000 0.005 0.000 0.009 0.000 0.005  
## 2 100 1 0.000 0.000 0.000 0.000 0.000 0.000  
## 3 200 1 0.000 0.000 0.000 0.000 0.000 0.000  
## 4 10 2 0.000 0.005 -0.121 0.008 0.015 0.005  
## 5 100 2 0.000 0.000 -0.017 0.000 0.000 0.000  
## 6 200 2 0.000 0.000 0.000 0.000 0.000 0.000  
## 7 10 10 0.003 0.003 -0.616 0.005 0.382 0.004  
## 8 100 10 0.000 0.000 -0.079 0.000 0.006 0.000  
## 9 200 10 0.000 0.000 0.000 0.000 0.000 0.000  
## Relative Efficiency  
## 1 0.974  
## 2 1.000  
## 3 1.000  
## 4 0.716  
## 5 0.895  
## 6 1.000  
## 7 0.381  
## 8 0.755  
## 9 1.000

1.2 Use t-test r-function to obtain p value Null hypothesis: The means of the two t-tests are the same Alternative Hypothesis: The means of the two t-tests are not the same

2.1 Null hypothesis: The means of the two t-tests are the same Alternative Hypothesis: The means of the two t-tests are not the same

2.2

n1 <- c(10, 100, 200)  
n2 <- 200  
var1 <- 2  
var2 <- c(1, 2, 10)  
mean1 <- 0  
mean2 <- 1  
conditions <- expand.grid(n1, var2)