ASSIGNMENT 1

1. Eitas Rimkus (u184503)
2. Nikodem Baehr (u459229)
3. Samuel Friedlaender (u848264)

Theoretical Exercises

Solution Theoretical Exercise 1

a) For $\hat{\mu} = \frac{1}{n+1}\sum_{i=1}^{n} y_i$:

- $E(\hat{\mu}) = E(\frac{1}{n+1}\sum_{i=1}^{n} y_i) = \frac{1}{n+1}E(\sum_{i=1}^{n} y_i) = \frac{n}{n+1} \cdot E(y_i) = \frac{n}{n+1} \cdot \mu \neq \mu \implies$ Hence, the estimator is `biased`.
- $\hat{\mu} = \frac{1}{n+1}\sum_{i=1}^{n} y_i = \frac{1}{n+1}\frac{n}{n}\sum_{i=1}^{n} y_i = \frac{n}{n+1}\frac{1}{n}\sum_{i=1}^{n} y_i \xrightarrow[Slutsky, LLN]{n\to\infty} 1 \cdot E(y_i) = \mu \implies$ Hence, the estimator is `consistent`.

b) For $\hat{\mu} = \frac{2}{n}\sum_{i=1}^{\frac{n}{2}} y_i$:

- $E(\hat{\mu}) = E(\frac{2}{n}\sum_{i=1}^{\frac{n}{2}} y_i) = \frac{2}{n}E(\sum_{i=1}^{\frac{n}{2}} y_i) = \frac{2}{n}\sum_{i=1}^{\frac{n}{2}} E(y_i) = \frac{2}{n} \cdot \frac{n}{2}E(y_i) = \mu \implies$ Hence, the estimator is `unbiased`.
- $\hat{\mu} = \frac{2}{n}\sum_{i=1}^{\frac{n}{2}} y_i \xrightarrow[LLN]{n\to\infty} E(y_i) = \mu \implies$ Hence, the estimator is `consistent`.

c) For $\hat{\mu} = \frac{0.1}{100}\sum_{i=1}^{100} y_i + \frac{0.9}{n-100}\sum_{i=101}^{n} y_i$:

- $E(\hat{\mu}) = E(\frac{0.1}{100}\sum_{i=1}^{100} y_i + \frac{0.9}{n-100}\sum_{i=101}^{n} y_i) = E(\frac{0.1}{100}\sum_{i=1}^{100} y_i) + E(\frac{0.9}{n-100}\sum_{i=101}^{n} y_i) = \frac{0.1}{100} \cdot 100E(y_i) + \frac{0.9}{n-100} \cdot (n-100)E(y_i) = 0.1\mu + 0.9\mu = \mu \implies$ Hence, the estimator is `unbiased`.
- $\hat{\mu} = \frac{0.1}{100}\sum_{i=1}^{100} y_i + \frac{0.9}{n-100}\sum_{i=101}^{n} y_i = \frac{0.1}{100}\sum_{i=1}^{100} y_i + \frac{0.9}{n-100}(\sum_{i=1}^{n} y_i - \sum_{i=1}^{100} y_i) = (\frac{0.1}{100} - \frac{0.9}{n-100})\sum_{i=1}^{100} y_i + \frac{0.9n}{n-100} \cdot \frac{1}{n}\sum_{i=1}^{n} y_i \xrightarrow[Slutsky, LLN]{n\to\infty} \frac{0.1}{100}\sum_{i=1}^{100} y_i + 0.9E(y_i) = \frac{0.1}{100}\sum_{i=1}^{100} y_i + 0.9\mu \neq \mu \implies$ Hence, the estimator is `inconsistent`.

Solution Theoretical Exercise 2

We have the OLS estimator $\hat{\beta}_{OLS} = (X^TX)^{-1}X^Ty$.

The Least Squares is defined as $S(\beta) = (y - X\beta)^T(y - X\beta) = y^Ty - 2\beta^TX^Ty + \beta^TX^TX\beta$.

To minimize it, we take its first and second order derivatives with respect to β:

FOC: $\frac{\partial}{\partial \beta} S(\beta) = -2X^T y + 2X^T X\beta$

SOC: $\frac{\partial^2}{\partial\beta\partial\beta^T} S(\beta) = 2X^T X$

Since SOC is positive definite as by the definition of OLS, $X^T X$ is invertible, so X has full rank.

Taking a FOC and given that SOC is > 0, we have that this is the global minimum - unique solution for β:

$0 = -2X^T y + 2X^T X\beta \Leftrightarrow 2X^T y = 2X^T X\beta \Leftrightarrow X^T y = X^T X\beta \longrightarrow \beta = (X^T X)^{-1} X^T y = \hat{\beta}_{OLS}$.

$\hat{\beta}_{OLS}$ minimises the sum of squared errors and $\hat{\beta}_{OLS} = argmin_{\beta\in\mathbb{R}^{\neg}} = (X^T X)^{-1} X^T y$

Solution Theoretical Exercise 3

Since $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ are orthogonal to each other, we have that they are independent, and $x_1 x_2 = 0$

The projection matrix M_{x_2} for x_2 is $M_{x_2} = I - P_{x_2} = I - x_2(x_2' x_2)^{-1} x_2'$

We have that

$$\hat{\beta}_1 = (x_1' M_{x_2} x_1)^{-1} x_1' M_{x_2} y$$

$$\hat{\beta}_1 = (x_1'(I - P_{x_2})x_1)^{-1} x_1'(I - P_{x_2})y$$

$$\hat{\beta}_1 = (x_1'(I - x_2(x_2'x_2)^{-1}x_2')x_1)^{-1} x_1'(I - x_2(x_2'x_2)^{-1}x_2')y$$

$$\hat{\beta}_1 = (x_1'x_1)^{-1} x_1' y$$

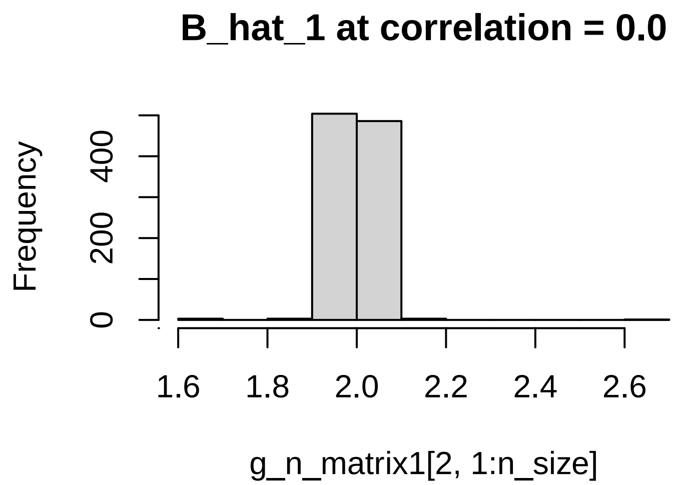
The implication of this result for the analysis of the 'partial' effect of x_1 on y is that we can be sure that x_2 does not contain any effect from x_1. This means that $\hat{\beta}_1$ accounts for the 'full' partial effect of x_1 on y without any of it being included in $\hat{\beta}_2$

Solution Theoretical Exercise 4

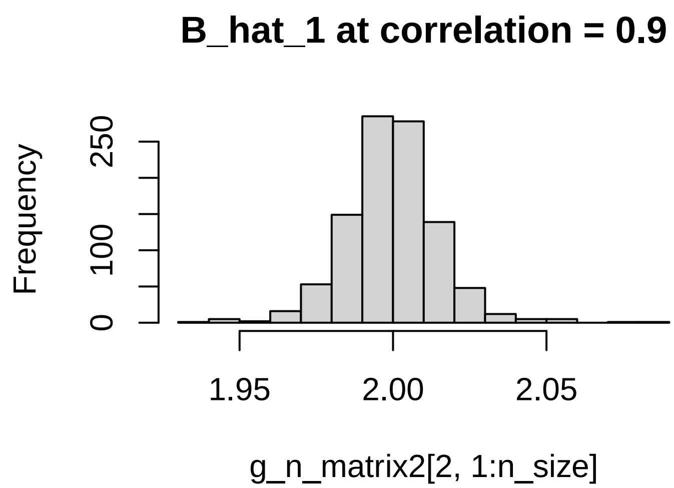
a) Let the true parameters be $\beta_0 = 0$, $\beta_1 = 2$, $\beta_2 = 3$. Consider two levels of correlation $\rho = 0.0$ and $\rho = 0.9$ between the explanatory variables x_1 and x_2. We run the Monte Carlo simulation 1000 times, and draw the histograms of the estimated coefficient $\hat{\beta}_1$.

```
library("mvtnorm")
means = c(0,0)
b_0 = 1
b_1 = 2
b_2 = 3
beta = c(b_0, b_1, b_2)

n_vec   = seq(10, 10000, by=10)
n_size  = length(n_vec)

g_n_matrix1 = matrix(ncol = n_size, nrow=3)
g_n_matrix2 = matrix(ncol = n_size, nrow=3)
```

```r
VCOV1 = matrix(c(1,0,0,1),nrow=2,byrow=T)
VCOV2 = matrix(c(1,0.9,0.9,1),nrow=2,byrow=T)

FunctionMonteCarlo = function(n)
{

  mvt1 = rmvnorm(n, mean = means, sigma = VCOV1);
  mvt2 = rmvnorm(n, mean = means, sigma = VCOV2);
  u = rnorm(n, mean=0, sd=1)
  X1 = cbind(rep(1,n),mvt1[,1],mvt1[,2])
  X2 = cbind(rep(1,n),mvt2[,1],mvt2[,2])
  y1 = X1%*%beta + u
  y2 = X2%*%beta + u

  B_hat_i1 = solve(t(X1)%*%X1)%*%t(X1)%*%y1
  B_hat_i2 = solve(t(X2)%*%X2)%*%t(X2)%*%y2
  b = rbind(B_hat_i1, B_hat_i2)
  return(b)
}

for (i in 1:n_size) {
  for (x in 1:3){
    g_n_matrix1[x,i] = FunctionMonteCarlo(n_vec[i])[x]
    g_n_matrix2[x,i] = FunctionMonteCarlo(n_vec[i]+n_size)[x]
  }
}

hist(g_n_matrix1[2, 1:n_size], main = "B_hat_1 at correlation = 0.0")
axis(side=1, at=seq(0.6, 1.4,by=0.1), seq(0.6, 1.4,by=0.1))
```




```
hist(g_n_matrix2[2, 1:n_size], main="B_hat_1 at correlation = 0.9")
```



Empirical Exercise

```
load(file = "C:/Users/eitas/OneDrive/Stalinis kompiuteris/Assignment_1_2023/data_Assignment1.Rdata")
```

b)

(i) Bias of OLS estimator

Effect: Larger correlation between regressors leads to more biased OLS estimators.

Explanation: Multicollinearity – correlation between regressors may cause biased OLS estimators since it is more difficult to distinguish the unique effects of the regressors on the dependent variable. As a result, the estimated coefficients might be altered and thus biased.

(ii) The Standard Error of the OLS estimator

Effect: Larger correlation between regressors leads to larger standard errors of OLS estimators.

Explanation: As explained above the correlation of the regressors (multicollinearity) may lead to altered estimators of the coefficients. Thus also the variance and standard error of the OLS estimator will be increased as the estimator gets less accurate.

(iii) The t-statistic

Effect: Larger correlation between regressors leads to a smaller t-statistic.

Explanation: The larger Standard error of the OLS estimator caused by a larger correlation between regressors also affects the t-statistic. More precisely it decreases the t-statistic. This implies that it might be more difficult to reject the null hypotheses (that the coefficient has no effect $\implies \beta = 0$) so the probability of Type II error increases.

Solution Empirical Exercise 1

(A1) $y = X\beta + \epsilon$ – linear model

Importance: The OLS model is based on the linear equation $y = ax+b$ so it assumes the relationship between regressors and the dependent variable is linear, otherwise the OLS estimators would not be relevant.

(A2*) $E[\epsilon X] = 0$ – zero mean assumption

Importance: The zero mean assumption is essential because it ensures that the the error terms multiplied by dependent variables are independent of each other when the sample size is large, i.e. goes to infinity, so that the OLS estimations are consistent.

(A6) $(Y_i, X_i) : i = 1, ..., n$ are i.i.d. – random sample

Importance: No correlation between the independent variables and the error terms. This assumption ensures there is no other unobserved variable that has some influence on the dependent variable but is not included in the OLS model which would cause unbiased and inconsistent estimates of the OLS coefficients.

(A7) $E[X_i X_i^\top]$ is a finite positive definite matrix.

Importance: It is needed for $E[X_i X_i^\top]$ to be invertible.

Solution Empirical Exercise 2

```r
#Cretating variables
iota = rep(1, nrow(educ))
X1 = educ
X2 = cbind(iota, exper, IQ, age, black)
I_two = diag(nrow(X2))

#Building the projection and residual maker matrices of X2
P_X2 = X2%*%solve(t(X2)%*%X2)%*%t(X2)
M_X2 = I_two - P_X2

#Calculating Beta 1 hat
B1_hat = solve(t(X1)%*%M_X2%*%X1)%*%t(X1)%*%M_X2%*%wage
drop(B1_hat)
```

```
## [1] 52.70889
```

Interpretation: given that education is increased by 1 year, ceteris paribus, the monthly nominal earnings increase by approximately 52.71 units.

Solution Empirical Exercise 3

```r
#Creating X matrix
X = cbind(iota, educ,exper,IQ,age,black)

#Estimating the coefficients and residuals
B_hat = solve(t(X)%*%X)%*%t(X)%*%wage
u_hat = wage - X%*%B_hat

#Calculating the variance of the regression error
n = nrow(X)
k = ncol(X)
s2 = (t(u_hat)%*%u_hat)/(n-k)

drop(s2)
```

```
## [1] 134930.4
```

Hence, $s^2 = 134930.4$.

Solution Empirical Exercise 4

```r
#The estimate of the variance-covariance matrix of the OLS coefficient estimates
B_hat_Variance =  drop(s2)*solve(t(X)%*%X)
B_hat_Variance
```

```
##              iota
## iota 27369.2482249 -221.415841 -0.9287369 -68.0274506 -517.3229576
##       -221.4158405   52.796387 12.8037063  -3.2840166   -9.1943896
##         -0.9287369   12.803706 13.6894122  -0.1618056   -9.4714319
##        -68.0274506   -3.284017 -0.1618056   0.9957525    0.3480598
##       -517.3229576   -9.194390 -9.4714319   0.3480598   21.5910012
##      -1783.7907535   -7.493332 -0.8700814  13.7372591    9.3037981
##
## iota -1783.7907535
##         -7.4933317
##         -0.8700814
##         13.7372591
##          9.3037981
##       1524.5335765
```

```r
#The estimated standard error of B3_hat
B3_hat_SEE = sqrt(B_hat_Variance[4,4])
drop(B3_hat_SEE)
```

```
## [1] 0.997874
```

The standard error of any coefficient measures the accuracy of the estimated coefficient. In terms of sampling distribution of the OLS estimator, the standard error represents the standard deviation of the sampling distribution of that coefficient - which is approximately the standard deviation of the normal distribution under the OLS Assumptions.

Solution Empirical Exercise 5

Assume that assumptions A1-A5 hold. $H_0 : \beta_3 \leq 0$ vs. $H_1 : \beta_3 > 0$

```r
tval=qt(0.95,n-k)
tstat=(B_hat[4,]-0)/B3_hat_SEE
tstat
```

```
##
## 4.127889
```

```
tstat>=tval
```

```
##
## TRUE
```

The test statistic is $t_{stat} = \frac{\hat{\beta_3} - 0}{SSE_{\hat{\beta_3}}} = 4.13$ Since $t_{stat} = 4.13 > 1.65 = t_{0.95,929}$, we reject the null hypothesis.

Solution Empirical Exercise 6

Assume assumptions A1-A5 hold. $H_0 : R\beta \neq r$ vs. $H_1 : R\beta = r$, where $R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$ and $r = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$.

```
R = rbind(c(0, 1, 0, 0, 0, 0), c(0, 0, 1, 0, 0, 0))
r = rbind(10, 5)
B = rbind(10,5)
q = 2

F_stat = t(R%*%B_hat-r)%*%solve(drop(s2)*R%*%solve(t(X)%*%X)%*%t(R))%*%(R%*%B_hat-r)/q
F_val = qf(0.95, q, n-k)
F_val
```

```
## [1] 3.005413
```

```
F_stat > F_val
```

```
##      [,1]
## [1,] TRUE
```

The F test statistic is $F = (R\hat{\beta} - r)^\top (s^2 R(X^\top X)^{-1} R^\top)^{-1} (R\hat{\beta} - r)/q$

Since $F = 18.1 > 3.0 = F_{0.95,2,929}$, we reject the null hypothesis.