

## Daniel Losada and Sergio Quintanilla

```
library(carData) #to load dataset
```

```
## Warning: package 'carData' was built under R version 4.4.2
```

```
library(dplyr)
library(MASS)
```

### Exercise 1

a)

```
model = lm(prestige~ income+education, Duncan)
n = nrow(Duncan)
k = 2 #2 or 3? not 100% sure
y_hat = model$fitted.values
y_bar = mean(Duncan$prestige)
E_hat = model$residuals

F_stat_func = function(y_hat,y_bar,E_hat,n,k){
  numerator = sum(((y_hat - y_bar)^2)/k)
  denominator = sum((E_hat^2)/(n-k-1))
  F_statistic = numerator/denominator
  F_statistic
}
F_stat_func(y_hat,y_bar,E_hat,n,k)
```

```
## [1] 101.2162
```

b) Bootstrap Pairs

```
N = 1000 #number of times to repeat the resampling and recalc shit

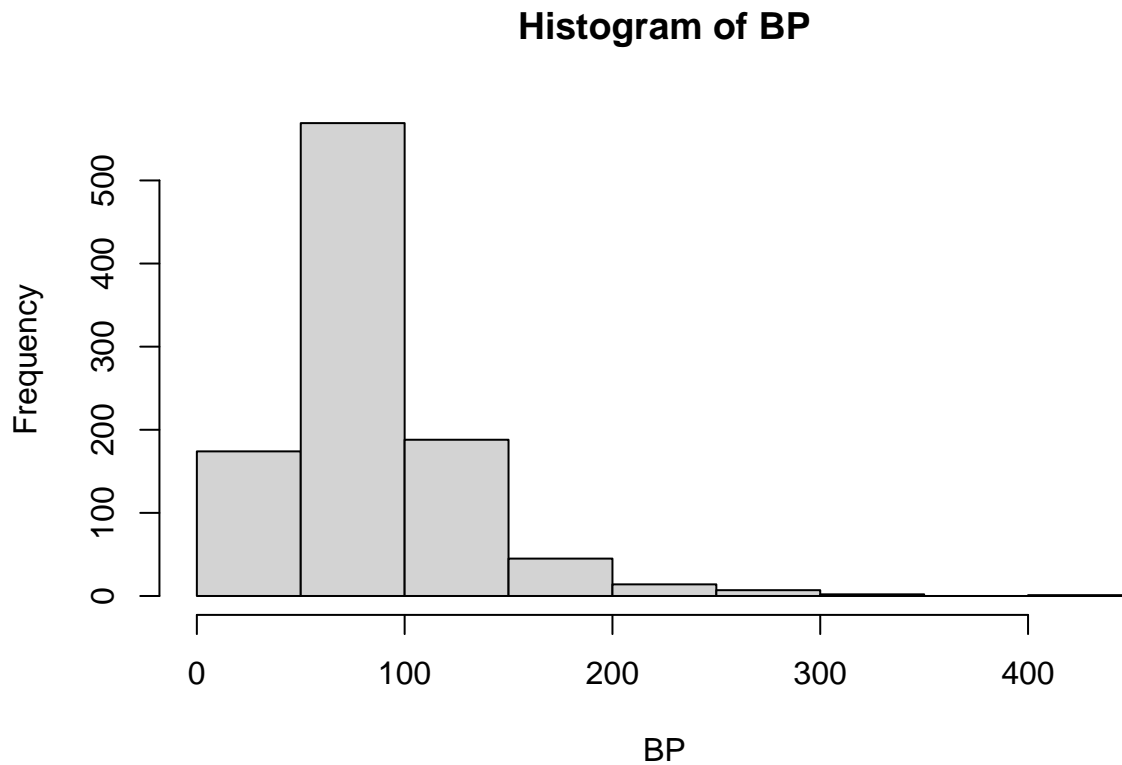
B_pairs = function(N){
  out = vector(length=N)
  for (i in 1:N){
    samp = sample_n(Duncan,30,replace=T)
    model = lm(prestige~ income+education, samp)
    n = nrow(samp)
    k = 2
    y_hat = model$fitted.values
    y_bar = mean(samp$prestige)
    E_hat = model$residuals
    out[i] = F_stat_func(y_hat,y_bar,E_hat,n,k)
  }
  out
}
```

```

}

BP = B_pairs(N)
hist(BP)

```



```

sd(BP) / sqrt(length(BP)) #standard error

```

```
## [1] 1.381
```

### c) Bootstrap Residuals

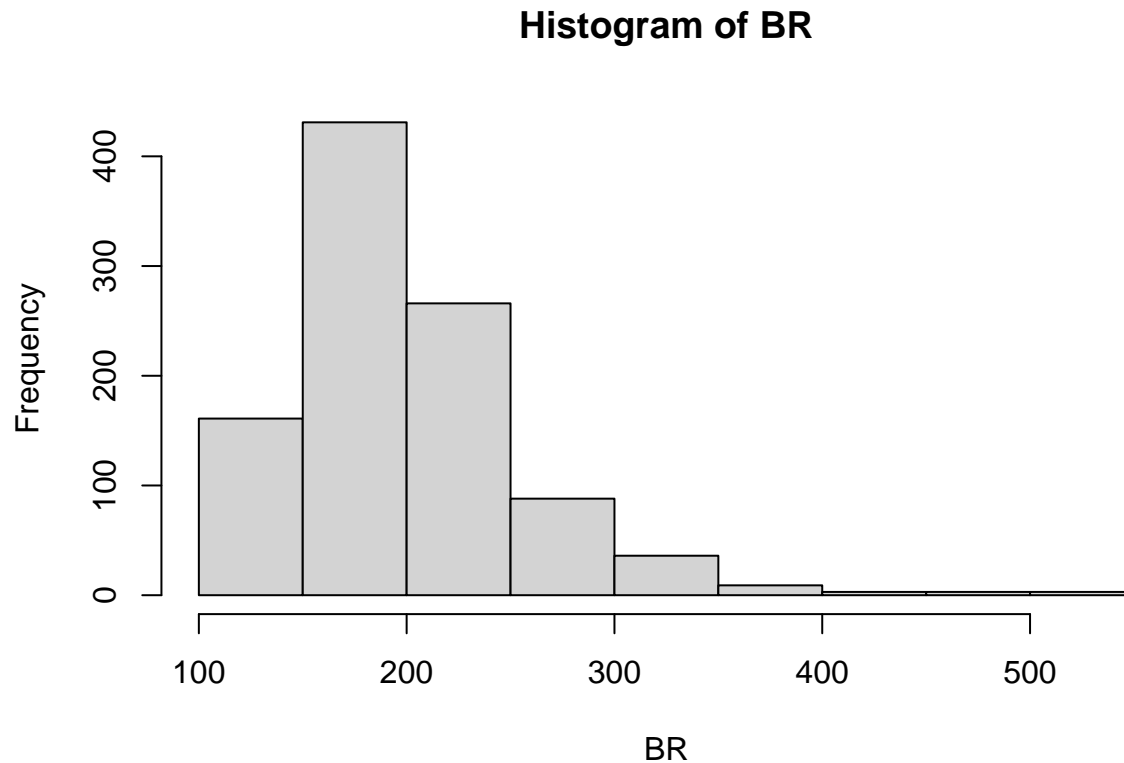
```

N = 1000 #number of times to repeat the resampling and recalc

B_residuals = function(N){
  out = vector(length=N)
  for (i in 1:N){
    E_samp = sample(E_hat,30,replace=T)
    Y_hat_star = y_hat + E_samp
    out[i] = F_stat_func(Y_hat_star,y_bar,E_samp,n,k)
  }
  out
}

```

```
BR = B_residuals(N)
hist(BR)
```



```
sd(BR) / sqrt(length(BR)) #standard error
```

```
## [1] 1.78349
```

#### d) RLM Standard

```
model = rlm(prestige~ income+education, Duncan)
summary(model)
```

```
##
## Call: rlm(formula = prestige ~ income + education, data = Duncan)
## Residuals:
##      Min       1Q   Median       3Q      Max
## -30.120  -6.889   1.291   4.592  38.603
##
## Coefficients:
##              Value Std. Error t value
## (Intercept) -7.1107   3.8813  -1.8320
## income       0.7014   0.1087   6.4516
```

```
## education    0.4854  0.0893    5.4380
##
## Residual standard error: 9.892 on 42 degrees of freedom
```

```
n = nrow(Duncan)
k = 2 #2 or 3? not 100% sure
y_hat = model$fitted.values
y_bar = mean(Duncan$prestige)
E_hat = model$residuals
F_stat_func(y_hat,y_bar,E_hat,n,k)
```

```
## [1] 104.1872
```

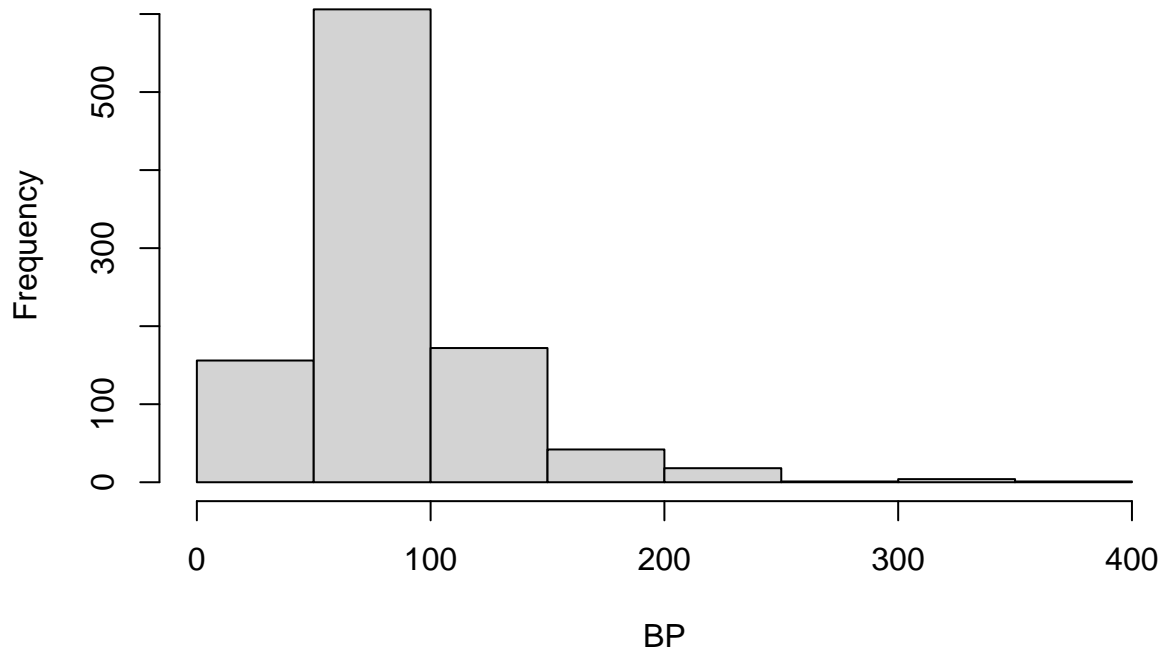
### RLM with Bootstrap Pairs

```
N = 1000 #number of times to repeat the resampling and recalc shit
```

```
B_pairs_rlm = function(N){
  out = vector(length=N)
  for (i in 1:N){
    samp = sample_n(Duncan,30,replace=T)
    model = rlm(prestige~ income+education, samp)
    n = nrow(samp)
    k = 2
    y_hat = model$fitted.values
    y_bar = mean(samp$prestige)
    E_hat = model$residuals
    out[i] = F_stat_func(y_hat,y_bar,E_hat,n,k)
  }
  out
}
```

```
BP = B_pairs_rlm(N)
hist(BP)
```

## Histogram of BP



```
sd(BP) / sqrt(length(BP)) #standard error
```

```
## [1] 1.299264
```

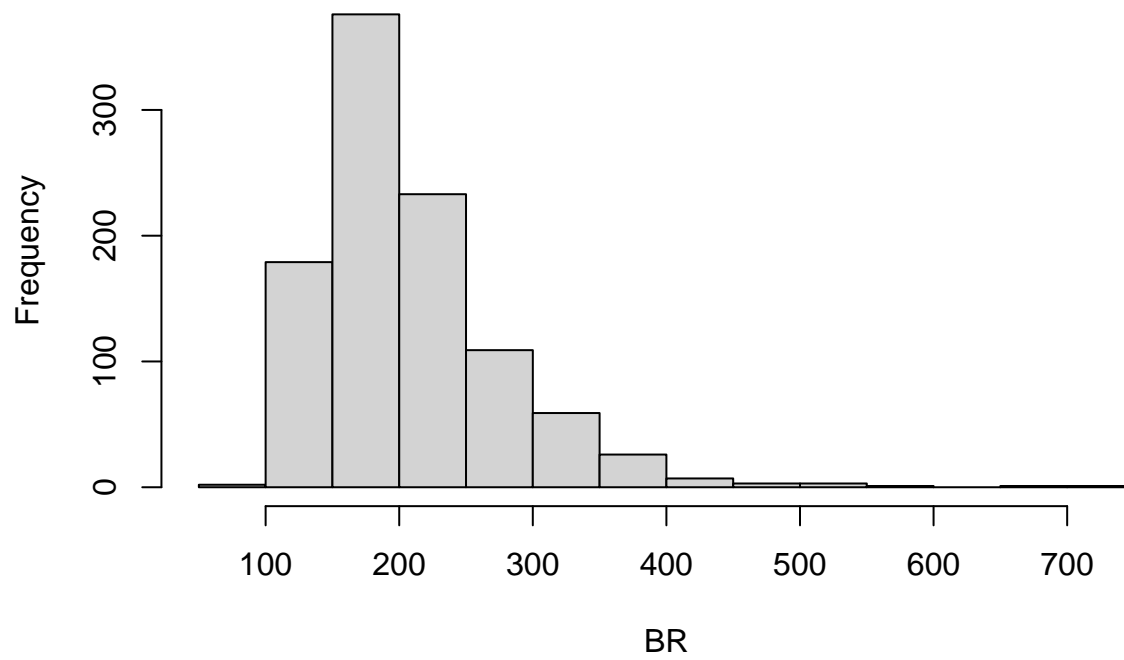
## RLM with Bootstrap Residuals

```
N = 1000 #number of times to repeat the resampling and recalculations
```

```
B_residuals = function(N){  
  out = vector(length=N)  
  for (i in 1:N){  
    E_samp = sample(E_hat,30,replace=T)  
    Y_hat_star = y_hat + E_samp  
    out[i] = F_stat_func(Y_hat_star,y_bar,E_samp,n,k)  
  }  
  out  
}
```

```
BR = B_residuals(N)  
hist(BR)
```

**Histogram of BR**



```
sd(BR) / sqrt(length(BR)) #standard error
```

```
## [1] 2.240174
```