Assignment

1. Question 1

a) Part a)

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 1000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = boot_obj, conf = 0.9, type = "perc")
##
## Intervals :
## Level      Percentile
## 90%    ( 4.713,  4.955 )
## Calculations and Intervals on Original Scale
```

b) Part b)

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 1000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = boot_obj, conf = 0.95, type = "perc")
##
## Intervals :
## Level      Percentile
## 95%    ( 4.690,  4.971 )
## Calculations and Intervals on Original Scale
```

c) Part c)

```
## [1] 6.757152e-30
```

Result:

- Since P-value < 0.05, so reject null hypothesis.

2. Question 2

a) Part a)

(b) Resample the 2×2 contingency table B $= 1000$ times. (Hint: Use the multinomial distribution and rmultinom() in R.)

->

```
Relevant_true <- c(123,27)
Irrelevant_true <- c(625,6703)
total_relevant_doc <- 150
total_irrelevant_doc <- 7328
Relevant_true <- rmultinom(1000, total_relevant_doc, Relevant_true)
Irrelevant_true <- rmultinom(1000, total_irrelevant_doc, Irrelevant_true)
```

(c) Find 90% and 95% confidence intervals for the true F2 for the complete database using Efron's percentile method.

-> The 90% confidence interval for F2 of database is:

```
##         5%        95%
## 0.4258304 0.4839970
```