Assignment 5

Emre Beray Boztepe

26 01 2024

Question 1:

Let X_{ij}, j = 1, . . . , m, be iid random variables from the exponential distribution with the expected value $\mu_i = E\left(X_{ij}\right)$. Construct a test for testing the hypothesis: $H_{0i}: \mu_i = 3$ vs $H_{Ai}: \mu_i > 3$.

Let us construct the structure and make it ready for applying the likelihood ratio $H_{0i}: \mu_i = 3$ -> $\lambda_i = \frac{1}{3}$ and $H_{Ai}: \mu_i > 3$ -> $\lambda_i < \frac{1}{3}$

Let us apply likelihood ratio by considering $\lambda_1 < \lambda_2 \in \Theta, \lambda_i = \frac{1}{\mu_i}$

Formula for likelihood ratio:

$$L(X) = \frac{\prod_{j=1}^{m} f\left(x_j, \lambda_2\right)}{\prod_{j=1}^{m} f\left(x_j, \lambda_1\right)} = \frac{\prod_{j=1}^{m} \lambda_2 \, e^{-x_j \lambda_2}}{\prod_{j=1}^{m} \lambda_1 \, e^{-x_j \lambda_1}} = (\frac{\lambda_2}{\lambda_1})^m e^{-(\lambda_2 - \lambda_1)\sum_{j=1}^{m} x_j}$$

Let $T = \sum_{i=1}^{n} x_i$, then we can say that

$$L(X) = (\frac{\lambda_2}{\lambda_1})^n e^{-(\lambda_2 - \lambda_1)T}$$

is a non-decreasing function of the statistic -T.

We reject $H_{0,i}$ if -T < c, for T such that $\mathbb{E}_{\mu=3}[-T < c] = \alpha$.

We know that under $H_{0,i}$ $T \sim Gamma\left(n, \frac{1}{3}\right)$. So, $-c = F^{-1}_{Gamma\left(n, \frac{1}{3}\right)}(1 - \alpha)$.

Finally, we can say that we reject $H_{0,i}$ if $T > F^{-1}_{Gamma\left(n, \frac{1}{3}\right)}(1 - \alpha)$.

The p-value of the test:

$$p = \mathbb{P}_0(t > T) = 1 - F_{Gamma\left(n, \frac{1}{3}\right)}(T)$$

## Question 2:

For all combinations of parameters - $m \in [20{,}100]$, - $n \in [200{,}1000]$, - $\epsilon \in [0{,}01; 0{,}05; 0{,}1; 0{,}2]$ consider the following mixture model: μ_i are iid random variables from the two-point distribution

$P(\mu_i = 3) = 1 - \epsilon = 1 - P(\mu_i = 5{,}5), i = 1, \dots, n$ and for j = 1 . . . m, X_{ij} are iid $Exp(\mu_i)$.

For each i = 1, . . . , n and $q \in \left[0{,}1; 0{,}1\sqrt{\frac{200}{m}}\right]$ use test from the Problem 1 with the following multiple testing corrections

i) Bonferroni procedure at the FWER level q,
ii) procedure controlling Bayesian FDR at the level q (find the respective critical values by using the command uniroot()),
iii) classical BH procedure at the FDR level q.

Use at least 1000 replicates to estimate FDR, Power (see, the List 4) and the expected value of the total experiment cost under the assumptions

i) $c_0 = c_A = 1$ (note that in this case the expected cost is equal to the expected number of misclassifications),
ii) $c_0 = 2, c_A = 1,$
iii) $c_0 = 1, c_A = 2,$

here c_0 is the cost for the type I error and c_A is the cost for the type II error. For the Bonferroni procedure and the BFDR controlling procedure calculate the exact values of the Power and of the expected cost

Import library and necessary definitions

```r
library(knitr)

epsilons = c(0.01,  0.05,  0.1,  0.2)
ns = c(200,  1000)
ms = c(20, 100)
costs = list(c(c0 = 1,  cA = 1),
              c(c0 = 2,  cA = 1),
              c(c0 = 1,  cA = 2))

list_all_results = list()
i = 1
set.seed(2021)
```

Functions for applying needed procedures

```r
# function for applying bonferroni procedure
bonferroni_procedure = function(p_values, alpha){
  n = length(p_values)
  return(p_values <= (alpha / n))
}

# function for applying benjamini-hochberg procedure
benjamini_hochberg_procedure = function(p_values, alpha){
  n = length(p_values)
  ord = order(p_values)
  ord2 = order(ord)
  res = (p_values[ord] <= (alpha * seq(n) / n))
  sapply(1:n, function(i) any(res[i:n]))[ord2]
}
```

Functions for calculating Bayesian FDR

```r
# this F formula is for calculating Bayesian FDR threshold based on the
formula given
apply_F_formula = function(c, m, q, eps) {
  (1 -  eps) * (1 - pgamma(c,  m, 1/3)) / (1 - (( 1 - eps) * pgamma(c, m,
1/3) + eps * pgamma(c, m,  1/5.5))) - q
}

# calculating the Bayesian FDR threshold for given parameters
BFDR_threshold = function(eps, m, q=0.05) {
  mm = if (m < 50) 200 else 600
  uniroot(apply_F_formula, c(10e-10, mm), m, q, eps)$root
}

# function for given T and threshold to decide if we should reject H0 or not
should_reject_H0 = function(T, threshold) {
  return(T > threshold)
}
```

Functions to calculate FDR, Power and Cost

```r
# Function to calculate False Discovery Rate
FDR = function(true_values, test_results) {
  fdr = sum(test_results[which(!true_values)]) / max(sum(test_results), 1)
  if(is.na(fdr))
    return(0)
  else
    return(fdr)
}

# Function to calculate Power
power = function(true_values, test_results) {
  pow = mean(test_results[which(true_values)])
  if(is.na(pow))
    return(0)
  else
    return(pow)
}

# Function to calculate the expected value of the total experiment cost under
the assumptions
cost = function(true_values, test_results, c0,  cA) {
  cost_val = sum((true_values != test_results) * (c0 * true_values +  cA *
test_results))
  if(is.na(cost_val))
    return(0)
  else
    return(cost_val)
}
```

Functions for calculating p_value and mu

```r
# Function to calculate p_value
p_value = function(T_statistics, m) {
  return(1 - pgamma(T_statistics, m, 1/3))
}

# Function to calculate mu_value
mu_val = function(n, eps) {
  return(1 / sample(c(3, 5.5), n, replace=T, prob=c(1 - eps, eps)))
}
```

This function is where the simulation starts. Basically generates random data and gives it to the related functions to apply procedures and returns values based on these calculations

```r
simulate_tests = function(mu_vector, n, m, eps, q, bayesian_fdr_c) {

  data = array(replicate(m, rexp(n, mu_vector)), c(n,  m))
  T_statistics = rowSums(data)
  p_values = p_value(T_statistics, m)

  bonf = bonferroni_procedure(p_values, q)
  bfdr = should_reject_H0(T_statistics, bayesian_fdr_c)
  bh = benjamini_hochberg_procedure(p_values, q)

  array(c(bonferroni = bonf, BFDR = bfdr, benjamini_hochberg = bh),  c(n, 3))
}
```

Function to start simulation with given replicate count and parameters for n, m, epsilon etc.

```r
start_simulating = function(mu_vector, repn, n, m, eps, q=0.05) {
  bayesian_fdr_c = BFDR_threshold(eps, m)
  array(replicate(repn, simulate_tests(mu_vector, n, m, eps, q,
bayesian_fdr_c)),  c(n,  3, repn))
}
```

Function that computes performance metrics (such as power, false discovery rate, and cost) based on the results of a simulation study

```r
evaluate_results = function(fun, result, mu_vector, c0=NULL,  cA=NULL) {
  if (is.null(c0) ||  is.null(cA)) {
    return(colMeans(array(sapply(1:(dim(result)[2]), function(test)
sapply(1:(dim(result)[3]),
                 function(iter) fun(mu_vector < 1 / 3,
result[,  test,  iter]))), dim(result)[3:2])))
  }
  else {
    return(colMeans(array(sapply(1:(dim(result)[2]), function(test)
sapply(1:(dim(result)[3]),
                                          function(iter) fun(mu_vector < 1 / 3,
result[,  test,  iter],c0, cA))), dim(result)[3:2])))
  }
}
```

Based on the evaluation, collect the results to a data frame and return it. Calculate all necessary assumptions

```
collect_all_results = function(mu_vector, result, costs) {
  power_values = evaluate_results(power, result, mu_vector)

  fdr_values = evaluate_results(FDR, result, mu_vector)

  cost_values_1 = evaluate_results(cost, result, mu_vector, costs[[1]]["c0"],
costs[[1]]["cA"])

  cost_values_2 = evaluate_results(cost, result, mu_vector, costs[[2]]["c0"],
costs[[2]]["cA"])

  cost_values_3 = evaluate_results(cost, result, mu_vector, costs[[3]]["c0"],
costs[[3]]["cA"])

  return(t(data.frame(power_values, fdr_values, cost_values_1, cost_values_2,
cost_values_3)))
}
```

Loop all epsilons; q, m and n values and write results as tables

```
for (eps  in epsilons) {
  for (m in ms) {
    for (n in ns) {
      mu_vector = mu_val(n, eps)
      for (q in c(0.1, 0.1 * sqrt(200 / m))) {
        all_results = start_simulating(mu_vector, 1000, n, m, eps, q)
        results = data.frame(collect_all_results(mu_vector, all_results,
costs))
        colnames(results) = c("Bonferroni", "control BFDR", "Benjamini-
Hochberg")
        list_all_results[[i]] = list(eps=eps, m=m, n=n, res = results, q=q)
        i = i+1
        show(kable(results, caption = paste0("eps: ", eps, ", m: ",  m, ",
n: ",  n, ", q: ", round(q, 3)), digits = 3))
      }
    }
  }
}

##
##
```

```
## Table: eps: 0.01, m: 20,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.414|        0.324|              0.456|
## |fdr_values    |      0.054|        0.018|              0.083|
## |cost_values_1 |      1.259|        1.375|              1.278|
## |cost_values_2 |      2.430|        2.726|              2.365|
## |cost_values_3 |      1.347|        1.399|              1.469|
##
##
## Table: eps: 0.01, m: 20,  n: 200, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.505|        0.308|              0.600|
## |fdr_values    |      0.153|        0.028|              0.309|
## |cost_values_1 |      1.285|        1.423|              2.014|
## |cost_values_2 |      2.275|        2.808|              2.814|
## |cost_values_3 |      1.580|        1.461|              3.228|
##
##
## Table: eps: 0.01, m: 20,  n: 1000, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.278|        0.316|              0.438|
## |fdr_values    |      0.024|        0.035|              0.098|
## |cost_values_1 |     10.207|        9.744|              8.700|
## |cost_values_2 |     20.313|       19.313|             16.575|
## |cost_values_3 |     10.308|        9.919|              9.525|
##
##
## Table: eps: 0.01, m: 20,  n: 1000, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.369|        0.316|              0.617|
## |fdr_values    |      0.054|        0.035|              0.318|
## |cost_values_1 |      9.151|        9.759|             10.128|
## |cost_values_2 |     17.983|       19.342|             15.486|
## |cost_values_3 |      9.470|        9.935|             14.898|
##
##
## Table: eps: 0.01, m: 100,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.998|        0.998|              0.998|
## |fdr_values    |      0.047|        0.048|              0.099|
## |cost_values_1 |      0.098|        0.102|              0.237|
## |cost_values_2 |      0.100|        0.104|              0.239|
## |cost_values_3 |      0.194|        0.202|              0.472|
##
##
## Table: eps: 0.01, m: 100,  n: 200, q: 0.141
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.999|        0.998|              0.999|
## |fdr_values    |      0.072|        0.055|              0.144|
## |cost_values_1 |      0.152|        0.117|              0.370|
## |cost_values_2 |      0.153|        0.119|              0.371|
## |cost_values_3 |      0.303|        0.232|              0.739|
##
##
## Table: eps: 0.01, m: 100,  n: 1000, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.993|        0.997|              0.999|
## |fdr_values    |      0.007|        0.033|              0.098|
## |cost_values_1 |      0.217|        0.579|              1.765|
## |cost_values_2 |      0.325|        0.618|              1.780|
## |cost_values_3 |      0.326|        1.119|              3.515|
##
##
## Table: eps: 0.01, m: 100,  n: 1000, q: 0.141
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.993|        0.998|              0.999|
## |fdr_values    |      0.010|        0.033|              0.137|
## |cost_values_1 |      0.264|        0.575|              2.568|
## |cost_values_2 |      0.366|        0.612|              2.581|
## |cost_values_3 |      0.426|        1.113|              5.123|
##
##
## Table: eps: 0.05, m: 20,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.403|        0.499|              0.580|
## |fdr_values    |      0.021|        0.050|              0.094|
## |cost_values_1 |      5.467|        4.768|              4.470|
## |cost_values_2 |     10.840|        9.275|              8.254|
## |cost_values_3 |      5.561|        5.029|              5.156|
##
##
## Table: eps: 0.05, m: 20,  n: 200, q: 0.316
```

```
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.511|        0.496|              0.755|
## |fdr_values    |      0.058|        0.054|              0.305|
## |cost_values_1 |      4.720|        4.825|              5.872|
## |cost_values_2 |      9.122|        9.365|              8.074|
## |cost_values_3 |      5.038|        5.110|              9.542|
##
##
## Table: eps: 0.05, m: 20,  n: 1000, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.281|        0.504|              0.595|
## |fdr_values    |      0.007|        0.048|              0.095|
## |cost_values_1 |     36.789|       26.599|             24.004|
## |cost_values_2 |     73.479|       51.884|             44.667|
## |cost_values_3 |     36.888|       27.913|             27.345|
##
##
## Table: eps: 0.05, m: 20,  n: 1000, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.367|        0.500|              0.767|
## |fdr_values    |      0.015|        0.049|              0.303|
## |cost_values_1 |     32.607|       26.867|             29.591|
## |cost_values_2 |     64.913|       52.392|             41.453|
## |cost_values_3 |     32.908|       28.209|             47.320|
##
##
## Table: eps: 0.05, m: 100,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.998|        0.999|              1.000|
## |fdr_values    |      0.010|        0.048|              0.097|
## |cost_values_1 |      0.133|        0.558|              1.188|
## |cost_values_2 |      0.155|        0.567|              1.192|
## |cost_values_3 |      0.244|        1.107|              2.372|
##
##
## Table: eps: 0.05, m: 100,  n: 200, q: 0.141
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.998|        0.999|              1.000|
## |fdr_values    |      0.012|        0.046|              0.136|
## |cost_values_1 |      0.155|        0.534|              1.744|
## |cost_values_2 |      0.180|        0.543|              1.746|
## |cost_values_3 |      0.285|        1.059|              3.486|
##
##
## Table: eps: 0.05, m: 100,  n: 1000, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.993|        0.999|              1.000|
## |fdr_values    |      0.001|        0.037|              0.094|
## |cost_values_1 |      0.592|        2.690|              7.150|
## |cost_values_2 |      1.084|        2.738|              7.170|
## |cost_values_3 |      0.692|        5.332|             14.280|
##
##
## Table: eps: 0.05, m: 100,  n: 1000, q: 0.141
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.994|        0.999|              1.000|
## |fdr_values    |      0.002|        0.036|              0.133|
## |cost_values_1 |      0.531|        2.626|             10.566|
## |cost_values_2 |      0.927|        2.671|             10.573|
## |cost_values_3 |      0.666|        5.207|             21.125|
##
##
## Table: eps: 0.1, m: 20,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.397|        0.588|              0.667|
## |fdr_values    |      0.009|        0.049|              0.088|
## |cost_values_1 |     11.529|        8.428|              7.694|
## |cost_values_2 |     22.984|       16.252|             14.028|
## |cost_values_3 |     11.603|        9.032|              9.054|
##
##
## Table: eps: 0.1, m: 20,  n: 200, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.516|        0.599|              0.837|
## |fdr_values    |      0.026|        0.049|              0.283|
## |cost_values_1 |      9.478|        8.230|              9.966|
## |cost_values_2 |     18.678|       15.848|             13.067|
## |cost_values_3 |      9.756|        8.842|             16.831|
##
##
## Table: eps: 0.1, m: 20,  n: 1000, q: 0.1
##
```

```
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.278|        0.594|              0.693|
## |fdr_values    |      0.003|        0.042|              0.087|
## |cost_values_1 |     83.822|       50.133|             43.451|
## |cost_values_2 |    167.562|       97.234|             79.066|
## |cost_values_3 |     83.904|       53.165|             51.287|
## 
## 
## Table: eps: 0.1, m: 20,  n: 1000, q: 0.316
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.366|        0.596|              0.852|
## |fdr_values    |      0.006|        0.041|              0.280|
## |cost_values_1 |     73.811|       49.852|             56.138|
## |cost_values_2 |    147.364|       96.722|             73.286|
## |cost_values_3 |     74.069|       52.834|             95.128|
## 
## 
## Table: eps: 0.1, m: 100,  n: 200, q: 0.1
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.997|        1.000|              1.000|
## |fdr_values    |      0.004|        0.057|              0.092|
## |cost_values_1 |      0.126|        1.093|              1.828|
## |cost_values_2 |      0.171|        1.099|              1.832|
## |cost_values_3 |      0.207|        2.180|              3.652|
## 
## 
## Table: eps: 0.1, m: 100,  n: 200, q: 0.141
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.998|        1.000|              1.000|
## |fdr_values    |      0.007|        0.056|              0.131|
## |cost_values_1 |      0.155|        1.070|              2.753|
## |cost_values_2 |      0.189|        1.072|              2.755|
## |cost_values_3 |      0.276|        2.138|              5.504|
## 
## 
## Table: eps: 0.1, m: 100,  n: 1000, q: 0.1
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.993|        1.000|              1.000|
## |fdr_values    |      0.001|        0.050|              0.092|
## |cost_values_1 |      0.827|        5.329|             10.097|
## |cost_values_2 |      1.554|        5.359|             10.113|
## |cost_values_3 |      0.927|       10.628|             20.178|
##
##
## Table: eps: 0.1, m: 100,  n: 1000, q: 0.141
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.994|        1.000|              1.000|
## |fdr_values    |      0.001|        0.051|              0.129|
## |cost_values_1 |      0.703|        5.361|             14.854|
## |cost_values_2 |      1.263|        5.388|             14.862|
## |cost_values_3 |      0.846|       10.695|             29.700|
##
##
## Table: eps: 0.2, m: 20,  n: 200, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.403|        0.695|              0.769|
## |fdr_values    |      0.003|        0.040|              0.074|
## |cost_values_1 |     26.914|       15.028|             13.264|
## |cost_values_2 |     53.768|       28.740|             23.655|
## |cost_values_3 |     26.974|       16.344|             16.137|
##
##
## Table: eps: 0.2, m: 20,  n: 200, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.514|        0.699|              0.908|
## |fdr_values    |      0.011|        0.043|              0.247|
## |cost_values_1 |     22.118|       14.980|             17.941|
## |cost_values_2 |     43.984|       28.524|             22.072|
## |cost_values_3 |     22.370|       16.416|             31.751|
##
##
## Table: eps: 0.2, m: 20,  n: 1000, q: 0.1
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.278|        0.696|              0.773|
## |fdr_values    |      0.001|        0.043|              0.076|
## |cost_values_1 |    166.103|       77.013|             67.047|
## |cost_values_2 |    332.126|      146.873|            119.340|
## |cost_values_3 |    166.183|       84.166|             81.801|
##
##
## Table: eps: 0.2, m: 20,  n: 1000, q: 0.316
##
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.365|        0.696|              0.908|
## |fdr_values    |      0.003|        0.042|              0.244|
## |cost_values_1 |    146.232|       77.022|             88.942|
## |cost_values_2 |    292.231|      147.024|            110.208|
## |cost_values_3 |    146.465|       84.042|            156.618|
## 
## 
## Table: eps: 0.2, m: 100,  n: 200, q: 0.1
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.997|        1.000|              1.000|
## |fdr_values    |      0.003|        0.052|              0.081|
## |cost_values_1 |      0.210|        2.149|              3.425|
## |cost_values_2 |      0.318|        2.156|              3.428|
## |cost_values_3 |      0.312|        4.291|              6.847|
## 
## 
## Table: eps: 0.2, m: 100,  n: 200, q: 0.141
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.998|        1.000|              1.000|
## |fdr_values    |      0.003|        0.052|              0.115|
## |cost_values_1 |      0.166|        2.121|              5.078|
## |cost_values_2 |      0.234|        2.125|              5.079|
## |cost_values_3 |      0.264|        4.238|             10.155|
## 
## 
## Table: eps: 0.2, m: 100,  n: 1000, q: 0.1
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.993|        1.000|              1.000|
## |fdr_values    |      0.000|        0.047|              0.080|
## |cost_values_1 |      1.643|       10.395|             18.348|
## |cost_values_2 |      3.199|       10.424|             18.363|
## |cost_values_3 |      1.730|       20.761|             36.681|
## 
## 
## Table: eps: 0.2, m: 100,  n: 1000, q: 0.141
## 
## |              | Bonferroni| control BFDR| Benjamini-Hochberg|
## |:-------------|----------:|------------:|------------------:|
## |power_values  |      0.994|        1.000|              1.000|
## |fdr_values    |      0.001|        0.047|              0.111|
## |cost_values_1 |      1.344|       10.398|             26.437|
## |cost_values_2 |      2.580|       10.425|             26.448|
## |cost_values_3 |      1.452|       20.769|             52.863|
```

Comments:

When we look into Table 5 and 6, when eps = 0.01, m = 100, n = 200, q = 0.1, we got power very close to 1 for all tests. When we change q into 0.141, we can see that the power results increases but not significantly there is a difference. But, that increase in q leads FDR and cost values to be higher. Also, for all tests, when m is 20, power values are so low no matter what the other values are. Increase in epsilon when m = 20 causes in cost values to be highest.

As far as I see, Best-performed ones: -Table 5, -Table 6, Only difference between these two tables is q = 0.1 for Table 5 and q = 0.141 for Table 6. As a difference, cost and FDR decreased for Bonferroni but increased for both BFDR and Benjamini-Hochberg.

-Table 13, -Table 14, Only difference between these two tables is q = 0.1 for Table 13 and q = 0.141 for Table 14. As a difference, cost and FDR increased for Bonferroni and Benjamini-Hochberg but decreased for BFDR.

-Table 21, -Table 22, Only difference between these two tables is q = 0.1 for Table 21 and q = 0.141 for Table 22. As a difference, cost and FDR increased for Bonferroni and Benjamini-Hochberg but decreased for BFDR.

-Table 29, -Table 30. Only difference between these two tables is q = 0.1 for Table 29 and q = 0.141 for Table 30. As a difference, cost and FDR decreased for Bonferroni and BFDR but increased for Benjamini-Hochberg .

Question 3:

For each of the above values of ε and combinations of c0 and cA derive the optimal Bayesian classifier and theoretically calculate its BFDR, Power and the expected value of the corresponding cost function.

$$\frac{f(T|H_A)}{f(T|H_0)} = (\frac{3}{5.5})^m e^{T\frac{5}{33}} \geq \frac{c_0\mathbb{P}(H_0)}{c_A\mathbb{P}(H_A)} = \frac{c_0(1-\epsilon)}{c_A\epsilon}$$

$$T \geq \frac{33}{5} ln\left((\frac{5.5}{3})^m \cdot \frac{c_0(1-\epsilon)}{c_A\epsilon}\right) = \tau$$

$$BFDR = \frac{\left(1 - F_{Gamma\left(m,\frac{1}{3}\right)}(\tau)\right)\cdot(1-\epsilon)}{\left(1 - F_{Gamma\left(m,\frac{1}{3}\right)}(\tau)\right)\cdot(1-\epsilon) + \left(1 - F_{Gamma\left(m,\frac{1}{5.5}\right)}(\tau)\right)\cdot\epsilon}$$

$$power = \mathbb{P}(rejected|H_1) = 1 - F_{Gamma\left(m,\frac{1}{5.5}\right)}(\tau)$$

Import library and necessary definitions

```
library(knitr)
```

```
epsilons = c(0.01,  0.05,  0.1,  0.2)
ms = c(20, 100)
costs = list(c(c0 = 1,  cA = 1),
             c(c0 = 2,  cA = 1),
             c(c0 = 1,  cA = 2))

set.seed(2021)
```

Function for calculating Tau value by the formula

```
calculate_tau = function(m, eps, c0, cA) {
  33 / 5 * log( (5.5 / 3) ^ m * (c0 * (1 - eps)) / (cA * eps))
}
```

Functions to calculate BFDR, Power, Cost by using calculated Tau value

```
# Function for calculating Bayesian FDR by its formula using tau
calculate_BFDR_tau = function(m, eps, c0, cA, tau_val) {
  return((1 - pgamma(tau_val, m, 1/3)) * (1 - eps) / ((1 - pgamma(tau_val, m,
1/3) * (1 - eps)) + ((1 - pgamma(tau_val, m, 1 / 5.5)) * eps)))
}

# Function for calculating power value by its formula using tau
calculate_power_tau = function(m, eps, c0, cA, tau_val) {
  1 - pgamma(tau_val,  m, 1 / 5.5)
}

# Function for calculating cost value by its formula using tau
calculate_cost_tau = function(m, eps, c0, cA, tau_val) {
  c0 * (1 - eps) * (1 - pgamma(tau_val, m, 1/3)) + cA * eps * (1 -
pgamma(tau_val, m, 1/5.5))
}
```

Data frame for storing all results (power, cost, BFDR using tau)

```
results = data.frame()
```

Loop all values and calculate BFDR, power, cost

```
for (eps  in epsilons) {
  for (m in ms) {
    for (cs in costs) {

      tau_val = calculate_tau(m, eps, cs["c0"], cs["cA"])
      results = rbind(results,
                      data.frame(eps = eps, m = m, c0 = cs["c0"], cA =
cs["cA"],
                                 tau = tau_val,
                                 BFDR = calculate_BFDR_tau(m, eps, cs["c0"],
cs["cA"], tau_val),
                                 power = calculate_power_tau(m, eps, cs["c0"],
cs["cA"], tau_val),
                              cost = calculate_cost_tau(m, eps, cs["c0"],
cs["cA"], tau_val)))
    }
  }
}
```

Show table for the results

```
kable(results, row.names = F, digits = 3)
```

eps	m	c0	cA	tau	BFDR	power	cost
0.01	20	1	1	110.338	0.061	0.465	0.006
0.01	20	2	1	114.912	0.030	0.393	0.005
0.01	20	1	2	105.763	0.117	0.540	0.013
0.01	100	1	1	430.377	0.003	0.990	0.010
0.01	100	2	1	434.952	0.002	0.987	0.010
0.01	100	1	2	425.803	0.004	0.992	0.020
0.05	20	1	1	99.443	0.061	0.644	0.037
0.05	20	2	1	104.018	0.032	0.569	0.034
0.05	20	1	2	94.868	0.109	0.716	0.082
0.05	100	1	1	419.483	0.002	0.995	0.050
0.05	100	2	1	424.058	0.001	0.993	0.050
0.05	100	1	2	414.908	0.003	0.996	0.100
0.10	20	1	1	94.512	0.057	0.721	0.083
0.10	20	2	1	99.086	0.031	0.650	0.076
0.10	20	1	2	89.937	0.100	0.787	0.177
0.10	100	1	1	414.551	0.001	0.996	0.100
0.10	100	2	1	419.126	0.001	0.995	0.100
0.10	100	1	2	409.977	0.002	0.997	0.200
0.20	20	1	1	89.159	0.052	0.797	0.179
0.20	20	2	1	93.734	0.029	0.733	0.167
0.20	20	1	2	84.585	0.088	0.853	0.377
0.20	100	1	1	409.199	0.001	0.997	0.200
0.20	100	2	1	413.774	0.001	0.996	0.200
0.20	100	1	2	404.624	0.001	0.998	0.400

Comments:

For eps = 0.01, m = 20, tau value is between 105-115 and got its best power value when c0 = 1 and cA = 2 by having 0.540. It can be said that when tau value decreases, power increases but in the meantime, FDR and cost become higher as well. Again, when m value is 100, power and tau value increases for all epsilon, c0 and cA values. Tau value is in 400s when m = 100 and that causes power value to become almost 1. When epsilon is 0.20 (m = 100), FDR value decreases to 0.001 and cost value decreases to 0.2-0.4 range.

Question 4:

Assume that you do not know ε and μ. For each replicate of the whole experiment estimate ε and μ using Expectation Maximization algorithm. Evaluate the accuracy of your estimators: draw histograms, calculate the bias, the variance and the mean squared error.

Likelihood for EM algorithm:

Z_i - latent indicators form binomial distribution ($\mathbb{P}(Z_i = 1) = \epsilon$)

$$L(T, Z|\mu, \epsilon) = L(T|Z, \mu, \epsilon) \cdot L(Z|\epsilon)$$

$$= \prod_{i=1}^{n} f_{Gamma\left(m,\frac{1}{\mu}\right)}(T)^{Z_i} \cdot \epsilon^{Z_i} \cdot f_{Gamma\left(m,\frac{1}{3}\right)}(T)^{1-Z_i} \cdot (1-\epsilon)^{1-Z_i}$$

$$L(T|Z, \mu, \epsilon) = \prod_{i=1}^{n} f_{Gamma\left(m,\frac{1}{\mu}\right)}(T)^{Z_i} \cdot f_{Gamma\left(m,\frac{1}{3}\right)}(T)^{1-Z_i}$$

$$= \prod_{i=1}^{n} [\frac{\left(\frac{1}{\mu}\right)^m}{\Gamma(m)} T_i^{m-1} e^{-\frac{1}{\mu}T_i}]^{Z_i} [\frac{\left(\frac{1}{3}\right)^m}{\Gamma(m)} T_i^{m-1} e^{-\frac{1}{3}T_i}]^{1-Z_i}$$

$$l(T|Z, \mu, \epsilon) = \sum_{i=1}^{n} Z_i \cdot ln\left(f_{Gamma\left(m,\frac{1}{\mu}\right)}(T)\right) + (1 - Z_i) \cdot ln\left(f_{Gamma\left(m,\frac{1}{3}\right)}(T)\right)$$

$$l(T, Z|\mu, \epsilon)$$

$$= \sum_{i=1}^{n} Z_i \cdot \left(ln\left(f_{Gamma\left(m,\frac{1}{\mu}\right)}(T_i)\right) + ln(\epsilon)\right) + (1 - Z_i) \cdot \left(ln\left(f_{Gamma\left(m,\frac{1}{3}\right)}(T_i)\right) + ln(1-\epsilon)\right)$$

Expectation step:

Replace ϵ with estimator

$$\pi_i^k = \mathbb{E}(Z_i|T, \mu_k) = \mathbb{P}(Z_i = 1|T, \mu_k)$$

$$\pi_i^k = \frac{f_{Gamma\left(m,\frac{1}{\mu_k}\right)}(T_i) \cdot \epsilon_k}{f_{Gamma\left(m,\frac{1}{\mu_k}\right)}(T_i) \cdot \epsilon_k + f_{Gamma\left(m,\frac{1}{3}\right)}(T_i) \cdot (1-\epsilon_k)}$$

$$Q(\mu, \epsilon|\mu_k, \epsilon_k) = \mathbb{E}_{Z|T,\mu_k,\epsilon_k} log\left(L(T, Z|\mu, \epsilon)\right)$$

$$
\begin{aligned}
&Q(\mu,\epsilon|\mu_k,\epsilon_k)\\
&=\sum_{i=1}^{n}\pi_i^k\cdot\left(ln\left(f_{Gamma\left(m,\frac{1}{\mu}\right)}(T_i)\right)+ln(\epsilon)\right)+\left(1-\pi_i^k\right)\\
&\qquad\cdot\left(ln\left(f_{Gamma\left(m,\frac{1}{3}\right)}(T_i)\right)+ln(1-\epsilon)\right)
\end{aligned}
$$

Maximization step

$$
Q_{k+1}=argmax_{\mu,\epsilon}Q(\mu,\epsilon|\mu_k,\epsilon_k)
$$

Libraries

```
library(knitr)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Parameters

```
epsilon_values = c(0.01,  0.05,  0.1,  0.2)
ns = c(200,  1000)
ms = c(20, 100)
costs = list(c(c0 = 1,  cA = 1),
             c(c0 = 2,  cA = 1),
             c(c0 = 1,  cA = 2))
set.seed(2023)
```

Function to generate true values of mu

```
generate_mu = function(n, eps) {
  1 / sample(c(3, 5.5), n, replace=T, prob=c(1 - eps, eps))
}
```

Function for calculating the conditional probabilities pi_k

```
calculate_pi_k = function(T_val, eps_k, mu_k) {
  dgamma(T_val, m, 1 / mu_k) * eps_k / (dgamma(T_val, m, 1 / mu_k) * eps_k +
dgamma(T_val, m, 1 / 3) * (1 - eps_k))
}
```

Function for the Q step in the EM algorithm

```
calculate_Q = function(Q_init, pik, T_val, m) {
  mu_val = Q_init["mu"]; eps = Q_init["eps"]
  sum((pik) * (log(dgamma(T_val, m, 1/mu_val)) + log(eps)) + (1 - pik) *
(log(dgamma(T_val, m, 1/3)) + log(1 - eps)))
}
```

Function for the maximization step in the EM algorithm

```
maximize_Q = function(T_val, eps_k, mu_k, m) {
  pik = calculate_pi_k(T_val, eps_k, mu_k)
  optim(c(mu = 5.5, eps = .5), calculate_Q, pik = pik, m=m, T_val = T_val,
control = list(fnscale = -1))$par
}
```

EM algorithm function

```
expectation_maximization = function(T_val, m, Q_init=NULL, i=0) {
  if (is.null(Q_init)) {
    Q_val = maximize_Q(T_val, .5, mean(T_val), m)
    expectation_maximization(T_val, m, Q_val, i+1)
  } else {
    Q_val = maximize_Q(T_val, Q_init["eps"], Q_init["mu"], m)
    if (all(Q_val - Q_init < 0.05) | (i > 10e4)) {
      c(Q_val["mu"], Q_val["eps"])
    } else {
      expectation_maximization(T_val, m, Q_val, i+1)
    }
  }
}
```

Main function for one experiment

```
perform_experiment = function(eps, m, n)  {
  mu_vec = generate_mu(n, eps)
  T_val = rowSums(array(replicate(m, rexp(n, mu_vec)), c(n,  m)))
  expectation_maximization(T_val, m)
}
```

List to store results

```
results_list = list()
experiment_count = 1
```

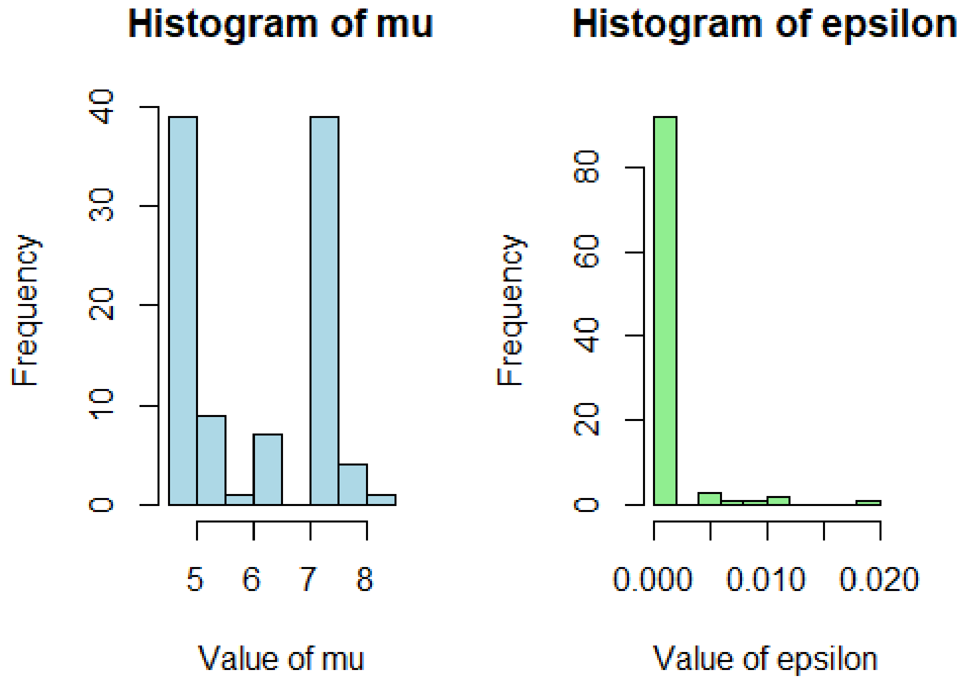
Loop through different combinations of parameters

```
for (eps  in epsilon_values) {
  for (m in ms) {
    for (n in ns) {
      results_list[[experiment_count]] = list(eps = eps, m = m, n = n, res =
replicate(100, perform_experiment(eps, m, n)))
      experiment_count = experiment_count+1
    }
  }
}
```

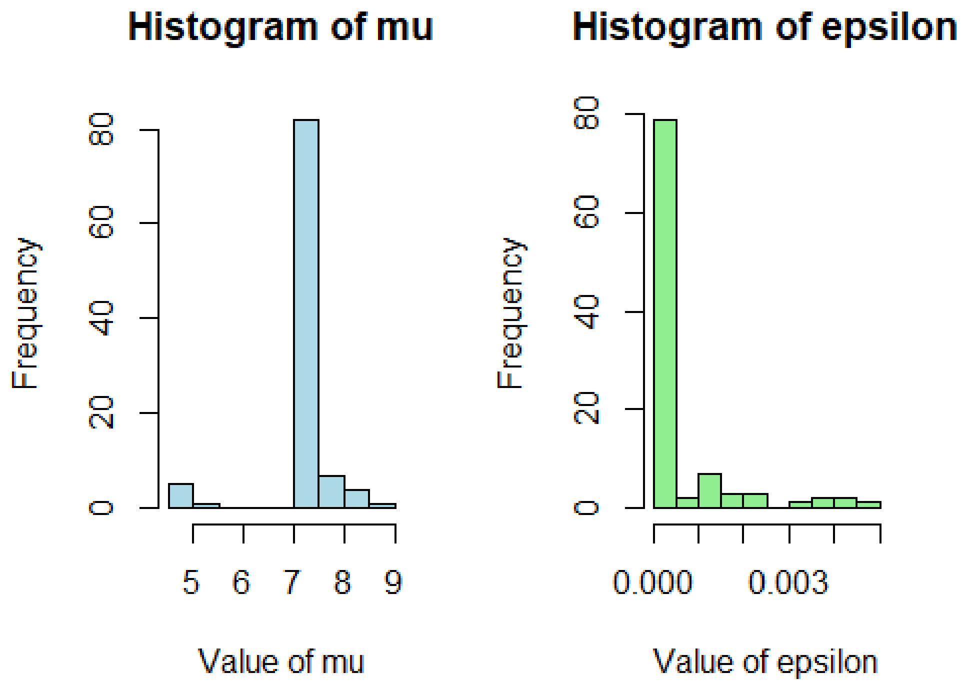
Loop through the results and display information

```
for (element in results_list) {
  bias = rowMeans(element$res - t(array(c(rep(5.5, 100), rep(element$eps,
100)), c(100, 2))))
  mse = c(mean((element$res["mu", ] - 5.5)^2), mean((element$res["eps", ] -
element$eps)^2))
  vars = c(var(element$res["mu", ]), var(element$res["eps", ]))

  # Calculate accuracy
  accuracy_mu = 1 - mean(abs(element$res["mu", ] - 5.5) > 0.5)
  accuracy_eps = 1 - mean(abs(element$res["eps", ] - element$eps) > 0.1)

  kable(t(data.frame(bias, mse, vars, accuracy_mu, accuracy_eps)),
        caption=paste0("epsilon: ", element$eps, ", m: ", element$m, ", n: ",
element$n),
        digits = 4) %>% show()

  # Plot histograms
  par(mfrow=c(1,2))
  hist(element$res["mu", ], main="Histogram of mu", xlab="Value of mu",
col="lightblue", border="black")
  hist(element$res["eps", ], main="Histogram of epsilon", xlab="Value of
epsilon", col="lightgreen", border="black")
  par(mfrow=c(1,1))
}
```

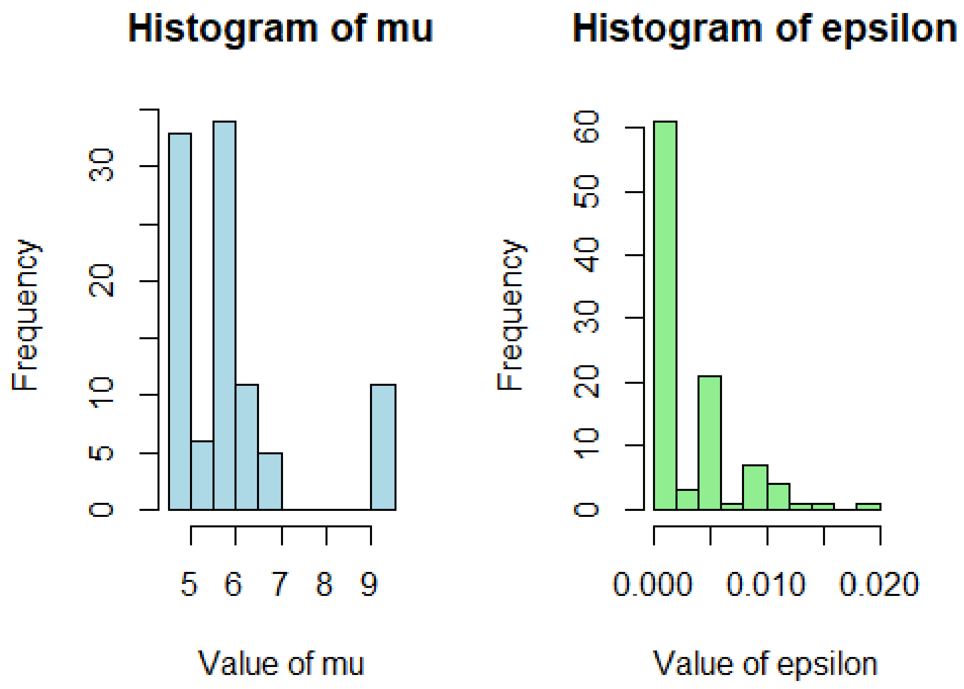
```
##
##
## Table: epsilon: 0.01, m: 20, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.5771| -0.0093|
## |mse          | 1.7002|  0.0001|
## |vars         | 1.3810|  0.0000|
## |accuracy_mu  | 0.1000|  0.1000|
## |accuracy_eps | 1.0000|  1.0000|
```



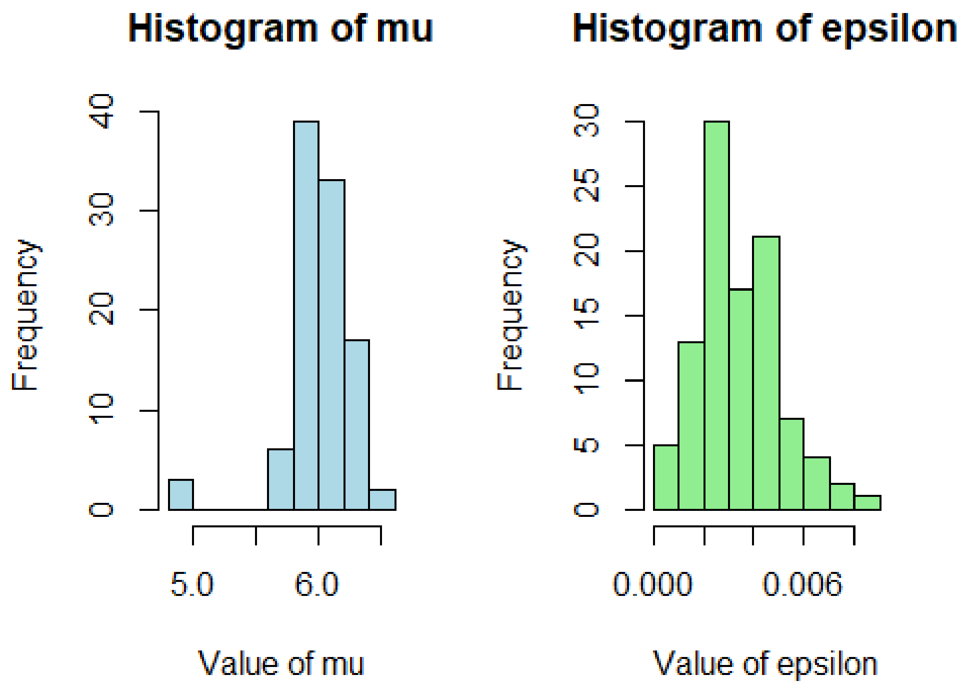
```
##
##
## Table: epsilon: 0.01, m: 20, n: 1000
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.7021| -0.0095|
## |mse          | 3.3166|  0.0001|
## |vars         | 0.4236|  0.0000|
## |accuracy_mu  | 0.0100|  0.0100|
## |accuracy_eps | 1.0000|  1.0000|
```



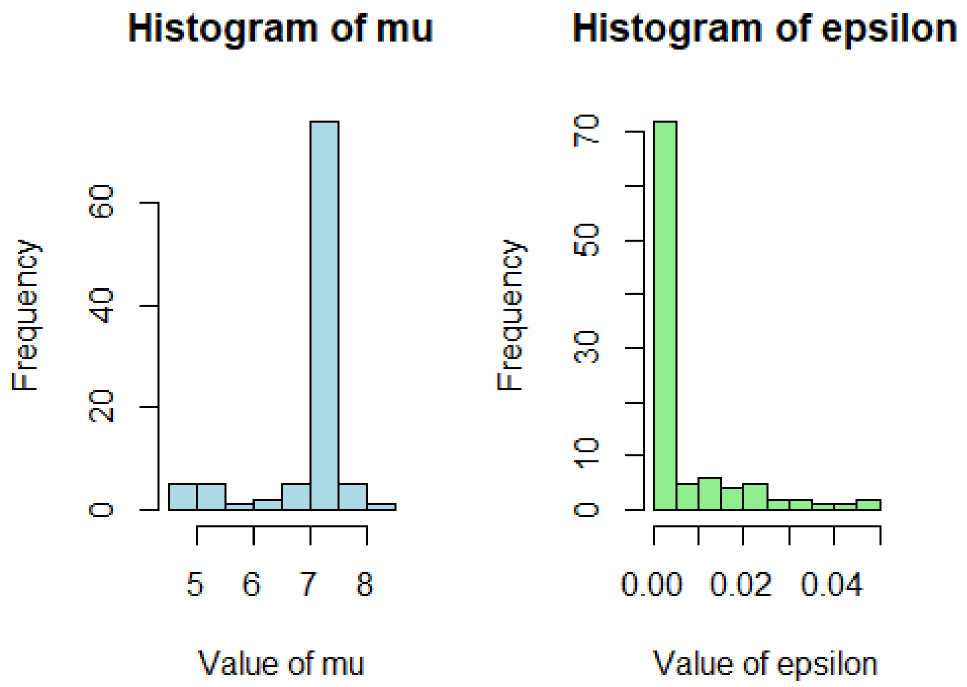
```
##
##
## Table: epsilon: 0.01, m: 100, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.3762| -0.0072|
## |mse          | 1.6921|  0.0001|
## |vars         | 1.5663|  0.0000|
## |accuracy_mu  | 0.4000|  0.4000|
## |accuracy_eps | 1.0000|  1.0000|
```



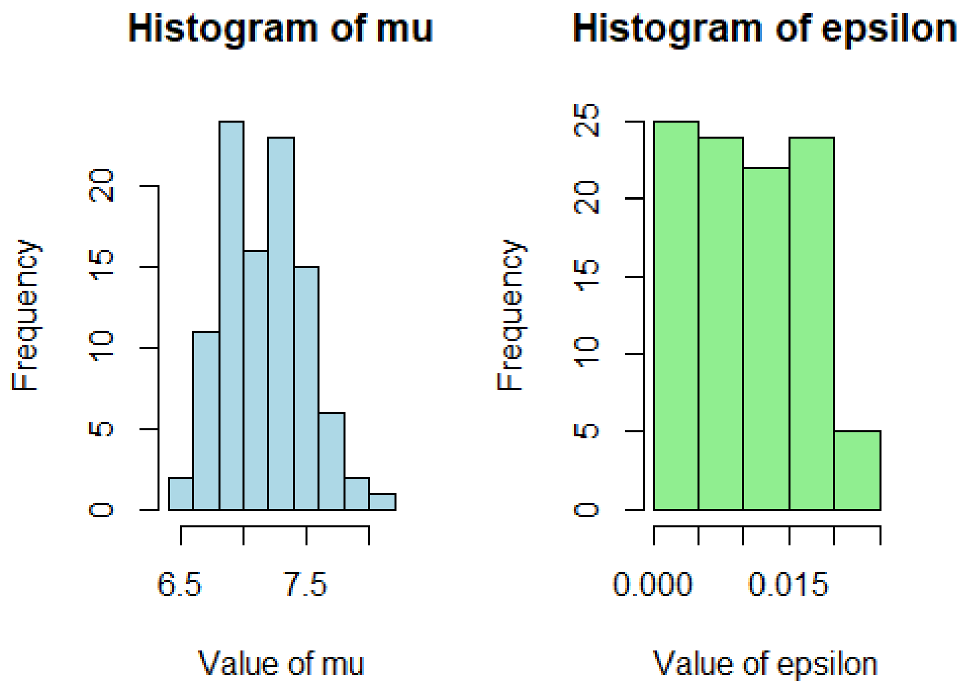
```
##
##
## Table: epsilon: 0.01, m: 100, n: 1000
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.5052| -0.0066|
## |mse          | 0.3260|  0.0000|
## |vars         | 0.0715|  0.0000|
## |accuracy_mu  | 0.4500|  0.4500|
## |accuracy_eps | 1.0000|  1.0000|
```



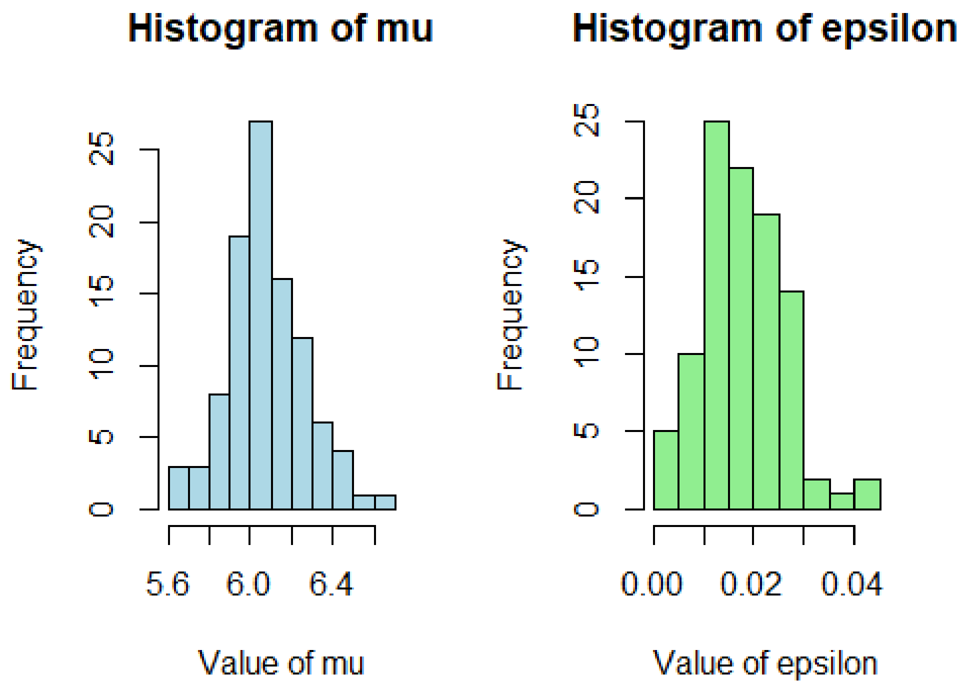
```
##
##
## Table: epsilon: 0.05, m: 20, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.4906| -0.0440|
## |mse          | 2.6957|  0.0021|
## |vars         | 0.4785|  0.0001|
## |accuracy_mu  | 0.0600|  0.0600|
## |accuracy_eps | 1.0000|  1.0000|
```



```
##
##
## Table: epsilon: 0.05, m: 20, n: 1000
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.6547| -0.0399|
## |mse          | 2.8405|  0.0016|
## |vars         | 0.1034|  0.0000|
## |accuracy_mu  | 0.0000|  0.0000|
## |accuracy_eps | 1.0000|  1.0000|
```



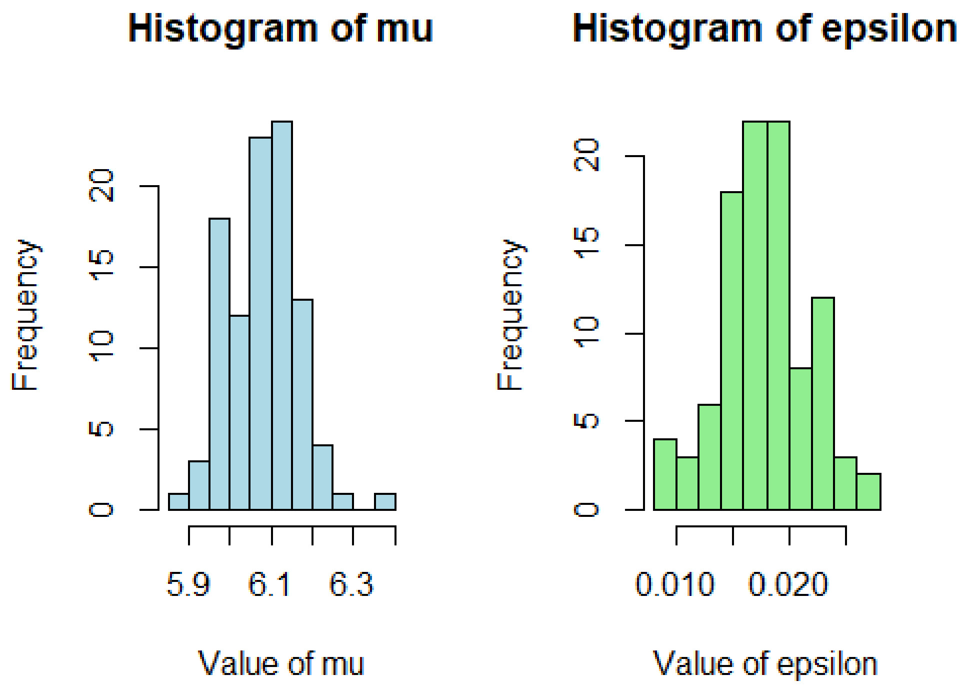
```
##
##
## Table: epsilon: 0.05, m: 100, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.5816| -0.0323|
## |mse          | 0.3731|  0.0011|
## |vars         | 0.0352|  0.0001|
## |accuracy_mu  | 0.3300|  0.3300|
## |accuracy_eps | 1.0000|  1.0000|
```



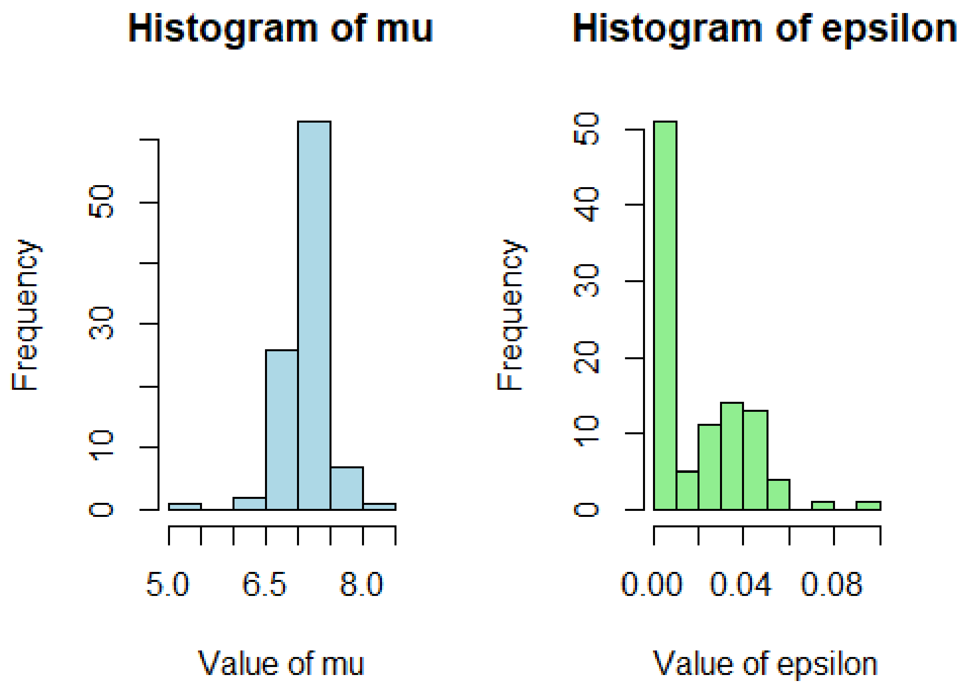
Histogram of mu

Histogram of epsilon

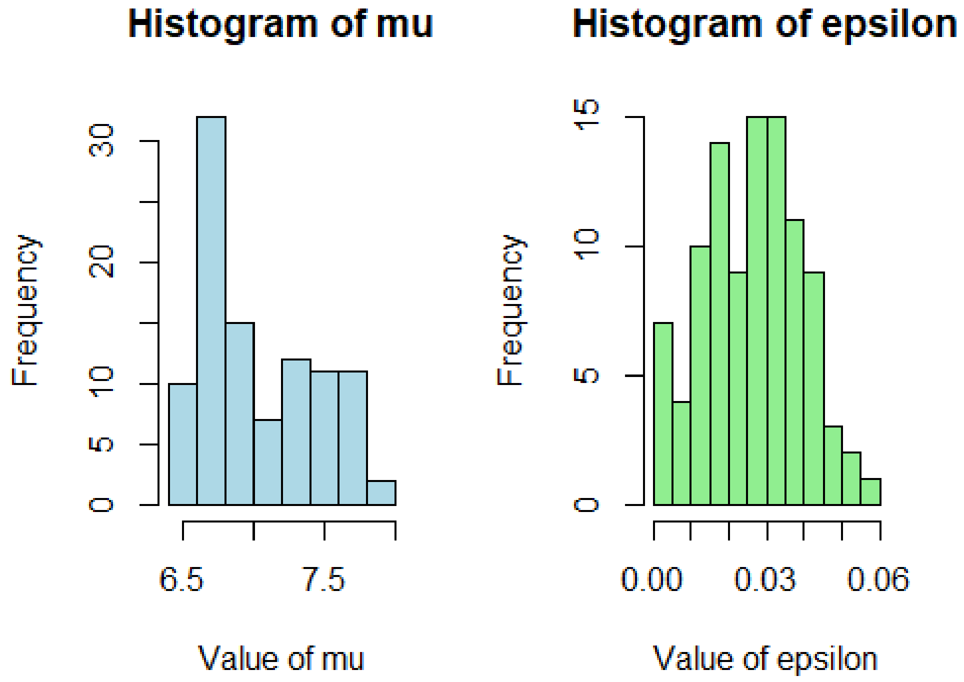
```
##
##
## Table: epsilon: 0.05, m: 100, n: 1000
##
## |             |     mu|    eps|
## |:------------|------:|------:|
## |bias         | 0.5810| -0.032|
## |mse          | 0.3443|  0.001|
## |vars         | 0.0068|  0.000|
## |accuracy_mu  | 0.2200|  0.220|
## |accuracy_eps | 1.0000|  1.000|
```



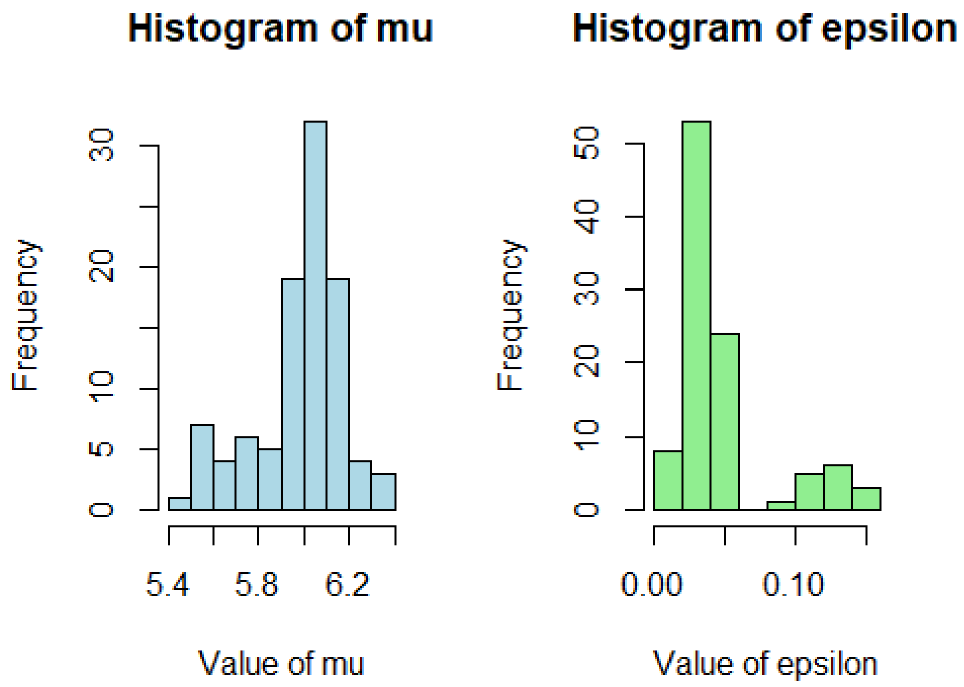
```
##
##
## Table: epsilon: 0.1, m: 20, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.6342| -0.0818|
## |mse          | 2.7932|  0.0071|
## |vars         | 0.1239|  0.0004|
## |accuracy_mu  | 0.0100|  0.0100|
## |accuracy_eps | 1.0000|  1.0000|
```



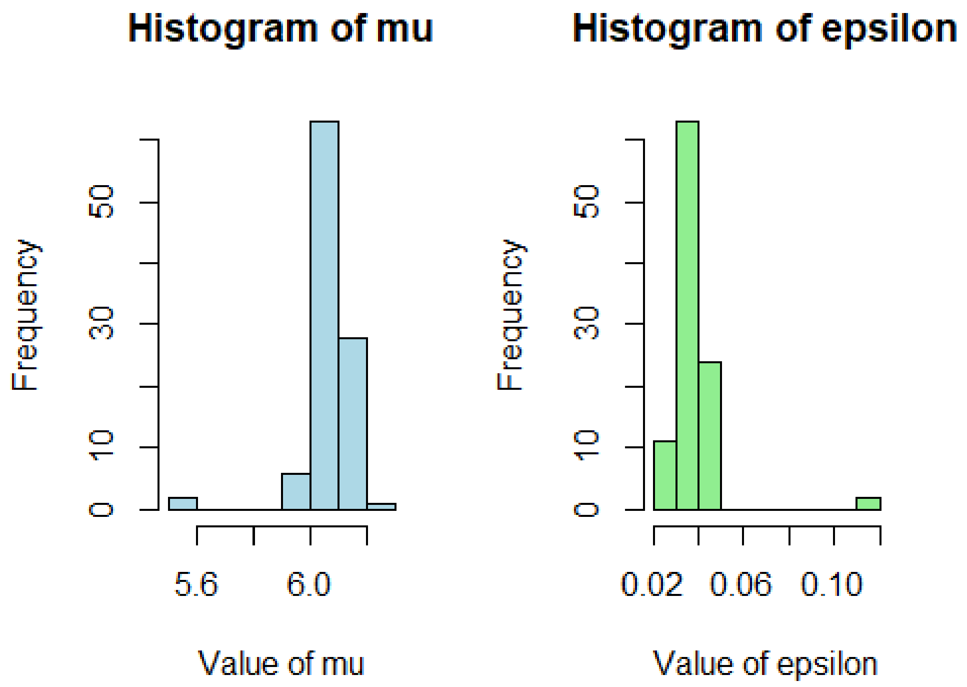
```
##
##
## Table: epsilon: 0.1, m: 20, n: 1000
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.5356| -0.0740|
## |mse          | 2.5156|  0.0057|
## |vars         | 0.1592|  0.0002|
## |accuracy_mu  | 0.0000|  0.0000|
## |accuracy_eps | 1.0000|  1.0000|
```



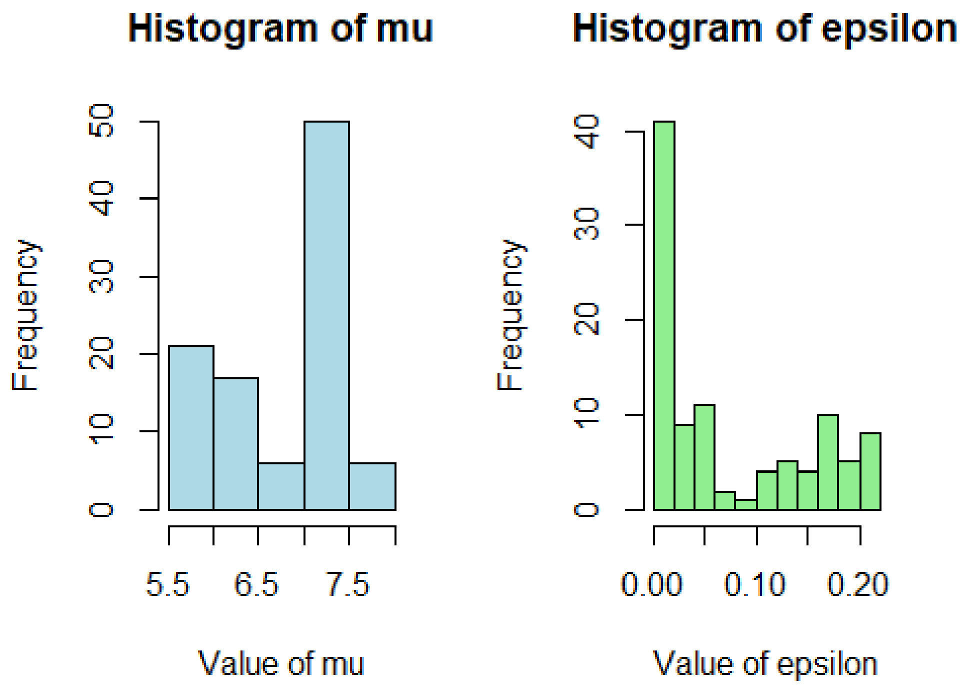
```
##
##
## Table: epsilon: 0.1, m: 100, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.4841| -0.0541|
## |mse          | 0.2724|  0.0041|
## |vars         | 0.0384|  0.0012|
## |accuracy_mu  | 0.4200|  0.4200|
## |accuracy_eps | 1.0000|  1.0000|
```



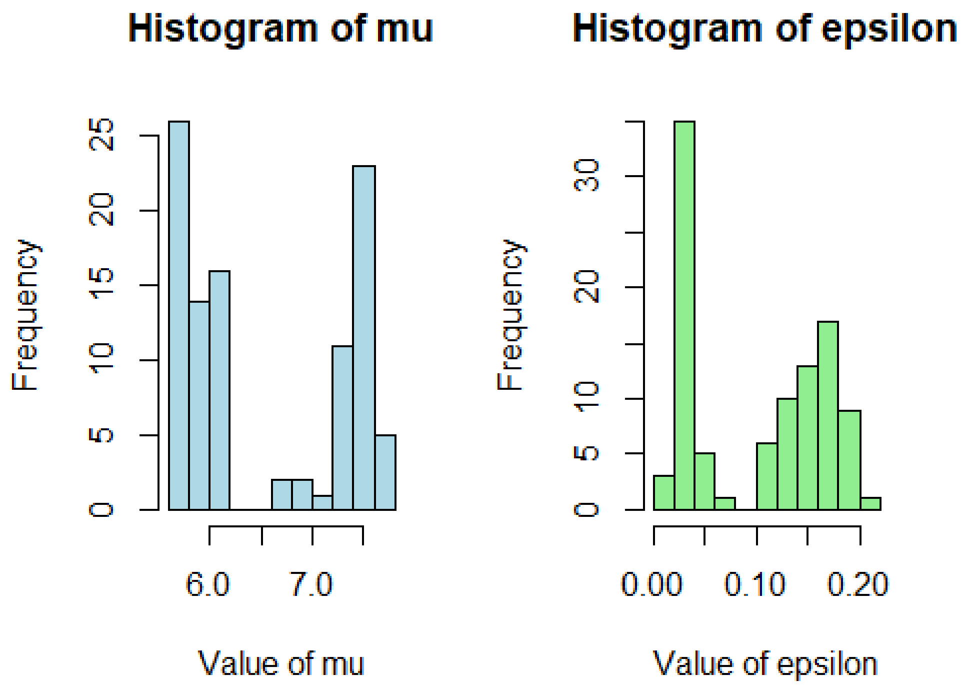
```
##
##
## Table: epsilon: 0.1, m: 100, n: 1000
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.5651| -0.0618|
## |mse          | 0.3269|  0.0040|
## |vars         | 0.0076|  0.0002|
## |accuracy_mu  | 0.0800|  0.0800|
## |accuracy_eps | 1.0000|  1.0000|
```



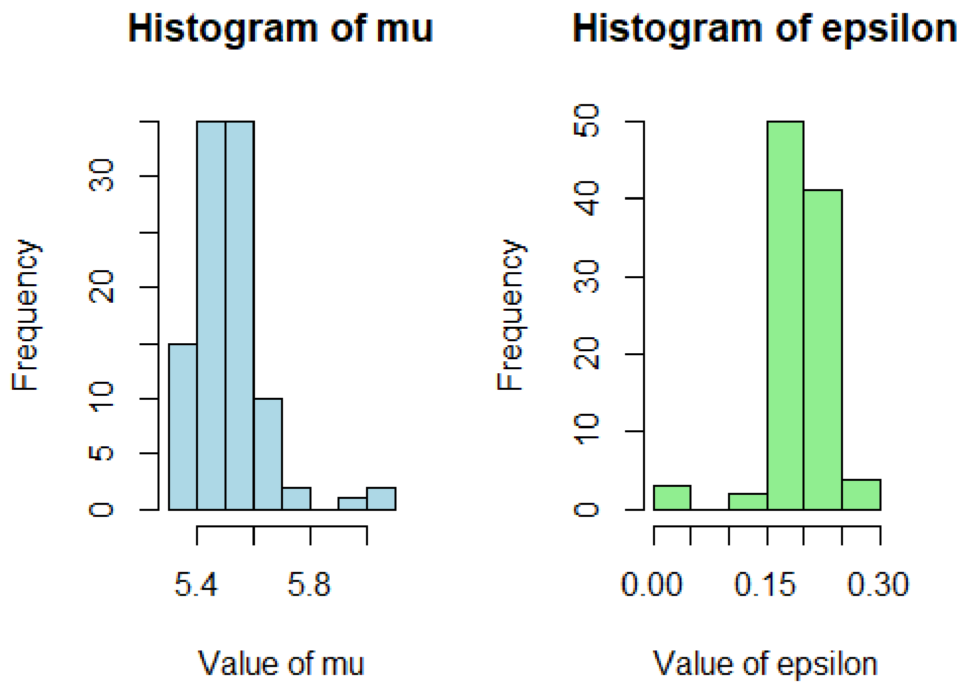
```
##
##
## Table: epsilon: 0.2, m: 20, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 1.2548| -0.1291|
## |mse          | 2.0068|  0.0226|
## |vars         | 0.4367|  0.0060|
## |accuracy_mu  | 0.2100|  0.2100|
## |accuracy_eps | 0.3600|  0.3600|
```



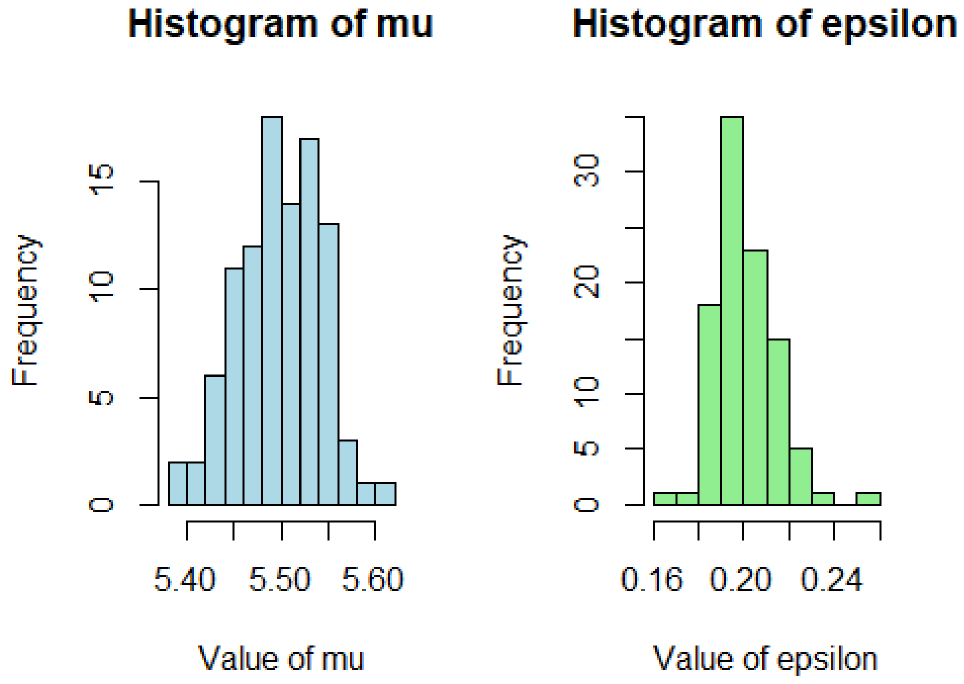
```
##
##
## Table: epsilon: 0.2, m: 20, n: 1000
##
## |              |     mu|     eps|
## |:-------------|------:|-------:|
## |bias          | 1.0439| -0.0987|
## |mse           | 1.6932|  0.0138|
## |vars          | 0.6097|  0.0041|
## |accuracy_mu   | 0.4000|  0.4000|
## |accuracy_eps  | 0.5600|  0.5600|
```



```
##
##
## Table: epsilon: 0.2, m: 100, n: 200
##
## |             |     mu|     eps|
## |:------------|------:|-------:|
## |bias         | 0.0172| -0.0040|
## |mse          | 0.0154|  0.0014|
## |vars         | 0.0153|  0.0014|
## |accuracy_mu  | 0.9800|  0.9800|
## |accuracy_eps | 0.9700|  0.9700|
```



```
##
##
## Table: epsilon: 0.2, m: 100, n: 1000
##
## |              |      mu|   eps|
## |:-------------|-------:|-----:|
## |bias          | -0.0015| 5e-04|
## |mse           |  0.0018| 2e-04|
## |vars          |  0.0018| 2e-04|
## |accuracy_mu   |  1.0000| 1e+00|
## |accuracy_eps  |  1.0000| 1e+00|
```



Comments:

- The bias values for mu estimation are generally small, indicating that, on average, the estimated mu values are relatively close to the true values.
- MSE values for mu are generally moderate, suggesting some variability in the accuracy of mu estimation across different settings.
- The variance of the estimated mu values provides insights into the spread or variability of the estimates. Larger variance values indicate greater variability in the accuracy of mu estimation.
- The accuracy in mu estimation is generally low, ranging from 0% to 45%. This suggests that the algorithm may struggle to precisely estimate mu within a certain threshold, especially for larger n.
- The accuracy in epsilon estimation is consistently high, being 100% in most cases. This suggests that the algorithm performs well in estimating epsilon within a certain threshold.

Question 5:

Given the above estimates

i) for each of the above values of ε and combinations of c0 and cA construct a plug-in Bayesian classifier (based on estimated values of ε and μ) and estimate its Power, FDR and the expected value of the corresponding cost function;

ii) For each of the above values of ε and q construct the plug-in version of the rule controlling BFDR (based on estimated values of ε and μ) and the modified version of

BH (define $i_0 = max\{i: p(i) \leq \frac{iq}{(1-\epsilon)n}\}$ and reject all hypothesis with p-values smaller or equal then p(i0)). Estimate Power, FDR and the expected values of the above cost functions for these modifications of the BFDR controlling rule and of BH.

Define a function for modified Benjamini-Hochberg procedure

```
mod_benjamini_hochberg = function(pvals, alpha, eps) {
  n = length(pvals)
  ord = order(pvals)
  ord2 = order(ord)
  # Apply the modified Benjamini-Hochberg procedure
  res =( pvals[ord] <= ((alpha * seq(n) / n / (1 - eps))))
  sapply(1:n, function(i) any(res[i:n]))[ord2]
}
```

Define a function for simulation

```
simulation = function(mu_vec, n, m, eps, q, c_bfdr, tau_val_1, tau_val_2,
tau_val_3) {

  # Generate data for each replication
  X = array(replicate(m, rexp(n, mu_vec)), c(n,  m))
  T_stat = rowSums(X)
  p_vals = p_value(T_stat, m)

  # Decide whether to reject H0 for different procedures
  bayes1 =should_reject_H0(T_stat, tau_val_1)
  bayes2 =should_reject_H0(T_stat, tau_val_2)
  bayes3 =should_reject_H0(T_stat, tau_val_3)
  bfdr = should_reject_H0(T_stat, c_bfdr)
  bh = mod_benjamini_hochberg(p_vals, q, eps)

  # Combine the results into an array
  array(c(bayes1 = bayes1, bayes2 = bayes2, bayes3 = bayes3, BFDR = bfdr,
mod_benjamini_hochberg = bh),  c(n, 5))
}
```

Define a function for conducting the simulation

```
start_simulation = function(mu_vec, repn, n, m, eps, q=0.05) {
  # Calculate the BFDR threshold and tau values
  c_bfdr = BFDR_threshold(eps, m)
  tau_val_1 = calculate_tau(m, eps, 1, 1)
  tau_val_2 = calculate_tau(m, eps, 2, 1)
  tau_val_3 = calculate_tau(m, eps, 1, 2)

  # Run the simulation for multiple replications
  array(replicate(repn, simulation(mu_vec, n, m, eps, q, c_bfdr, tau_val_1,
tau_val_2, tau_val_3)),  c(n,  5, repn)) # [i, test, replication]
}
```

Initialize a list to store all results

```
all_res_plugin = list()
i = 1
```

Loop through different parameter settings

```
for (element in results_list) {

  # Extract parameters
  m = element$m
  n = element$n
  mu_A = mean(element$res["mu", ])
  eps =  mean(element$res["eps", ])
  mu_vec = 1 / sample(c(3, mu_A), n, replace=T, prob=c(1 - eps, eps))

  # Loop through different q values
  for (q in c(0.1, 0.1 * sqrt(200 / m))) {
    # Run the simulation
    all_results = start_simulation(mu_vec, 1000, n, m, eps, q)
    results = data.frame(collect_all_results(mu_vec, all_results, costs))
    colnames(results) = c("Bayes c0=cA=1", "Bayes c0=2, cA=1", "Bayes c0=1,
cA=2", "plug-in control BFDR", "modified Benjamini-Hochberg")
    all_res_plugin[[i]] = list(eps=eps, m=m, n=n, res=results, q=q)
    i = i+1
    show(kable(results, caption = paste0("eps: ", eps, ", m: ",  m, ",  n: ",
n, ", q: ", round(q, 3)), digits = 3))
  }
}

##
##
## Table: eps: 0.000726457946131968, m: 20,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.375|            0.312|            0.437|
0.230|                       0.581|
## |fdr_values    |         0.005|            0.002|            0.020|
0.001|                       0.099|
## |cost_values_1 |         0.631|            0.691|            0.589|
0.771|                       0.605|
## |cost_values_2 |         1.256|            1.379|            1.152|
1.541|                       1.024|
## |cost_values_3 |         0.637|            0.694|            0.615|
0.772|                       0.791|
##
##
## Table: eps: 0.000726457946131968, m: 20,  n: 200, q: 0.316
```

```
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.391|            0.336|            0.456|
0.271|                       0.728|
## |fdr_values    |         0.009|            0.003|            0.018|
0.001|                       0.332|
## |cost_values_1 |         0.620|            0.669|            0.568|
0.731|                       1.342|
## |cost_values_2 |         1.229|            1.333|            1.112|
1.460|                       1.614|
## |cost_values_3 |         0.631|            0.674|            0.592|
0.733|                       2.412|
##
##
## Table: eps: 0.000466555288642138, m: 20,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.641|            0.576|            0.698|
0.486|                       0.740|
## |fdr_values    |         0.019|            0.010|            0.038|
0.003|                       0.113|
## |cost_values_1 |         0.387|            0.438|            0.362|
0.518|                       0.503|
## |cost_values_2 |         0.746|            0.862|            0.664|
1.032|                       0.763|
## |cost_values_3 |         0.415|            0.452|            0.422|
0.522|                       0.746|
##
##
## Table: eps: 0.000466555288642138, m: 20,  n: 1000, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.649|            0.594|            0.707|
0.499|                       0.847|
## |fdr_values    |         0.010|            0.004|            0.025|
0.001|                       0.302|
## |cost_values_1 |         0.366|            0.412|            0.330|
0.503|                       1.192|
## |cost_values_2 |         0.717|            0.818|            0.623|
1.004|                       1.345|
## |cost_values_3 |         0.381|            0.418|            0.367|
0.505|                       2.231|
##
##
## Table: eps: 0.00280155089846519, m: 100,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.997|            0.996|            0.997|
0.998|                       1.000|
## |fdr_values    |         0.001|            0.001|            0.002|
0.008|                       0.093|
## |cost_values_1 |         0.014|            0.016|            0.016|
0.037|                       0.420|
## |cost_values_2 |         0.024|            0.029|            0.025|
0.042|                       0.420|
## |cost_values_3 |         0.018|            0.019|            0.023|
0.069|                       0.840|
##
##
## Table: eps: 0.00280155089846519, m: 100,  n: 200, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.997|            0.995|            0.998|
0.999|                       1.000|
## |fdr_values    |         0.002|            0.001|            0.002|
0.008|                       0.147|
## |cost_values_1 |         0.014|            0.019|            0.015|
0.036|                       0.730|
## |cost_values_2 |         0.022|            0.033|            0.021|
0.039|                       0.731|
## |cost_values_3 |         0.020|            0.024|            0.024|
0.069|                       1.459|
##
##
## Table: eps: 0.00336976689876937, m: 100,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.999|            0.999|            0.999|
1.000|                       1.000|
## |fdr_values    |         0.008|            0.005|            0.012|
0.058|                       0.104|
## |cost_values_1 |         0.025|            0.017|            0.039|
0.183|                       0.365|
## |cost_values_2 |         0.027|            0.020|            0.041|
0.184|                       0.366|
## |cost_values_3 |         0.048|            0.031|            0.076|
0.365|                       0.729|
##
##
## Table: eps: 0.00336976689876937, m: 100,  n: 1000, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            0.999|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.006|            0.003|            0.013|
0.054|                       0.135|
## |cost_values_1 |         0.020|            0.012|            0.040|
0.167|                       0.487|
## |cost_values_2 |         0.021|            0.015|            0.041|
0.167|                       0.487|
## |cost_values_3 |         0.039|            0.021|            0.079|
0.334|                       0.974|
##
##
## Table: eps: 0.00600232773808412, m: 20,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.793|            0.747|            0.838|
0.677|                       0.846|
## |fdr_values    |         0.027|            0.014|            0.060|
0.005|                       0.096|
## |cost_values_1 |         0.717|            0.803|            0.704|
0.982|                       0.855|
## |cost_values_2 |         1.339|            1.561|            1.191|
1.950|                       1.318|
## |cost_values_3 |         0.812|            0.848|            0.921|
0.996|                       1.247|
##
##
## Table: eps: 0.00600232773808412, m: 20,  n: 200, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.804|            0.756|            0.846|
0.688|                       0.925|
## |fdr_values    |         0.035|            0.015|            0.069|
0.006|                       0.309|
## |cost_values_1 |         0.702|            0.777|            0.706|
0.950|                       2.067|
## |cost_values_2 |         1.290|            1.508|            1.167|
1.886|                       2.292|
## |cost_values_3 |         0.816|            0.823|            0.951|
0.964|                       3.909|
##
##
## Table: eps: 0.0101105547367041, m: 20,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.851|            0.807|            0.888|
0.752|                       0.833|
## |fdr_values    |         0.126|            0.065|            0.230|
0.031|                       0.102|
## |cost_values_1 |         2.013|            1.801|            2.860|
1.926|                       1.970|
## |cost_values_2 |         3.053|            3.153|            3.646|
3.663|                       3.140|
## |cost_values_3 |         2.986|            2.250|            4.934|
2.115|                       2.770|
##
##
## Table: eps: 0.0101105547367041, m: 20,  n: 1000, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.850|            0.806|            0.891|
0.753|                       0.908|
## |fdr_values    |         0.124|            0.063|            0.228|
0.026|                       0.314|
## |cost_values_1 |         1.998|            1.784|            2.812|
1.888|                       4.229|
## |cost_values_2 |         3.047|            3.142|            3.578|
3.620|                       4.876|
## |cost_values_3 |         2.947|            2.210|            4.858|
2.044|                       7.811|
##
##
## Table: eps: 0.0176864354053699, m: 100,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
```

```
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            1.000|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.003|            0.002|            0.005|
0.032|                       0.101|
## |cost_values_1 |         0.019|            0.014|            0.031|
0.197|                       0.688|
## |cost_values_2 |         0.021|            0.016|            0.033|
0.197|                       0.688|
## |cost_values_3 |         0.036|            0.026|            0.060|
0.394|                       1.376|
##
##
## Table: eps: 0.0176864354053699, m: 100,  n: 200, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.999|            0.999|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.003|            0.001|            0.005|
0.031|                       0.139|
## |cost_values_1 |         0.024|            0.014|            0.031|
0.190|                       1.008|
## |cost_values_2 |         0.027|            0.019|            0.032|
0.190|                       1.008|
## |cost_values_3 |         0.045|            0.023|            0.061|
0.380|                       2.016|
##
##
## Table: eps: 0.0179565834192206, m: 100,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.999|            0.999|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.004|            0.002|            0.006|
0.041|                       0.101|
## |cost_values_1 |         0.109|            0.073|            0.160|
0.981|                       2.604|
## |cost_values_2 |         0.124|            0.092|            0.171|
0.985|                       2.605|
## |cost_values_3 |         0.203|            0.127|            0.309|
1.958|                       5.207|
##
```

```
##
## Table: eps: 0.0179565834192206, m: 100,  n: 1000, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            0.999|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.004|            0.002|            0.006|
0.041|                       0.144|
## |cost_values_1 |         0.105|            0.068|            0.153|
0.978|                       3.896|
## |cost_values_2 |         0.112|            0.080|            0.159|
0.979|                       3.896|
## |cost_values_3 |         0.203|            0.124|            0.300|
1.955|                       7.792|
##
##
## Table: eps: 0.0182146052130434, m: 20,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.883|            0.845|            0.917|
0.802|                       0.888|
## |fdr_values    |         0.080|            0.041|            0.145|
0.018|                       0.102|
## |cost_values_1 |         0.850|            0.802|            1.096|
0.866|                       0.994|
## |cost_values_2 |         1.317|            1.422|            1.426|
1.658|                       1.443|
## |cost_values_3 |         1.233|            0.984|            1.862|
0.940|                       1.539|
##
##
## Table: eps: 0.0182146052130434, m: 20,  n: 200, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.875|            0.832|            0.908|
0.787|                       0.941|
## |fdr_values    |         0.074|            0.037|            0.140|
0.017|                       0.319|
## |cost_values_1 |         0.848|            0.835|            1.096|
0.920|                       2.732|
## |cost_values_2 |         1.346|            1.508|            1.464|
1.773|                       2.968|
## |cost_values_3 |         1.198|            0.997|            1.824|
0.987|                       5.228|
##
##
## Table: eps: 0.0259610558998947, m: 20,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.885|            0.847|            0.916|
0.803|                       0.873|
## |fdr_values    |         0.121|            0.064|            0.211|
0.031|                       0.104|
## |cost_values_1 |         5.304|            4.696|            7.430|
4.916|                       5.143|
## |cost_values_2 |         7.836|            8.070|            9.271|
9.245|                       7.931|
## |cost_values_3 |         8.076|            6.018|           13.019|
5.503|                       7.498|
##
##
## Table: eps: 0.0259610558998947, m: 20,  n: 1000, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.887|            0.849|            0.918|
0.806|                       0.940|
## |fdr_values    |         0.117|            0.062|            0.208|
0.032|                       0.317|
## |cost_values_1 |         5.181|            4.621|            7.303|
4.874|                      11.596|
## |cost_values_2 |         7.675|            7.952|            9.104|
9.132|                      12.915|
## |cost_values_3 |         7.868|            5.911|           12.805|
5.490|                      21.873|
##
##
## Table: eps: 0.0458856960735628, m: 100,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            0.999|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.002|            0.002|            0.003|
0.035|                      0.094|
## |cost_values_1 |         0.031|            0.030|            0.049|
0.471|                      1.356|
## |cost_values_2 |         0.037|            0.039|            0.052|
0.472|                      1.356|
## |cost_values_3 |         0.056|            0.051|            0.095|
0.941|                      2.712|
##
##
## Table: eps: 0.0458856960735628, m: 100,  n: 200, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            1.000|            1.000|
1.000|                      1.000|
## |fdr_values    |         0.002|            0.001|            0.004|
0.035|                      0.140|
## |cost_values_1 |         0.030|            0.022|            0.049|
0.465|                      2.143|
## |cost_values_2 |         0.034|            0.027|            0.052|
0.465|                      2.143|
## |cost_values_3 |         0.056|            0.039|            0.095|
0.930|                      4.286|
##
##
## Table: eps: 0.0381900944927197, m: 100,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         1.000|            1.000|            1.000|
1.000|                      1.000|
## |fdr_values    |         0.003|            0.002|            0.004|
0.039|                      0.101|
## |cost_values_1 |         0.158|            0.106|            0.225|
2.011|                      5.490|
## |cost_values_2 |         0.174|            0.129|            0.236|
2.013|                      5.490|
## |cost_values_3 |         0.300|            0.189|            0.439|
4.020|                     10.980|
##
##
## Table: eps: 0.0381900944927197, m: 100,  n: 1000, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
```

```
-------------:|---------------------------:|
## |power_values  |         1.000|            1.000|            1.000|
1.000|                       1.000|
## |fdr_values    |         0.003|            0.002|            0.004|
0.040|                       0.142|
## |cost_values_1 |         0.151|            0.099|            0.234|
2.027|                       8.150|
## |cost_values_2 |         0.166|            0.116|            0.247|
2.030|                       8.152|
## |cost_values_3 |         0.287|            0.181|            0.455|
4.051|                      16.298|
##
##
## Table: eps: 0.0709224310285799, m: 20,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.907|            0.873|            0.935|
0.842|                       0.905|
## |fdr_values    |         0.104|            0.059|            0.178|
0.034|                       0.106|
## |cost_values_1 |         2.863|            2.603|            3.912|
2.653|                       2.960|
## |cost_values_2 |         4.162|            4.386|            4.824|
4.867|                       4.293|
## |cost_values_3 |         4.427|            3.423|            6.912|
3.092|                       4.587|
##
##
## Table: eps: 0.0709224310285799, m: 20,  n: 200, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.908|            0.877|            0.936|
0.842|                       0.961|
## |fdr_values    |         0.097|            0.053|            0.170|
0.032|                       0.313|
## |cost_values_1 |         2.733|            2.455|            3.746|
2.626|                       7.359|
## |cost_values_2 |         4.014|            4.182|            4.636|
4.842|                       7.900|
## |cost_values_3 |         4.185|            3.183|            6.602|
3.036|                      14.177|
##
##
## Table: eps: 0.101253232025777, m: 20,  n: 1000, q: 0.1
```

```
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.904|            0.868|            0.933|
0.837|                       0.898|
## |fdr_values    |         0.109|            0.061|            0.186|
0.039|                       0.100|
## |cost_values_1 |        19.807|           18.030|           26.764|
18.775|                      19.329|
## |cost_values_2 |        28.972|           30.603|           33.154|
34.300|                      28.995|
## |cost_values_3 |        30.449|           23.487|           47.138|
22.025|                      28.992|
##
##
## Table: eps: 0.101253232025777, m: 20,  n: 1000, q: 0.316
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.908|            0.871|            0.937|
0.839|                       0.964|
## |fdr_values    |         0.109|            0.060|            0.184|
0.037|                       0.320|
## |cost_values_1 |        19.415|           17.644|           26.285|
18.421|                      47.159|
## |cost_values_2 |        28.185|           29.934|           32.300|
33.709|                      50.609|
## |cost_values_3 |        30.060|           22.998|           46.555|
21.554|                      90.868|
##
##
## Table: eps: 0.196005109713602, m: 100,  n: 200, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.998|            0.997|            0.998|
1.000|                       1.000|
## |fdr_values    |         0.002|            0.001|            0.003|
0.051|                       0.101|
## |cost_values_1 |         0.171|            0.181|            0.192|
2.141|                       4.477|
## |cost_values_2 |         0.262|            0.305|            0.265|
2.143|                       4.477|
## |cost_values_3 |         0.251|            0.238|            0.311|
4.280|                       8.954|
##
##
## Table: eps: 0.196005109713602, m: 100,  n: 200, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.998|            0.997|            0.999|
1.000|                       1.000|
## |fdr_values    |         0.002|            0.001|            0.003|
0.049|                       0.145|
## |cost_values_1 |         0.168|            0.179|            0.184|
2.068|                       6.778|
## |cost_values_2 |         0.256|            0.302|            0.239|
2.074|                       6.779|
## |cost_values_3 |         0.248|            0.235|            0.313|
4.130|                      13.555|
##
##
## Table: eps: 0.200544528958453, m: 100,  n: 1000, q: 0.1
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.997|            0.996|            0.998|
1.000|                       1.000|
## |fdr_values    |         0.002|            0.001|            0.003|
0.050|                       0.100|
## |cost_values_1 |         0.942|            0.977|            0.971|
10.570|                      22.426|
## |cost_values_2 |         1.491|            1.710|            1.361|
10.593|                      22.433|
## |cost_values_3 |         1.335|            1.221|            1.552|
21.117|                      44.845|
##
##
## Table: eps: 0.200544528958453, m: 100,  n: 1000, q: 0.141
##
## |              | Bayes c0=cA=1| Bayes c0=2, cA=1| Bayes c0=1, cA=2| plug-
in control BFDR| modified Benjamini-Hochberg|
## |:-------------|-------------:|----------------:|----------------:|-------
-------------:|---------------------------:|
## |power_values  |         0.997|            0.996|            0.998|
1.000|                       1.000|
## |fdr_values    |         0.002|            0.001|            0.003|
0.050|                       0.142|
## |cost_values_1 |         0.911|            0.961|            0.971|
10.599|                      33.391|
## |cost_values_2 |         1.463|            1.695|             1.369|
10.627|                      33.397|
## |cost_values_3 |         1.270|            1.188|             1.544|
21.170|                      66.776|
```

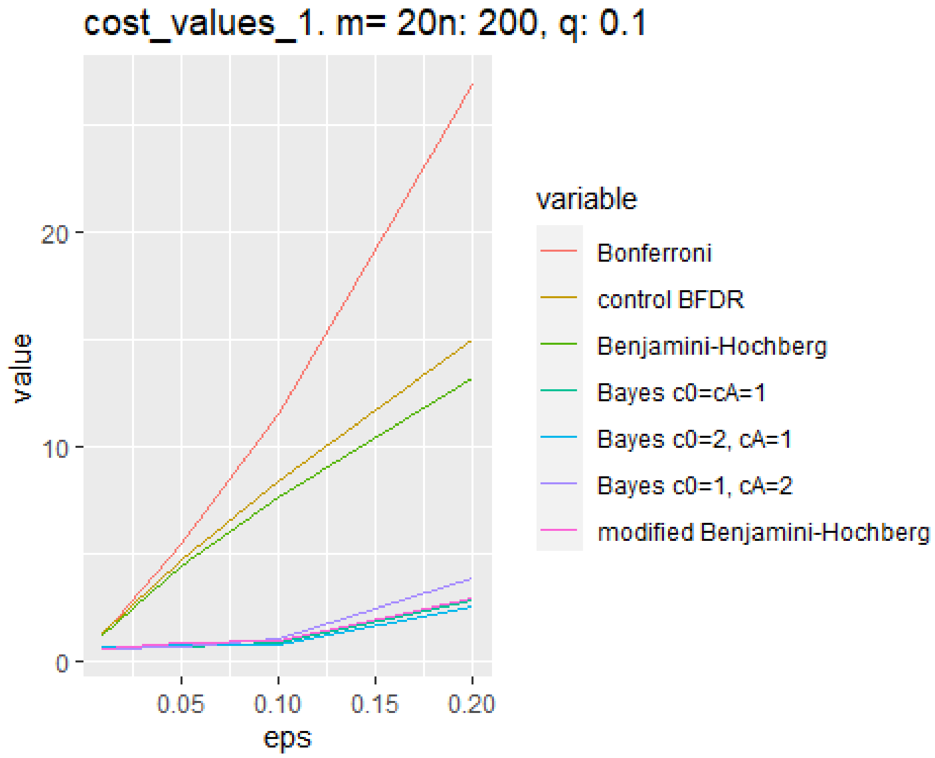
Comments:

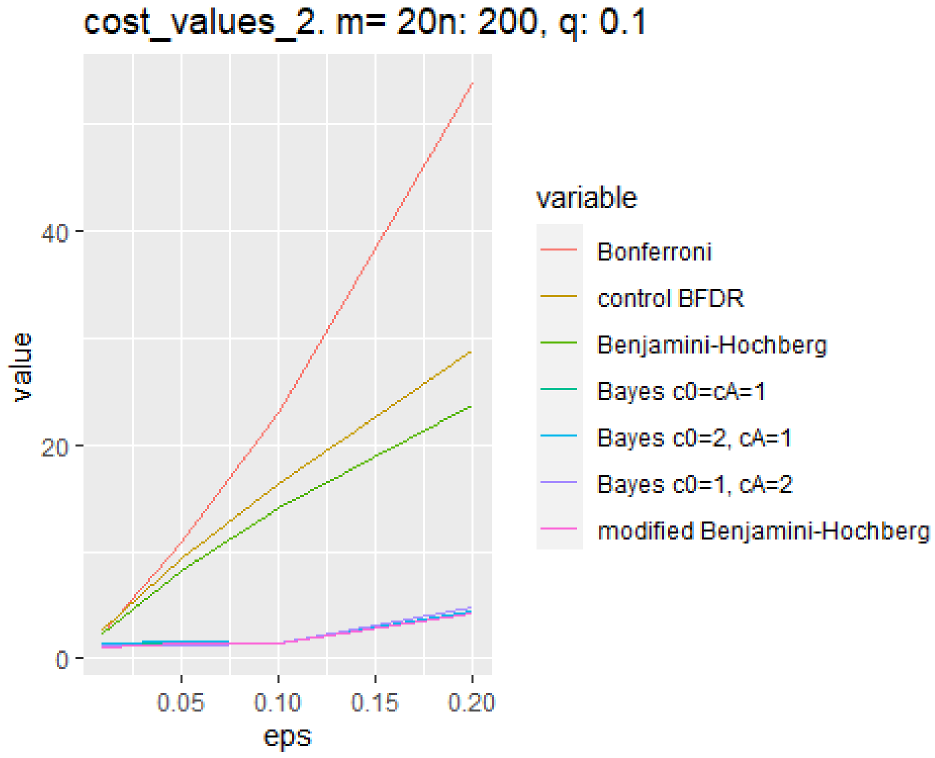
- As epsilon increases, the power of the tests generally increases across different scenarios. This is expected, as a larger epsilon allows for a wider range of true positives to be identified. Conversely, the false discovery rate (FDR) tends to decrease with higher epsilon values, indicating a better control over false positives.
- Power increases as both sample size (n) and group size (m) increase. This is consistent with statistical intuition, where larger sample sizes provide more information, leading to better detection of true positives. Similarly, FDR tends to decrease with larger sample sizes and group sizes, indicating better control over false discoveries.
- The different methods for controlling the False Discovery Rate (plug-in control BFDR, modified Benjamini-Hochberg) show varying performance across scenarios. The modified Benjamini-Hochberg method tends to have higher power compared to plug-in control BFDR, especially as the threshold (q) increases. However, this improvement in power often comes at the cost of a higher FDR.
- The cost values provide a comprehensive evaluation, considering both power and FDR. It appears that, in some scenarios, modified Benjamini-Hochberg achieves higher power at the expense of increased costs, especially as q increases.
- Generally, the trends observed across different scenarios are consistent. Higher epsilon, larger sample sizes, and group sizes contribute to increased power and improved control over false discoveries.

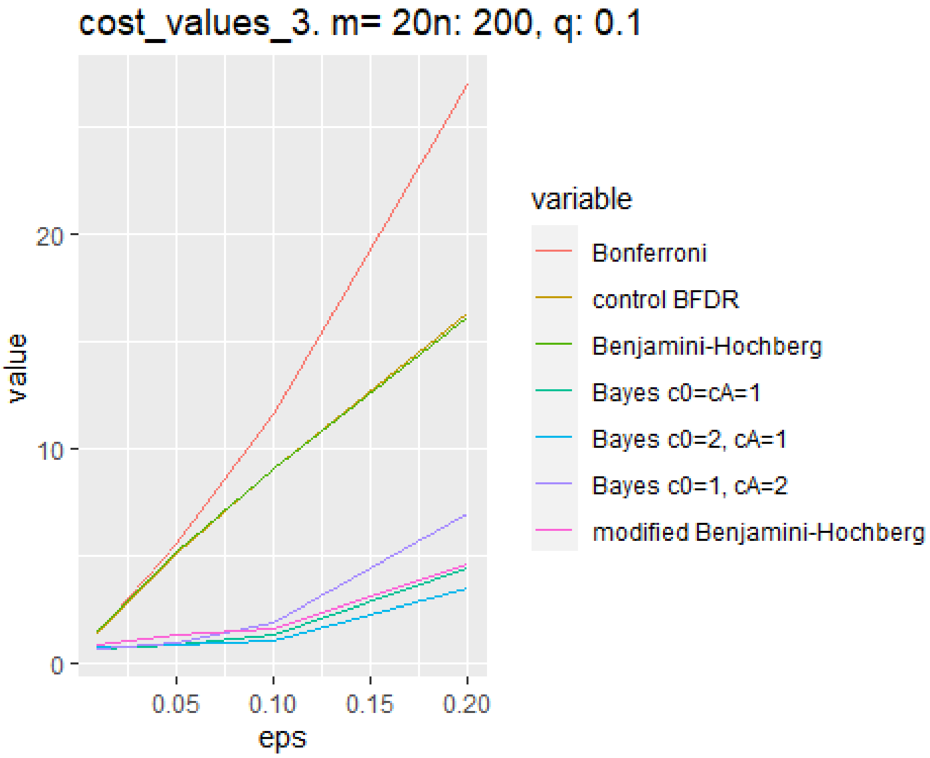
Question 6:

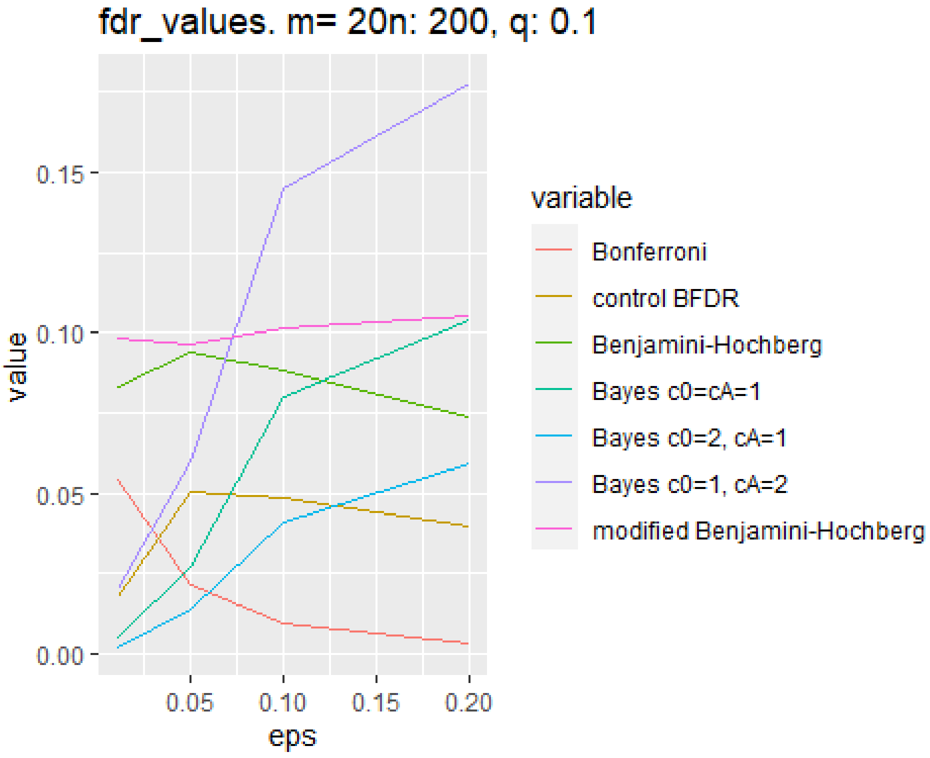
Present values of each of the above characteristics (FDR, Power, expected cost) as a function of ε. Compare results of direct and plug-in procedures on the same graph.

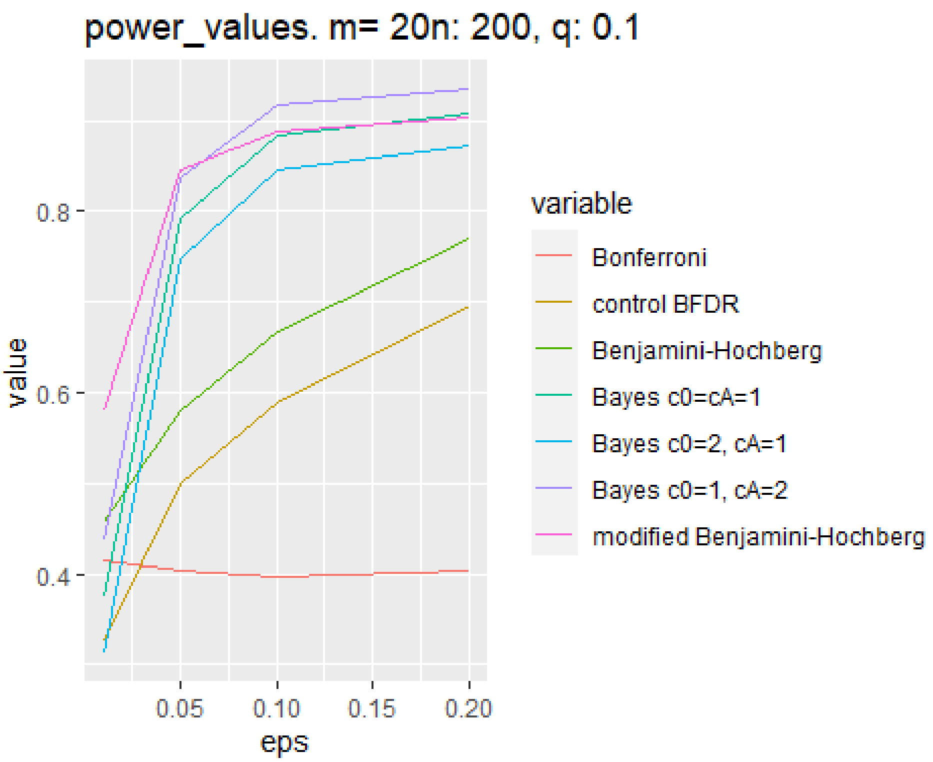
```r
# Load required libraries
library(reshape2)
library(ggplot2)
library(stringr)

# Initialize an empty data frame for direct power
direct_power = data.frame()

# Loop through the list of all results
for (i in seq(length(list_all_results))) {
  # Extract the relevant data from the list
  tmp = list_all_results[[i]][[4]]

  # Add a new column 'metric' with row names
  tmp$metric = row.names(tmp)

  # Combine and append the data to the direct power data frame
  direct_power = rbind(direct_power,
cbind(data.frame(list_all_results[[i]][1:3]), tmp))
}

# Initialize an empty data frame for plugin power
plugin_power = data.frame()

# Loop through the list of plugin results
for (i in seq(length(list_all_results))) {
  # Set the 'eps' column for each plugin result
  all_res_plugin[[i]]$eps = results_list[[ceiling(i/2)]]$eps

  # Extract the relevant data from the plugin result
  tmp = all_res_plugin[[i]][[4]]

  # Add a new column 'metric' with row names
  tmp$metric = row.names(tmp)

  # Combine and append the data to the plugin power data frame
  plugin_power = rbind(plugin_power,
cbind(data.frame(all_res_plugin[[i]][c(1:3, 5)]), tmp))
}

# Set 'q' column for direct power equal to the 'q' column in plugin power
direct_power$q = plugin_power$q

# Merge the direct power and plugin power data frames
data = plyr::join(direct_power, plugin_power, by=c("m", "n", "eps", "q",
"eps", "metric"))

# Order the combined data frame
data = data[with(data, order(m, n, q, metric, eps)), ]

# Create plots for each combination of parameters
for (i in seq(40)) {
  # Extract data for the current combination
  tmp = data[((i-1)*4 + 1): (i*4), ]

  # Extract specific values for the plot title
  m = tmp[1, "m"]
  n = tmp[1, "n"]
  q = tmp[1, "q"]
  metric = tmp[1, "metric"]

  # Reshape the data for plotting
  tmp = melt(tmp, "eps", c("Bonferroni", "control BFDR", "Benjamini-
Hochberg",
                            "Bayes c0=cA=1", "Bayes c0=2, cA=1", "Bayes c0=1,
cA=2",
                            "control BFDR", "modified Benjamini-Hochberg"))

  # Display the plot using ggplot
  show(ggplot(tmp, aes(x=eps, y=value, color=variable)) +
         geom_line() +
         labs(title = str_c(metric, ". m= ", m, "n: ", n, ", q: ", round(q,
4)))
  )
}
```

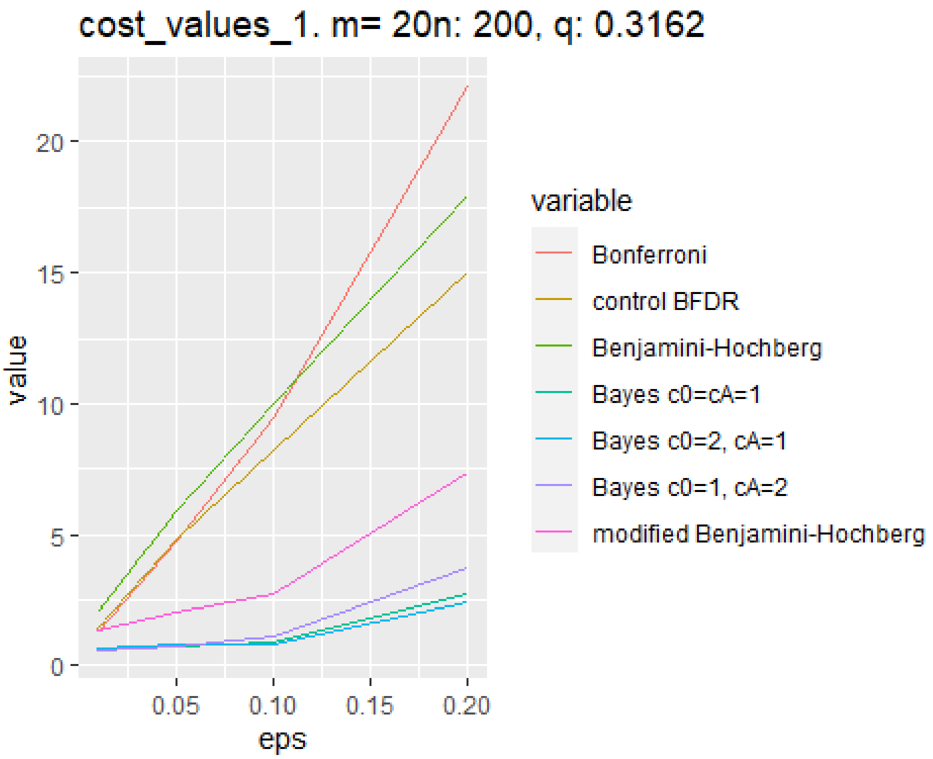

cost_values_1. m= 20n: 200, q: 0.1
value
20
10
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

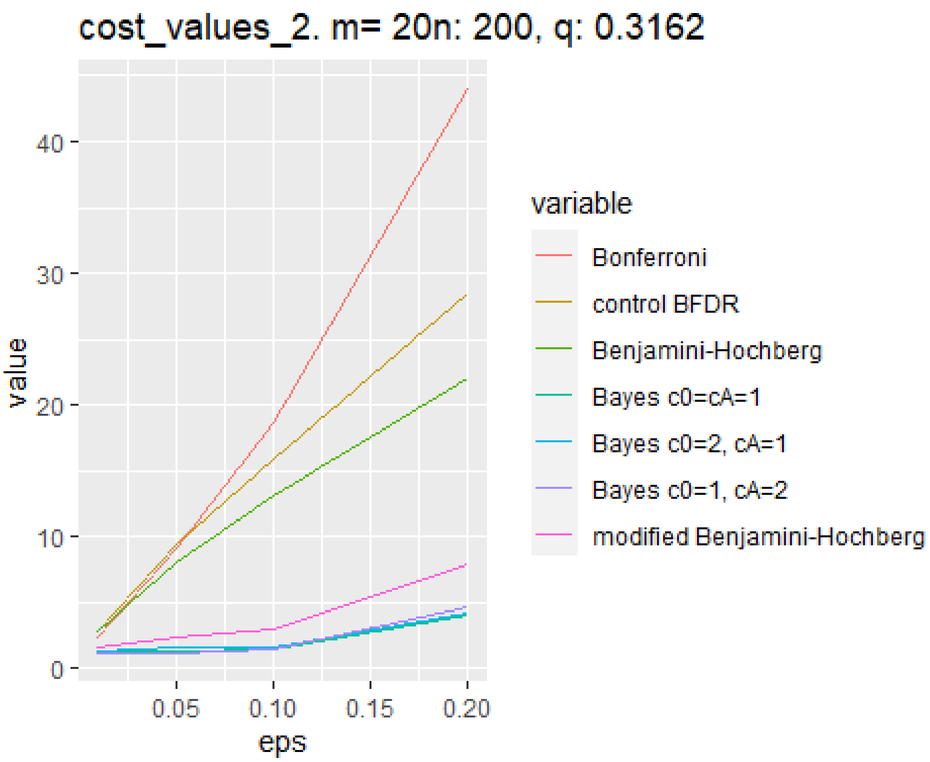

cost_values_2. m= 20n: 200, q: 0.1
value
40
20
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

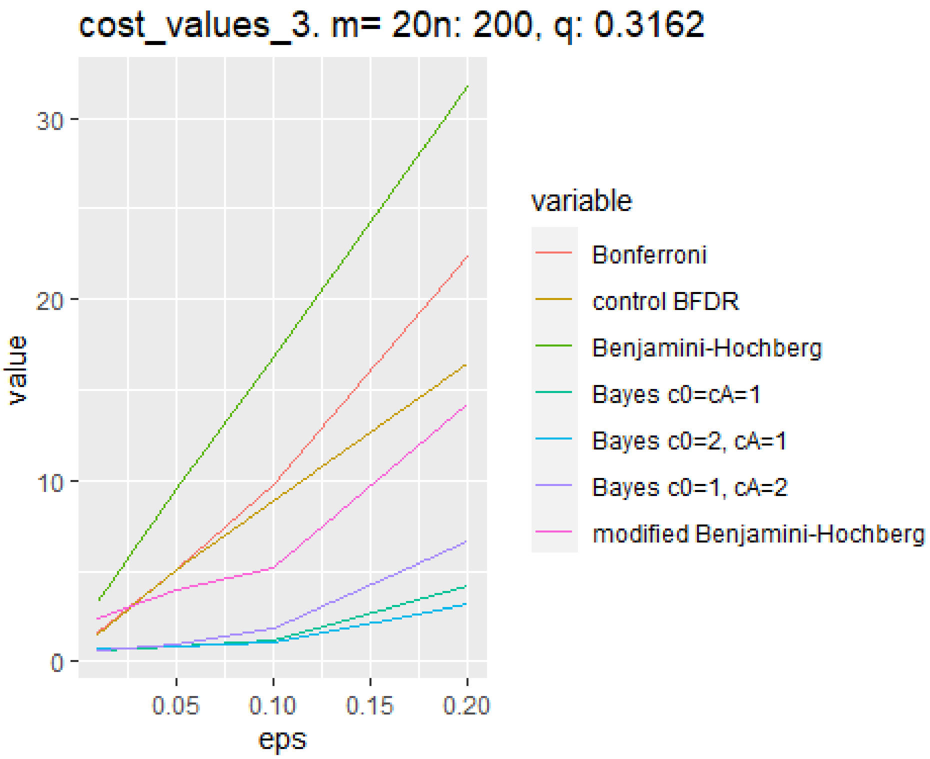

cost_values_3. m= 20n: 200, q: 0.1
value
0
10
20
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

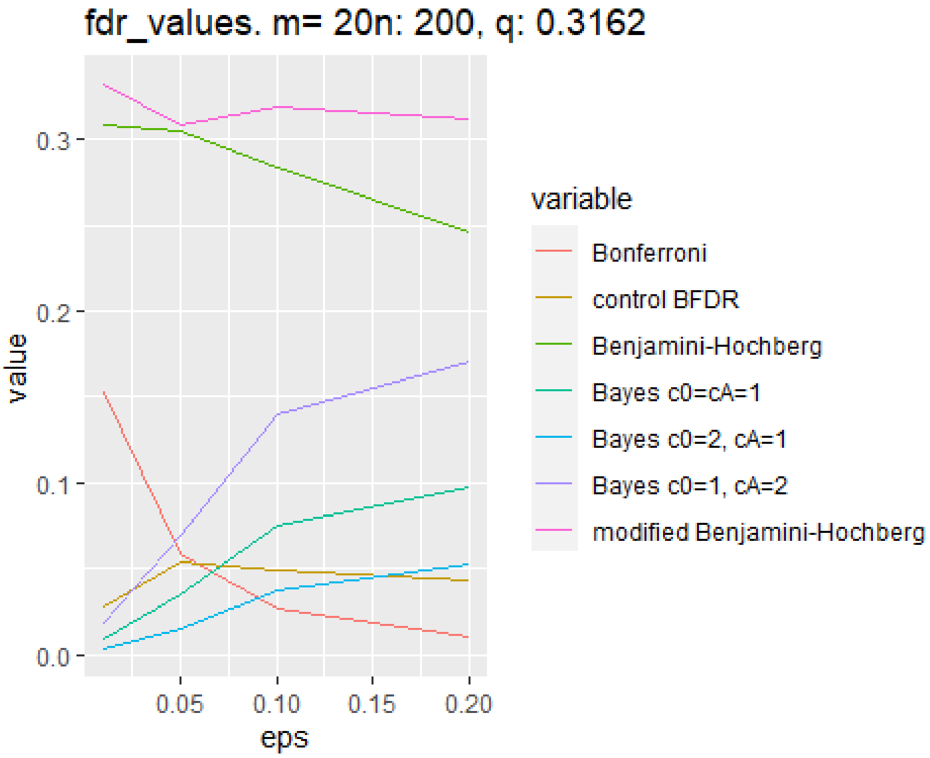

fdr_values. m= 20n: 200, q: 0.1
value
0.00
0.05
0.10
0.15
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


power_values. m= 20n: 200, q: 0.1
value
0.4
0.6
0.8
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

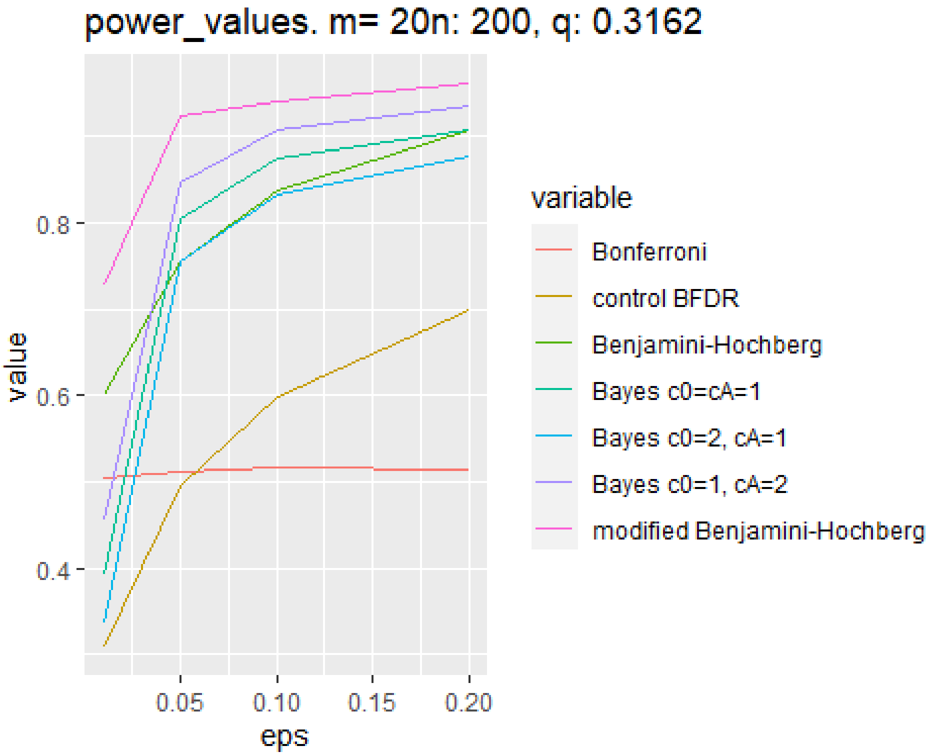

cost_values_1. m= 20n: 200, q: 0.3162
value
0
5
10
15
20
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_2. m= 20n: 200, q: 0.3162
value
0
10
20
30
40
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_3. m= 20n: 200, q: 0.3162
value
0
10
20
30
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

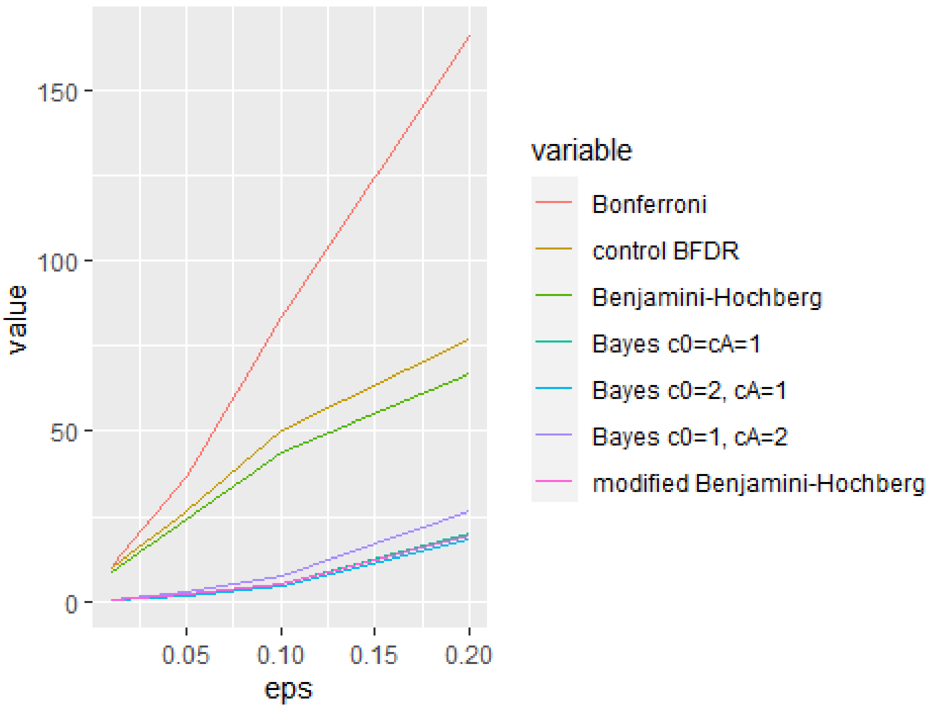


fdr_values. m= 20n: 200, q: 0.3162

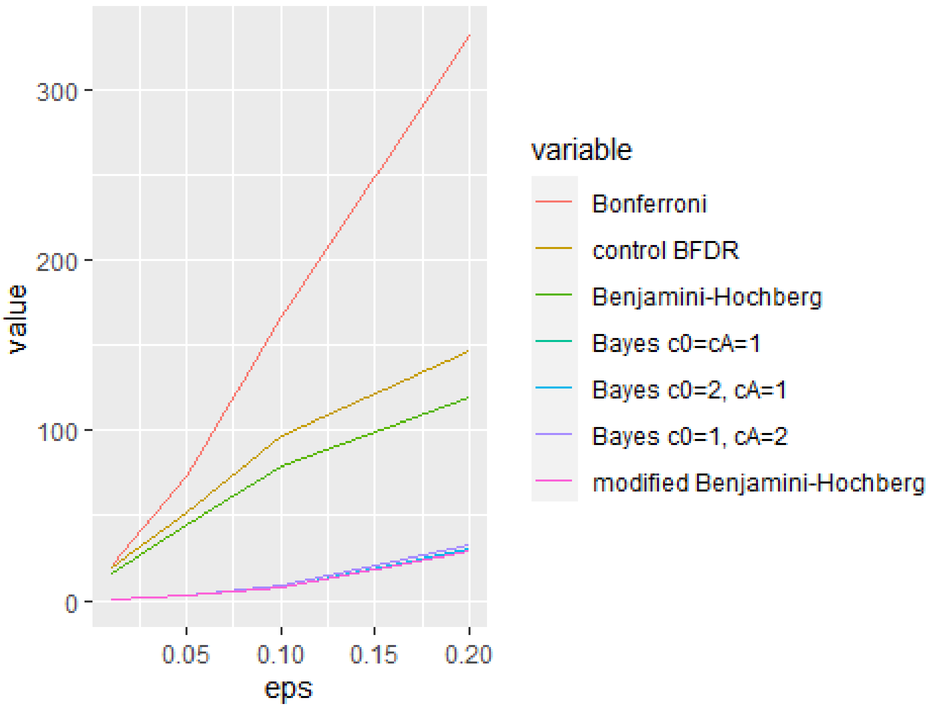


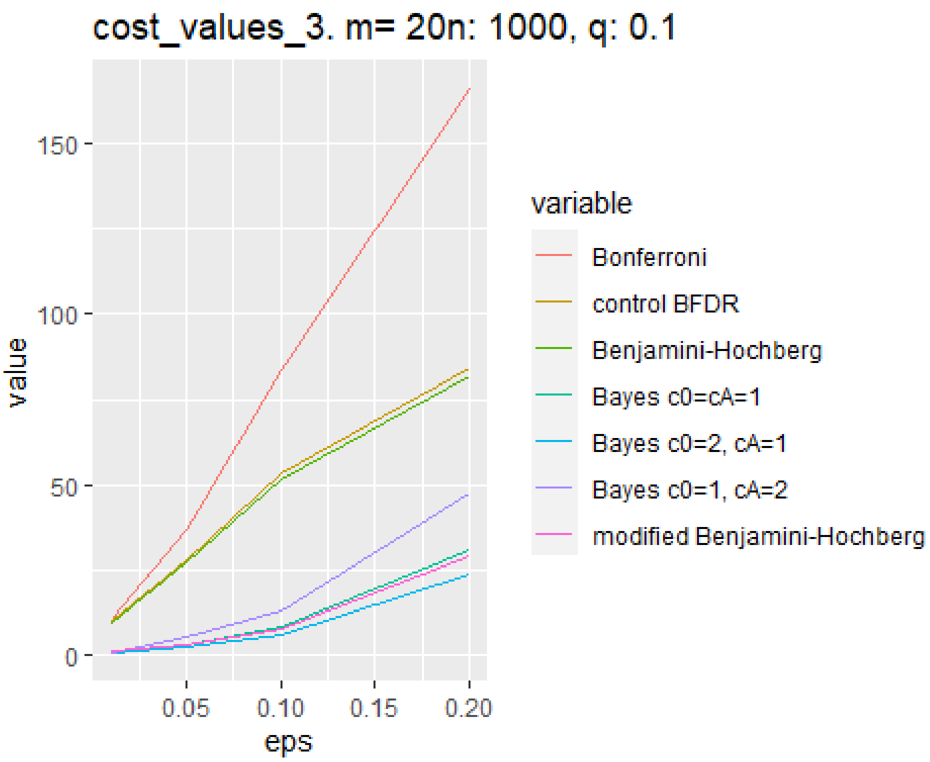
power_values. m= 20n: 200, q: 0.3162

cost_values_1. m= 20n: 1000, q: 0.1

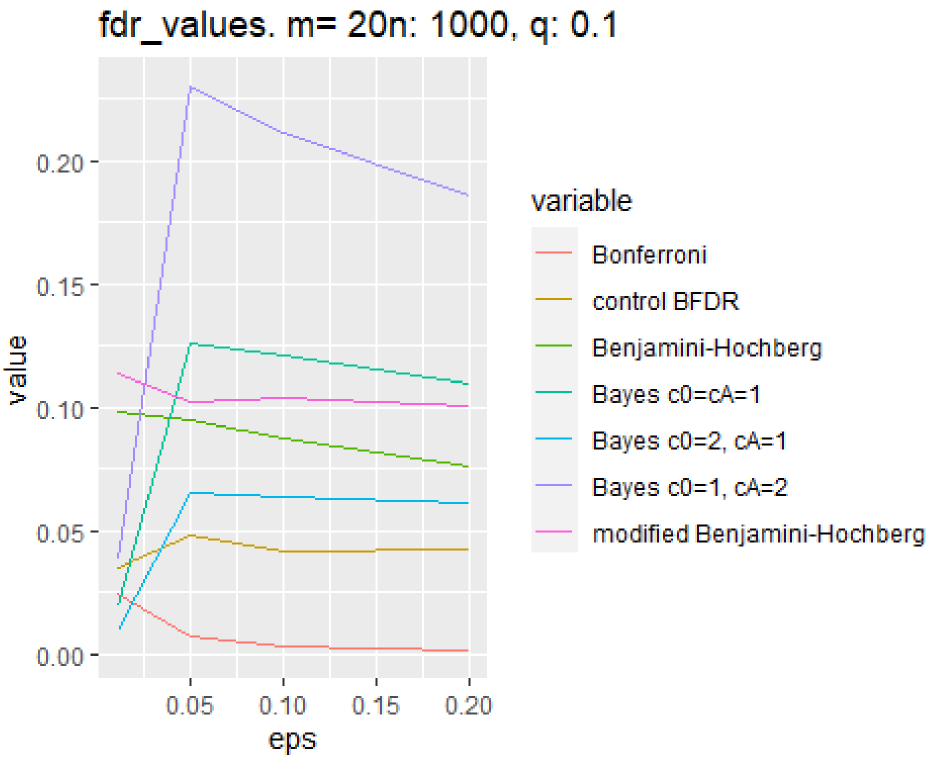


cost_values_2. m= 20n: 1000, q: 0.1

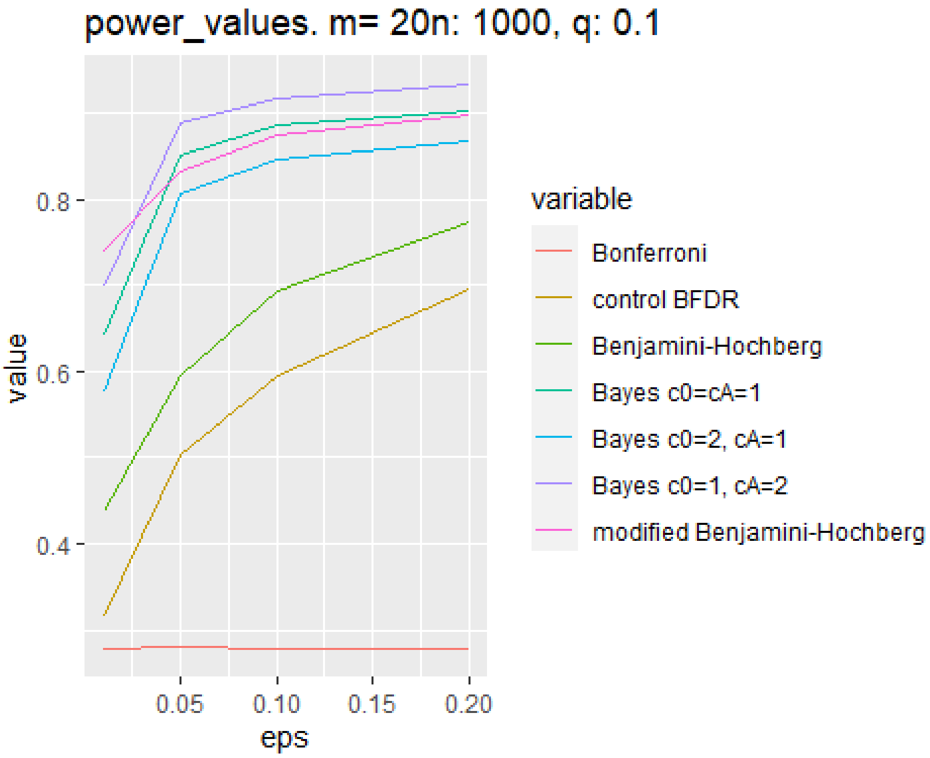


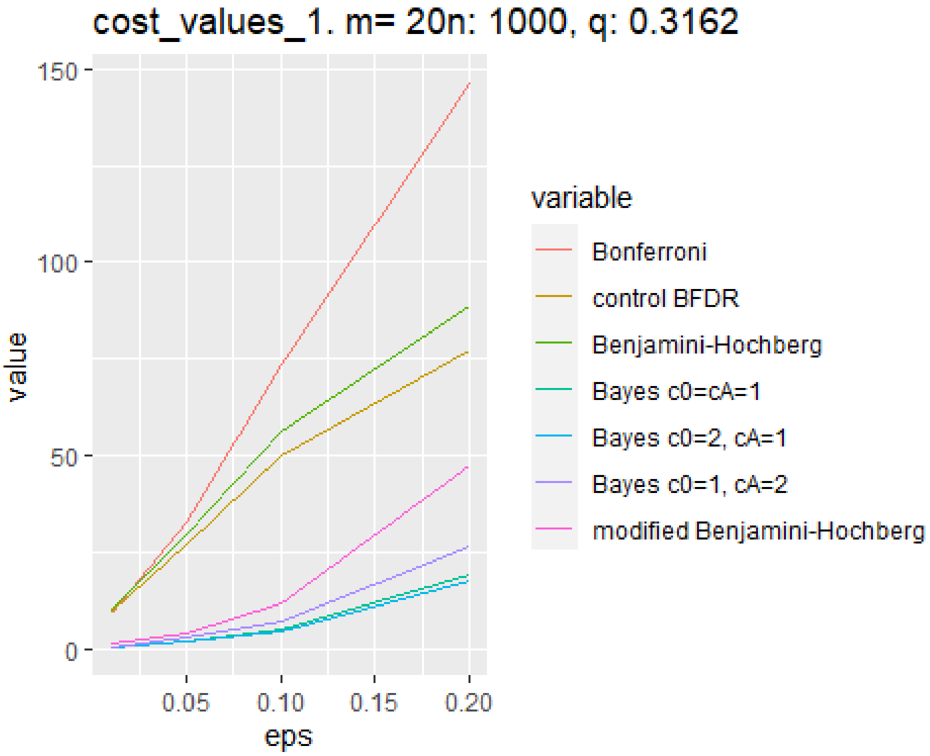


cost_values_3. m= 20n: 1000, q: 0.1

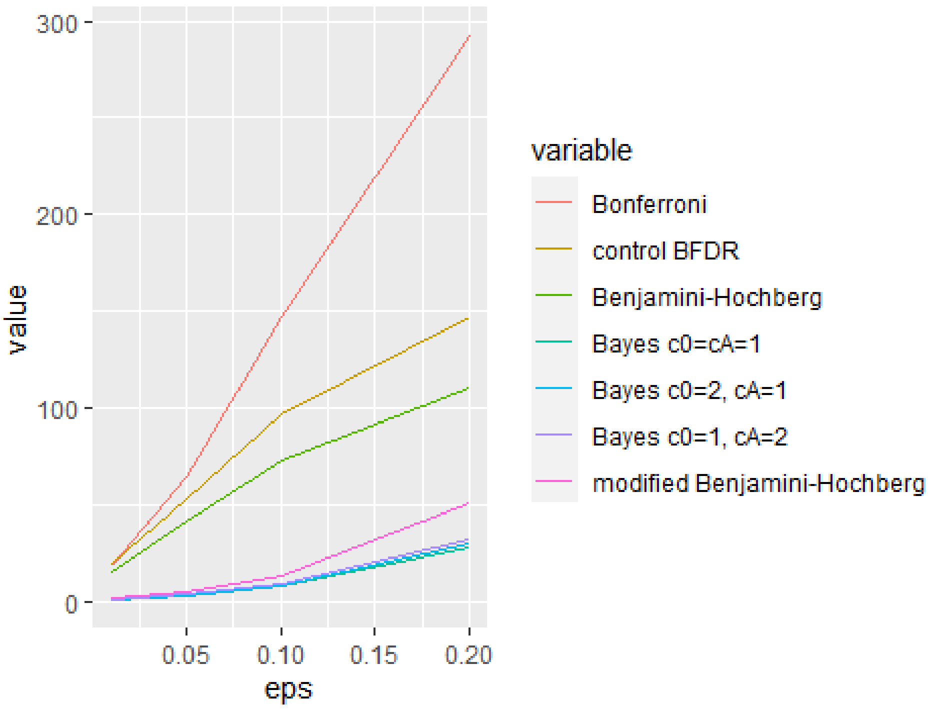


fdr_values. m= 20n: 1000, q: 0.1

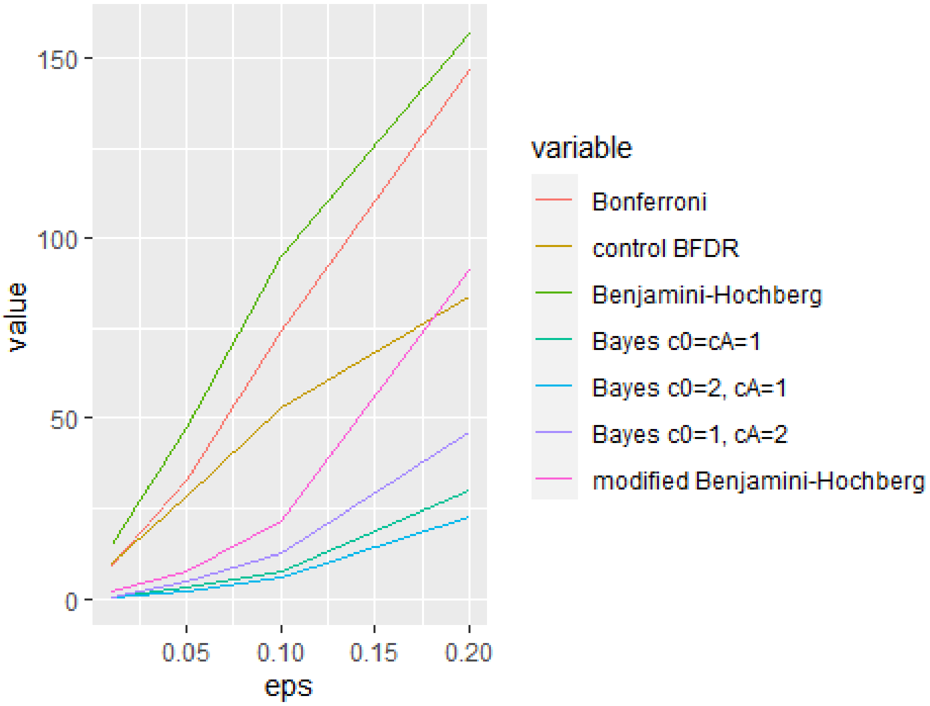

power_values. m= 20n: 1000, q: 0.1
value
0.4
0.6
0.8
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

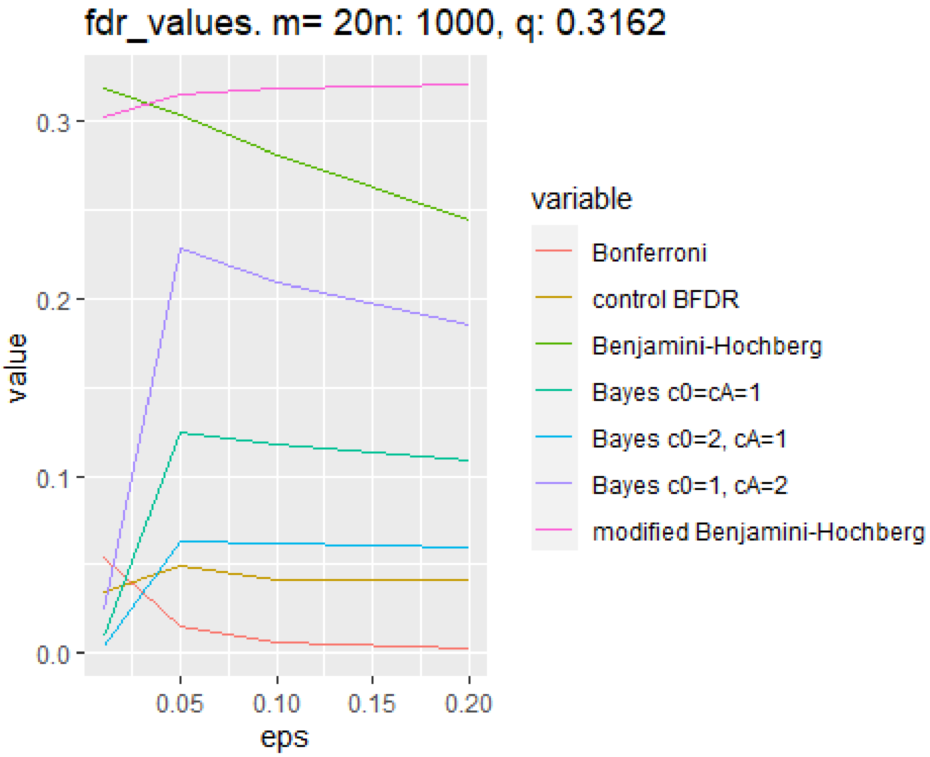

cost_values_1. m= 20n: 1000, q: 0.3162
value
0
50
100
150
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

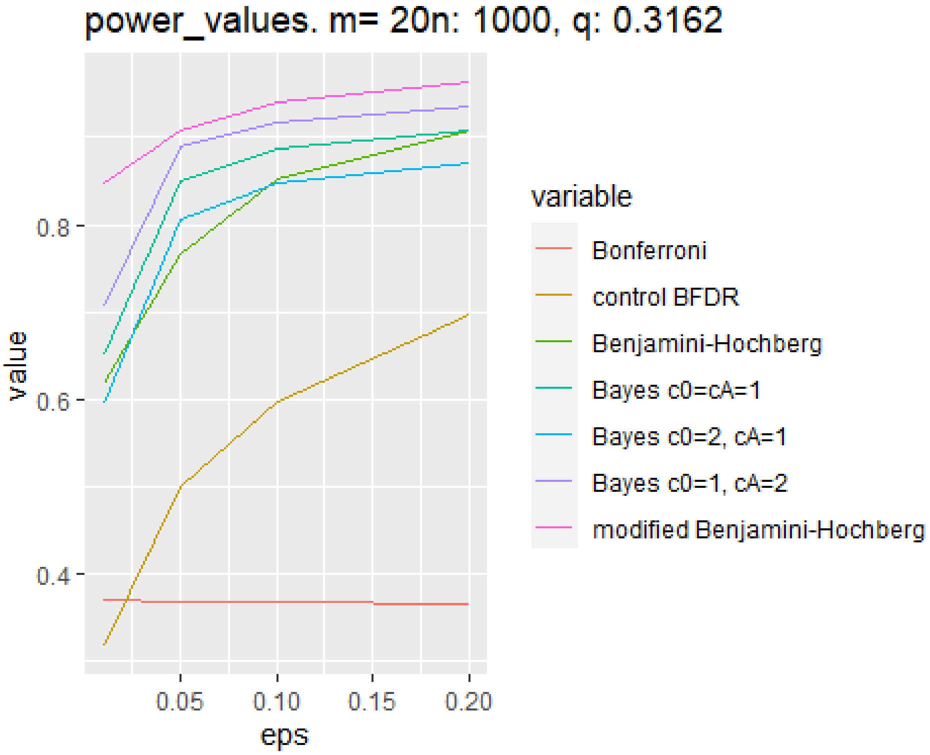
cost_values_2. m= 20n: 1000, q: 0.3162

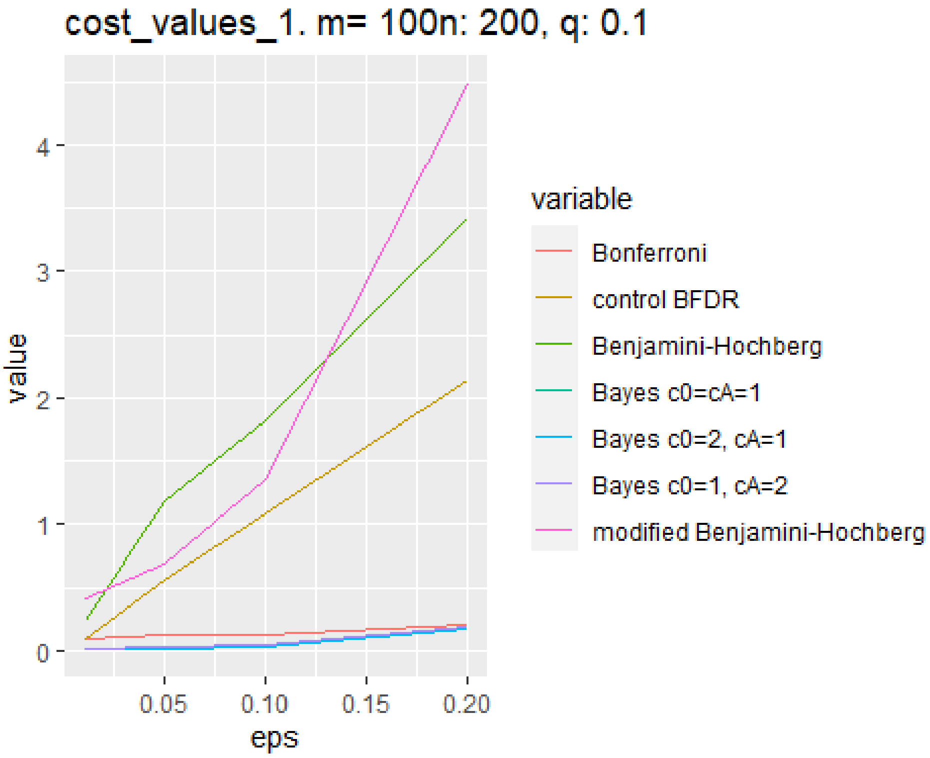


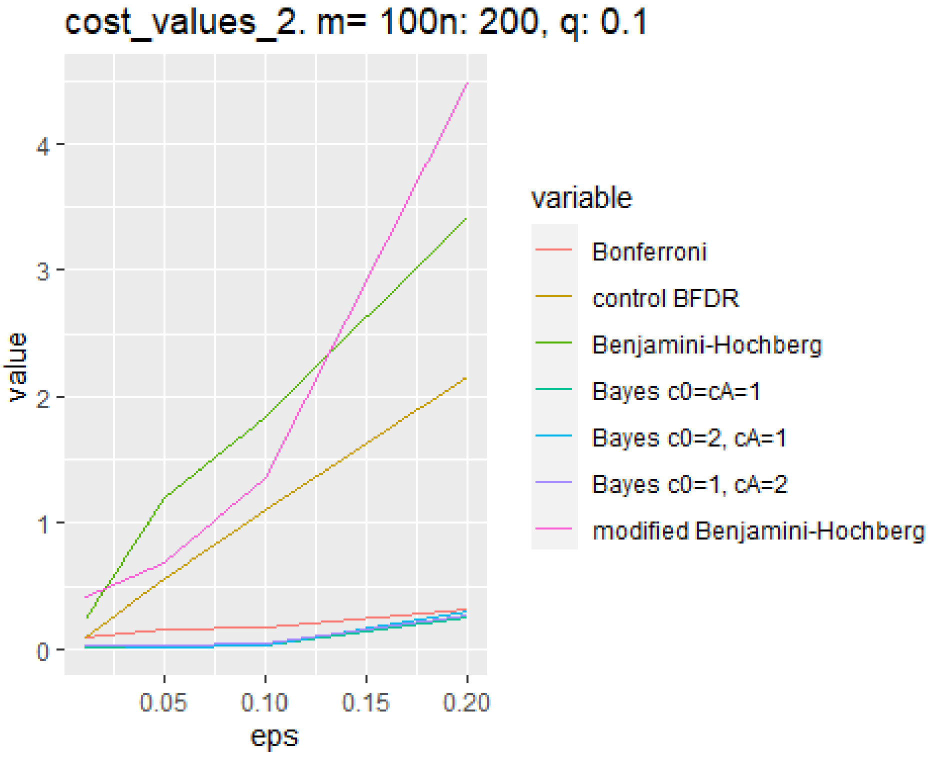
cost_values_3. m= 20n: 1000, q: 0.3162

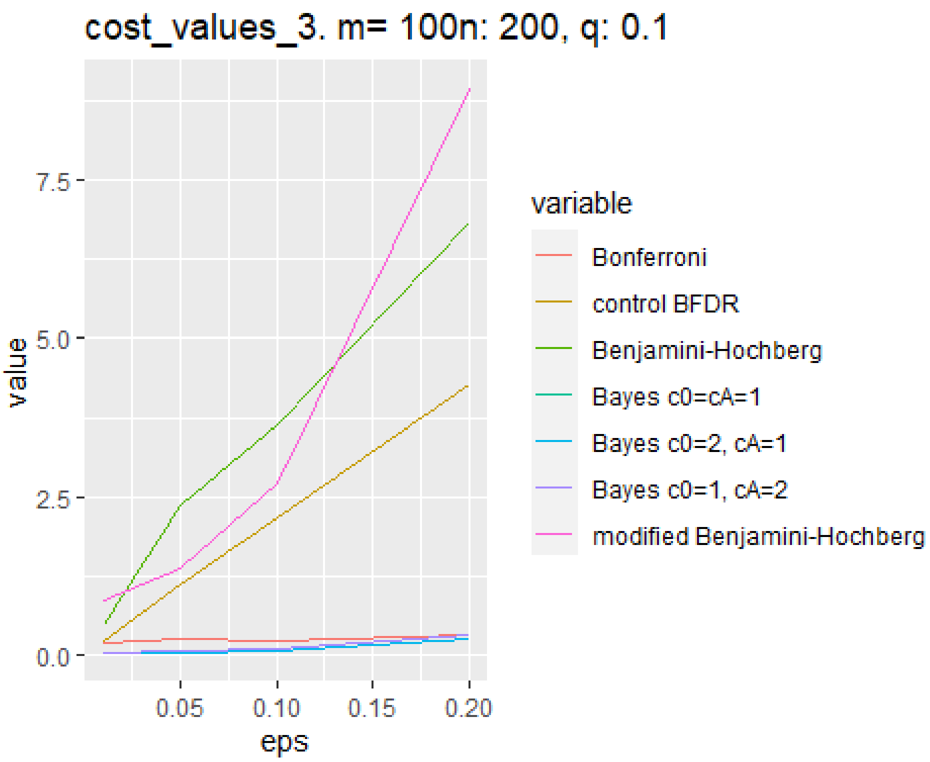



fdr_values. m= 20n: 1000, q: 0.3162
value
0.3
0.2
0.1
0.0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

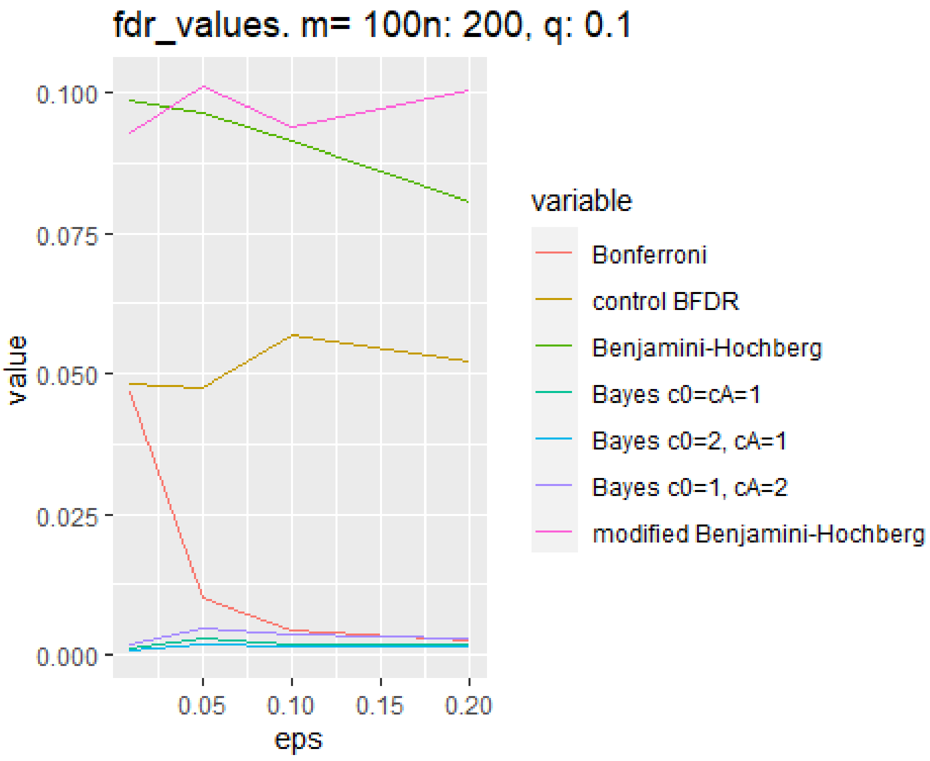

power_values. m= 20n: 1000, q: 0.3162
value
0.8
0.6
0.4
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_1. m= 100n: 200, q: 0.1
value
0
1
2
3
4
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

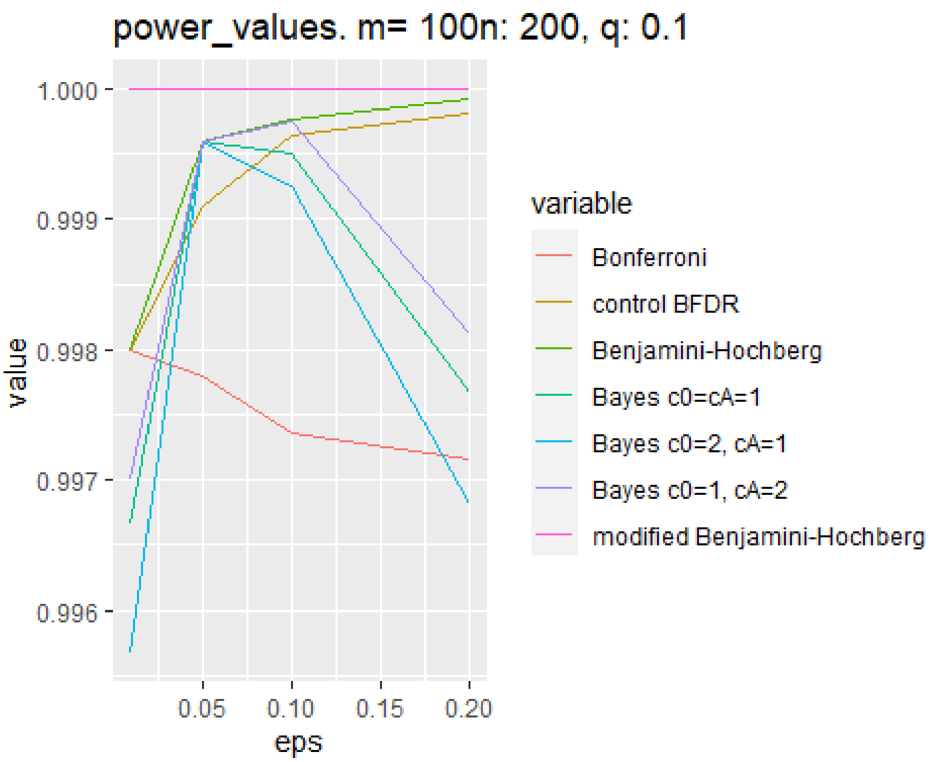

cost_values_2. m= 100n: 200, q: 0.1
value
0
1
2
3
4
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

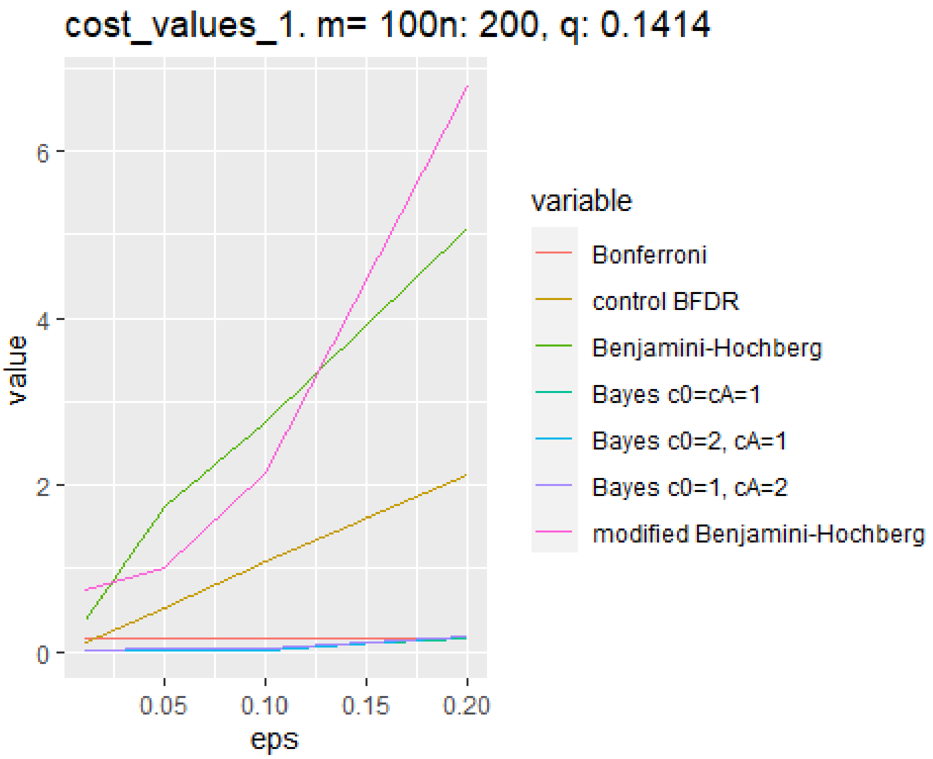


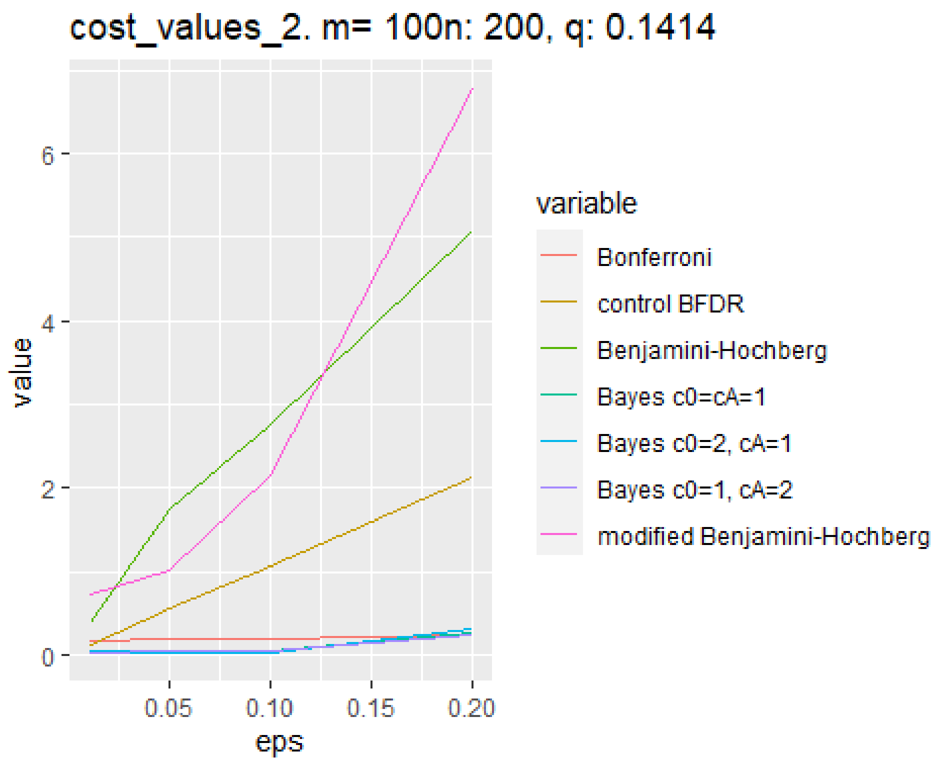
cost_values_3. m= 100n: 200, q: 0.1

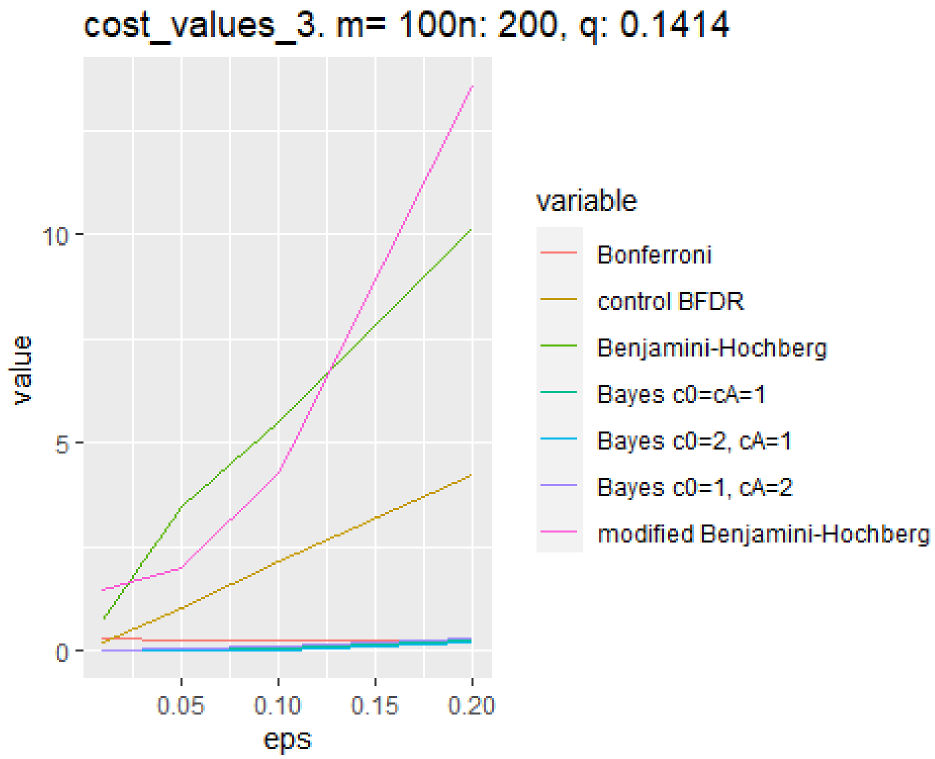


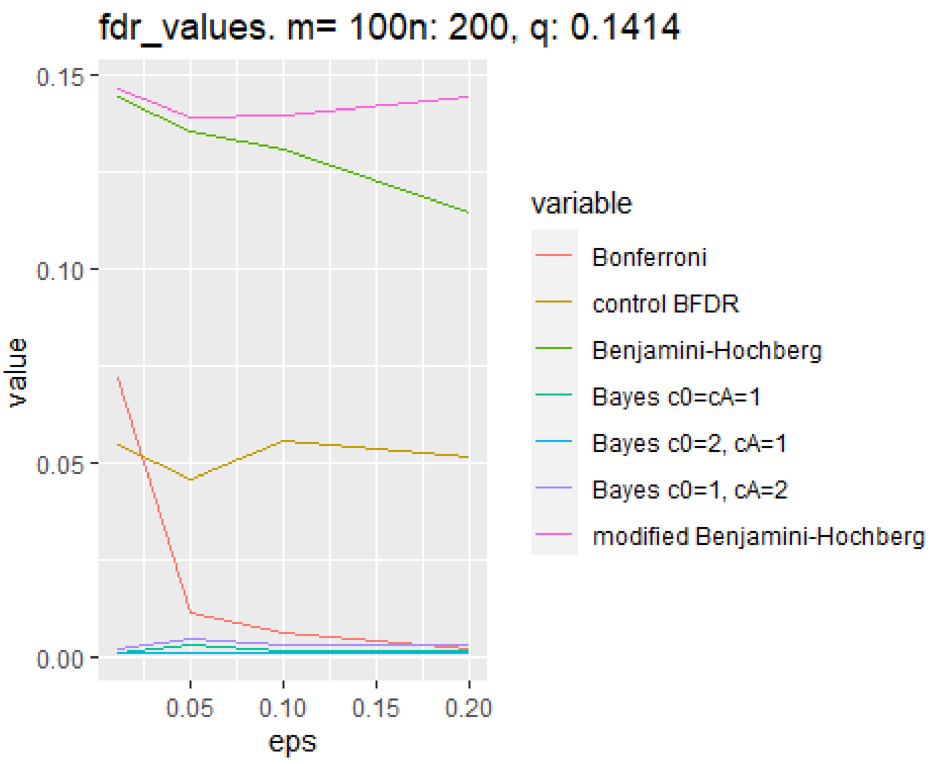
fdr_values. m= 100n: 200, q: 0.1

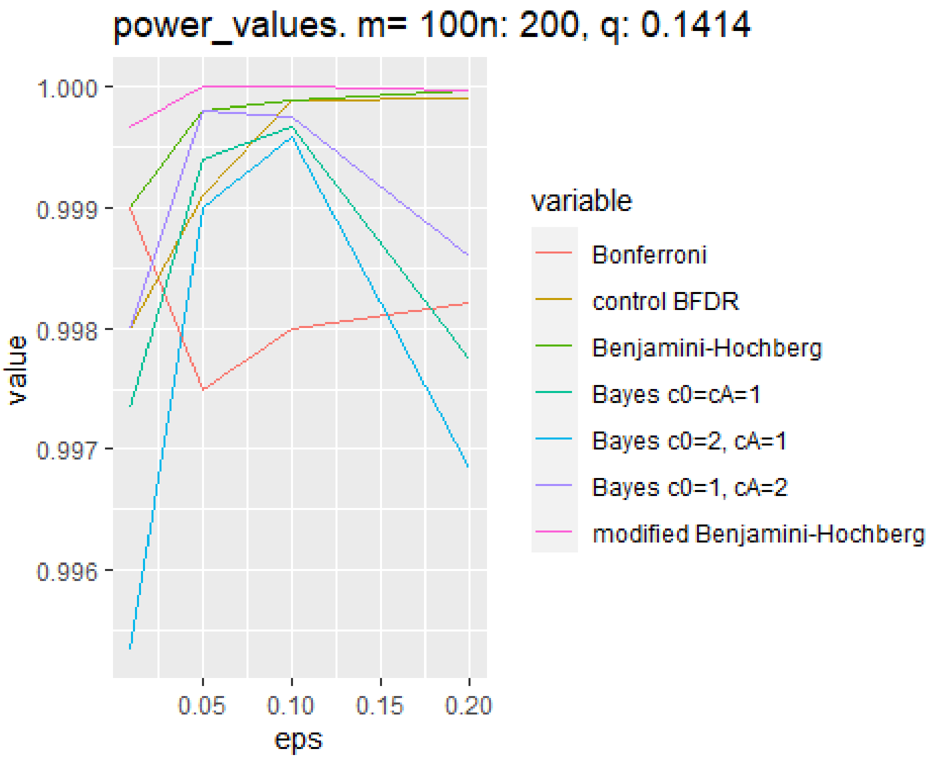

power_values. m= 100n: 200, q: 0.1
1.000
0.999
0.998
0.997
0.996
value
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

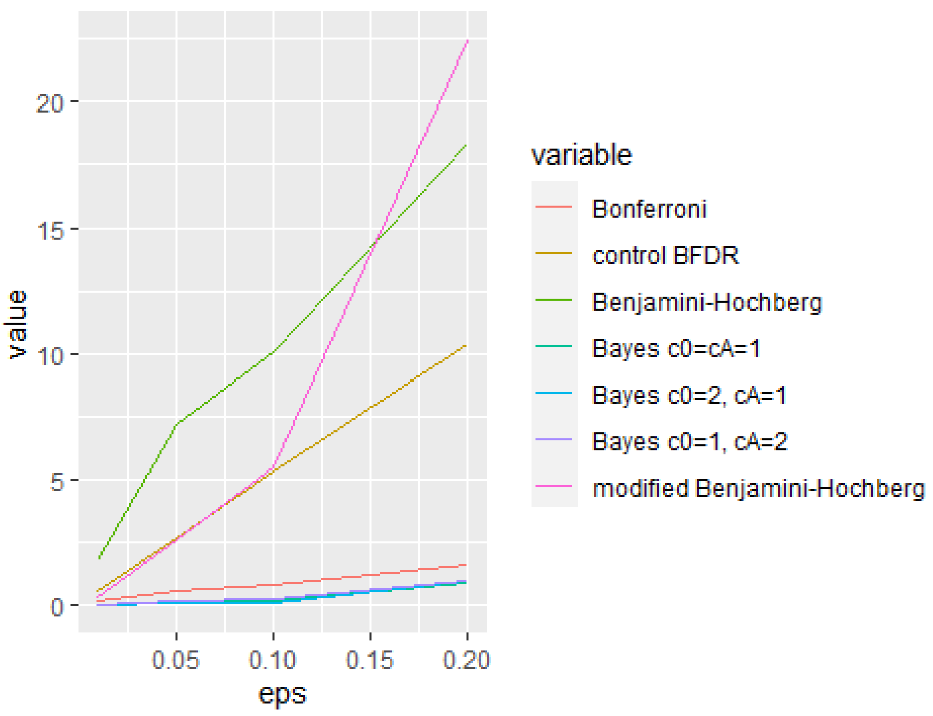

cost_values_1. m= 100n: 200, q: 0.1414
6
4
2
0
value
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

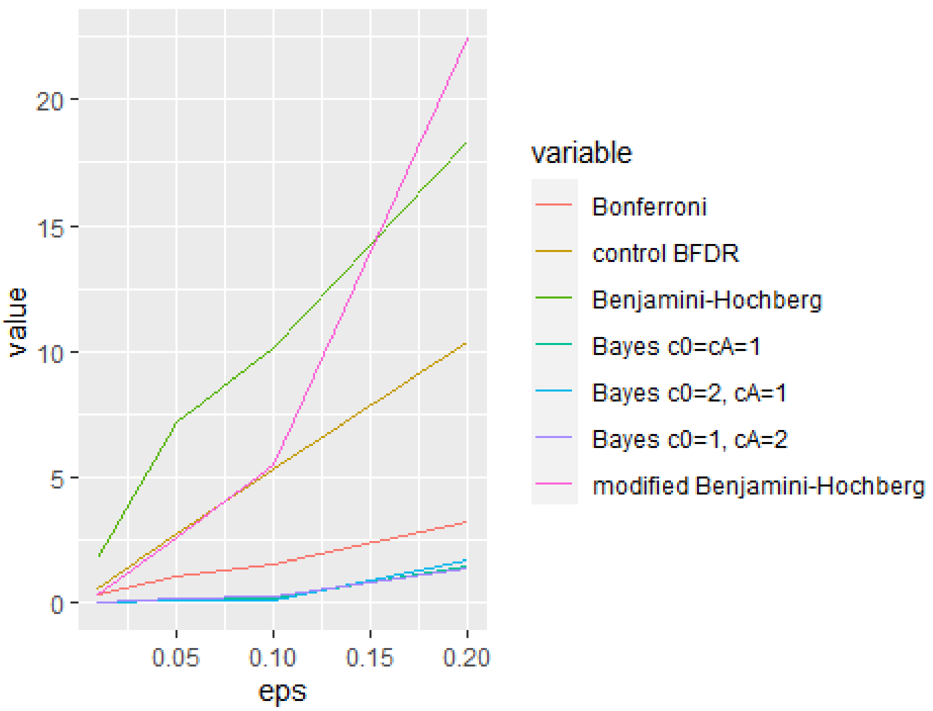

cost_values_2. m= 100n: 200, q: 0.1414

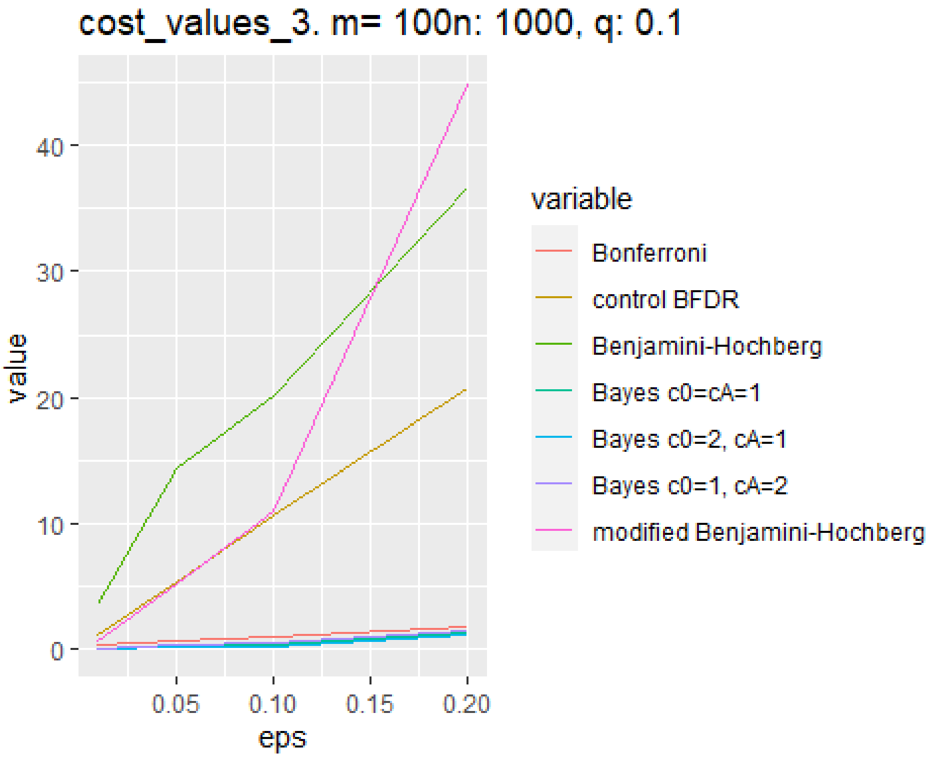

cost_values_3. m= 100n: 200, q: 0.1414

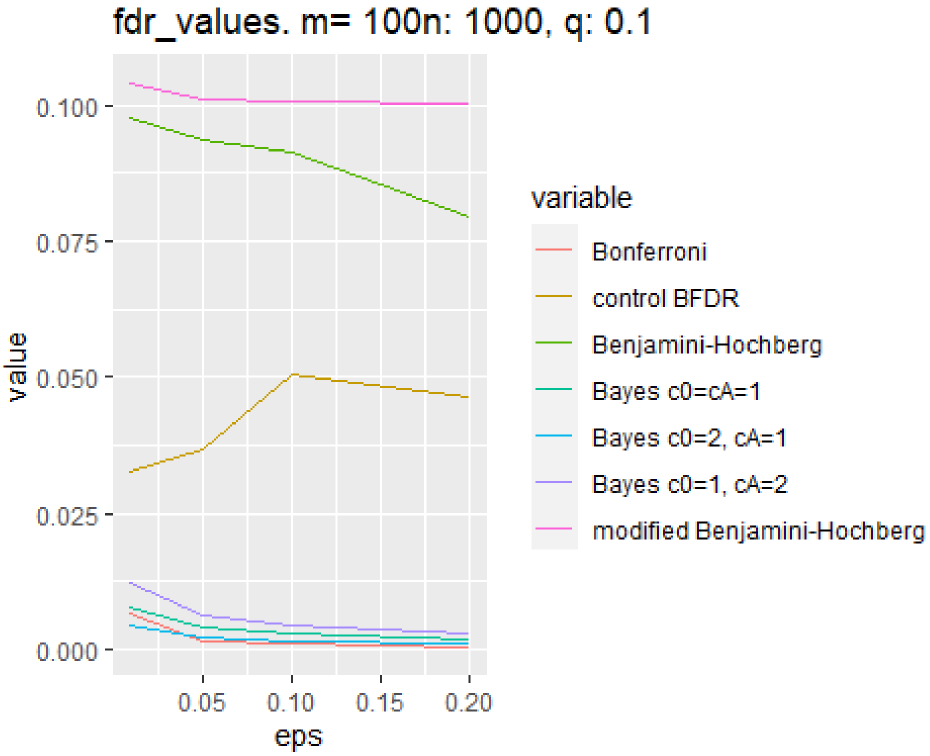

fdr_values. m= 100n: 200, q: 0.1414
value
0.15
0.10
0.05
0.00
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

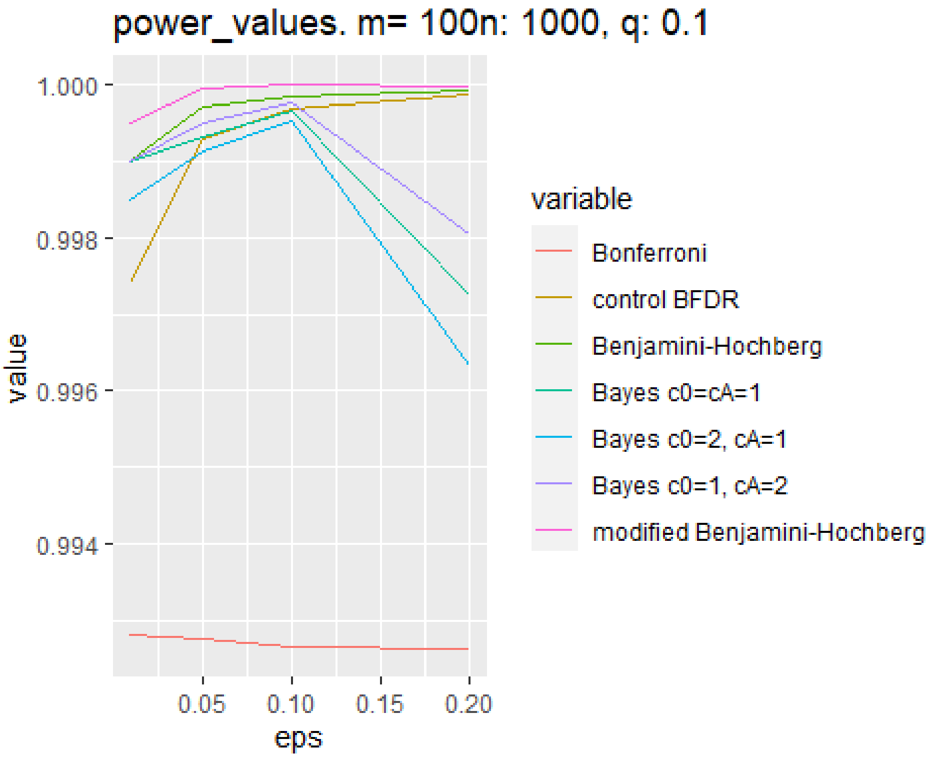

power_values. m= 100n: 200, q: 0.1414
value
1.000
0.999
0.998
0.997
0.996
eps
0.05
0.10
0.15
0.20
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

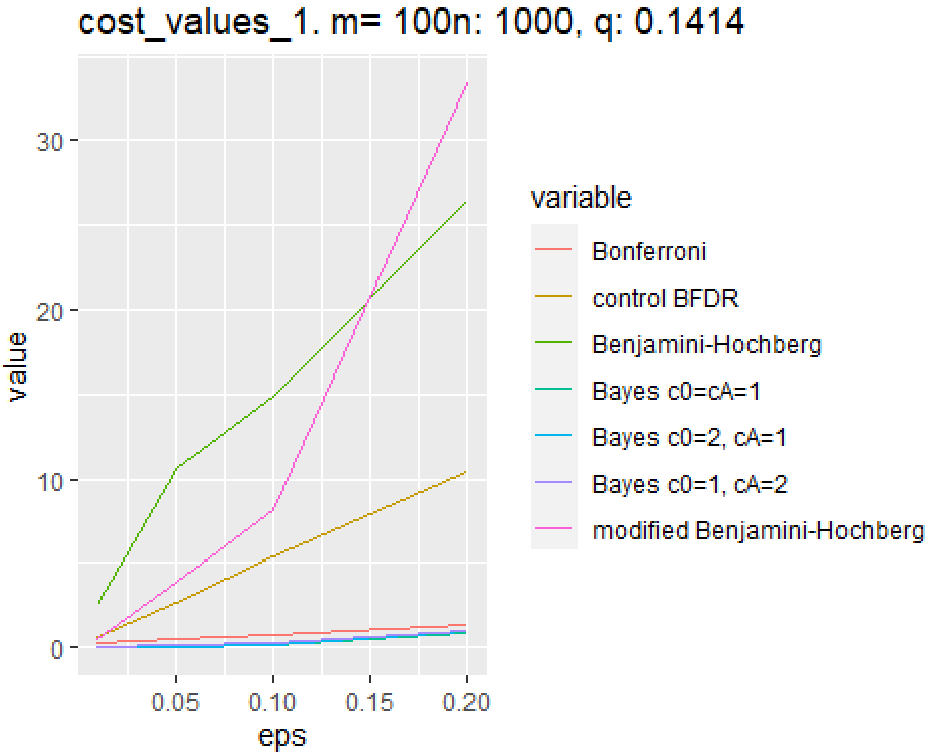

cost_values_1. m= 100n: 1000, q: 0.1

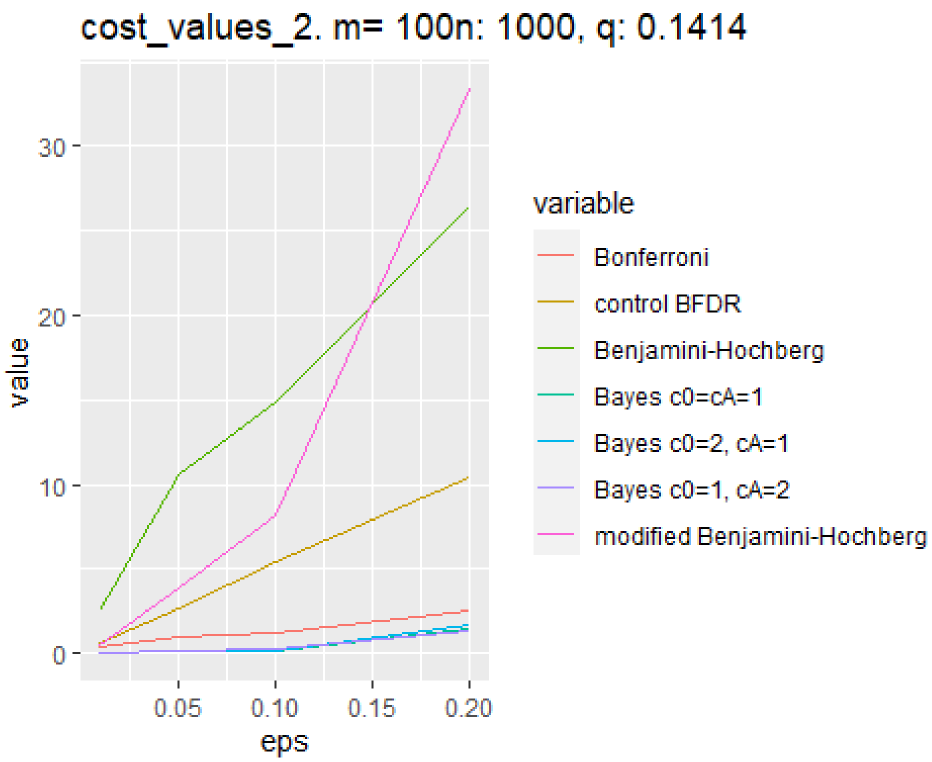

cost_values_2. m= 100n: 1000, q: 0.1

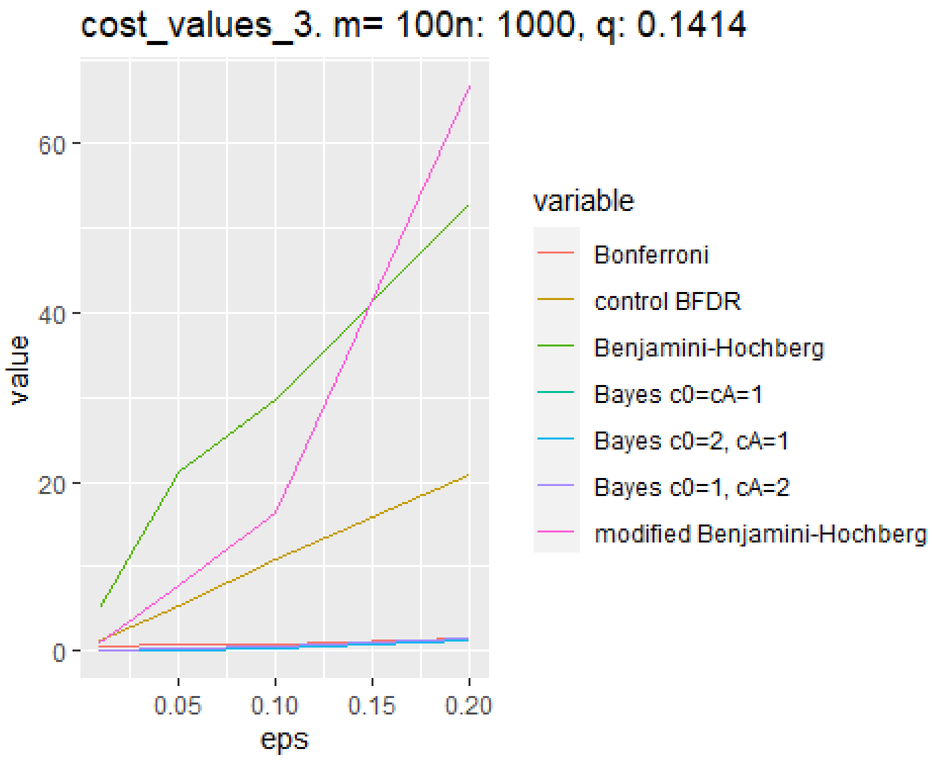

cost_values_3. m= 100n: 1000, q: 0.1
value
40
30
20
10
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

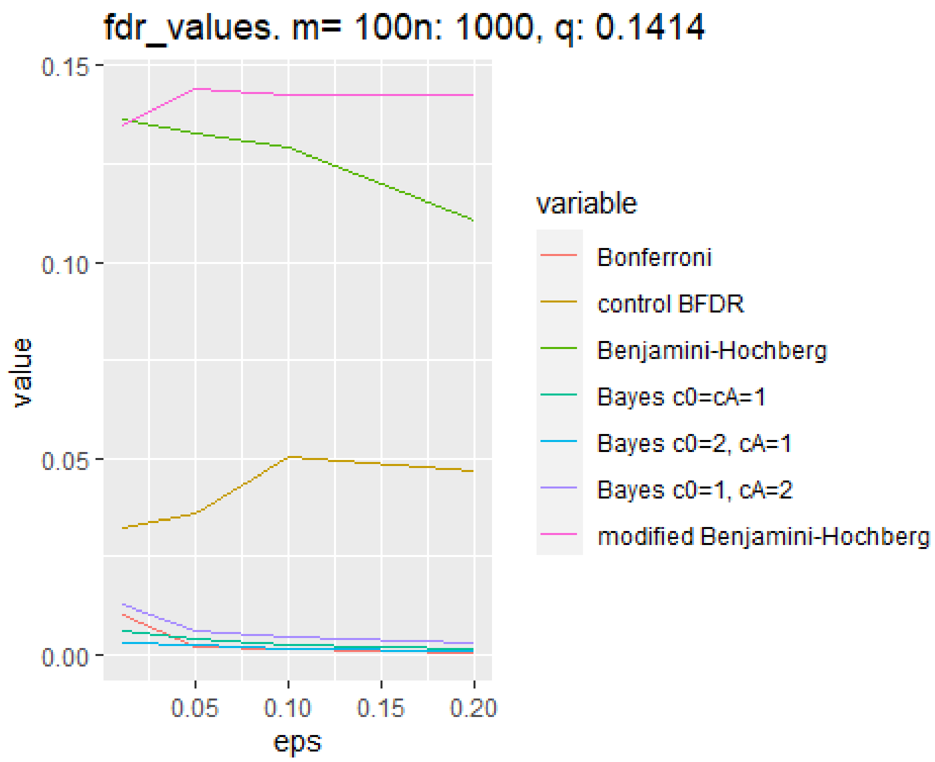

fdr_values. m= 100n: 1000, q: 0.1
value
0.100
0.075
0.050
0.025
0.000
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg

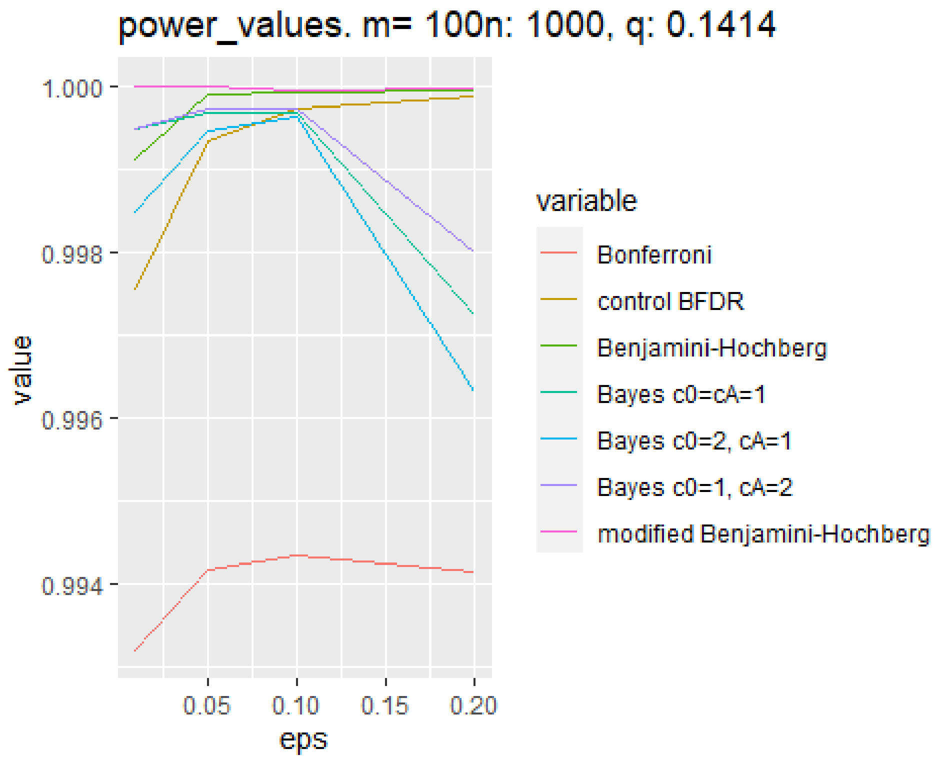

power_values. m= 100n: 1000, q: 0.1
value
1.000
0.998
0.996
0.994
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_1. m= 100n: 1000, q: 0.1414
value
30
20
10
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_2. m= 100n: 1000, q: 0.1414
value
30
20
10
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg


cost_values_3. m= 100n: 1000, q: 0.1414
value
60
40
20
0
0.05
0.10
0.15
0.20
eps
variable
Bonferroni
control BFDR
Benjamini-Hochberg
Bayes c0=cA=1
Bayes c0=2, cA=1
Bayes c0=1, cA=2
modified Benjamini-Hochberg





Comments:

- According to the graphs, cost values (1, 2, 3) Bonferroni has the highest value and BFDR, BH follows it. Value also increases when q, m and n increase.

- FDR for Bonferroni decreases and gets the lowest when epsilon increases. BH and BH mod have the highest value for FDR and the value gets higher a bit every time epsilon increases. Also, BH-mod value decreases a bit. It is the same for every q, n, m values.
- Power for Bonferroni is the lowest. For all bayes values, power value increases when epsilon increases and they have the highest values. For higher q, n, m values, all bayes values starts to decrease when epsilon increases.