R-Output Presenattion

Chenghui Zheng

2024-05-31

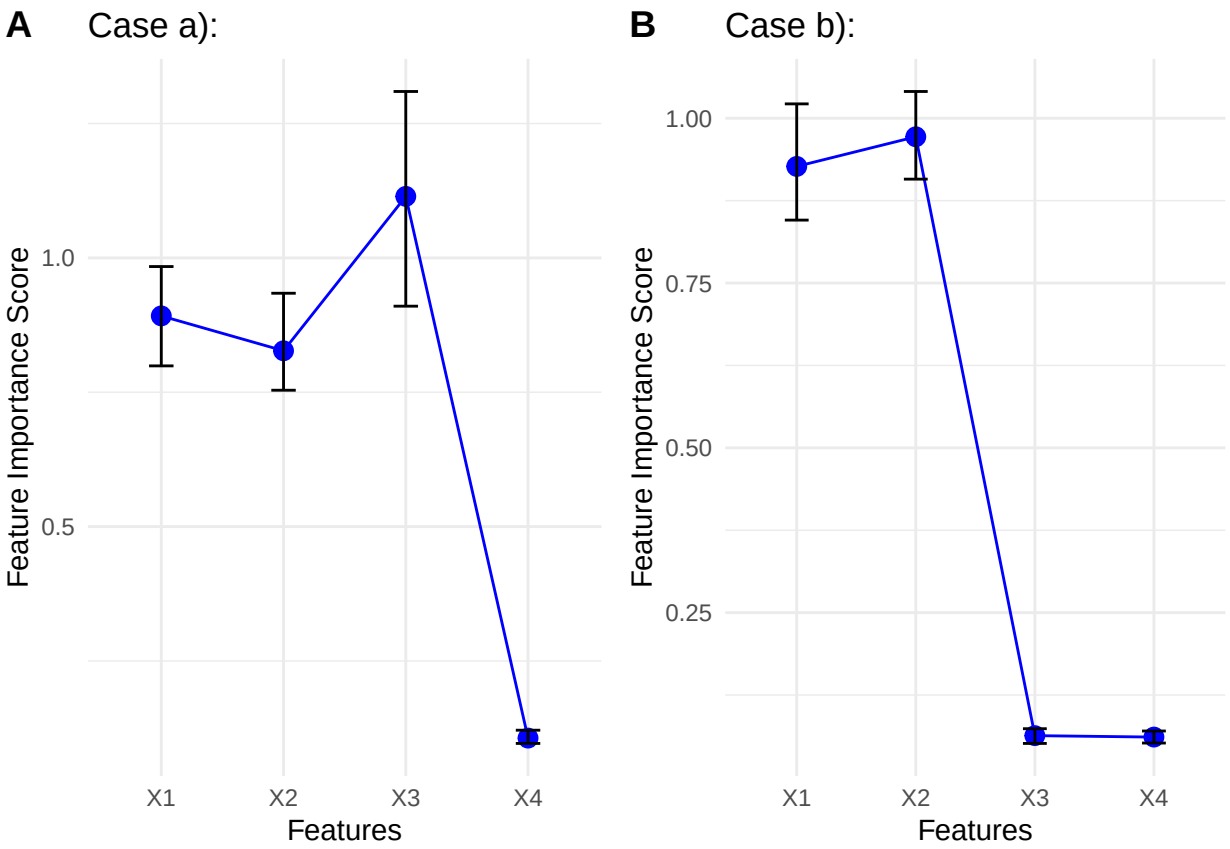
(a) $Y_1 \sim X_1 + X_2 + X_3 + \epsilon$, where X_i are mutually independent.

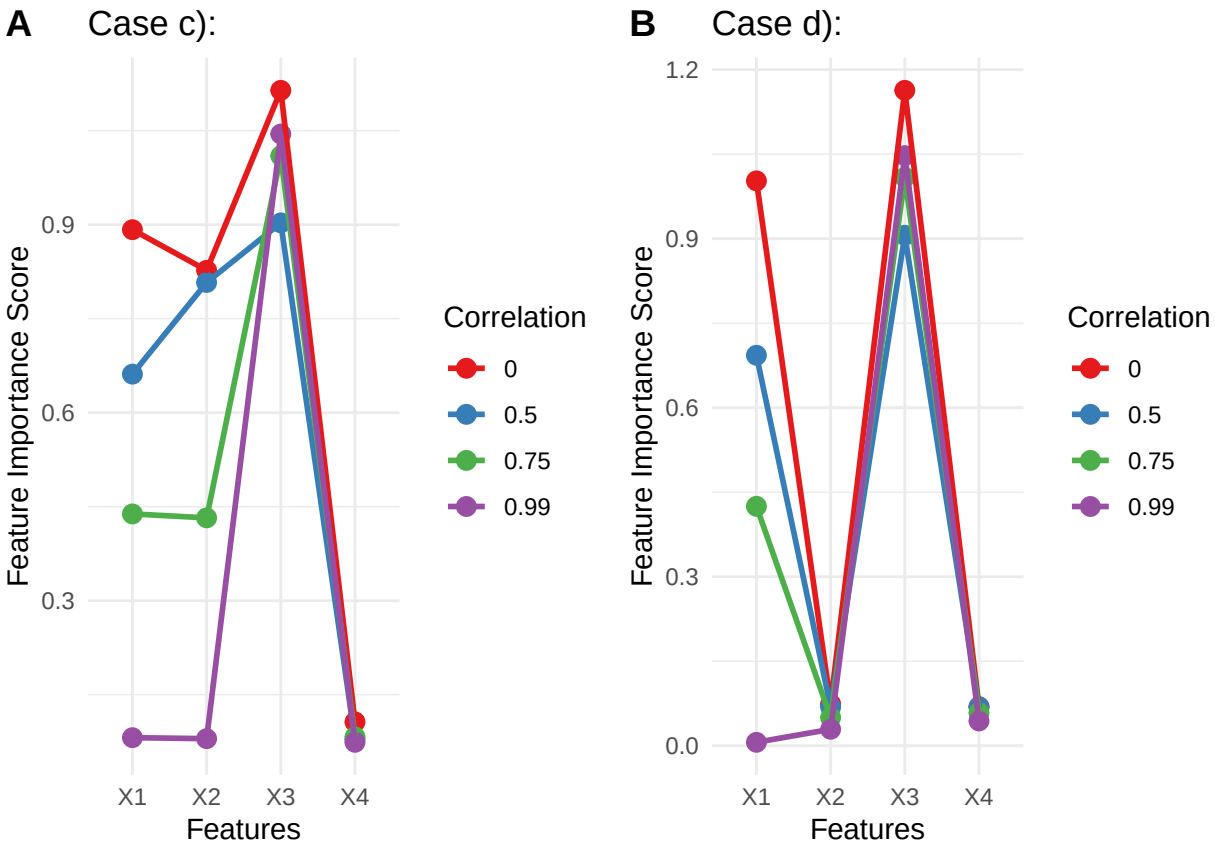
(b) $Y_2 \sim X_1 + X_2 + 0.001X_3 + \epsilon$, where X_i are mutually independent.

(c) $Y_3 \sim X_1 + X_2 + X_3 + \epsilon$, where $X_1 \not\perp\!\!\!\perp X_2$.

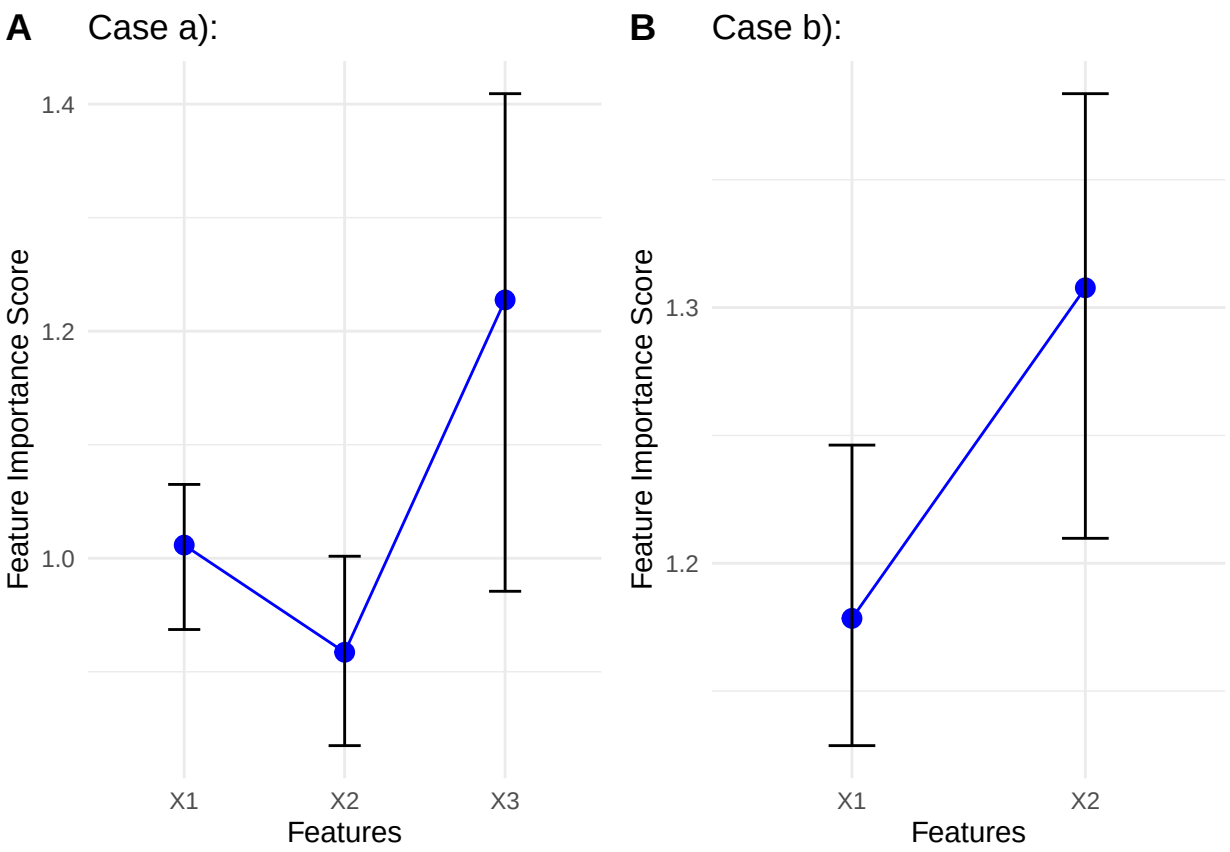
(d) $Y_4 \sim X_1 + X_3 + \epsilon$, where $X_1 \not\perp\!\!\!\perp X_2$.

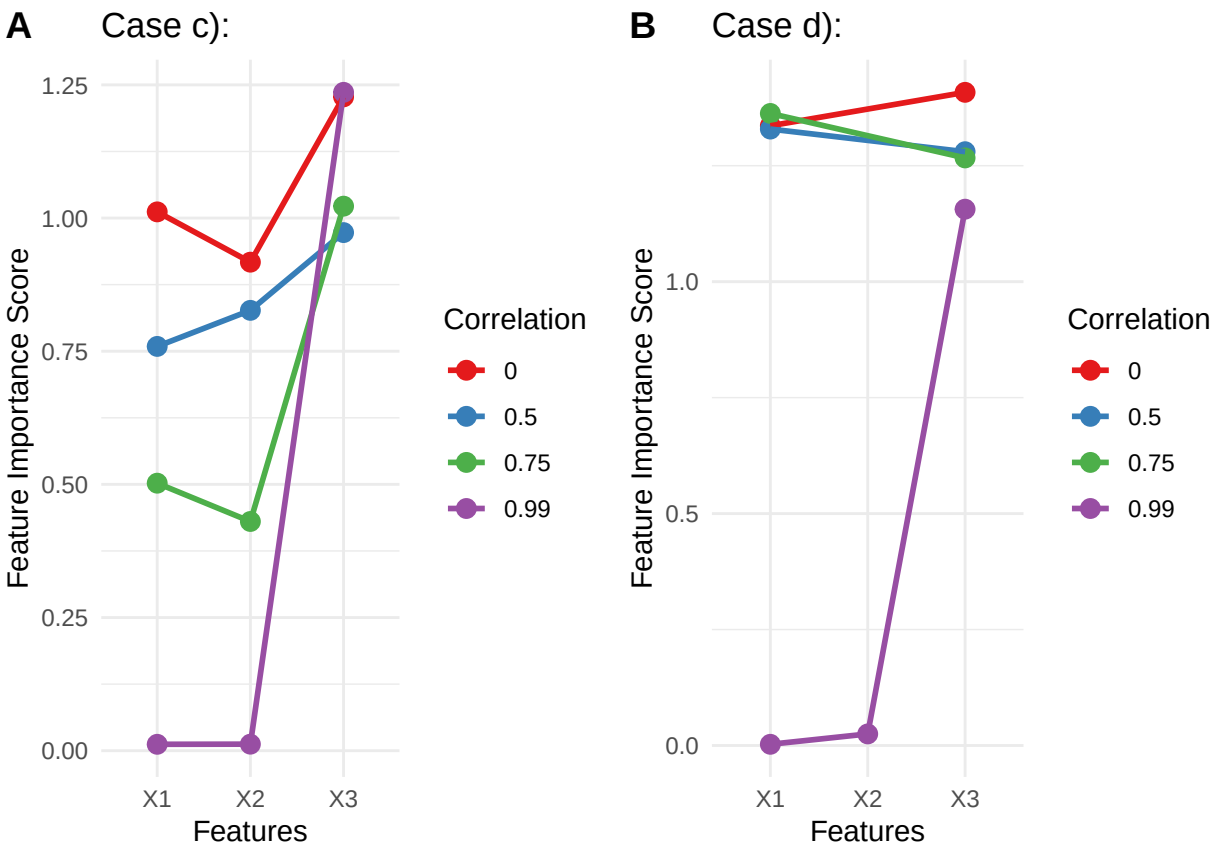
LOCO only



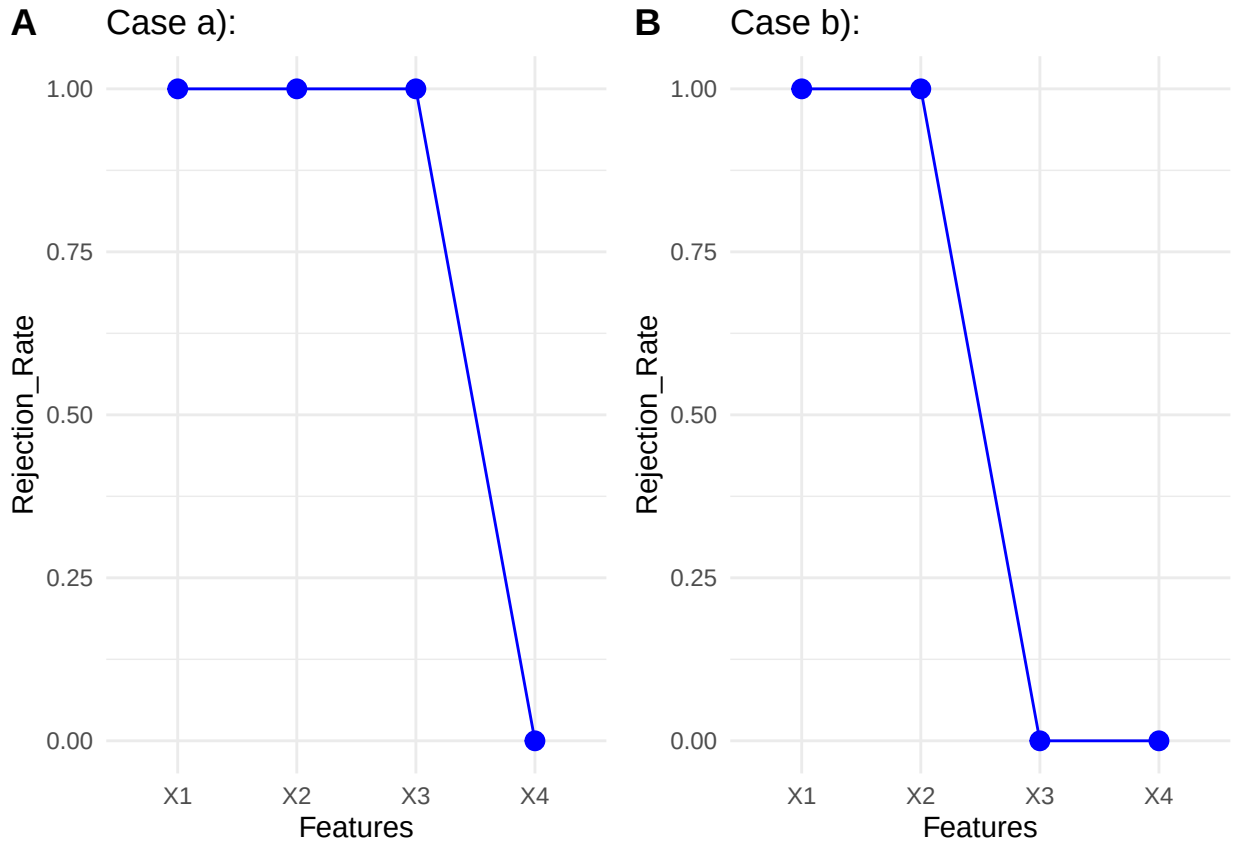

A
Case c):
Feature Importance Score
0.9
0.6
0.3
X1
X2
X3
X4
Features
Correlation
0
0.5
0.75
0.99
B
Case d):
Feature Importance Score
1.2
0.9
0.6
0.3
0.0
X1
X2
X3
X4
Features
Correlation
0
0.5
0.75
0.99

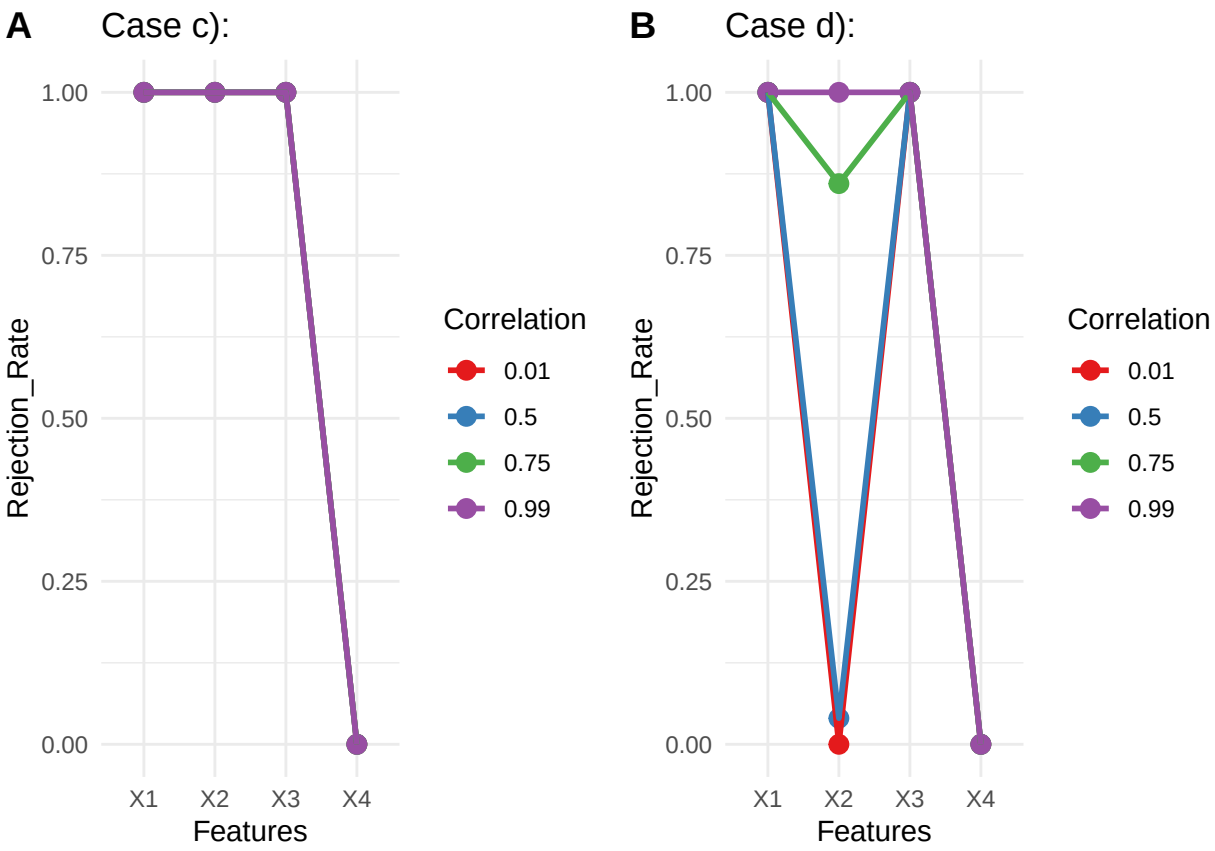
GCM filter first + LOCO




A
Case c):
1.25
1.00
0.75
0.50
0.25
0.00
Feature Importance Score
X1
X2
X3
Features
Correlation
0
0.5
0.75
0.99
B
Case d):
1.0
0.5
0.0
Feature Importance Score
X1
X2
X3
Features
Correlation
0
0.5
0.75
0.99

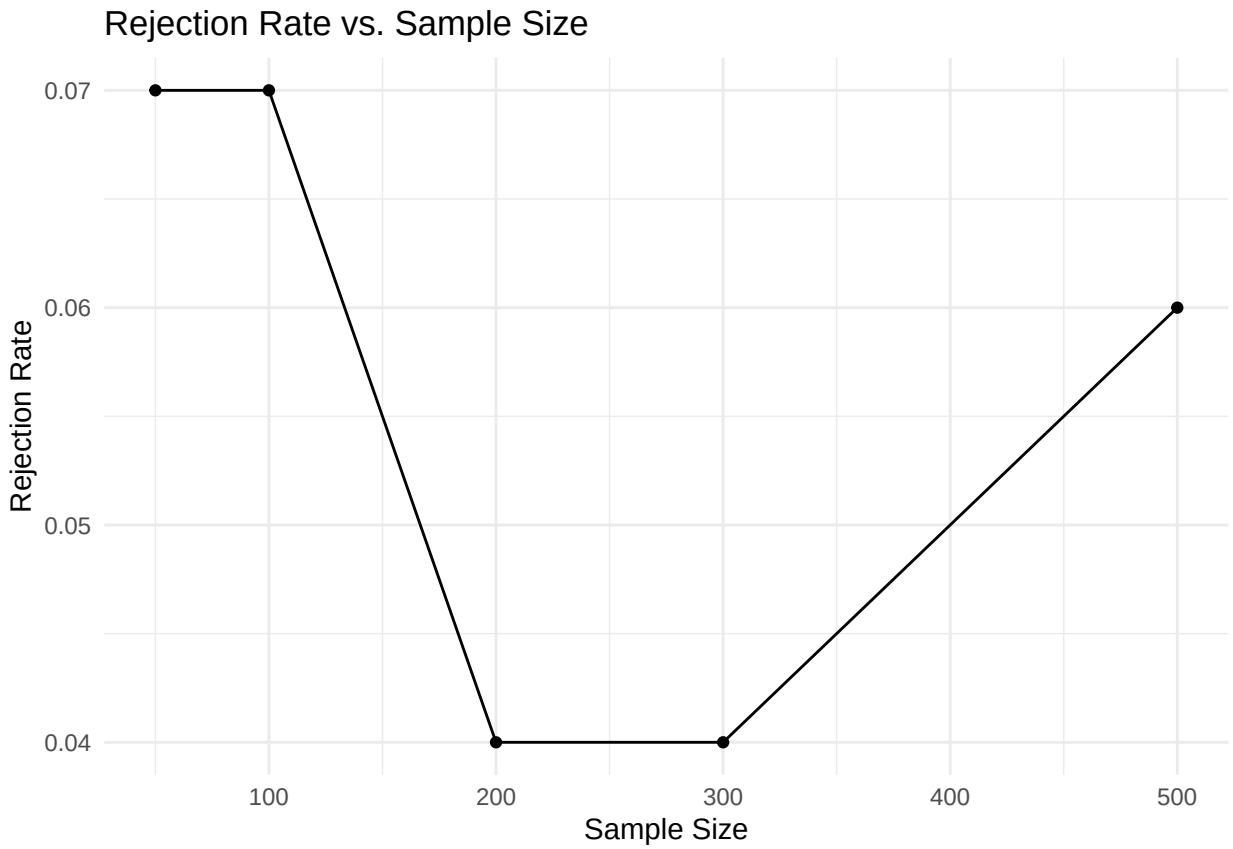
GCM only


A
Case a):
1.00
0.75
0.50
0.25
0.00
Rejection_Rate
X1
X2
X3
X4
Features
B
Case b):
1.00
0.75
0.50
0.25
0.00
Rejection_Rate
X1
X2
X3
X4
Features

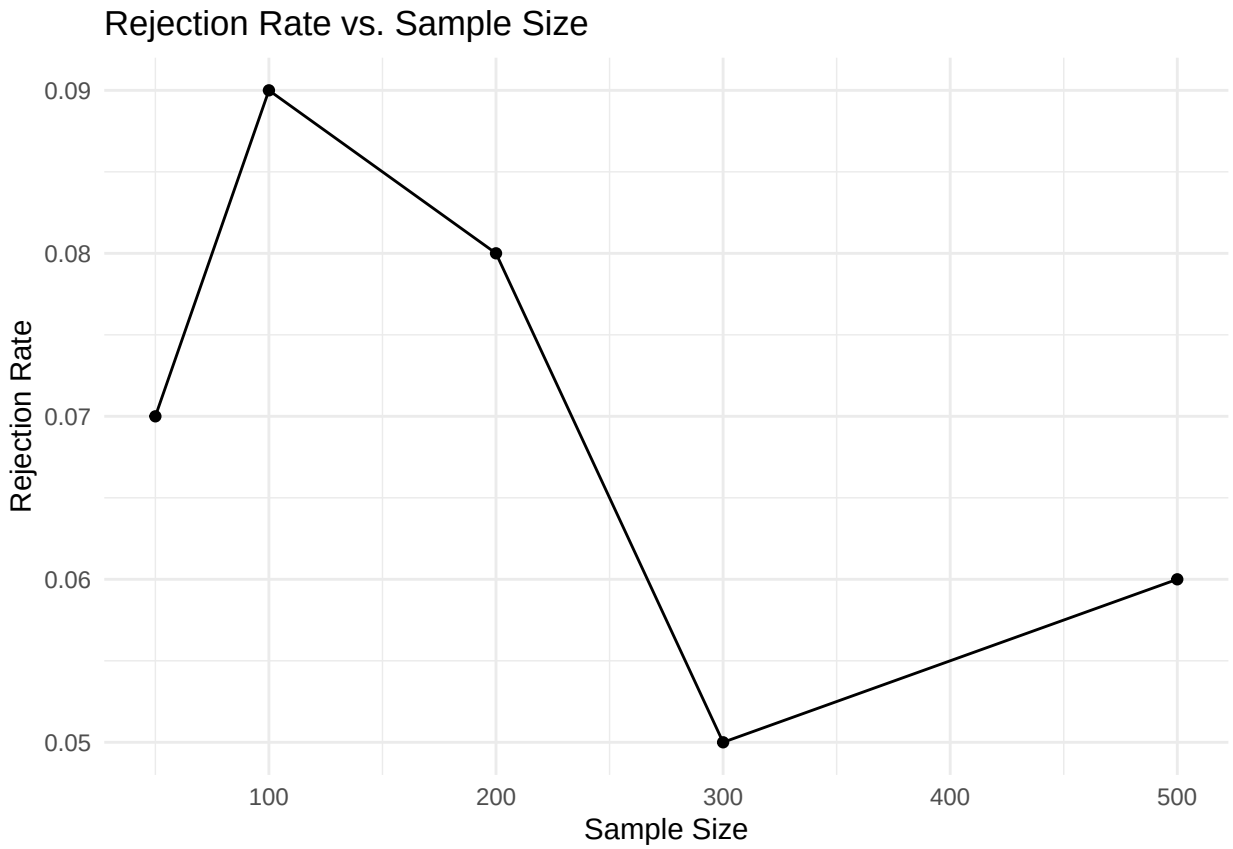


####################GCM Simulation##################################

$a) Z \sim N(0,1), X = 2 * sin(Z) + 0.1 * N(0,1), Y = 2 * sin(Z) + 0.1 * N(0,1)$



$b) Z \sim N(0,1), X = 2 * Z + 0.1 * N(0,1), Y = 2 * Z + 0.1 * N(0,1)$


Rejection Rate vs. Sample Size
Rejection Rate
0.09
0.08
0.07
0.06
0.05
100
200
300
400
500
Sample Size