CS4341: Assignment #3

Gordon Gao, Thomas Lourenco, Kevin Ouellette, Monika Danielewicz

Group 16

# Part 1: Warmup

Given the following probabilities:

P(good quiz grades) = 0.8

P(pass) = 0.9

P(good quiz grades | pass) = 0.8

P(good quiz grades | pass) = P(good quiz grades ^ pass) / P(pass)

P(good quiz grades ^ pass) = P(good quiz grades | pass) \* P(pass)

P(good quiz grades ^ pass) = P(pass ^ good quiz grades) = 0.8 \* 0.9

**P(pass ^ good quiz grades) = 0.72**

P(pass | good quiz grades) = P(pass ^ good quiz grades) / P(good quiz grades)

P(pass | good quiz grades) = 0.72 / 0.8

**P(pass | good quiz grades) = 0.9**

# Part 2: Rejection Sampling

For rejection sampling we created the 4 test cases listed below and created graphs for the estimated probability and confidence intervals out of the data our program created. From the graphs we can determine that the first three test cases stabilize around 1,000 iterations, but the fourth tests case stabilizes around 10,000,000. The time taken to execute the code for 1,000 iterations is 0.008 seconds, while the time taken for 1,000,000 is 6.30407 seconds.

For the confidence interval, we only show the absolute margin of error (AME) based on the equation:

95% Confidence Interval = Mean +/- AME

where

AME = 2\*SD/sqrt(N)

For each of the test cases listed below, we created trend lines based on the data in the AME graphs below. Reversing the equations of these lines allowed us to calculate the estimated number of sample needed to get 90% confidence intervals of 0.2, 0.1, 0.05, and 0.01. These are listed in Table 1. The trendline equations used to get this data are also shown below.

## Test Cases

No Observed Nodes: sample exams=true [iterations]

2 Observed Nodes: sample snow=false [iterations] stress=high icy=true

4 Observed Node: sample icy=true [iterations] humidity=high day=weekday cloudy=false

exams=true

All Observed Nodes: sample stress=high [iterations] humidity=medium temp=mild icy=false

snow=true day=weekday exams=false cloudy=true

## Tables and Graphs

X-axis is scaled logarithmically in the graphs.

In these equations, y is AME and x is the number of iterations. We solved for x to obtain the estimated number of iterations necessary.

TC1: y = 0.6131 \* x-0.502

TC2: y = 9.4864 \* x-0,503

TC3: y = 10.395 \* x-0.542

TC4: y = 7.848 \* x-0.292

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Estimated Number of Iterations to reach 95% Confidence | | | | |
| AME for 95% | TC1: Iterations | TC2: Iterations | TC3: Iterations | TC4:Iterations |
| 0.2 | 10 | 2149 | 1465 | 287125 |
| 0.1 | 38 | 8524 | 5262 | 3083156 |
| 0.05 | 148 | 33814 | 18901 | 33107018 |
| 0.01 | 3638 | 829261 | 368212 | 8196962611 |

Table 1: Estimated Number of Iterations to reach 95% Confidence

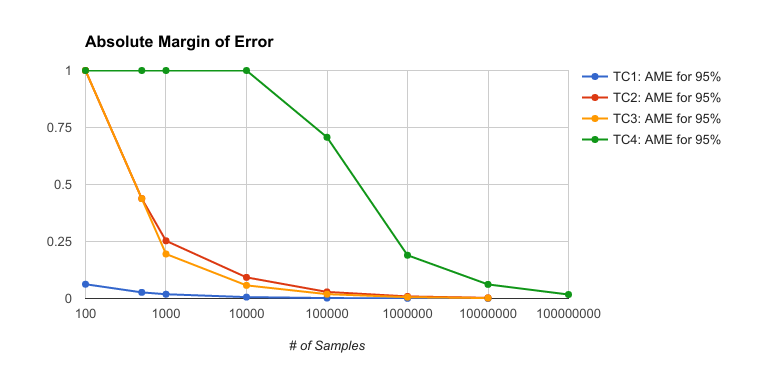
## 

## 

## 

## Chart

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Estimated Probability | | | | |
| # of Samples | TC1: No Observed Nodes | TC2: 2 Observed Nodes | TC3: 4 Observed Nodes | TC4: All Observed Nodes |
| 100 | 0.11 | 0 | 0 | 0 |
| 500 | 0.1 | 0.6 | 0.4 | 0 |
| 1000 | 0.097 | 0.2 | 0.1875 | 0 |
| 10000 | 0.1021 | 0.24419 | 0.16363 | 0 |
| 100000 | 0.10098 | 0.23005 | 0.15957 | 0.5 |
| 1000000 | 0.10171 | 0.23853 | 0.15635 | 0.1 |
| 10000000 | 0.10129 | 0.23627 | 0.15763 | 0.10309 |
| 100000000 |  |  |  | 0.08241 |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Absolute Margin of Error | | | | |
| # of Samples | TC1: AME for 95% | TC2: AME for 95% | TC3: AME for 95% | TC4: AME for 95% |
| 100 | 0.06258 | 1 | 1 | 1 |
| 500 | 0.02683 | 0.43818 | 0.43818 | 1 |
| 1000 | 0.01872 | 0.25298 | 0.19516 | 1 |
| 10000 | 0.00606 | 0.09265 | 0.0576 | 1 |
| 100000 | 0.00191 | 0.02884 | 0.01904 | 0.70711 |
| 1000000 | 0.0006 | 0.00911 | 0.00591 | 0.18974 |
| 10000000 | 0.00019 | 0.00289 | 0.00189 | 0.06175 |
| 100000000 |  |  |  | 0.01744 |