ELEN90054 Probability and Random Models

MATLAB Workshop 5 Hypothesis Testing; week 10 = week of 7 May 2018

Prepare this workshop individually, in particular questions Q1-Q6, before you come to your workshop session.

This workshop is worth 5% of the overall subject assessment and should be completed in pairs. Allocation will be random by your demonstrator, so you should not choose your project partner yourself. The next workshop will have a new allocation. Be aware that seeking or providing detailed assistance from/to people other than your workshop partner is collusion - see
http://academichonesty.unimelb.edu.au/plagiarism.html.

Each group is expected to upload two files:

1. A pdf (scanned/typed) file containing their worked solutions;
   a. Only one member of the group needs to upload the pdf file.
   b. The naming convention of the file should include workshop number, day, time and the assigned group number, e.g. Workshop_2_Mon09_Gp5.pdf for group 5 in the Monday 0900 hrs workshop slot.
   c. You also need to **read and attach** the Engineering cover-sheet (see the ELEN90054 LMS site Workshops) signed by both of you.
2. A single zip file, containing all the required functions and **a single main script** which calls these functions to generate the required outputs as outlined in the workshop questions.
   a. Only one member of the group needs to upload the zip file.
   b. The following naming convention should strictly be followed:
   Workshop_2_Mon12_Gp5_Matlab.zip for group 5 in the Monday 1200 hrs workshop slot.

The workshop times are: Mon09, Mon12, Tue10, Tue17, Wed13, Wed17, Wed18, Wed19, Thu11, Thu19, Fri10, Fri13. The group numbers are randomly assigned by your demonstrator at the start of the workshop.

Both submissions should be made before the start of the next workshop of week 11. **This is a strict deadline**.

In special circumstances you can email the pdf/code to your demonstrator. For demonstrator email address information, see "staff info" on the ELEN90054 LMS site.

Note: The main purpose of this workshop is to theoretically derive decision rules and error probabilities for hypothesis testing and then simulate their performance via simulations.

Hypothesis Testing

(total=36 marks + 4 on-time attendance marks)

The Open Sesame company makes swipe-card readers for controlling access to buildings. Under normal operating conditions, there is a small probability of 0.02 that swiping a card fails to open the entrance, even with no fault in the reader (e.g. due to noise in the reader's circuits and random variations in the speed of each swipe-through). However, some of these readers have a particular fault. If this fault occurs, then the probability of swipe failure rises to 0.1.

From exhaustive tests at its factory, Open Sesame estimates that 22% of all the readers it has sold so far have this fault. It decides to send technicians around to test all the readers it has sold, and replace the faulty ones. You are tasked with designing a rule to help the technicians decide whether or not a card reader is faulty. They are to test each reader by swiping a card through it repeatedly in successive independent trials, until a failure occurs on the N-th trial. Based on N, the technician must decide whether or not the tested reader is faulty.

Questions

NOTE: To get full marks in Q1-6, all decision rules must be written as a range of values of N corresponding to each hypothesis, e.g. "decide not faulty if $N \geq \ldots$, otherwise decide faulty"

Q1) (2 marks) Define the two main random variables of this inference problem, as well as the mathematical formula for the conditional distribution that is described in words in the text above.

Q2) (2 marks) Find the decision rule that minimises the conditional probability of declaring a reader to be faulty given that it is not faulty. What is this probability called, P_{FA} or P_{MISS}?

Q3) (4 marks) Find the maximum likelihood (ML) decision rule.

Q4) (4 marks) Find the maximum a posterior (MAP) decision rule. Is your MAP decision threshold different from your ML threshold? If yes, explain why. If no, explain why they are the same.

Now assume the following:

- every reader that is **incorrectly** declared faulty costs Open Sesame \$10 to replace. Replacing units which are indeed faulty costs nothing, due to a subsidy from the factory where they are built.
- every reader that is **incorrectly** declared to be not faulty costs them \$50, due to the cost of a return visit when the affected customer eventually complains, as well as the damage to their reputation.

Q5) (7 marks) Find the decision rule that minimises the expected cost per test. Also find the corresponding values of the expected cost, the miss rate P_{miss} and the false alarm rate P_{FA}. Is your minimum-cost decision threshold different from your MAP threshold? If yes, explain why. If no, explain why they are the same.

Q6) (7 marks) Find the decision rule that minimises the false alarm rate while ensuring a miss rate of no more than 2%. Also find the corresponding values of the expected cost per test, the miss rate P_{miss} and the false alarm rate P_{FA}. Briefly explain why your decision threshold(s) might be different from the minimum-cost threshold(s).

Q7) (10 marks) Using MATLAB:

a) Write a function which simulates a single test and returns the correct indicator of whether the reader was faulty or not, as well as the decisions made by the methods from Q5 and Q6.

b) Using the previous function, simulate 2,000,000 tests. For each of the two decision rules display empirical values of the expected cost, the miss rate and the false alarm rate. Discuss whether your results are consistent with Q5 and Q6.

End of MATLAB Workshop 5 Questions

Q1) (2 marks) Define the two main random variables of this inference problem, as well as the mathematical formula for the conditional distribution that is described in words in the text above.

Priors

H_1 is the hypothesis that reader is faulty

$$P(R=1) = P(H_1) = 0.22$$

H_0 is the hypothesis that reader is not faulty

$$P(R=0) = P(H_0) = 0.78$$

Likelihoods

$N|H_1 \sim \text{geometric}(0.1)$

$$P_{N|H_1}(N=n) = (0.9)^{n-1} 0.1$$

PMF of number of trials until a swipe failure occurs given that the reader is faulty

$N|H_0 \sim \text{geometric}(0.02)$

$$P_{N|H_0}(N=n) = (0.98)^{n-1} 0.02$$

PMF of number of trials until a swipe failure occurs given that the reader is not faulty

Accept that the reader is not faulty if the probability of the posterior $P(H_0|N=n) \geq P(H_1|N=n)$,

$$P(H_0|N=n) \geq P(H_1|N=n)$$

$$\frac{P(N=n|H_0)\,P(H_0)}{P(N=n)} \geq \frac{P(N=n|H_1)\,P(H_1)}{P(N=n)}$$

$$P(N=n|H_0)P(H_0) \geq P(N=n|H_1)\,P(H_1)$$

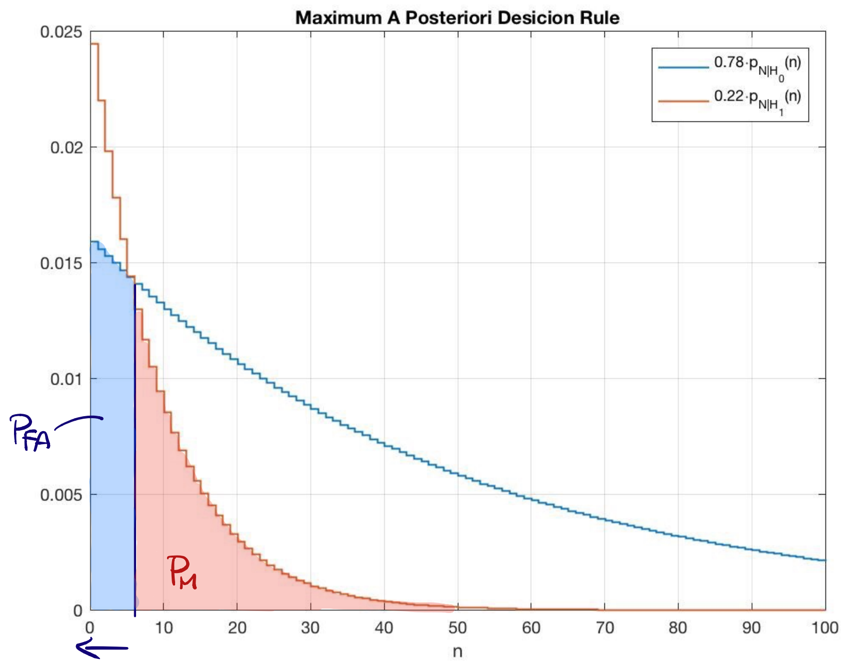
$$\frac{P(N=n|H_0)\,P(H_0)}{P(N=n|H_1)\,P(H_1)} \geq 1$$

Q2) (2 marks) Find the decision rule that minimises the conditional probability of declaring a reader to be faulty given that it is not faulty. What is this probability called, P_{FA} or P_{MISS}?

The conditional probability of declaring a reader to be faulty given that it is not faulty is P_FA, which is the probability of a False Alarm.

To minimise the probability of a False Alarm, always declare that the reader is not faulty regardless of the number of swipes before a swipe failure occurs. In this way, the probability of a false alarm will never occur, and as such, it is minimised to zero.

Declare that the reader is NOT FAULTY if reader hasn't failed after N ≥ 0 swipes.



To minimise P_FA, move decision rule threshold to 0!

Q3) (4 marks) Find the maximum likelihood (ML) decision rule.

Declare reader is not faulty iff likelihood functions $P_{N|H_0}(n) \geq P_{N|H_1}(n)$,

$$\frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq 1$$

$$\frac{(0.98)^{n-1}\,0.02}{(0.9)^{n-1}\;0.1} \geq 1$$

$$0.2\,\frac{(0.98)^{n-1}}{(0.9)^{n-1}} \geq 1$$

$$(1.09)^{n-1} \geq 5$$

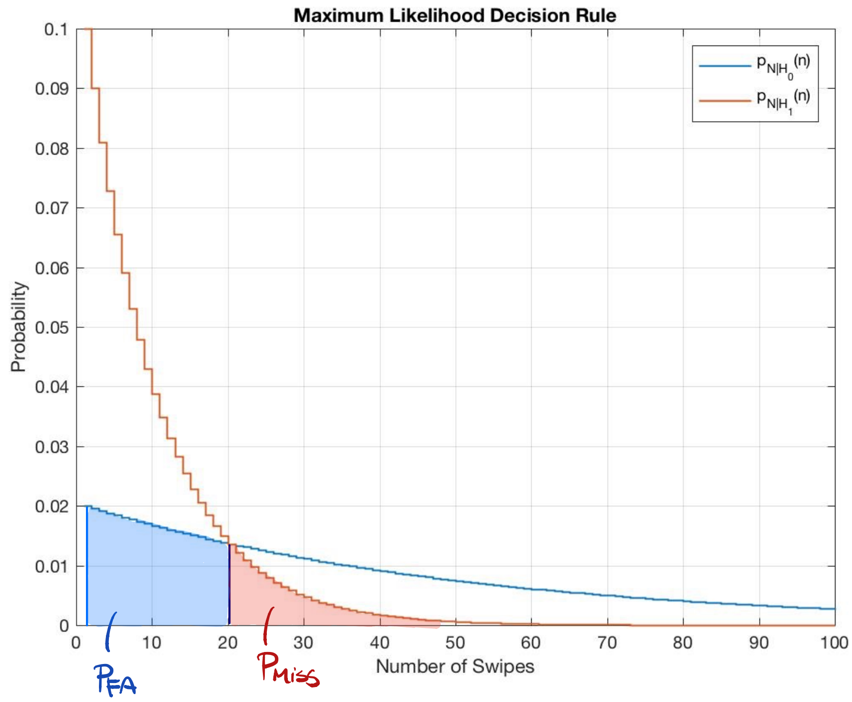
$$(n-1)\ln(1.09) \geq \ln(5)$$

$$n \geq \frac{\ln(5)}{\ln(1.09)} + 1$$

$$n \geq 19.899$$

n has to be the next integer up,

$$\therefore n \geq 20$$



Declare that the reader is NOT FAULTY if reader hasn't failed after N ≥ 20 swipes, otherwise, declare that the reader is faulty.

Q4) (4 marks) Find the maximum a posterior (MAP) decision rule. Is your MAP decision threshold different from your ML threshold? If yes, explain why. If no, explain why they are the same.

Declare reader is not faulty iff $P(H_0|N=n) \geq P(H_1|N=n)$,

$$\frac{P(H_0|N=n)}{P(H_1|N=n)} \geq 1$$

$$\frac{P_{N|H_0}(n) \cdot P(H_0)}{P_{N|H_1}(n) \cdot P(H_1)} \geq 1$$

$$\frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq \frac{P(H_1)}{P(H_0)}$$

$$\frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq 0.282$$

$$\frac{(0.98)^{n-1}(0.02)}{(0.9)^{n-1}(0.1)} \geq 0.282$$

$$\left(\frac{0.98}{0.9}\right)^{n-1} \geq 1.41$$

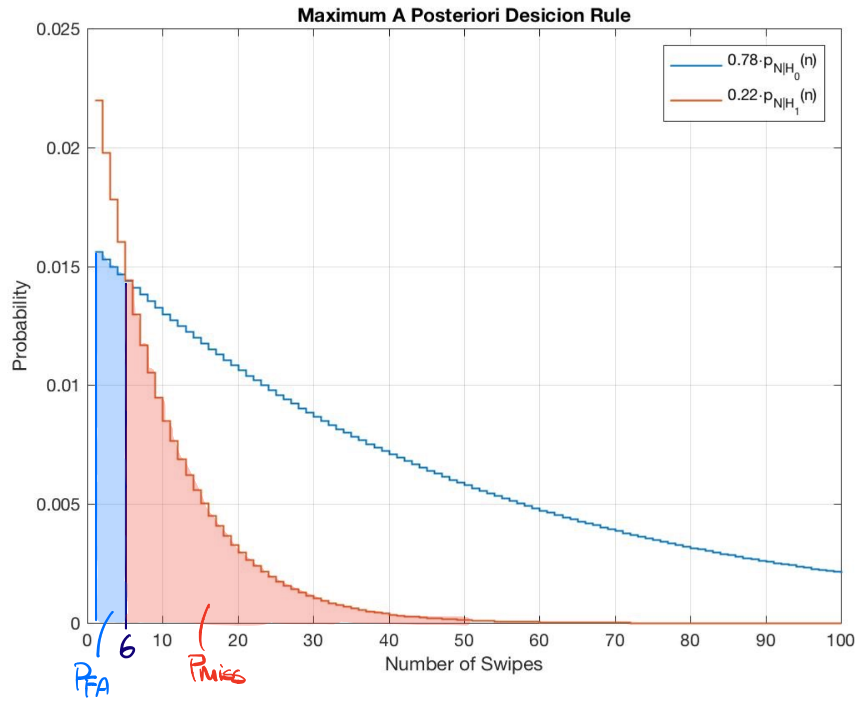
$$(n-1)\ln\left(\frac{0.98}{0.9}\right) \geq \ln(1.41)$$

$$n \geq \frac{\ln(1.41)}{\ln\left(\frac{0.98}{0.9}\right)} + 1$$

$$n \geq 5.0369$$

n has to be the next integer up,

$\therefore n \geq 6$



Declare that the reader is NOT FAULTY if reader hasn't failed after N ≥ 6 swipes, otherwise, declare that the reader is faulty.

The Maximum A Posteriori decision threshold is different to the Maximum Likelihood threshold because it takes into account the probabilities of the priors P[H0] and P[H1].

$P_{N|H_0}(n) \cdot P(H_0) \geq P_{N|H_1}(n) \cdot P(H_1)$

where $P(H_0) = 0.78$

$P(H_1) = 0.22$

Likelihood functions are weighted by the probability of the prior.

Q5) (7 marks) Find the decision rule that minimises the expected cost per test. Also find the corresponding values of the expected cost, the miss rate P_{miss} and the false alarm rate P_{FA}. Is your minimum-cost decision threshold different from your MAP threshold? If yes, explain why. If no, explain why they are the same.

Declare reader is not faulty iff $C_{FA} \cdot P(H_0|N=n) \geq C_M \cdot P(H_1|N=n)$

$C_{miss} = \$50$

$C_{FA} = \$10$

$$\frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq \frac{C_M \cdot P(H_1)}{C_{FA} \cdot P(H_0)}$$

$$\frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq 1.41$$

$P_{N|H_0}(n) = (0.98)^{n-1}\,0.02$ $\quad$ $P_{N|H_1}(n) = (0.9)^{n-1}\,0.1$

$F_{N|H_0}(n) = 1-(0.98)^n$ $\quad$ $F_{N|H_1}(n) = 1-(0.9)^n$

$$\frac{0.02\,(0.98)^{n-1}}{0.1\,(0.9)^{n-1}} \geq 1.41$$

$$\left(\frac{0.98}{0.9}\right)^{n-1} \geq 7.0513$$

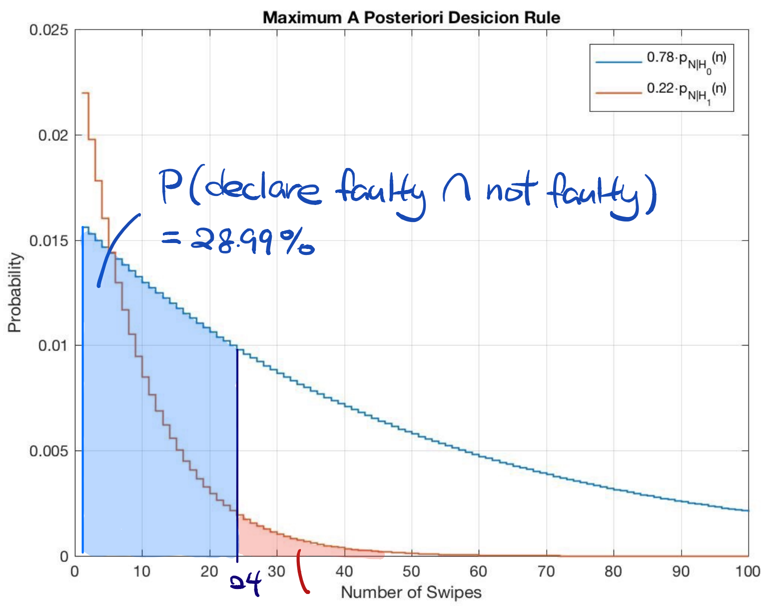
$$(n-1)\ln\left(\frac{0.98}{0.9}\right) \geq \ln(7.0513)$$

$$n \geq \frac{\ln(7.0513)}{\ln\left(\frac{0.98}{0.9}\right)} + 1$$

$$n \geq 23.936$$

n has to be the next integer up,

$\therefore n \geq 24$



P(declare not faulty ∩ faulty)
= 1.95%

Declare that the reader is NOT FAULTY if reader hasn't failed after N ≥ 24 swipes, otherwise, declare that the reader is faulty.

$P_{FA} = P(\text{declare faulty}\,|\,\text{not faulty})$

$= P_{N|H_0}(N<24|H_0)$

$= F_{N|H_0}(23)$

$= 1-(0.98)^{23}$

$= 0.3717$

$P_{miss} = P(\text{declare not faulty}\,|\,\text{faulty})$

$= P_{N|H_1}(N \geq 24|H_1)$

$= 1 - F_{N|H_1}(23)$

$= (0.9)^{23}$

$= 0.0886$

$P(\text{declare faulty} \cap \text{not faulty})$

$= P(\text{declare faulty} \mid \text{not faulty}) \cdot P(\text{not faulty})$

$= P_{N|H_0}(N<24 \mid H_0)\, P(H_0)$

$= 0.3717\,(0.78)$

$= 0.2899$

$P(\text{declare not faulty} \cap \text{faulty})$

$= P(\text{declare not faulty} \mid \text{faulty}) \cdot P(\text{faulty})$

$= P_{N|H_1}(N \geq 24 \mid H_1)\, P(H_1)$

$= 0.0886\,(0.22)$

$= 0.0195$

Minimum expected cost per test of declaring a reader is faulty but it is not faulty,

$P(\text{declare faulty} \cap \text{not faulty}) \cdot C_{FA}$

$= P_{N|H_0}(N<24 \mid H_0)\, P(H_0) \cdot C_{FA}$

$= \$\, 0.975$

Minimum expected cost per test of declaring a reader is not faulty but it is faulty,

$P(\text{declare not faulty} \mid \text{faulty}) \cdot C_{miss}$

$= P_{N|H_1}(N \geq 24 \mid H_1)\, P(H_1)\, C_{Miss}$

$= \$\, 2.899$

Total minimum expected cost per test,

$P(\text{declare not faulty} \mid \text{faulty}) \cdot C_{miss} + P(\text{declare faulty} \cap \text{not faulty}) \cdot C_{FA} = \$\, 3.874$

The minimum cost decision threshold is different to that of the Maximum A Posteriori threshold because it takes into account the cost of a Miss and a False Alarm.

$C_{FA} \cdot P(H_0 \mid N=n) \geq C_M \cdot P(H_1 \mid N=n)$

where $C_{FA} = \$10$

$C_M = \$50$

Posteriors are weighted by the cost of a false alarm and miss.

Because P[H1 | N=n] weights more, it becomes more important to detect faulty readers, hence it only make sense that the probability of Missing a faulty reader is reduced.

Q6) (7 marks) Find the decision rule that minimises the false alarm rate while ensuring a miss rate of no more than 2%. Also find the corresponding values of the expected cost per test, the miss rate P_{miss} and the false alarm rate P_{FA}. Briefly explain why your decision threshold(s) might be different from the minimum-cost threshold(s).

$P_{miss} \leq 0.02$

$(0.9)^n \leq 0.02$

This is a negative number

$n \ln(0.9) \leq \ln(0.02)$

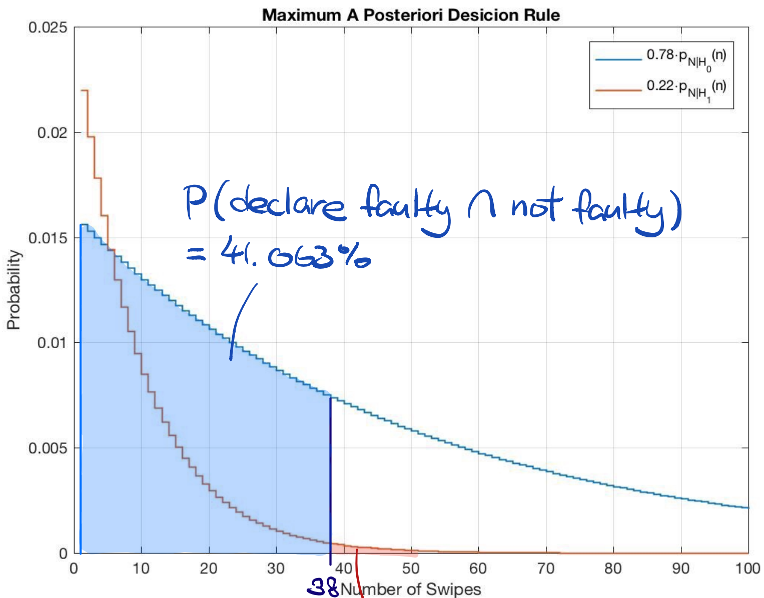
$\therefore n \geq \dfrac{\ln(0.02)}{\ln(0.9)}$

$n \geq 37.13$

n has to be the next integer up,

$\therefore n \geq 38$

Declare that the reader is NOT FAULTY if reader hasn't failed after N ≥ 38 swipes, otherwise, declare that the reader is faulty.



$P(\text{declare not faulty} \cap \text{faulty})$
$= 0.446\%$

The decision threshold is now different to the minimum-cost threshold because the False Alarm rate is now to be kept at minimum given that the Miss rate cannot exceed 2%. As such, the decision threshold has to move to the right to keep probability of Miss ≤2%, thereby increasing the probability of False Alarm.

Using $n^* = 38$,

$P_{FA} = P_{N|H_0}(N<38 \mid H_0)\, P(H_0)$

$= 1-(0.98)^{37}$

$= 0.5265$

Expected cost per test of declaring a reader is faulty but it is not faulty,

$P_{N|H_0}(N<38 \mid H_0)\, P(H_0)\, C_{FA} = \$0.223$

$P_{miss} = P_{N|H_1}(N \geq 38 \mid H_1)$

$= (0.9)^{37}$

$= 0.0203$

Expected Cost per test of declaring a reader is not faulty but it is faulty,

$P_{N|H_1}(N \geq 38 \mid H_1)\, P(H_1)\, C_{Miss} = \$4.1063$

Total expected cost per test,

$P_{N|H_0}(N<38 \mid H_0)\, P(H_0)\, C_{FA} + P_{N|H_1}(N \geq 38 \mid H_1)\, P(H_1)\, C_{Miss}$

$= \$4.3294$

Q7) (10 marks) Using MATLAB:

a) Write a function which simulates a single test and returns the correct indicator of whether the reader was faulty or not, as well as the decisions made by the methods from Q5 and Q6.

b) Using the previous function, simulate 2,000,000 tests. For each of the two decision rules display empirical values of the expected cost, the miss rate and the false alarm rate. Discuss whether your results are consistent with Q5 and Q6.

The two parts of question 7a are executed on MATLAB.

Please kindly note that the simulation for the two parts of question 7b can take a while to finish its execution because 2 million trails are being executed.
The results of the two parts of question 7b is as follows:

```
Decision Rule:
Declare reader is not faulty if N ≥ 24
Otherwise, reader is faulty.
```

Probability_of	Empirical_Value	Theoretical_Value
'False Alarm'	'37.2227%'	'37.1653%'
'Miss'	'8.8089%'	'8.8629%'
[]	[]	[]
'Expected cost per test of False Alarm'	'$0.9687'	'$0.97492'
'Expected cost per test of Miss'	'$2.9036'	'$2.8989'
'Total Expected cost per test'	'$3.8723'	'$3.8738'

All empirical probabilities obtained from simulation adheres with calculated theoretical probabilities

```
Decision Rule:
Declare reader is not faulty if N ≥ 38
Otherwise, reader is faulty.
```

Probability_of	Empirical_Value	Theoretical_Value
'False Alarm'	'52.6811%'	'52.6451%'
'Miss'	'2.0084%'	'2.0276%'
[]	[]	[]
'Expected cost per test of False Alarm'	'$0.22102'	'$0.22303'
'Expected cost per test of Miss'	'$4.1086'	'$4.1063'
'Total Expected cost per test'	'$4.3296'	'$4.3294'

All empirical probabilities obtained from simulation adheres with calculated theoretical probabilities