

Application of Monte-Carlo Simulation Towards a Better Understanding of Bayes' Theorem in Engineering Education

Rui Assis¹, Pedro Carmona Marques², Raphaela Vidal³

¹ElGeS - Research Centre in Industrial Engineering, Management and Sustainability, Lusófona University, 1749-024 Lisbon, Portugal (p1028@ulusofona.pt) ORCID [0000-0002-1697-055X](https://orcid.org/0000-0002-1697-055X);

²ElGeS - Research Centre in Industrial Engineering, Management and Sustainability, Lusófona University, 1749-024 Lisbon, Portugal (p4803@ulusofona.pt) ORCID [0000-0003-4891-1754](https://orcid.org/0000-0003-4891-1754);

³ElGeS - Research Centre in Industrial Engineering, Management and Sustainability, Lusófona University, 1749-024 Lisbon, Portugal (raphaela.vidal@ulusofona.pt) ORCID [0000-0002-9769-9166](https://orcid.org/0000-0002-9769-9166)

Abstract

Bayes' Theorem (BT) is treated in probability theory and statistics. The BT shows how to change the probabilities a priori in view of new evidence, to obtain probabilities a posteriori. With the Bayesian interpretation of probability, the BT is expressed as the probability of an event (or the degree of belief in the occurrence of an event) should be changed, after considering evidence about the occurrence of that event. Bayesian inference is fundamental to Bayesian statistics. An example of practical application of this theorem in Health Systems is to consider the existence of false positives and false negatives in diagnoses. At the Academy, the theme of BT is exposed almost exclusively in its analytical form. With this article, the authors intend to contribute to clarify the logic behind this theorem, and get students better understanding of its important fields of application, using three methods: the classic analytical (Bayesian inference), the frequentist (frequency inference) and the numerical simulation of Monte-Carlo. Thus, it intends to explain BT on a practical and friendly way that provides understanding to students avoiding memorizing the formulas. We provide a spreadsheet that is accessible to any professor. Moreover, we highlight the methodology could be extended to other topics.

Author Keywords. Bayes, Monte-Carlo Simulation, False Positives, False Negatives, Engineering Education, Computation.

Type: Research Article

 Open Access  Peer Reviewed  CC BY

1. Introduction

Bayes' Theorem (alternatively Bayes' law or Bayes' rule) has played since 18th century. It refers to a mathematical formula used to calculate the probability of an event based on the data about another event that has already occurred, which is called conditional probability ([Joyce 2021](#)).

Bayes' Theorem (BT) is still strongly popular due to its applications in many fields, helping people to make decisions under conditions of uncertainty ([Fienberg 1992](#)). In the domain of scientific knowledge disciplines, such as Physics and Chemistry, the first studies were essentially experimental, and the observed effects were later explained with the support of mathematical logic ([Jeffreys 1973](#)). The experiences carried out by its precursors are now routinely reproduced in the laboratories of academic institutions, to provide students with a better understanding of the phenomena and their laws that those disciplines develop ([Daveedu Raju et al. 2019](#)). In the field of Statistics, however, experimentation in academic

circles is not common; explained perhaps by the difficulty in obtaining appropriate didactic materials or by its inexistence, given the means and the necessarily long time to reproduce the experiences that the authors of the many laws of this discipline, certainly carried out to formulate them (Warren 2018).

Nowadays, these difficulties are overcome, since computers, equipped with numerical simulation programs, can easily and quickly reproduce the same experiments that those authors have carried out with a lot of effort, treat the resulting simulated data (samples), and verify the correction of those laws, for better understanding and satisfaction on the part of the students (Cook 2002; Gerow et al. 2018).

Teaching methods vary accordingly to each course's objectives. There are some teaching techniques used in engineering classes that involve lectures, demonstrating concepts, focusing on principles and/or on applications, making analogies and lastly, problem-solving sessions to increase students' capability of understanding and applying theory and solve problems. However, we have noticed, in Probabilities and Statistics discipline, BT is most taught in an analytical way causing that a large part of students chooses to memorize the formulas rather than understand the inherent logic - since they do not get the whole understanding (deductive logic).

The rapid technological advancements in recent years have led to the development of diverse tools and infrastructure of integrating science, education, and technology. This in turn has expanded the variety of novel methods for learning and communication. The use of computer numerical simulation can thus contribute decisively so that Statistics is no longer considered by most students to be something abstract and boring, being limited only to the description of its laws and carrying out exercises based on the consequent form. Simulations can provide powerful pedagogical tools to explain difficult statistical concepts in probability and statistics classes (Christou and Dinov 2010). The logic in the construction of numerical simulation models, reproducing statistical laws will contribute decisively to the improvement of cognition in their learning (Gigerenzer and Hoffrage 1995). Hence, the purpose of this study is to demonstrate how to use Monte-Carlo Simulation to get engineering students a better understanding of BT.

This paper is organized as follows: next section presents the materials and methods used, focusing on learning the Bayesian's Theorem. The following section – results – shows practical applications. Discussion highlights some important aspects from the practical application to students and final remarks are made.

2. Materials and Methods

To present a computational tool for applying Monte-Carlo Simulation, it is necessary to understand how active learning happens, exploring this subject. Additionally, it is important to understand how the probability and statistics discipline has been taught over time, to suggest a new and effective tool that help students to understand what BT all is about and what it is used for. With this aim, it was performed an online search about methodologies used in the discipline Probability and Statistics at Portuguese faculties of engineering, with special attention on BT. We have found eight high education institutions (FEUP-University of Oporto, ISEL-Polytechnic of Lisbon, Lusófona University, NOVA University of Lisbon, Polytechnic of Tomar, Técnico-University of Lisbon, UBI-University of Beira Interior, and UTAD-University of Alto Douro) that provide the content of Probability and Statistics discipline on their websites. Most of the courses do not highlight methodologies used and others have focus on teaching how to use software like Software R and STATISTICA. No innovative tool was

identified, what made us to believe that they use conventional methods and motivated us to develop a new approach using Monte-Carlo simulation, aiming to facilitate engineering students' understanding of BT. Thus, we intend to provide a computational tool to teach BT in engineering classes instead of a passive transfer of classic knowledge based on analytical methods. Besides, MS EXCEL is the only software tool needed which is widely available in highest educational facilities. A simulation tool allows to engage students in deep learning as opposed to merely memorizing - the act of memorization does not last long. Three different ways to use BT are presented: a traditional solution by Bayesian inference (Figure 1), followed by the frequentist method (Figure 2) and finally we apply the Monte-Carlo simulation method in MS EXCEL (Figure 3). The spreadsheets are programmed in such a way that enable readers to solve their own problems. The spreadsheet is available in a DROPBOX link.

2.1. Active learning

Any engineering course aims to prepare students towards a solid background. However, there is a sum of challenges to face daily at universities. Engineering professors find quite often bored and demotivated students that do not pay attention to classes and consequently achieve low tests grades. Nowadays, teachers must reinvent themselves, follow the technological tendencies of teaching (Auster and Wylie 2006; D'Andrea and Gosling 2005; Diamond, Koernig, and Iqbal 2008) and change their teaching style to match student's interest. The fast changes in the business world push the academia to adopt new ways to achieve learning's goals. In this dynamic context, students build their knowledge through an educational approach called "active learning" that places students at the center of the learning process, requiring an active participation of students on building their own knowledge. In that case, professors assume the role of supervisors on building knowledge and are open to its exchange. In fact, teaching tendencies mostly make use of computers since it is part of the daily activities at workplace (Konopka, Adaime, and Mosele 2015; Wang 2020). The use of MS EXCEL in practical classes help students to solve problems and improves the learning process in the case of engineering students. In addition, it enables to provide dynamics practical classes, increasing students' motivation.

2.2. Probability and statistics

Probability and Statistics is an essential subject to teach engineering students. It has become a vital tool to engineers (Johnson 2017). Statistics is the art of learning from the data (Ross 2009) providing the basis for making decisions or choosing actions. It is applied in so many areas such as manufacturing, development of food products, computer software, energy source and others (Walpole et al. 2012). Related to quality improvement, statistics plays a major role and engineers with basic statistical skills can be succeeded in attaining this goal (Johnson 2017). Different teaching approaches have been studied since the 70s concerning Probability and Statistics (Simon, Atkinson, and Shevokas 1976). Recently, research has shown interest to make mathematics closer to student's experience. They recommend the use the real-life context in teaching, turning it in a prominent method to use (Budimir 2016). Thus, the application of Monte Carlo simulation will highlight a real-life health problem. The problem is considered to be transversal, since the concepts and technics that are explored during the case are within the scope of Industrial and Engineering Management's issues (e.g. the may be applied in a hospital or any health service provider company). The example is also adequate for Biomedical Engineering, as now-a-days, lots of studies have been developed in health field relating Covid-19 to mathematics and statistics models. Moreover, we consider presenting another example as a challenge for the readers to use and prove the results.

2.3. Concepts

BT describes how to get the probability (*a posteriori*) of an event, based on the probability *a priori*, with new evidence. One of the many applications of BT is Bayesian Inference, an approach to statistical inference (Box and Tiao 1973). Consider two natural states: *F* (Favorable) meaning that a certain hypothesis under consideration is true and (Unfavorable) meaning the opposite, that is, that a certain hypothesis under consideration is false. Now, suppose that we have:

- The probability *a priori* that anyone inside a population is in a favorable state *F* is $P(F)$, and the probability that he or she is in an unfavorable state *U* is $P(U) = 1 - P(F)$.
- The probability of a *T* test being positive (+) when state Unfavorable applies is $P(T + |U)$ (this is a false positive).
- The probability of a *T* test being negative (-) when state Favorable applies is $P(T - |F)$ (this is a false negative).

In these circumstances, it is possible to know the probability that anyone inside this population, who got a positive result in that test, is actually in a Favorable state $P(F|T +)$ (meaning that it is true), or in an Unfavorable state $P(U|T +)$ (meaning that it is false) or when anyone who got a negative result in that test is actually in an Unfavorable state $P(F|T -)$ (meaning that it is false), otherwise $P(U|T -)$ (meaning that it is true). The problem will be solved by three approaches: Bayesian inference, Frequentist method, and lastly, Monte-Carlo simulation (Law 2015). In the analytical method, the expressions that teachers disseminate, and students memorize are as follows (Ross 2009; Walpole et al. 2012):

$$P(F|T +) = \frac{P(T + |F) \times P(F)}{P(T + |F) \times P(F) + P(T + |U) \times P(U)} \quad (1)$$

$$P(U|T +) = \frac{P(T + |U) \times P(U)}{P(T + |U) \times P(U) + P(T + |F) \times P(F)} \quad (2)$$

$$P(F|T -) = \frac{P(T - |F) \times P(F)}{P(T - |F) \times P(F) + P(T - |U) \times P(U)} \quad (3)$$

$$P(U|T -) = \frac{P(T - |U) \times P(U)}{P(T - |U) \times P(U) + P(T - |F) \times P(F)} \quad (4)$$

In this case, we consider that these expressions, once their logic is explained, should be programmed in a spreadsheet by the students themselves with the support of the teacher. Students should then concentrate on solving many application exercises. That is, instead of students spending time on the mechanics of resolution, they concentrate preferentially on the selection of the appropriate expression and on the interpretation of the results. The automatic method of calculation used also allows the performance of a quick sensitivity analysis of the results of each problem to different values of each of the input variables, helping to consolidate the concepts. The case described in the next paragraph, programmed in EXCEL, and named "Bayes test" illustrates this idea and can be downloaded here:

https://www.dropbox.com/s/a8vl5enkbh46mhh/Bayes_test.xlsx?dl=0

2.4. Application Case I

Enter the application "Bayes_test.XLSX". This application enhances the results of applying Bayes' theorem to the results of a test. The case solved on the three sheets, whose statement

is on the sheet "App title", is based on (Gigerenzer 2015) with the following example: "The probability of any woman between 40 and 50 years of age to be sick from breast cancer is 0.008 (prior probability). If any woman, who is ill from this disease is submitted to a test, the probability that the test results positive is 0.9 (true positive). If the woman does not suffer from this disease, the probability that the test results positive is 0.07 (false positive). Under these circumstances, what is the probability that any woman, whose test has resulted positive, suffers from cancer? And in the case the test has resulted negative?"

2.4.1. Analytical (or Bayesian) method

In the "Bayesian" sheet (Figure 1), the posterior probabilities (or *a posteriori*) are calculated from the direct application of the Expressions given by Bayes' theorem. Cell D4 introduces the prior (or *a priori*) probability that any woman, picked up randomly from a specific population, suffers from breast cancer or is ill (natural state favorable (*F*) to the hypothesis). Cell E4 calculates the probability of not having cancer or being healthy (natural state unfavorable (*U*) to the hypothesis). Cell D8 introduces the probability that the test will be negative (*T* −) when the hypothesis of being sick is true (false negative). Cell D7 calculates the probability that the test will be positive (*T* +) when the hypothesis of being sick is true (true positive). Cell E7 introduces the probability that the test will be positive (*T* −) when the hypothesis of being sick is false (false positive). Cell E8 calculates the probability that the test will be negative (*T* +) when the hypothesis of being sick is false (true negative).

	A	B	C	D	E	F	G	H
1								
2		In probability terms:		Natural states				
3				Favourable (F)	Unfavourable (U)			
4		Prior probability		0.008	0.992		F - Favourable means that a certain hypothesis under consideration is true (ill in this case)	
5								
6		Test results		Favourable (F)	Unfavourable (U)			
7		Positive (T+)		0.9	0.07			
8		Negative (T-)		0.1	0.93		U - Unfavourable means that a certain hypothesis under consideration is false (healthy in this case)	
9								
10		Test results		Favourable (F)	Unfavourable (U)			
11		Positive (T+)		P(T+ F) = 0.9	P(T+ D) = 0.07			
12		Negative (T-)		P(T- F) = 0.1	P(T- D) = 0.93			
13								
14			Posterior probabilities given the test result is positive (T+)					
15								
16		Natural states	Conditional probabilities	Prior probability	Joint probabilities		Posterior probability	
17		Favourable (F)	0.9	0.008	0.0072	P(F T+) =	0.09394572	True positive
18		Unfavourable (U)	0.07	0.992	0.06944	P(U T+) =	0.90605428	False positive
19				1	0.07664		1	
20								
21			Posterior probabilities given the test result is negative (T-)					
22								
23		Natural states	Conditional probabilities	Prior probability	Joint probabilities		Posterior probability	
24		Favourable (F)	0.1	0.008	0.0008	P(F T-) =	0.000866401	False negative
25		Unfavourable (U)	0.93	0.992	0.92256	P(U T-) =	0.999133599	True negative
26				1	0.92336		1	
27								
28							...confused? Go to next sheet.	
29								

Figure 1: Screen of the "Bayesian" sheet of the "Bayes" application

Cells in columns C, D and E show the next calculation steps to obtain the posterior probabilities in column G. That is:

- Cell G17 calculates the probability that a woman will have cancer when the test is positive $P(F|T+) = 0.09395$ (true positive).
- Cell G18 shows the probability of a woman not having cancer after all when the test was positive $P(U|T+) = 0.90605$ (false positive).
- Cell G24 calculates the probability that a woman will eventually have cancer when the test is negative $P(F|T-) = 0.00087$ (false negative).
- The G25 cell calculates the probability that a woman will not actually have cancer when the test is negative $P(U|T-) = 0.999913$ (true negative).

2.4.2. Frequentist method

On the sheet "Frequentist" (Figure 2), the natural frequencies are calculated according to the following logical reasoning: In every 1,000 women (cell E5), 8 will effectively have cancer ($0.008 \times 1,000 = 8$, cell C6) and 992 will be healthy ($0.992 \times 1,000 = 992$, cell G6). Once the test, which presents 10% of false negatives, applied to 8 women who are sick with cancer, will give the following results: The existence of cancer is confirmed in $0.9 \times 8 = 7.2$ women (cell B8) and is not confirmed in $0.1 \times 8 = 0.8$ women (cell D8). In turn, the same test that presents 7% of false positives, applied to 992 women who are healthy, will give the following results: The absence of cancer will be confirmed in $0.93 \times 992 = 922.56$ women (cell H8) and $0.07 \times 992 = 69.44$ women (cell F8) will receive a (false) result of having cancer. So, $7.2 / (7.2 + 69.44) \times 100 = 9.39\%$ of women (cell E12) who received a positive mammogram will have cancer and $0.8 / (0.8 + 922.56) \times 100 = 0.09\%$ (cell E13) will be healthy.

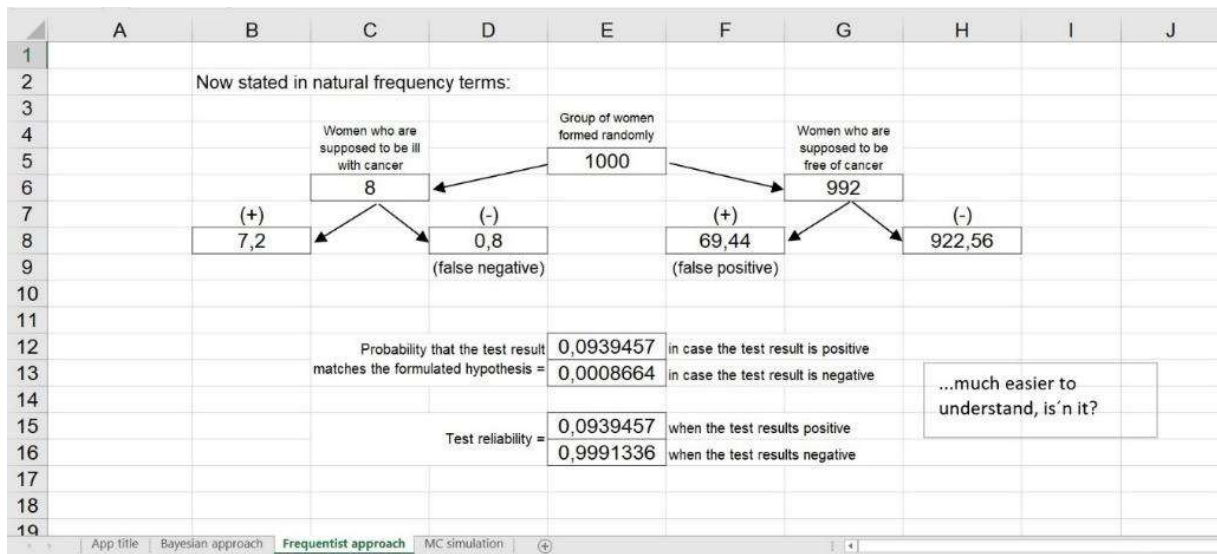


Figure 2: Screen of the "Frequentist" sheet of the "Bayes" application

2.4.3. Monte-Carlo numerical simulation method

On the "Simulation" sheet (Figure 3), the results of $n = 1,000$ random tests are simulated, and the Monte-Carlo method calculates the posterior probabilities (or *a posteriori*). Let us now leave the previous example and refer to a generic case of a population that has a certain characteristic, or property or is in a certain state (immunity to a disease, for example). Suppose that this characteristic exists in the proportion of 1 in 4 elements – or in 25% of the population (cell D4 of the "Bayesian" sheet = 0.25). This "frequency" is now referred to as "*a priori* probability" for the purposes of inference. Suppose further that the test / examination method used is not entirely reliable and that it generates a proportion of 20% false negatives (cell D8 of the "Bayesian" sheet = 0.2) and 10% false positives (cell E7 of the sheet "Bayesian" = 0.1).

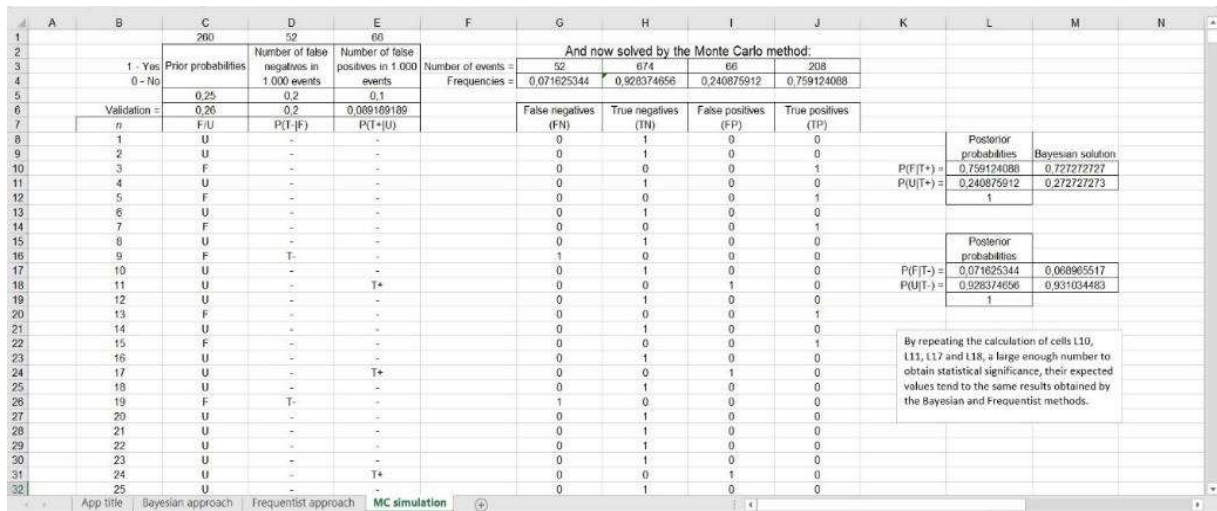


Figure 3: Screen of the "Simulation" sheet of the "Bayes" application

Once on the "Simulation" sheet:

Column C creates a large sample¹, for example, 1,000 elements between lines 8 and 1007, taken from this population that has a proportion of elements immune to the disease (favorable natural state) equal to 0.25 and, therefore, a proportion of non-immune elements (unfavorable natural state) equal to $(1 - 0.25) = 0.75$.

Each cell in this column C was programmed with the following logic (line 8, for example): $= \text{IF}(\text{RAND}() \leq \$C\$5; "F"; "U")$. That is, we generate a random number between 0 and 1 (RAND() function) and if it is less than or equal to 0.25, the function returns "F" (favorable, i.e., the element is immune to the disease), if it is greater than 0.25 and less than 1, the function returns "U" (unfavorable, i.e., the element is not immune). Cell C5 shows the *a priori* probability (0.25) entered in the "Bayesian" sheet. C1 and C6 cells validate the existence of F in range C8: C1007. Cell C1 counts how many F's there are in the sample and cell C6 calculates the correspondent proportion.

C5, D5 and E5 cells were assigned, respectively, to the *a priori* probability, false negative probability and false positive probability, whose values were entered in the "Bayesian" sheet in cells D4, D8 and E7, respectively.

In columns D and E, we simulate the results of a test / exam that is not entirely reliable. In fact, the diagnosis was sometimes correct: $(T + | F)$ (true positive) or $(T - | U)$ (true negative); other times incorrect: $(T - | F)$ (false negative - column D) or $(T + | U)$ (false positive - column E). Cells D1 and D6 validate the existence of T - in range D8: D1007. Cell D1 counts how many T - there are in the sample and cell D6 calculates the correspondent proportion. E1 and E6 cells validate the existence of T + in the E8: E1007 field. Cell E1 counts how many T + are in the sample and cell E6 calculates the correspondent proportion.

In column D, to simulate false negatives, cell D8, e.g., was programmed with the following logic:

$= \text{IF}(C8 = "U"; "-"; \text{IF}(\text{RAND}() \leq \$D\$5; "T -"; "-"))$

¹ The larger the sample, the smaller the number of runs in a Repeater will be necessary to obtain the same degree of precision (sampling error) and vice versa.

That is, if an element is not immune to the disease ("U" in column C), the function returns "-". Otherwise, if the element is in fact immune to the disease ("F" in column C), the function generates a random number between 0 and 1 (RAND () function). If the result is ≤ 0.2 (cell D5), the function returns $T -$ (false negative). Otherwise, the function returns "-".

In column E, to simulate false positives, cell E8, e.g., was programmed with the following logic: = IF (C8 = "F"; "-"; IF (RAND () <= \$ E \$ 5; "T +"; "-"))

That is, if an element is immune to the disease ("F" in column C), the function returns "-". Otherwise, if the element is in fact not immune to the disease ("U" in column C), the function generates a random number between 0 and 1 (RAND () function). If the result is ≤ 0.1 (cell E5), the function returns $T +$ (false positive). Otherwise, the function returns "-".

In columns G, H, I and J the test results are obtained:

Cell G8, e.g., was programmed with the following logic: = IF (AND (C8 = "F"; D8 = "T -"; E8 = "-"); 1; 0). That is, if the element is immune to the disease (cell C8 = "F") and if the test is negative (cell D8 = $T -$) and not positive (cell E8 = "-"), the function returns 1 (FN - False Negative) and 0 otherwise.

Cell H8, e.g., was programmed with the following logic: = IF (AND (C8 = "U"; D8 = "-"; E8 = "-"); 1; 0). That is, if the element is not immune to the disease (cell C8 = "U") and if the test result is correct (cell D8 = "-" and cell E8 = "-"), the function returns 1 (TN - True Negative) and 0 otherwise.

Cell I8, e.g., was programmed with the following logic: = IF (AND (C8 = "U"; E8 = "T +"; D8 = "-"); 1; 0). That is, if the element is not immune to the disease (cell C8 = "U") and if the test is positive (cell E8 = $T +$) and non-negative (cell D8 = "-"), the function returns 1 (FN - False Negative) and 0 otherwise.

Cell J8, e.g., it was programmed with the following logic: = IF (AND (C8 = "F"; E8 = "-"; D8 = "-"); 1; 0). That is, if the element is immune to the disease (cell C8 = "F") and if the test result is correct (cell E8 = "-" and cell D8 = "-"), the function returns 1 (FN - False Negative) and 0 otherwise.

Cells G3 to J3 count how many ones were obtained in columns G, H, I and J, respectively. Cells G4 to J4 count the proportions of False Negatives (FN) or $P(F|T -)$, True Negatives (TN) or $P(U|T -)$, False Positives (FP) or $P(U|T +)$ and True Positives (TP) or $P(F|T +)$. These results can also be read in yellow cells L17, L18, L11 and L10, respectively, to be easily compared with the results obtained by the analytical method (cells M17, M18, M11 and M10).

By repeating the calculation of the yellow cells, a large enough number to obtain statistical significance, their expected values tend to the same results obtained by both the Bayesian and Frequentist methods.

A second application case, showing the importance of the subject to engineering is described in the next section.

3. Application Case II

A complex electronic system fails in several ways. One of them - the failure mode X - occurs in the proportion of 1/4. It is intended to carry out its monitoring by means of a self-diagnostic module, which, during reliability tests, correctly diagnosed 98% of the occurrences of this failure mode X (2% false negatives, Bayesian sheet D8 cell) and misdiagnosed 1% of other failures as failure mode X (false positives, Bayesian sheet cell E7). How reliable is the module? Answers: The probability of the module correctly diagnosing failure mode X when it occurs is

0.9703 (cell G17). The probability of the module diagnosing (or not) another failure mode when in fact it is not failure mode X is 0.9933 (G25 cell of the Bayesian sheet).

4. Discussion and Conclusions

It is demonstrated throughout this article that a theorem with a complicated statement, and an even more complicated mathematical representation, can be described in a simpler way using pure logic through two methods: the frequentist and the numerical simulation – the later even more intuitive than the former. This approach to a law that we observe in so many circumstances of the Engineering activity, also present in natural sciences, and which causes so much difficulty to young students during their academic career, can after all, be described using the same logic that was certainly followed by authors, when they interpreted the results of their experiments, although taking a lot of time and effort, in contrast to the ease today provided by computational media. We highlight the method of logic applied in the construction of a numerical simulation model can be used in the demonstration of other statistical laws, for example, the analytical expressions that describe the Central Limit Theorem or the Confidence Intervals of the mean and standard deviation of continuous or discrete variables, or the Covariance between two random variables. In conclusion, we strongly recommend the use of the numerical simulation technique for the construction of models that imitate the real world with just-enough precision and save time and money. While this technique is already trivial in many areas of Engineering and Management, it is not, however, in Statistics and hence our interest in the dissemination of this article.

References

- Auster, E. R., and K. K. Wylie. 2006. "Creating active learning in the classroom: A systematic approach". *Journal of Management Education* 30, no. 2 (april): 333-53. <https://doi.org/10.1177/1052562905283346>.
- Box, G., and G. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. John Wiley and Sons Ltd.
- Budimir, I. 2016. "Modelling in teaching Probability and Statistics at the Faculties of Technology in the Republic of Croatia". *The holistic approach to environment* 6, no. 4: 141-57. <https://hrcak.srce.hr/162094>.
- Christou, N., and I. D. Dinov. 2010. "A study of students' learning styles, discipline attitudes and knowledge acquisition in technology-enhanced probability and statistics education". *MERLOT Journal of Online Learning and Teaching* 6, no. 3 (september). https://jolt.merlot.org/vol6no3/dinov_0910.htm.
- Cook, T. D. 2002. "Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them". *Educational Evaluation and Policy Analysis* 24, no. 3 (Fall): 175-99. <https://doi.org/10.3102/01623737024003175>.
- D'Andrea, V., and D. Gosling. 2005. *Improving teaching and learning in higher education: A whole institution approach*. Open University Press.
- Daveedu Raju, A., C. H. Gaayathre, G. Leela Deepthi, K. Madhuri, and D. Maheswari. 2019. "Prediction of students' performance for a multi class problem using naïve bayes classifier". *International Journal of Innovative Technology and Exploring Engineering* 8, no. 7 (may): 2209-14. <https://www.ijitee.org/download/volume-8-issue-7/>.
- Diamond, N., S. K. Koernig, and Z. Iqbal. 2008. "Uniting active and deep learning to teach problem-solving skills: Strategic tools and the learning spiral". *Journal of Marketing Education* 30, no. 2 (august): 116-29. <https://doi.org/10.1177/0273475308317707>.

- Fienberg, S. E. 1992. "A brief history of statistics in three and one-half chapters: A review essay". *Statistical Science* 7, no. 2 (may): 208-25. <http://www.jstor.org/stable/2246307>.
- Gerow, A., Y. Hu, J. Boyd-Graber, D. M. Blei, and J. A. Evans. 2018. "Measuring discursive influence across scholarship". *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 13 (march): 3308-13. <https://doi.org/10.1073/pnas.1719792115>.
- Gigerenzer, G. 2015. *Risk Savvy: How to make good decisions*. Penguin Putnam Inc.
- Gigerenzer, G., and U. Hoffrage. 1995. "How to improve Bayesian reasoning without instruction: Frequency formats". *Psychological Review* 102, no. 4: 684-704. <https://doi.org/10.1037/0033-295x.102.4.684>.
- Jeffreys, H. 1973. *Scientific Inference*. 3rd ed. Cambridge University Press.
- Johnson, R. 2017. *Miller & Freund's Probability and Statistics for Engineers*. 9th ed. Pearson.
- Joyce, J. 2021. "Bayes' Theorem". In *Stanford Encyclopedia of Philosophy Archive (Fall 2021 Edition)*, edited by E. N. Zalta. Article published June 28, 2003; last modified September 30, 2003. <https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/>.
- Konopka, C. L., M. B. Adaime, and P. H. Mosele. 2015. "Active teaching and learning methodologies: Some considerations". *Creative Education* 06, no. 14 (august): 1536-45. <http://doi.org/10.4236/ce.2015.614154>.
- Law, A. M. 2015. *Simulation Modeling and Analysis*. 5th ed. McGraw-Hill, Inc.
- Ross, S. M. 2009. *Introduction to Probability and Statistics for Engineers and Scientists*. 4th ed. Boston: Academic Press. <https://doi.org/10.1016/B978-0-12-370483-2.00006-0>.
- Simon, J. L., D. T. Atkinson, and C. Shevokas. 1976. "Probability and Statistics: Experimental Results of a Radically Different Teaching Method". *American Mathematical Monthly* 83, no. 9: 733-39.
- Walpole, R., R. Myers, S. Myers, and K. Ye. 2012. *Probability & Statistics for Engineers & Scientists*. 9th ed. Pearson
- Wang, Y. H. 2020. "Design-based research on integrating learning technology tools into higher education classes to achieve active learning". *Computers and Education* 156 (october): Article number 103935. <https://doi.org/10.1016/j.compedu.2020.103935>.
- Warren, A. R. 2018. "Quantitative critical thinking: Student activities using Bayesian updating". *American Journal of Physics* 86, no. 5 (may): 368-80. <https://doi.org/10.1119/1.5012750>.