Journal of Advances in Medicine and Medical Research

26(1): 1-21, 2018; Article no.JAMMR.40784
ISSN: 2456-8899
(Past name: British Journal of Medicine and Medical Research, Past ISSN: 2231-0614, NLM ID: 101570965)

Common Fallacies of Probability in Medical Context: A Simple Mathematical Exposition

Rufaidah Ali Rushdi[1] and Ali Muhammad Rushdi[2*]

[1]Department of Pediatrics, Kasr Al-Ainy Faculty of Medicine, Cairo University, Cairo, 11562, Arab Republic of Egypt.
[2]Department of Electrical and Computer Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia.

Authors' contributions

This work was carried out in collaboration between the both authors. Author RAR envisioned and designed the study, performed the analyses, stressed the medical context for the fallacies considered, and solved the detailed examples. Author AMR managed the literature search and wrote the preliminary manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMMR/2018/40784
Editor(s):
(1) Angelo Giardino, Professor, Texas Children's Hospital, Houston, Texas and Pediatrics, Baylor College of Medicine, Houston, TX, USA.
Reviewers:
(1) Maria A. Ivanchuk, Bucovinian State Medical University, Ukraine.
(2) Grienggrai Rajchakit, Maejo University, Thailand.
(3) Jeffrey Clymer, USA.
Complete Peer review History: http://www.sciencedomain.org/review-history/24107

Review Article

Received 16th January 2018
Accepted 4th April 2018
Published 11th April 2018

ABSTRACT

This paper presents and explores the most frequent and most significant fallacies about probability in the medical fields. We allude to simple mathematical representations and derivations as well as to demonstrative calculations to expose these fallacies, and to suggest possible remedies for them. We pay a special attention to the evaluation of the *posterior* probability of disease given a positive test. Besides exposing fallacies that jeopardize such an evaluation, we offer an approximate method to achieve it under justified typical assumptions, and we present an exact method for it *via* the normalized two-by-two contingency matrix. Our tutorial exposition herein might hopefully be helpful for our intended audience in the medical community to avoid the detrimental effects of probabilistic fallacies. As an offshoot, the pedagogical nature of the paper might allow probability educators to utilize it in helping their students to learn by unraveling their private misconceptions about probability.

**Corresponding author: E-mail: arushdi@kau.edu.sa, arushdi@ieee.org, alirushdi@gmail.com, arushdi@yahoo.com;*

Keywords: Fallacy; misconception; paradox; conditional probability; posterior probability; medical context; normalized contingency table.

1. INTRODUCTION

Knowledge of elementary probability concepts and capability to perform simple probability calculations are indispensable to and essential for both medical students and practitioners [1-6]. By contrast, full command of deep and profound probability concepts and mastery of sophisticated probability operations are not warranted for the general body of medical practitioners, despite being a must for those among them conducting medical research [7-19]. This paper is a sequel of our earlier recent publications [20-23] concerned with applications of probability and statistics in medicine, and is part of our ongoing efforts to enhance understanding of, simplify calculations with, and facilitate reasoning about probability, in general, and conditional probability, in particular. However, this paper differs from its predecessors in that it views the subject matter from the negative side by exploring fallacies and misconceptions that might jeopardize and degrade sound learning.

The fallacies discussed herein are mistakes or errors (pertaining to probability), which are committed so frequently by many members of the medical community to warrant the cost of labeling, classifying, and exposing them. Knowledge of such fallacies might arm physicians against faulty decision making (in real-life situations), which, nevertheless, sounds deceptively agreeable and correct. Unfortunately, most of probability fallacies seem to be somewhat incorrigible, tenacious and highly resistant to attempts of correction or reform. Experimental studies show that a notable number of medical practitioners adhere to their earlier misconceptions about probability, and persevere in erring in probability calculations even after being shown ways to bypass fallacies while performing such calculations [2].

The purpose of this paper is to present and explore the simplest forms of the most frequent and most significant fallacies of probability that are spread in the medical circles. We do not assume too much knowledge of probability for our readers, and we hope not to lose any reader by alluding to simple mathematics without giving and exposing such mathematics. We apologize to advanced readers who might find some of the material presented herein elementary, simplistic, redundant, obvious, and even easy to dispense with. Besides exposing probabilistic fallacies that arise in medical contexts, the paper serves as a tutorial on typical calculations encountered with diagnostic testing. The paper offers approximate calculations that are valid under mild typical assumptions. These approximate calculations are formally justified *via* simple mathematics, and are found to be in excellent agreement with exact calculations. The paper also uses a normalized contingency table to perform exact calculations with an extra step to check correctness of the calculations. Both approximate and exact methods are welcome additions to the arsenal of methods reported in [21-23] to facilitate calculations associated with diagnostic testing.

The literature of fallacies (and associated misconceptions and paradoxes) in probability is extensive, indeed [24-52], but most of it is devoted to legal and judicial issues [24,31,46,50]. There is also a plethora of articles where the medical and legal domains overlap (*e.g.,* on forensic science [37,38,41,45,52]). However, there is obviously some gap when the issue of fallacies concerns clinical medicine *per se*, and we hope our current contribution might bridge this gap, at least partially.

It is well known that probability theory could be problematic and challenging for all users (and not just for laypersons) [53-57]. The subject matter of probability might be difficult to access for reasons other than fallacies, misconceptions, and paradoxes. There is no agreed-upon heuristic to translate a novel word problem of probability into a concrete mathematical model. Mathematical knowledge might not suffice for solving probability problems because these problems require insight, deep understanding, lengthy contemplation as well as patience and perseverance. Moreover, solutions of these problems are frequently counter-intuitive, hard to accept, and difficult to swallow. Another source of difficulty is that conditionality might be interpreted as causality. Occasionally, conditional events might be thought of unnecessarily as sequential events. Some puzzling, paradoxical, and notorious problems labeled as "*teasers*" [54, 57] are also frequently encountered. These are either inherently-ambiguous problems such that they admit no solution, or ultimately-solvable ones but only after their mysteries are unraveled through "proper" partitioning of the sample space [56]. We stress that we are not dealing herein with the solution of general probability problems in this paper. We restrict ourselves to simple well-posed probability problems of medical context that have already been known for some time and have well-established correct solutions in the literature, but could be mishandled by medical students and practitioners due to some inherent misconceptions or as a result of lack of adequate training.

Due to space limitations, we restrict our discussion to fallacies encountered frequently in the medical circles. We have to leave out many fallacies that are well-known in other walks of life and other scientific disciplines. These include (albeit not restricted to) the Equi-probability Fallacy [58], the Gambler's Fallacy [33], the Fallacy of the Impartial Expert [24], the Efficient Breach Fallacy [32], the Individualization Fallacy [45], the Association Fallacy [46], the Defense-Attorney Fallacy [31,51], the Uniqueness Fallacy [45], and the Disjunction Fallacy [59,60].

The organization of the rest of this paper is as follows. Section 2 explores the Multiplication and Addition Fallacies, and as an offshoot, comments on basic probability formulas. Section 3 is the main contribution of this paper. Besides exposing the Inverse Fallacy, it presents a method to estimate P(Disease|positive test) approximately under typical assumptions, and also offers another method to evaluate this probability exactly *via* the normalized two-by-two contingency table. Section 3 also presents three illustrative examples, which are computational in nature and medical in context. Moreover, Section 3 reflects on certain comments available in the literature on the validity of the pertaining model itself. Section 4 reviews the concept of an event being favorable to another and discusses the Fallacy of the Favorable Event. Section 5 investigates the Conditional-Marginal Fallacy and discusses the relations among conditional and marginal probabilities. Section 6 investigates Simpson's Paradox through it is not really a fallacy as such, but, being a paradox, it shares the problematic nature of a fallacy. Section 7 demonstrates the Conjunction Fallacy *via* a medical example while Section 8 discusses the Appeal-to-Probability Fallacy. Section 9 illustrates the drastic effects of the Base-Rate Fallacy by considering its effect on one of the examples of Section 3. Section 10 covers the Representative-Sampling Fallacy. Section 11 adds some useful observations, while Section 12 concludes the paper. To make the paper self-contained, an appendix (Appendix A) on "conditional probability" is included. Any equation we present herein that is not generally true will be identified as such (in an admittedly harsh way) by labeling it as "WRONG."

2. MULTIPLICATION AND ADDITION FALLACIES

The Multiplication and Addition Fallacies for two general events A and B amount to mathematically expressing the probabilities of the intersection and union of these two events (see Appendix A) as the product and sum of their probabilities, namely

$$P(A \cap B) = P(A)P(B), \qquad \text{(WRONG)} \quad (1)$$

$$P(A \cup B) = P(A) + P(B). \qquad \text{(WRONG)} \quad (2)$$

The wide-spread prevalence of these fallacies is perhaps mainly due to the way probability is introduced in pre-college education. In fact, these fallacies are appealing because they simply replace set operations in the event domain by their arithmetic counterparts in the probability domain. Another possible reason for the popularity of these fallacies is the inadvertent neglect or disregard of the conditions under which they become valid. Equation (1) is correct provided the events A and B are (statistically) independent, while Equation (2) is exactly correct when the events A and B are mutually exclusive ($A \cap B = \emptyset$). In particular, the addition formula (2) is correct if A and B are primitive outcomes or singletons, *i.e.*, events comprising individual points of the pertinent sample space. Moreover, Equation (2) is approximately correct when the events A and B are independent and the probabilities $P(A)$ and $P(B)$ are particularly very small. The correct versions for (1) and (2) are the elementary formulas [1,61].

$$P(A \cap B) \;=\; P(A|B)P(B) \;=\; P(A)P(B|A) \quad (3)$$

$$P(A \cup B) \;=\; P(A) + P(B) - P(A \cap B\,) \qquad (4)$$

Equation (3) asserts that the fundamental concept of conditional probability is *unavoidable*, indeed, since statistical independence is not always guaranteed. Equation (3) provides definitions for the conditional probabilities $P(A|B)$ and $P(B|A)$ provided $P(B) \neq 0$, and $P(A) \neq 0$, respectively. The probability $P(A \cap B)$ can be neglected (considered almost 0) in (4) so as to approximate (4) by (2) when the events A and B are independent and $P(A)$ and $P(B)$ are very small. However, this same probability cannot be neglected in (3), even when $P(A)$ and $P(B)$ are extremely small since such an action leads to a catastrophic error (100% relative error), or to a fallacy of its own called the Rare-Event Fallacy.

3. THE INVERSE FALLACY

The Inverse Fallacy [7,29,31,36,40-42,47], is also called the Confusion-of-the-Inverse Fallacy, the Fallacy of the Transposed Probability, the Conditional-Probability Fallacy, or the Prosecutor's Fallacy. In this fallacy, the event *A* given *B* is confused with the event *B* given *A*, or the conditional probability $P(A|B)$ is considered (exactly or approximately) equal to the conditional probability $P(B|A)$, which is called the inverse or transpose of the former probability $P(A|B)$. This fallacy is very common in medical circles [2,4,23,29,40]. Let *A* and *B* denote {Disease is present} and {Test says that disease is present}, then the Inverse Fallacy, is manifested in believing that the test Positive Predictive Value (PPV_{ij}) given by

$$P(A|B) = P\begin{pmatrix} Disease\ is\ present | Test\ says\ that \\ disease\ is\ present \end{pmatrix} \quad (5)$$

is the same as the test Sensitivity $(Sens_{ij})$.

$$P(B|A) = P\begin{pmatrix} Test\ says\ that\ disease\ is\ present | \\ Disease\ is\ present \end{pmatrix} \quad (6)$$

Since the former probability is typically substantially smaller than the latter one, this fallacy has grave consequences, as it means misinterpreting false positive test results (which are already bad and alarming besides being misleading) to make them even more disturbing and threatening.

The two conditional probabilities $P(A|B)$ and $P(B|A)$ are not related by the equality relation fallaciously assumed, but are related by Bayes' formula expressed by Equation (3). Therefore, the ratio of these two conditional probabilities is equal to the ratio of the unconditional or marginal probabilities, namely

$$\frac{P(A|B)}{P(B|A)} = \frac{P(A)}{P(B)} \quad (7)$$

With our earlier designation of *A* and *B* as {Disease is present} and {Test says disease is present}, the ratio in (7) is not 1 as the fallacy demands, but it is the ratio of *True Prevalence* P(A) (true probability of disease presence or such a probability according to a gold standard) to P*erceived or Apparent Prevalence* P(B) (probability of disease presence according to the test). The Perceived Prevalence P(B) is given by the Total Probability Formula [1, 61] as

$$P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) \quad (8)$$

In typical situations, a test has a nearly perfect Sensitivity $P(B|A)$, and hence we can approximately assume that

$$P(B|A) \approx 1, \quad (9)$$

Also the true prevalence is usually very low, and although we cannot assume P(A) to be zero, we might safely assume that

$$P(\overline{A}) = 1 - P(A) \approx 1 \quad (10)$$

Therefore, Equation (8) can be rewritten approximately as

$$P(B) \approx P(A) + P(B|\overline{A}) \quad (11)$$

The approximation (11) does not violate the probability axiom $\{P(B) \leqslant 1\}$, since both *P(A)* and $P(B|\overline{A})$ are known to be small compared to 1. However, the probability $P(B|\overline{A})$ (called the False Positive Rate, FPR_{ij}) (albeit small) could be significantly larger than *P(A)*. Hence, the perceived P(B) is greater (or even much greater than) the *true prevalence* P(A). This makes the ratio in (7) definitely smaller (usually much smaller) than 1. In other words, $P(B|A) > P(A|B)$ (typically $P(B|A) \gg P(A|B)$). This means that in many cases, a test *PPV* is (significantly) smaller than its Sensitivity, and should not be mistaken as being equal to it. Under typical mild assumptions, we can assess the *PPV* $P(A|B)$ approximately through a combination of (7) and (11) as

$$P(A|B) \approx P(B|A)\frac{P(A)}{\left(P(A)+P(B|\overline{A})\right)} \quad (12)$$

$$\approx \frac{P(A)}{\left(P(A)+P(B|\overline{A})\right)} \quad (12a)$$

To give a concrete example, we quote a celebrated problem of Gigerenzer, et al. [2],

"Assume you conduct breast cancer screening using mammography in a certain region. You know the following information about the women in this region:

(a) The probability that a woman has breast cancer is 1% (True Prevalence)
(b) If a woman has breast cancer, the probability that she truly tests positive is 90% (Sensitivity)
(c) If a woman does not have breast cancer, the probability that she nevertheless tests positive is 9% (False-Positive Rate)

A woman tests positive. She wants to know from you whether that means that she has breast cancer for sure, or what the chances are. What is the best answer?"

In this problem, we identify the given information in our notation as:

(a) $P(A) = 0.01$ (True Prevalence) (13a)

(b) $P(B|A) = 0.90$ (Sensitivity or *TPR*) (13b)

(c) $P(B|\overline{A}) = 0.09$ (*FPR*) (13c)

and recognize the required unknown in this problem as the *PPV* or $P(A|B)$. We observe that the assumptions we made above are all valid, namely

1. $P(A) = 0.01 \ll 1,\ P(\overline{A}) = 0.99 \simeq 1$ (14a)
2. $P(B|A) = 0.90 \simeq 1$ (14b)
3. $P(B|\overline{A}) = 0.09$ is small but is (much) larger than $P(A)$, (14c)

Our approximate answer in (11) is

$$P(B) \simeq 0.01 + 0.09 = 0.10, \quad (15)$$

while the exact answer computed by Rushdi, et al. [23] (for the same problem) is 0.0981. Correspondingly, our approximate answer in (12a) is

$$P(A|B) \approx \frac{0.01}{0.01+0.09} = 0.1 \quad (16)$$

In an experiment conducted by Gigerenzer, et al. [2], 160 gynecologists were requested to choose the *best* value for $P(A|B)$ among four given values

(a) 0.81, (b) 0.90, (c) 0.10, (d) 0.01

where incorrect answers were spaced about one order of magnitude away from the *best* answer. Only 21% of the gynecologists found the *best* answer of 0.10 in (c) while 47% and 13% of them grossly overestimated the answer as 0.90 and 0.81, respectively, perhaps falling victim to (or at least being influenced by) the Inversion Fallacy. Only 19% of the respondents underestimated the answer.

Another example reported by Eddy [29] runs as follows:

"The prior probability, P(ca), 'the physician's subjective probability', that the breast mass is malignant is assumed to be 1%. To decide whether to perform a biopsy or not, the physician orders a mammogram and receives a report that in the radiologist's opinion the lesion is malignant. This is new information and the actions taken will depend on the physician's new estimate of the probability that the patient has cancer. This estimate also depends on what the physician will find about the accuracy of mammography. This accuracy is expressed by two figures: sensitivity, or true-positive rate P(+ | ca), and specificity, or true-negative rate P(- | benign). They are respectively 79.2% and 90.4%."

We choose to give a detailed analysis of this example *via* the normalized contingency table of Fig. 1, which summarizes our earlier findings in [21-23]. Substituting for the symbolic notation in Fig. 1, we produce a complete solution for the aforementioned example in Fig. 2. The results obtained indicate that the particularly-required result of P(cancer | positive test), or in our notation PPV_{ij} = P(j=+1|i=+1) = $P(A|B)$ is 0.076923 or approximately 7.7%. According to Eddy [29], most physicians interviewed estimated this *posterior* probability to be about 75%, *i.e.*, almost ten times larger. He attributed this to the Inverse Fallacy, which led them to believe that the required probability is approximately equal to its transpose of P(Test positive|cancer) = P(i=+1| j=+1) which is given as 0.792.

Boumans [62-64] criticizes the above findings, by questioning the validity of the model on which they are based. He constructs a different double-threshold model that ultimately justifies why physicians tend to give high estimates for the *posterior* probability P(cancer|positive test). It seems that this tendency among physicians is excused on the grounds that two wrongs (a wrong model and a wrong method of calculations) can possibly make one right. Boumans [62-64] asserts that decision making in real-life situations is different from decision making in a laboratory controlled experiment. "A model of a decision

$TP_{ij}/N = P(i = +1 \cap j = +1)$ $=TPR_{ij} * TP$ $=PPV_{ij} * PP$	$FP_{ij}/N = P(i = +1 \cap j = -1)$ $=FPR_{ij} * (1 - TP)$ $=FOR_{ij} * PP$	$P(i = +1)$ $=Percieved$ or $Apparent$ $Prevalence$ $=PP$
$FN_{ij}/N = P(i = -1 \cap j = +1)$ $=FNR_{ij} * TP$ $=FOR_{ij} * (1 - PP)$	$TN_{ij}/N = P(i = -1 \cap j = -1)$ $=TNR_{ij} * (1 - TP)$ $=NPV_{ij} * (1 - PP)$	$P(i = -1)$ $=1 - PP$
$P(j = +1)$ $=True\ Prevalence$ $=TP$	$P(j = -1)$ $=1 - TP$	1

Fig. 1. The normalized two-by-two contingency matrix in medical context. Symbols used are taken from [21-23]

(0.792) (0.01) = 0.00792 = (0.10296) (0.076923)	(1 – 0.904) (0.99) = (0.096) (0.99) = 0.09504	0.00792 + 0.09504 = 0.102964
0.01 – 0.00792 = 0.00208	0.99 – 0.09504 = 0.89496	0.00208 + 0.8996 = 0.89740
0.01	1 – 0.01 = 0.99	1.00 = 0.10296 + 0.89704 ✓

Fig. 2. Complete solution of the second example of Sec. 2 with the aid of the normalized contingency table introduced in Fig. 1. Initially known entries are highlighted in red

0.001 – 0 = 0.001	(0.05) (0.999) = 0.04995	0.001 + 0.04995 = 0.05095
0 (Assumption)	0.999 – 0.04995 = 0.94905	0 + 0.94905 = 0.94905
0.001	1 – 0.001 = 0.999	1.00 = 0.05095 + 0.94905 ✓

Fig. 3. Complete solution of the third example in Sec. 2 via a normalized contingency table. Given data is shown in red. The assumption of FNR=0 (TPR=1) was added (to the original formulation by Casscells et al. [7]) by subsequent authors such as Westbury [4] and Sloman et al. [65]

problem frames that problem in three dimensions: sample space, target probability and information structure. Each specific model imposes a specific rational decision. As a result, different models may impose different, even contradictory, rational decisions, and create choice 'anomalies' and 'paradoxes'." Boumans also calls for a new planner called "the normative statistician, the expert in reasoning with uncertainty par excellence." Boumans also argues that "rationality should be model-based, which means that not only the isolated decision-making process should take a Bayesian updating process as its norm, but should also model the acquisition of evidence (priors and test results) as a rational process." Essentially, statisticians are needed to understand medicine better, rather than physicians are requested to make a better mastery of statistics. In our opinion, unity of science is a must, and a sound reconciliation of differences among statisticians and physicians is highly urged. Further exploration of this subject is definitely warranted, albeit it seems somewhat beyond the scope of this paper.

We close this section with a third example due to Casscells, et al. [7]. This example reads as follows:

"If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5 per cent, what is the chance that the person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?"

Fig. 3 indicates immediately that the data given is not complete and must be supplemented by something else. The solution given in Fig. 3 is based on neglecting the False Negative Rate (*FNR*), and is the solution intended by those who posed the problem. However, the missing information in the example defeats the purpose of the experiment done by Casscells, et al. [7]. The physicians who resorted during that experiment to the Inverse Fallacy are only partially to be blamed, since they were forced to seek a means to fill in an inadvertent and unnecessary gap. We noted that subsequent authors who referred to the problem of Casscells, et al. [7] augmented the original formulation above by adding {*FNR* = 0} [4], or equivalently, {*TPR* = 1} [65].

4. FALLACY OF THE FAVORABLE EVENT

An event A is called favorable to another B [25, 42] when the occurrence of A implies an increase in the chances B occurs, *i.e.,*

$$P(B|A) > P(B) \tag{17}$$

The Fallacy of the Favorable Event is to infer from the fact that the conditional probability $P(A|B)$ is "large" that the conditional (conditioned) event A is favorable to the conditioning event B. This fallacy is so problematic that it is not even amenable to precise mathematical description, since one does not really know how "large" is "large."

Krämer and Gigerenzer [42] cite many examples in which this fallacy occurs in various contexts, and suggest that it is "possibly the most frequent logical error that is found in the interpretation of statistical information." A newspaper headline stating that "Boys more at risk on bicycles" is cited [42] to be based on the report that "among children involved in bicycle accidents the majority were boys." The writer(s) of the headline, in fact, observed that

$$\text{P(Boys|bicycle accident) is "large"} \tag{18}$$

and went on to conclude that

$$P(bicycle\ accident|Boys) > P(bicycle\ accident) \tag{19}$$

The statement (19) is only possibly unwarranted, *i.e.*, it is not necessarily false, but the fact is that it cannot be logically inferred from (18).

The concept of "favorableness" discussed in this section is also involved in Simpson's Paradox [25, 42]. In general, Simpson's paradox describes a phenomenon in which a trend appears in individual groups of data, but disappears or reverses when these groups are combined, or amalgamated [25,26,28,35,42,48]. We will discuss this reversal-upon-amalgamation paradox in Sec. 6.

5. THE CONDITIONAL-MARGINAL FALLACY

In the Conditional-Marginal Fallacy, the conditional probability $P(A|B)$ is mistaken for the marginal unconditional probability $P(A)$, or, equivalently (according to Eq. (7)), the inverse conditional probability $P(B|A)$ is equivocated with $P(B)$. We note that this is generally fallacious unless the two events *A* and *B* are (statistically) independent. In fact, the very definition of independence of event *A* from event *B* is the requirement that $P(A|B)$ be equal to $P(A)$. Similarly, independence of event B from event A is the requirement that $P(B|A)$ be equal to $P(B)$. These two definitions are equivalent, and hence, we do not need to refer to independence of one event from another, but to independence between the two events. Any of the following equivalent twelve relations can be used to denote (statistical) independence between events *A* and *B*, and can be used to mathematically deduce any of the other relations

$$P(A) = P(A|B) = P(A|\overline{B}) \tag{20a}$$

$$\mathrm{P(B) = P(B|A) = P(B|\overline{A})} \tag{20b}$$

$$\mathrm{P(\overline{A}) = P(\overline{A}|B) = P(\overline{A}|\overline{B})} \tag{20c}$$

$$\mathrm{P}(\overline{\mathrm{B}}) = \mathrm{P}(\overline{\mathrm{B}}|\mathrm{A}) = \mathrm{P}(\overline{\mathrm{B}}|\overline{\mathrm{A}}) \quad (20d)$$

$$\mathrm{P}(\mathrm{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{A})\,\mathrm{P}(\mathrm{B}) \quad (20e)$$

$$\mathrm{P}(\mathrm{A} \cap \overline{\mathrm{B}}) = \mathrm{P}(\mathrm{A})\,\mathrm{P}(\overline{\mathrm{B}}) \quad (20e)$$

$$\mathrm{P}(\overline{\mathrm{A}} \cap \mathrm{B}) = \mathrm{P}(\overline{\mathrm{A}})\,\mathrm{P}(\mathrm{B}) \quad (20g)$$

$$\mathrm{P}(\overline{\mathrm{A}} \cap \overline{\mathrm{B}}) = \mathrm{P}(\overline{\mathrm{A}})\,\mathrm{P}(\overline{\mathrm{B}}) \quad (20h)$$

In passing, we note that the multiplication rule (20e) is used by many authors as a definition of (statistical) independence of two events *A* and *B*. However, we stress that any of the eight relations in (20a) – (20d) are more intuitively appealing as definitions. They all convey the message that conditioning on an independent event is irrelevant, or, equivalently, that assessment of the probability of an event is not affected by the presence or absence of information about occurrence or non-occurrence of an independent event. The two relations (among these eight relations) that are without event complementation ($P(A|B) = P(A)$ or $P(B|A) = P(B)$) are the preferred defining methods (see, e.g., Trivedi [1] or Rushdi & Talmees [22]). The actual general relation between P(A) and $P(A|B)$ is that either

$$P(A|B) \leqslant P(A) \leqslant P(A|\overline{B}) \quad (21a)$$

Or

$$P(A|B) \geqslant P(A) \geqslant P(A|\overline{B}) \quad (21b)$$

This follows from the fact that it is certain that either $\{P(A) \geqslant P(A|B)\}$ or $\{P(A) \leqslant P(A|B)\}$, when this fact is combined with the two implications

$$\{\, P(A) \geqslant P(A|B)\} \Rightarrow \{\, P(A) \leqslant P(A|\overline{B})\} \quad (22a)$$

$$\{\, P(A) \leqslant P(A|B)\} \Rightarrow \{\, P(A) \geqslant P(A|\overline{B})\} \quad (22b)$$

For example, the first implication can be ascertained by applying $\{\, P(A|B) \leqslant P(A)\}$ to the Total Probability Formula

$$P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B}) \quad (8a)$$

to obtain

$$\{P(A|B) \leqslant P(A)\} \Rightarrow \{P(A) \leqslant P(A)P(B) + P(A|\overline{B})P(\overline{B})\} \quad (23a)$$

$$\{P(A|B) \leqslant P(A)\} \Rightarrow \{P(A)\big(1 - P(B)\big) \leqslant P(A|\overline{B})P(\overline{B})\,\} \quad (23b)$$

Nothing that $P(\overline{B}) \neq 0$ (as guaranteed implicitly due to the existence of $P(A|\overline{B})$), we can divide both sides of the inequality in the implied part of Eq. (23b) by $\big(1 - P(B)\big) = P(\overline{B})$ to obtain (22a).

The two inequalities (21a) and (21b) state that P(A) is located in an interval bounded by $P(A|B)$ and $P(A|\overline{B})$, irrespective of which is the upper bound and which is the lower bound. In fact, $P(A|B)$ is the upper bound if B is favorable to A. By contrast, $P(A|\overline{B})$ is the upper bound when $\overline{B}$ (rather than B) is favorable to A. When neither B nor $\overline{B}$ is favorable to A, A is independent of B (and consequently B is independent of A) and the interval bounded by $P(A|B)$ and $P(A|\overline{B})$ reduces to [P(A), P(A)] so that the equalities (20a)-(20h) hold. In a sense, the larger the interval bounded by $P(A|B)$ and $P(A|\overline{B})$, the more dependent event *A* on event *B* is. When this interval collapses to a single point, the two events *A* and *B* are independent.

We now revisit the first detailed example considered in Sec. 3. In this example, $P(A) = 0.01$ while $P(A|B) \approx 0.1$, which means that committing the Conditional-Marginal Fallacy while assessing $P(A|B)$ amounts to underestimating $P(A|B)$ by one order of magnitude. Similarly, $P(B|A) = 0.90$ while $P(B) \simeq 0.10$, which means that using this fallacy to assess $P(B)$ leads to a value overestimated, again by almost one order of magnitude. Similar comments might be deduced by viewing the two other examples in Sec. 3 and observing their solutions in Figs. 2 and 3.

Some authors use Berkson's Fallacy (Berkson's Bias or Berkson's Paradox) as a name for the Marginal-Conditional Fallacy. However, it seems that Berkson's Paradox is a much more involved fallacy than the Marginal-Conditional Fallacy [66]. Berkson's Paradox asserts that two diseases which are independent in the general population may become 'spuriously' associated in hospital-based case-control studies [66].

We devote the remaining part of this section for a *novel* visual consolidation of some of the notions and derivations reported herein. Fig. 4 is used to explore the possible relations among the events A, B, $\overline{A}$, and $\overline{B}$, by representing these events on area-proportional Venn diagrams. In these diagrams, the probability of an event is proportional to the area allotted for it. Contrarily to common practice, we do not depict the (noncomplementary) events A and B as circles or ellipses, but draw them as rectangles or trapezoids. The resulting straight-line-only diagrams look much similar to Karnaugh maps, and allow areas to be assessed readily and exactly. Figure 4(a) represents the case when events A and B are mutually favorable (and consequently when events $\overline{A}$ and $\overline{B}$ are also mutually favorable, while events A and $\overline{B}$ are mutually unfavorable and events $\overline{A}$ and B are also mutually unfavorable). By contrast, Fig. 4(b) denies both favorableness and unfavorableness within any of the four pairs of events $\{A, B\}$, $\{\overline{A}, B\}$, $\{A, \overline{B}\}$, and $\{\overline{A}, \overline{B}\}$, and thereby asserts mutual (statistical) independence between members of these pairs. Finally, Fig. 4(c) represents the case when events A and $\overline{B}$ are mutually favorable (and consequently when events $\overline{A}$ and B are also mutually favorable, while events A and B are mutually unfavorable and events $\overline{A}$ and $\overline{B}$ are also mutually unfavorable). Mutual exclusiveness between events A and B might be viewed as a case of extreme unfavorableness. These results are detailed in Table 1.

6. SIMPSON'S PARADOX

Simpson's Paradox occurs when the two events A and B enjoy the following characteristics

a) They are conditionally positively correlated given a third event C,
b) They are also conditionally positively correlated given the complement $\overline{C}$ of that third event,
c) They are, however, unconditionally negatively correlated.

These characteristics are expressed mathematically as [48]

$$P(A \cap B|C) \geqslant P(A|C)\, P(B|C) \tag{24a}$$

$$P(A \cap B|\overline{C}) \geqslant P(A|\overline{C})\, P(B|\overline{C}) \tag{24b}$$

$$P(A \cap B) \leqslant P(A)\, P(B) \tag{24c}$$

with at least one of the three inequalities being strict. Equations (24a)-(24c) constitute a Positive Simpson's Reversal. Their opposites, namely

$$P(A \cap B|C) \leqslant P(A|C)\, P(B|C) \tag{25a}$$

$$P\left(A \cap B|\overline{C}\right) \leqslant P\left(A|\overline{C}\right) P\left(B|\overline{C}\right) \tag{25b}$$

$$P(A \cap B) \geqslant P(A)\, P(B) \tag{25c}$$

constitute a Negative Simpson's Reversal (again with at least one strict inequality) [48]. We use Fig. 5 for a Karnaugh-map demonstration of a particular case of a Negative Simpson's Reversal. In this figure, the Karnaugh map serves as a convenient and natural sample space, and represents a pseudo-Boolean function rather than a Boolean one [21]. The full sample space involving the three variables A, B, and C in Fig. 5(a) is compacted *via* additive elimination [21] to the reduced one in Fig. 5(b), in which variable C is eliminated. Every two cells looped together in Fig. 5(a) are merged into a single cell (sharing the common values of the variables A and B in the two parent cells) in Fig. 5(b). The entries in the two parent cells are added to produce the entry in the corresponding merged cell of Fig. 5(b). Fig. 5 offers a good exercise in applying the definitions in Appendix A to deduce various probabilities from two versions of the same sample space. First, we note that $P(A \cap B|C) = \frac{1}{8}$, $P(A|C) = \frac{3}{8}$, and $P(B|C) = \frac{3}{8}$, so inequality (25a) is satisfied strictly since $\frac{1}{8} = \frac{8}{64} < \left(\frac{3}{8}\right)\left(\frac{3}{8}\right) = \frac{9}{64}$. Also $P\left(A \cap B|\overline{C}\right) = \frac{5}{11}$, $P\left(A|\overline{C}\right) = \frac{8}{11}$, and $P\left(B|\overline{C}\right) = \frac{7}{11}$, so inequality (25b) is satisfied strictly since $\frac{5}{11} = \frac{5\,5}{12\,1} < \left(\frac{8}{11}\right)\left(\frac{7}{11}\right) = \frac{5\,6}{12\,1}$. Now, $P(A \cap B) = \frac{6}{19}$, $P(A) = \frac{11}{19}$, and $P(B) = \frac{10}{19}$, so that inequality (25c) is satisfied strictly since $\frac{6}{19} = \frac{114}{361} > \left(\frac{11}{19}\right)\left(\frac{10}{19}\right) = \frac{110}{361}$. Therefore, the situation depicted by Fig. 5 is a Negative Simpson's Reversal. We include Simpson's paradox in our current study though it is simply a paradox rather than a fallacy *per se*, since its intriguing nature contributes to the troubles (and agony!) of medical personnel (and even statisticians) in their attempts to grasp concepts of probability. The terminology of Simpson's Paradox can also be confused with those of some of the fallacies discussed herein such as the Favorable-Event Fallacy and the Conjunctive Fallacy. Explorations of Simpson's paradox are based on the confounding or non-collapsibility phenomena or on realizing the need to use different analyses for identical data arising from different causal structures [67]. Many examples of Simpson's Paradox are available in the medical literature. In a now classical example, Julious and Mullee [35] report a study of two treatments of kidney stones in which the first treatment is more effective for both large and small stones and appears less effective when the data are aggregated (amalgamated) over the two types of stones.

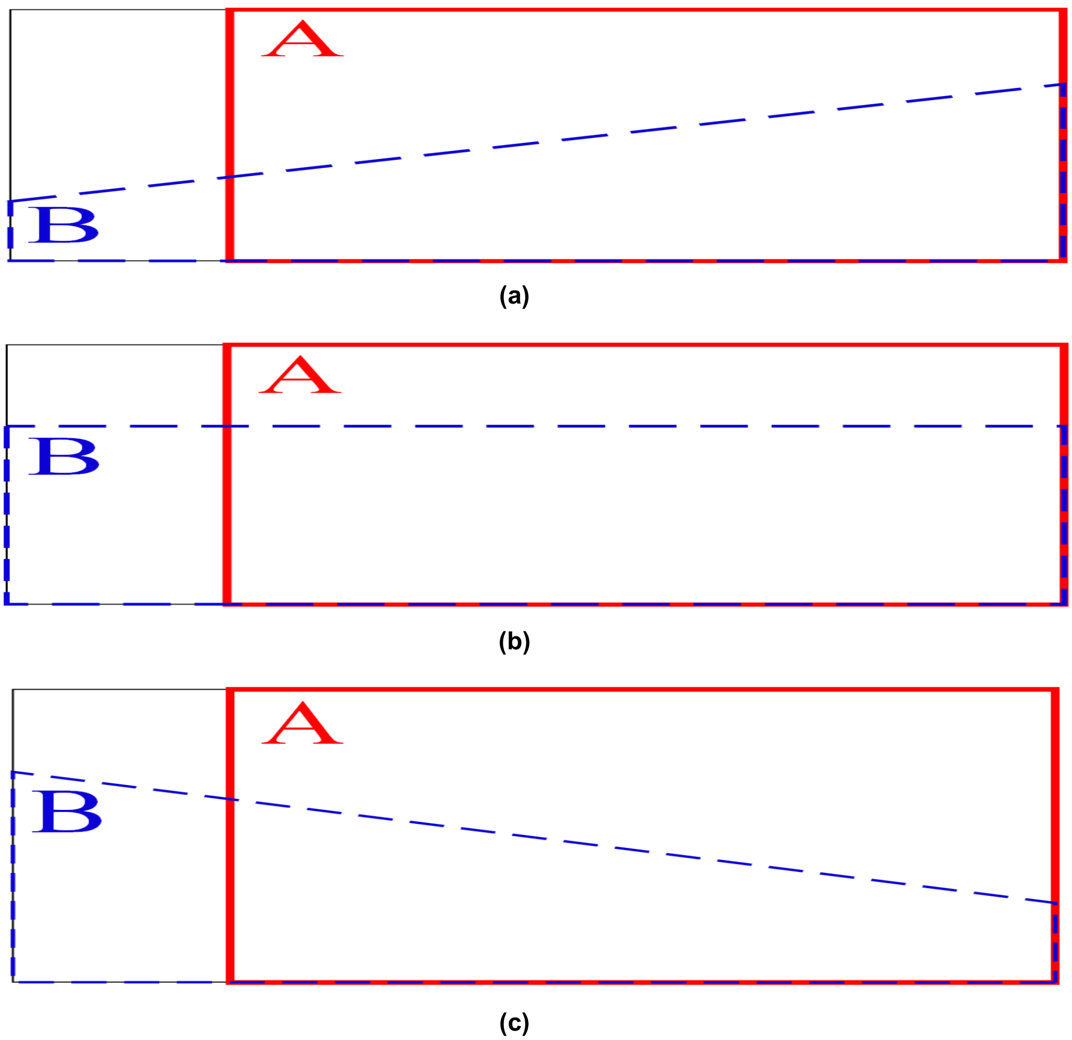

Fig. 4. Representation of events A and B *via* area-proportional Venn diagrams (The areas allotted to an event is proportional to its probability)

7. THE CONJUNCTION FALLACY

The Conjunction Fallacy considers the probability of the intersection of two events greater than that of one of the events.

$$P(A \cap B) > P(A) \quad \text{(WRONG)} \qquad (26)$$

The statement in (26) is obviously wrong since a measure for a subset cannot be strictly larger than that associated with a superset. Many people commit this fallacy by tending to ascribe a higher likelihood to a combination of events, "erroneously associating quantity of events with quantity of probability." In an experimental study of the Conjunctive Fallacy using medical stuff as the subject matter, and testing its spread among beginning medical students, Rao [68] presented the following vignette to the students.

"Amelia is a 23-year-old medical student who comes to your office for help. You suspect she has a common cold. In the blank spaces below, based on your knowledge and experience with the common cold, estimate the probability that Amelia would experience each of the following symptoms or symptom combinations. For example, if you believe Amelia has a 100% chance of experiencing "b" and a 90% chance of experiencing "c," put 100% and 90% in the respective blanks." Options given were (a) runny nose and diarrhea, (b) fatigue, (c) diarrhea, (d) ear pain and shortness of breath, (e) sore throat, and (f) headache."

Table 1. Possible cases of favorableness between two events and their complements

Fig.	Situation	Equivalent verbal descriptions	Equivalent mathematical descriptions
4(a)	$P(A \cap B) > P(A)\,P(B)$ $P(A \cap \overline{B}) < P(A)\,P(\overline{B})$ $P(\overline{A} \cap B) < P(\overline{A})\,P(B)$ $P(\overline{A} \cap \overline{B}) > P(\overline{A})\,P(\overline{B})$	A is favorable to B $\overline{A}$ is unfavorable to B B is favorable to A $\overline{B}$ is unfavorable to A $\overline{A}$ is favorable to $\overline{B}$ A is unfavorable to $\overline{B}$ $\overline{B}$ is favorable to $\overline{A}$ B is unfavorable to $\overline{A}$	$P(B\|A) > P(B) > P(B\|\overline{A})$ $P(A\|B) > P(A) > P(A\|\overline{B})$ $P(\overline{B}\|\overline{A}) > P(\overline{B}) > P(\overline{B}\|A)$ $P(\overline{A}\|\overline{B}) > P(\overline{A}) > P(\overline{A}\|B)$
4(b)	$P(A \cap B) = P(A)\,P(B)$ $P(A \cap \overline{B}) = P(A)\,P(\overline{B})$ $P(\overline{A} \cap B) = P(\overline{A})\,P(B)$ $P(\overline{A} \cap \overline{B}) = P(\overline{A})\,P(\overline{B})$	A is neither favorable nor unfavorable to B ($\overline{B}$) B is neither favorable nor unfavorable to A ($\overline{A}$) $\overline{A}$ is neither favorable nor unfavorable to B ($\overline{B}$) $\overline{B}$ is neither favorable nor unfavorable to A ($\overline{A}$) ***A* and *B* are independent**	Equations (20)
4(c)	$P(A \cap B) < P(A)\,P(B)$ $P(A \cap \overline{B}) > P(A)\,P(\overline{B})$ $P(\overline{A} \cap B) > P(\overline{A})\,P(B)$ $P(\overline{A} \cap \overline{B}) < P(\overline{A})\,P(\overline{B})$	A is favorable to $\overline{B}$ $\overline{A}$ is unfavorable to $\overline{B}$ $\overline{B}$ is favorable to A B is unfavorable to A $\overline{A}$ is favorable to B A is unfavorable to B B is favorable to $\overline{A}$ $\overline{B}$ is unfavorable to $\overline{A}$	$P(\overline{B}\|A) > P(\overline{B}) > P(\overline{B}\|\overline{A})$ $P(A\|\overline{B}) > P(A) > P(A\|B)$ $P(B\|\overline{A}) > P(B) > P(B\|A)$ $P(\overline{A}\|B) > P(\overline{A}) > P(\overline{A}\|\overline{B})$

Table 2. Listing of famous fallacies pertaining to probability and comparing fallacious formulas to correct one

Fallacy	Fallacious Formula	Correct formula
Multiplication Fallacy	$P(A \cap B) = P(A)P(B)$, A and B are general	$P(A \cap B) = P(A\|B)P(B)$ $=P(B\|A)P(A)$ ...... $P(A \cap B) = P(A)P(B)$, A and B are statistically independent
Addition Fallacy	$P(A \cup B) = P(A) + P(B)$, A and B are general	$P(A \cup B)$ $=P(A) + P(B) - P(A \cap B)$ ...... $P(A \cup B) = P(A) + P(B)$, A and B are mutually exclusive
Inverse Fallacy	$P(A\|B) = P(B\|A)$	$P(A\|B) = \frac{P(A)}{P(B)} P(B\|A)$
Conditional-Marginal Fallacy	$P(A\|B) = P(A)$ A and B are general	$P(A\|B) = P(A)$ A and B are statistically independent
Conjunction Fallacy	$P(A \cap B) > P(A)$ or $P(A \cap B) > P(B)$	$P(A \cap B) \leqslant min(P(A), P(B))$
The Appeal-to-Probability Fallacy	$\{P(A) > 0\} \Rightarrow \{P(A) = 1\}$	$\{P(A) > 0\} \Rightarrow \{P(A) \in (0,1]\}$ trivially
The Base-Rate Neglect	$P(A) = P(\overline{A}) = 0.5$	P(A) is not necessarily equal to $P(\overline{A})$ Typically For medical applications when P(A) denotes disease prevalence $P(A) \ll 1,\ P(\overline{A}) \simeq 1$

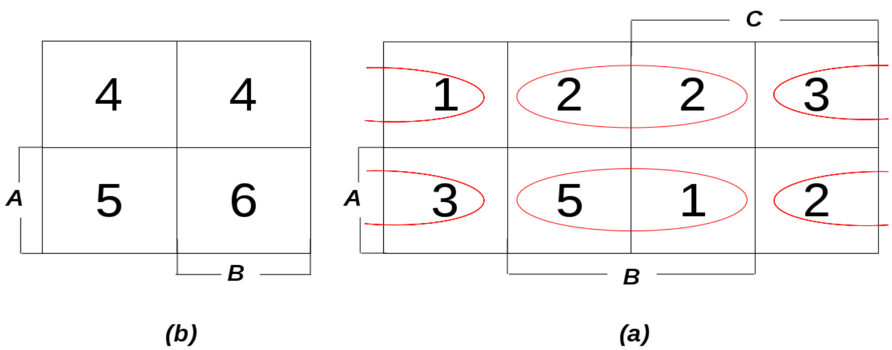

Fig. 5. Utilization of the Karnaugh map as a convenient and natural sample space to demonstrate Simpson's Paradox. The full sample space in (a) is compacted via additive elimination [21] to the reduced one in (b)

(0.792) (0.5) = 0.369 =(0.444) (0.89189)	(1– 0.904) (0.5) = (0.096) (0.5) = 0.048	0.396 + 0.048 = 0.444
0.5 – 0.396 = 0.104	0.5 – 0.048 = 0.452	0.104 + 0.452 = 0.556
0.5	0.5	1.0 = 0.444 + 0.556

Fig. 6. Complete solution of the second example of Sec. 2 under the Base-Rate Fallacy ignoring the prior knowledge of true prevalence. An extremely exaggerated value of 89.2% for the *PPV* is obtained

The common cold was chosen for the vignette above, since entering medical students have little or no clinical experience but are assumed at one point or another to have suffered themselves from the common cold and would have some knowledge of typical and atypical symptoms. Runny nose is widely known to be a common symptom; diarrhea is not [68]. A violation of the conjunction rule (i.*e*., Conjunction Fallacy) was recorded if diarrhea was assigned a lower probability than the combination of runny nose and diarrhea, regardless of the absolute assigned probability value or the values recorded for the other options [68]. In the exercise, the mean estimate of the probability of diarrhea was 17.2%. The mean estimate of the probability of the combination of runny nose and diarrhea was 31.6%. Overall, 47.8% of the students violated the conjunction rule by assigning a higher probability to runny nose and diarrhea than to diarrhea alone [68]. The moral of the study in [68] is that teaching medical students about the Conjunction Fallacy and other biases in assessment of probability has, in theory at least, the potential to improve students' decision making.

In passing, we note that while most people tend to overestimate a conjunctive probability [68]. a majority of people are also more likely to underestimate a disjunctive probability, which is a phenomenon referred to as the Disjunction Fallacy [59, 60].

8. THE APPEAL-TO-PROBABILITY FALLACY

The Appeal-to-Probability Fallacy (sometimes called the Appeal-to-Possibility Fallacy) equates a probable event to a certain one, *i.e*., it asserts that if A is not the impossible event, then it is the certain event

$$\{ A \neq \emptyset\} \Rightarrow \{A = S\} \qquad \text{(WRONG)} \quad (27a)$$

or

$$\{ P(A) > 0\} \Rightarrow \{P(A) = 1\} \qquad \text{(WRONG)} \quad (27b)$$

This fallacy is perhaps committed by patients rather than physicians. It is particularly misleading when its fallacious reasoning is preceded by some sort of wild guessing. A meticulous, doubting and suspicious person believes for sure that he/she definitely has a certain disease when he/she senses or imagines having some of its symptoms. Having the disease might (probably) be the case but cannot be taken for granted, and should not be assumed as a matter of fact. That is basically why a patient should seek medical help, consultation, and diagnosis, and why a physician should be well-trained to meet the expectations and needs of the patients. The essence of the diagnosis process is to employ scientific methodology to attribute correctly-observed symptoms to their genuine causes.

The Appeal-to-Probability Fallacy is one of notoriously-many logical flaws or errors [69-74] that might be collectively called appeal-to fallacies. These include Appeal to Accomplishment, Anger, Authority, Coercion, Coincidence, Common Belief, Common Sense, Equality, Emotion, Evidence Neglect, Expert Opinion, Extremes, Faith, Fear and Threat, Force, Human Nature, Ignorance, Intuition, Miracles, Misused Language, Money, Normality, Novelty, Pity, Popular Opinion, Ridicule, Self-Evidence, Stupidity, Tradition, Trust, or Wrong Reason.

9. THE BASE-RATE NEGLECT

Neglect of the Base Rate means substituting $P(A) = P(\overline{A}) = 0.5$ in the Total Probability Formula (8), so that this formula is inadvertently replaced by

$$P(B) = \left(P(B|A) + P(B|\overline{A})\right)/2 \quad \text{(WRONG)(28)}$$

To illustrate the grave consequences of the Base-Rate Fallacy, let us inadvertently disregard the important information given to us indicating a very low true prevalence of 0.01 in the first problem of Section 3. We instead use an arbitrary "true" prevalence of 0.5. Fig. 6 represents a normalized contingency Table detailing the solution steps in this case. The answer obtained for P(Cancer|positive test) now becomes 0.89189 or approximately 89.2% which is 11.59 times the correct answer of 7.7% obtained earlier in Fig. 2.

10. THE REPRESENTATIVE-SAMPLING FALLACY

One of the prevailing erroneous intuitions about probability is the belief that a sample randomly drawn from a population is highly representative of the population, i.e., similar to the population in all its "essential" characteristics [75]. This leads to the expectation that any two samples drawn from a particular population to be more similar to one another and to the population than sampling theory predicts, at least for small samples. In fact, the law of large numbers guarantees that very large samples will indeed be highly representative of the population from which they are drawn. The aforementioned intuitions about random sampling appear to follow an alleged law of small numbers [75, 76], which asserts that the law of large numbers applies also to small numbers (through a presumed self-corrective tendency).

Results of diagnostic testing, or other types of general experimental endeavor, are less "appealing" to those who obtain them when they are inconclusive and insignificant. By contrast, highly significant (and probably surprising) results are more informative and more desirable (albeit being frequently suspected to be too good to be true, and occasionally being thought of as fraudulent or fabricated). The credibility of these latter results, therefore, needs to be enhanced by replication. Contrary to a widespread belief, a replication sample might be required to be larger than the original one, and is (unreasonably) expected by skeptical users to be independently significant [75].

The Representative-Sampling Fallacy is only mentioned briefly herein. It is intimately related to fallacies and misconceptions of P values [77-90], which are significance levels that measure the strength of the evidence against the null hypothesis; the smaller the P value, the stronger the evidence against the null hypothesis [81]. These fallacies and misconceptions are probably the most ubiquitous, frequently misunderstood or misinterpreted, and occasionally miscalculated indices in biomedical research [86]. The topic of P values belongs to somewhat advanced statistics and is beyond the domain of elementary probability, and hence lies outside the scope of the current paper.

11. DISCUSSION

The naming, definition, and classification of fallacies vary according to the pertinent subject matter, adopted framework, and intended audience. It is beyond the capacity of any author to develop a complete coverage of all types of fallacies. Therefore, we limited our treatment of fallacies herein to medical subject matter, and restricted our framework to elementary (even simplistic) mathematics, and tailored our exposition to address a medical audience. Out of the extensive multitude of existing fallacies, we strived to cover a sample of the most frequently-encountered (and hopefully, the most representative). We hope our work might have some modest contribution towards the ultimate desirable goal of minimization, elimination, removal, suppression, and eradication of fallacious reasoning. Without sincere correcting and remedial efforts, perpetuated and unchallenged fallacies may proliferate so as to comprise a dominant portion of applicable knowledge.

Our strategy to confront fallacies herein is simply to unravel them from the point of view of elementary mathematics. There is a long history of research challenging such fallacies [91-96]. In particular, we note that Arkes [94] counts influence of preconceived notions among five impediments to accurate clinical judgment, and discusses possible ways to minimize their impact.

In passing, we discuss some other problematic notions and abbreviated rules of diagnostic testing, which are claimed even to be counterintuitive and misleading or to suffer from definitional arbitrariness. It is desirable that a diagnostic test possess high values for both Sensitivity and Specificity. Sensitivity is the probability of a positive test, given the presence of the disease $\{P(i=+1|j=+1)\}$, while specificity is the probability of a negative test, given the absence of the disease $\{P(i=-1|j=-1\}$. The natural inclination among many people is to think that a highly-sensitive test is effective at identifying persons who have the disease, and that a highly-specific test is effective at identifying those without the disease. By contrast, a highly-sensitive test is effective at ruling out the disease (when it is really absent), while a highly-specific test is effective at ruling in the disease (when it is really present). The following acronyms are used as mnemonics to help remember the aforementioned fact [97-100]

SnOUT : If **Sen**sitivity is high, a negative test will rule the disorder **OUT**.

SpIN : If **Sp**ecificity is high, a positive test will rule the disorder **IN**.

These two mnemonics might be sometimes misleading since they seem to be concerned with test characteristics only and do not stress enough the need to know the status of the patient. Therefore, they are being replaced [101-105] by the following more explicit forms, in which both test properties and patient status are specified.

SnNOUT : If **Sen**sitivity is high, a **N**egative test will rule the disorder **OUT**.
(For a highly-sensitive test, a positive test result is not very helpful, but a negative result is useful for asserting disorder absence).

SpPIN : If **Sp**ecificity is high, a **P**ositive test will rule the disorder **IN**.
(For a highly-specific test, a negative test result is not very helpful, but a positive result is useful for asserting disorder presence).

The assertion that: “If a test has high Sensitivity, a Negative result helps rule out the disease” might be mathematically understood as follows. If a person actually does have the disease $\{P(j=+1)=1\}$, we would expect a highly-sensitive test $\{P(i=+1|j=+1)\simeq 1\}$ to be positive with high probability $\{P(i=+1)\simeq 1\}$. Therefore, when a highly-sensitive test is negative, we can confidently assume disease absence (rule out the disease). Likewise, we interpret the assertion: “If a test has a high Specificity, a Positive result helps rule in the disease” mathematically as follows. If a person actually does not have the disease $\{P(j=-1)=1\}$, we would with high probability expect a highly-specific test $\{P(i=-1|j=-1)\simeq 1\}$ to be negative $\{P(i=-1)\simeq 1\}$. Therefore, when a highly-specific test is positive, we can confidently assume disease presence (rule in the disease).

12. SUMMARY AND CONCLUSIONS

We have studied probabilistic fallacies in medicine using simple mathematical representations and derivations. We summarize our results in Table 2, which shows the wrong proposition of each fallacy as well as an appropriate correction for it. The study made herein should hopefully be of significant help to medical students and medical practitioners alike. It might ensure that they acquire the necessary knowledge of elementary probability, but it does not demand that they gain too much knowledge that might distract them from their genuine (vital and critical) subject matter. It also attempts to remedy the notorious and grave ramifications of probabilistic fallacies residing as permanent misconceptions in their “private” knowledge databases. The material presented herein could also be of benefit to probability educators who deliberately want to engage their students in the learning process, *i.e.*, to guide them to be active learners. There are many reasons why 'active learning' in beneficial [106-111]. However, we believe that the single most important reason why it is so is the fact that it is the most effective method for unraveling misconceptions and eradicating fallacies.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Trivedi KS. Probability & statistics with reliability, queuing and computer science applications. John Wiley & Sons, New York, NY, USA; 2008.
2. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. Psychological Science in the Public Interest. 2007;8(2):53-96.
3. Swift L, Miles S, Price GM, Shepstone L, Leinster SJ. Do doctors need statistics? Doctors' use of and attitudes to probability and statistics. Statistics in Medicine. 2009;28(15):1969-1981.
4. Westbury CF. Bayes' rule for clinicians: An introduction. Frontiers in Psychology. 2010;1(192):1-7.
5. Thomas R, Mengersen K, Parikh RS, Walland MJ, Muliyil J. Enter the reverend: Introduction to and application of Bayes' theorem in clinical ophthalmology. Clinical & Experimental ophthalmology. 2011; 39(9):865-870.
6. Sanfilippo PG, Hewitt AW, Mackey DA. The Importance of conditional probability in diagnostic reasoning and clinical decision making: A primer for the eye care practitioner. Ophthalmic Epidemiology. 2017;24(2):81-89.
7. Casscells W, Schoenberger A, Grayboys T. Interpretation by physicians of clinical laboratory results. New England Journal of Medicine. 1978;299(18):999-1000.
8. Glantz SA. Biostatistics: How to detect, correct and prevent errors in the medical literature. Circulation. 1980;61(1):1-7.
9. Falk R. Inference under uncertainty via conditional probabilities. Studies in mathematics education: The Teaching of Statistics. 1989;7:175-184.
10. Anderson TW, Finn JD. Summarizing multivariate data: Association between categorical variables, chapter 6 in the new statistical analysis of data. Springer Science & Business Media; 1996.
11. Scherokman B. Selecting and interpreting diagnostic tests. The Permanente Journal. 1997;1(2):4-7.
12. Dunson DB. Commentary: Practical advantages of Bayesian analysis of epidemiologic data. American Journal of Epidemiology. 2001;153(12):1222-1226.
13. Fidler F, Thomason N, Cumming G, Finch S, Leeman J. Editors can lead researchers to confidence intervals, but can't make them think: Statistical reform lessons from medicine. Psychological Science. 2004; 15(2):119-126.
14. Lang T. Twenty statistical errors even you can find in biomedical research articles, Croatian Medical Journal 2004;45(4):361-370.
15. Strasak AM, Zaman Q, Marinell G, Pfeiffer KP, Ulmer H. The use of statistics in medical research: A comparison of The New England Journal of Medicine and Nature Medicine. The American Statistician. 2007;61(1):47-55.
16. Senn S. Three things that every medical writer should know about statistics. The Write Stuff. 2009;18(3):159-162.
17. Zhou XH, McClish DK, Obuchowski NA. Statistical Methods in Diagnostic Medicine. John Wiley & Sons. 2009;569.
18. Leeflang MMG. Systematic reviews and meta-analyses of diagnostic test accuracy. Clinical Microbiology and Infection. 2014;20(2):105-113.
19. Iwamoto K, Tajiri R, Fujikawa K, Yanagawa T. Is the second independent diagnostic test in medical diagnosis useful?. Japanese Journal of Biometrics. 2016;36(2):53-62.
20. Rushdi RA. Fetal malnutrition: Assessment by the CANS score versus anthropometry and impact on early neonatal morbidities,

unpublished master thesis, department of pediatrics, faculty of medicine at Kasr Al-Ainy, Cairo University, Cairo, Egypt; 2017. Available:https://www.researchgate.net/profile/Rufaidah_Rushdi/contributions
21. Rushdi RA, Rushdi AM. Karnaugh-map utility in medical studies: The case of Fetal Malnutrition. International Journal of Mathematical, Engineering and Management Sciences (IJMEMS). 2018; 3(3):220-244. Available:www.ijmems.in/ijmems—volumes.html
22. Rushdi AM, Talmees FA. An exposition of the eight basic measures in diagnostic testing using several pedagogical tools. Journal of Advances in Mathematics and Computer Science. 2018;26(3):1-17.
23. Rushdi RA, Rushdi AM, Talmees FA. Novel pedagogical methods for conditional-probability computations in medical disciplines, Journal of Advances in Medicine and Medical Research. 2018;25(10):1-15.
24. Diamond BL. The fallacy of the impartial expert. Archives of Criminal Psychodynamics. 1959;3(2): 221-236.
25. Blyth CR. Simpson's paradox and mutually favorable events. Journal of the American Statistical Association. 1973;68(343):746-746.
26. Gardner M. Fabric of inductive logic, and some probability paradoxes. Scientific American. 1976;234(3):119-122.
27. Fryback DG. Bayes' theorem and conditional non-independence of data in medical diagnosis. Computers and Biomedical Research. 1978;11(5):423-434.
28. Wagner CH. Simpson's paradox in real life. The American Statistician. 1982;36(1):46-48.
29. Eddy DM. Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic and A. Tversky (Eds.) Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press. 1982;249–267.
30. Tversky A, Kahneman D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review. 1983;90(4):293-315.
31. Thompson WC, Schumann EL. Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. Law and Human Behavior. 1987;11(3):167-187.
32. Friedmann D. The efficient breach fallacy. The Journal of Legal Studies. 1989;18(1): 1-24.
33. Clotfelter CT, Cook PJ. The gambler's fallacy in lottery play. Management Science. 1993;39(12):1521-1525.
34. Falk R, Greenbaum CW. Significance tests die hard: The amazing persistence of a probabilistic misconception. Theory & Psychology. 1995;5(1):75-98.
35. Julious SA, Mullee MA. Confounding and Simpson’s paradox. British Medical Journal, 1994;209(6967):1480–1481.
36. Evett IW. Avoiding the transposed conditional. Science and Justice. 1995;35(3):127-131.
37. Balding DJ. Interpreting DNA evidence: Can probability theory help? In Statistical science in the courtroom. Springer, New York, NY. 2000;51-70.
38. Jowett C. Lies, damned lies, and DNA statistics: DNA match testing, Bayes' Theorem, and the criminal courts. Medicine, Science and the Law. 2001;41(3):194-205.
39. Pfannkuch M, Seber GA, Wild CJ. Probability with less pain. Teaching Statistics. 2002;24(1):24-30.
40. Villejoubert G, Mandel DR. The inverse fallacy: An account of deviations from Bayes’ theorem and the additivity principle. Memory & cognition. 2002;30(2):171-178.
41. Leung WC. The prosecutor's fallacy—A pitfall in interpreting probabilities in forensic evidence. Medicine, Science and the Law. 2002;42(1):44-50.
42. Krämer W, Gigerenzer G. How to confuse with statistics or: The use and misuse of conditional probabilities. Statistical Science. 2005;20(3):223-230.
43. Batanero C, Sanchez E. What is the nature of high school students' conceptions and misconceptions about probability? In Exploring Probability in school Springer, Boston, MA. 2005;241-266
44. Kuklin B. Probability misestimates in medical care. Arkansas Law Review. 2006; 59:527-554.
45. Saks MJ, Koehler JJ. The individualization fallacy in forensic science evidence. Vanderbilt Law Review. 2008;61(1):199-219.
46. De Macedo C. Guilt by statistical association: Revisiting the prosecutor's fallacy and the interrogator's fallacy. The Journal of Philosophy. 2008;105(6):320-332.

47. Kalinowski P, Fidler F, Cumming G. Overcoming the inverse probability fallacy: A comparison of two teaching interventions. Methodology. 2008;4(4):152-158.
48. Pavlides MG, Perlman MD. How likely is Simpson's paradox? The American Statistician. 2009;63(3):226-233.
49. Groth RE. Teachers' construction of learning environments for conditional probability and independence. International Electronic Journal of Mathematics Education. 2010;5(1):1-55.
50. Fenton NE, Neil M. Avoiding probabilistic reasoning fallacies in legal practice using Bayesian networks, Australian Journal of Legal Philosophy. 2011;36:114-151.
51. Prakken H. On direct and indirect probabilistic reasoning in legal proof. Law, Probability and Risk. 2014;13(3-4):327-337.
52. Koehler JJ. Forensic fallacies and a famous judge. Jurimetrics. 2014;54(3): 211-219.
53. Mosteller F. Fifty challenging problems in probability with solutions. Courier Corporation; 1965.
54. Bar-Hillel M, Falk R. Some teasers concerning conditional probabilities. Cognition. 1982;11(2):109-122.
55. Falk R. Conditional probabilities: Insights and difficulties. In Proceedings of the Second International Conference on Teaching Statistics. 1986;292-297.
56. Nathan A. How not to solve it. Philosophy of Science. 1986;53(1):114-119.
57. Bar-Hillel M. How to solve probability teasers. Philosophy of Science. 1989; 56(2):348-358.
58. Maher N, Muir T. I don't really understand probability at all": Final year pre-service teachers' understanding of probability. Mathematics Education Research Group of Australasia. 2014;437-444. Sydney: MERGA.
59. Bar-Hillel M, Neter E. How alike is it versus how likely is it: A disjunction fallacy in probability judgments. Journal of Personality and Social Psychology. 1993; 65(6):1119-1131.
60. Costello FJ. Fallacies in probability judgments for conjunctions and disjunctions of everyday events. Journal of Behavioral Decision Making. 2009;22(3): 235-251.
61. Rushdi AM, Rushdi MA. Switching-algebraic analysis of system reliability, chapter 6 in Ram, M. and Davim, P. (Editors). Advances in Reliability and System Engineering. Management and Industrial Engineering Series. Springer International Publishing, Cham, Switzerland. 2017;139-161.
62. Boumans M. Battle in the planning office: Field experts versus normative statisticians. Social Epistemology. 2008;22(4):389-404.
63. Boumans M. The two-model problem in rational decision making. Rationality and Society. 2011;23(3):371-400.
64. Boumans M. Science outside the laboratory: Measurement in field science and economics. Oxford University Press; 2015.
65. Sloman SA, Over D, Slovak L, Stibel JM. Frequency illusions and other fallacies. Organizational Behavior and Human Decision Processes. 2003;91(2):296-309.
66. Snoep JD, Morabia A, Hernández-Díaz S, Hernán MA, Vandenbroucke JP. Commentary: a structural approach to Berkson's fallacy and a guide to a history of opinions about it. International Journal of Epidemiology. 2014;43(2):515-521.
67. Hernán MA, Clayton D, Keiding N. The Simpson's paradox unraveled. International Journal of Epidemiology. 2011;40(3):780-785.
68. Rao G. Probability error in diagnosis: The conjunction fallacy among beginning medical students. Family Medicine. 2009; 41(4):262-265.
69. Massey GJ. The fallacy behind fallacies. Midwest Studies in Philosophy. 1981;6(1): 489-500.
70. Jungwirth E. Avoidance of logical fallacies: A neglected aspect of science-education and science-teacher education. Research in Science & Technological Education. 1987;5(1):43-58.
71. Walton D. The appeal to ignorance, or argumentum ad ignorantiam. Argumentation. 1999;13(4):367-377.
72. Rushdi AM, Ba-Rukab OM. The modern syllogistic method as a tool for engineering problem solving. Journal of Qassim University: Engineering and Computer Sciences. 2008;1(1):57-70.
73. Floridi L. Logical fallacies as informational shortcuts. Synthese. 2009;167(2):317-325.
74. Rushdi AM, Ba-Rukab OM. An exposition of the modern syllogistic method of propositional logic. Umm Al-Qura University Journal: Engineering and Architecture. 2009;1(1):17-49.

75. Tversky A, Kahneman D. Belief in the law of small numbers. Psychological Bulletin. 1971;76(2):105-110.
76. Rabin M. Inference by believers in the law of small numbers. The Quarterly Journal of Economics. 2002;117(3):775-816.
77. Rozeboom WW. The fallacy of the null-hypothesis significance test. Psychological Bulletin. 1960;57(5):416-428.
78. Cohen J. The earth is round (p<. 05). American Psychologist. 1994;49(12):997-1003.
79. Kirk RE. Practical significance: A concept whose time has come. Educational and Psychological Measurement. 1996;56(5): 746-759.
80. Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. Annals of Internal Medicine. 1999;130(12): 995-1004.
81. Sterne JA, Smith GD. Sifting the evidence—what's wrong with significance tests? Physical Therapy. 2001;81(8):1464-1469.
82. Dixon P. The p-value fallacy and how to avoid it. Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale. 2003;57(3): 189-202.
83. Fidler F, Geoff C, Mark B, Neil T. Statistical reform in medicine, psychology and ecology. The Journal of Socio-Economics. 2004;33(5):615-630.
84. Moyé LA. Statistical reasoning in medicine: The intuitive P-value primer. Springer Science & Business Media; 2006.
85. Wagenmakers EJ. A practical solution to the pervasive problems of p values. Psychonomic Bulletin & Review. 2007; 14(5):779-804.
86. Goodman S. A dirty dozen: Twelve p-value misconceptions. In Seminars in Hematology, Elsevier. 2008;45(3):135-140.
87. Li XR, Li XB. Common fallacies in applying hypothesis testing. In Information Fusion, 2008 11th International Conference on (pp. 1-8). IEEE; 2008.
88. Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: An explanation for new researchers. Clinical Orthopaedics and Related Research. 2010; 468(3):885-892.
89. Verdam MG, Oort FJ, Sprangers MA. Significance, truth and proof of p values: Reminders about common misconceptions regarding null hypothesis significance testing. Quality of Life Research. 2014; 23(1):5-7.
90. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG. Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. European Journal of Epidemiology. 2016;31(4):337-350.
91. Cheever DW. The value and the fallacy of statistics in the observation of disease. The Boston Medical and Surgical Journal. 1861; 63(24):476-483.
92. Ingle DJ. Fallacies and errors in the wonderlands of biology, medicine, and Lewis Carroll. Perspectives in Biology and Medicine. 1972;15(2):254-283.
93. Christensen-Szalanski JJ, Bushyhead JB. Physicians' use of probabilistic information in a real clinical setting. Journal of Experimental Psychology: Human Perception and Performance. 1981;7(4): 928-935.
94. Arkes HR. Impediments to accurate clinical judgment and possible ways to minimize their impact. Journal of Consulting and Clinical Psychology. 1981;49(3):323-330.
95. Gigerenzer G. How to make cognitive illusions disappear: Beyond “heuristics and biases”. European Review of Social Psychology. 1991;2(1):83-115.
96. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Academic Medicine. 2003; 78(8):775-780.
97. Davidson M. The interpretation of diagnostic tests: A primer for physiotherapists. Australian Journal of Physiotherapy. 2002;48(3):227-232.
98. Suss R. Sensitivity and specificity: alien edition: A light-hearted look at statistics. Canadian Family Physician. 2007;53(10): 1743-1744.
99. Hayden SR. Basic statistics: Assessing the impact of a diagnostic test; choosing a gold standard, sensitivity, specificity, PPV, NPV, and likelihood ratios. Doing Research in Emergency and Acute Care: Making Order out of Chaos. 2015;205-222.
100. Williams KA, Harrild D, Williams DN. Statistics in nuclear cardiology. Annals of Nuclear Cardiology. 2016;2(1):174-177.
101. Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with “SpPIn” and “SnNOut”: A note of caution. BMJ. 2004;329(7459):209-213.

102. Aslan D, Sandberg S. Simple statistics in diagnostic tests. Journal of Medical Biochemistry. 2007;26(4):309-313.
103. Baudrillard J, Peat J, Barton B. Categorical and continuous variables: Diagnostic statistics. Medical Statistics: A Guide to Data Analysis and Critical Appraisal. 2008; 278-295.
104. Hegedus EJ, Stern B. Beyond SpPIN and SnNOUT: Considerations with dichotomous tests during assessment of diagnostic accuracy. Journal of Manual & Manipulative Therapy. 2009;17(1):1E-5E.
105. Denegar CR, Cordova ML. Application of statistics in establishing diagnostic certainty. Journal of Athletic Training. 2012; 47(2):233-236.
106. Garfield J, Ahlgren A. Difficulties in learning basic concepts in probability and statistics: Implications for research. Journal for Research in Mathematics Education. 1988;19(1):44-63.
107. Lord TR. 101 reasons for using cooperative learning in biology teaching. The American Biology Teacher. 2001; 63(1):30-38.
108. Prince M. Does active learning work? A review of the research. Journal of Engineering Education. 2004;93(3):223-231.
109. Modell H, Michael J, Wenderoth MP. Helping the learner to learn: The role of uncovering misconceptions. The American Biology Teacher. 2005;67(1):20-26.
110. Michael J. Where's the evidence that active learning works? Advances in Physiology Education. 2006;30(4):159-167.
111. Gürbüz R, Birgin O. The effect of computer-assisted teaching on remedying misconceptions: The case of the subject "probability". Computers & Education. 2012; 58(3):931-941.
112. Da Costa NC. Pragmatic probability. Erkenntnis. 1986;25(2):141-162.
113. Paolella MS. Fundamental probability: A computational approach. Wiley, New York, NY. USA; 2006.
114. Shafer G. Conditional probability. International Statistical Review/Revue Internationale de Statistique. 1985;53(3): 261-275.
115. Tomlinson S, Quinn R. Understanding conditional probability. Teaching Statistics. 1997;19(1):2-7.
116. Utts J. What educated citizens should know about statistics and probability. The American Statistician. 2003;57(2):74-79.
117. Carlton M. Pedigrees, prizes, and prisoners: The misuse of conditional probability. Journal of Statistics Education. 2005;13(2).
Available:http://ww2.amstat.org/publications/jse/v13n2/carlton.html
118. Ancker JS. The language of conditional probability. Journal of Statistics Education. 2006;14(2).
Available:http://ww2.amstat.org/publications/jse/v14n2/ancker.html
119. Manage AB, Scariano SM. A classroom note on: student misconceptions regarding probabilistic independence vs. mutual exclusivity. Mathematics and Computer Education. 2010;44(1):14-21.
120. Böcherer-Linder K, Eichler A, Vogel M. Understanding conditional probability through visualization. In Proceedings of the International Conference Turning Data into Knowledge: New Opportunities for Statistics Education. 2015;14-23.

APPENDIX A: ON CONDITIONAL PROBABILITY

There are three important and differing interpretations of probability, namely: the empirical, the logical, and the subjectivistic. Despite this disagreement on the meaning of probability, there is a widespread agreement on the basic axioms of the probability calculus and its mathematical structure [112]. We use the empirical (frequentistic or common-sense) interpretation, and base our probability notions on a probability "sample space" which constitutes the set of all possible (equally-likely) outcomes or primitive events of the underlying random "experiment." We describe events as subsets of the sample space, and hence get $\mathrm{N} = 2^n$ events for a sample space of n outcomes or sample points [1, 113]. Since events are sets, we can also describe events *via* elementary set operations (complementation, intersection, union, and set difference), namely:

The complement $\overline{A}$ of a set A is the set of all elements of the universal set S that are not elements of A. The intersection $A \cap B$ of two sets A and B is the set that contains all elements of A that also belong to B (or equivalently, all elements of B that also belong to A), but no other elements. It is also the set of all elements of the universal set S that are not elements of $\overline{A} \cup \overline{B}$.

The union $A \cup B$ of two sets A and B is the set of elements which are in A alone, in B alone, or in both A and B. It is also the set of all elements of the universal set S that are not elements of $\overline{A} \cap \overline{B}$.

The difference $A - B$ of two sets A and B is the set that contains all elements of A that are not elements of B. It is also the set of elements in $A \cap \overline{B}$.

Since the outcomes in a sample space are equally likely, the probability $P(A)$ of an event A is defined as the number of outcomes constituting A (favoring A) divided by the total number of outcomes in the sample space. The concept of conditional probability of event A given event B is based on the notion that event B replaces the universal set S as a certain event. Hence, this conditional probability is given by [1,6,9,21-23,113-120]

$$P(A|B) = P\,(A \cap B)/P\,(B), \qquad P(B) \neq 0 \tag{A1}$$

We utilize the Venn diagram and the Karnaugh map of Fig. A1 to visualize the concept of conditional probability. In this figure, $P(A|B)$ can be interpreted as the ratio of two areas, provided the diagram and map be drawn as area-proportional. Table A1 lists the values $P(A|B)$ for some special cases. Understanding these special cases is helpful for grasping the basic concepts of (and consolidating knowledge about) conditional probability. The case $A = \{i\}$ is typically used as a readily-comprehensible starting point for the introduction of conditional probability. Its aggregation over a multitude of sample points comprising A immediately produces the general definition (A1).

Conditional probability is just a probability; it satisfies the axioms of probability and it is a dimensionless quantity, and hence it is given by a numerical value with no unit associated with it. An Unconditional probability might be thought of as a probability without any restrictions, or a probability of an event conditioned on the certain event. We prefer to define the conditional probability $P(A|B)$ as the chance that an event A occurs given that an event B occurs. Some authors might paraphrase this statement as "Conditional probability represents the chance that one event *will occur* given that a second event *has already occurred*." This paraphrasing imposes an unwarranted sense of "sequentiality" rather than that of pure "conditionality". Unfortunately, conditional events in statistics sometimes become confusing if conceptualized as sequential [118].

Table A1. Value of the conditional probability $P(A|B)$ in important special cases

Special Case	Mathematical Description	Value of $P(A\|B)$ is
B is the impossible event	$B = \emptyset$ $P(B) = P(\emptyset) = 0$	not defined
A is the impossible event	$A = \emptyset$ $P(A \cap B) = P(\emptyset) = 0$	0
B is the certain event	$B = S$ $P(B) = P(S) = 1$	$P(A)$
A is the certain event	$A = S$ $P(A \cap B) = P(B)$	1
B is a singleton (primitive event) $\{j\}$ where j is a single outcome	$(A \cap B)$ equals $B = \{j\}$ if $j \in A$ and equals $\emptyset$ otherwise	1 if $j \in A$ and 0 otherwise
A is a singleton (primitive event) $\{i\}$ where i is a single outcome	$(A \cap B)$ equals $A = \{i\}$ if $i \in B$ and equals $\emptyset$ otherwise	$P(\{i\})/P(B)$ if $i \in B$ and 0 otherwise
A and B are mutually exclusive	$A \cap B = \emptyset$ $P(A \cap B) = 0$	0
B is a subset of A	$A \cap B = B$ $P(A \cap B) = P(B)$	1
A is a subset of B	$A \cap B = A$ $P(A \cap B) = P(A)$	$P(A)/P(B)$
A and B are independent	$P(A \cap B) = P(A)P(B)$	$P(A)$

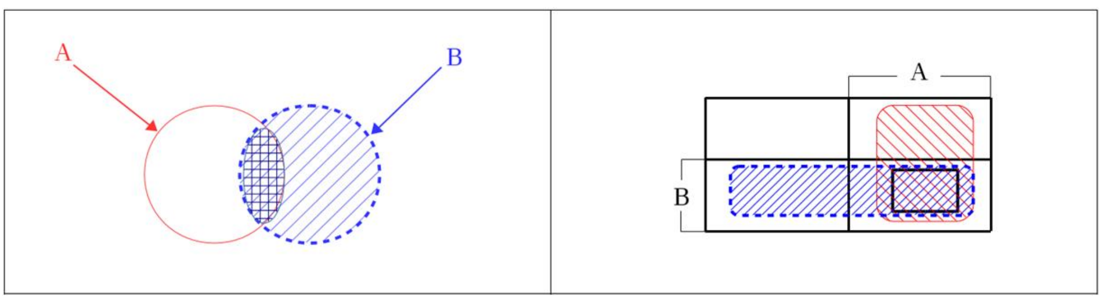

Fig. A1. Definition of the conditional probability $P(A|B)$ as the black area common to A and B, divided by the blue area of B (both the Venn diagram and the Karnaugh map are assumed to be area-proportional)

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/24107