A Study on the Concepts and Interpretations of Probability in Mathematical Analysis

*Dr. Durgesh Pareek

Abstract

The paper highlights the key distinctions between the measurement process and the dice-throwing procedure, and it makes some inferences about how the probability notion is used and understood in each scenario.

Keywords: measurement process, dice-throwing procedure, probability notion, uncertainty interpretations, physical probability, evidential probability, stochastic process, experimental framework, systematic effects, Bayesian inference

1. Introduction

Everyone agrees that probability plays a role in statistics. However, since the Tower of Babel, there has rarely been such a total breakdown of communication and debate over what probability is and how it relates to statistics. Undoubtedly, a large portion of the dispute is purely terminological and would vanish with careful examination.

Following a brief introduction to the various meanings of the terms "uncertainty" and "probability," this paper tries to restrict the example to a few different interpretations of probability and discuss whether applying this idea to dice throwing has the same meaning as applying it to measurement results, or if there are differences.

The idea of probability originated historically from conjecture on predictions in gambling difficulties, such as whether a particular face of a fair coin will appear in subsequent tosses or a fair dice in later throws.

When researchers in experimental science understood that measurements could never provide ideal, complete information, the concept of probability emerged as the most obvious solution to get around the problem and simulate chance.

By examining several viewpoints, the study demonstrates why and how, generally speaking, the conditions of its use are essentially different. There are some repercussions.

2. UNCERTAINTY INTERPRETATIONS

"Uncertainty" can mean different things. "Uncertainty is pervasive in the majority of the fields of science and technology (as it is in real life and typical thinking), and in all cases it seems to me that what really matters and worries are the quantification of the 'amount' of uncertainty in some sense attributable to the to consider statements or events, whether they are susceptible to repetition (random events) or not (singular events)," is a discussion of this topic from an epistemological point of view. The term "uncertainty" is generally used with a somewhat nebulous definition; however, it is nearly always done with the goal of measuring the variables that are impacted by uncertainty in a variety of scenarios.

In line with such method, "the linguistic term uncertainty has a very wide applicative spectrum in both ordinary life and science, in addition to being imprecise." In order to approach its possible meaning scientifically, it appears necessary to first document the various contexts and reasons in which their mother-predicate, U = unsure, which is the reverse of certain, is used. As a result, it can be represented by a fuzzy set.

In the realm of science, it not only introduced the idea of "probability," but also a plethora of other areas of chance, commonly referred to as "imprecise probability." These include, but are not limited to, belief operates possibility and necessity regulations lower and upper previsions, sets of desirable gambles, p-boxes, robust Bayes methods, lower and upper probabilities, or interval probabilities, as well as the possibility and fuzzy type of reasoning.

Specifically, it is not always the case that uncertainty measures are additive.

The purpose of measurement uncertainty is to define the intended meaning by adding that attribute: "measurement uncertainty: non-negative parameter characterising the dispersion of the number of values being assigned to a measurand, based on the information used." Here, it is also helpful to review Note 1: "Measurement uncertainty comprises definitional uncertainty as well as components resulting from regular effects, such as components related to corrections and the given quantity values of measurement standards. In certain cases, estimated systematic impacts are not adjusted for; instead, related uncertainty in measurement components is included (emphasis added), which is helpful for the discussion that follows.

3. INTERPRETATIONS OF PROBABILITY

Rather than using the word "certainty," probability was developed to include the idea of uncertainty, which is inherent in chance.

The two main types of probability interpretations are known as "physical" and "evidential" probabilities, respectively.

Physical probabilities are related to random physical systems like roulette wheels, rolling dice, and radioactive atoms. They are also known as objective or frequency probabilities. A particular kind of occurrence (like the dice producing a six) tends to happen at a consistent rate, or "relative frequency," across an extended period of trials in such systems. These steady frequencies can be explained or are explained by physical probability. Therefore, discussing physical probability only makes sense in the context of clearly described random experiments. Propensity accounts, like those of Popper, Miller, Giere, and Fetzer, and frequentist accounts, like those of Venn, Reichenbach, and von Mises, are the two primary categories of theories of physical probability.

Even in the absence of a random process, evidential probability, also known as Bayesian probability or subjectivist probability, can be applied to any statement to indicate its subjective plausibility—that is, how well the evidence supports the statement. Evidential probabilities are generally seen as degrees of belief, expressed in terms of propensities to wager at particular odds. The classical interpretation (e.g., Laplace's), the subjective interpretation (de Finetti and Savage), an epistemic or inductive interpretation (e.g., Ramsey, Cox), and the logical interpretation (e.g., Keynes, Carnap) are the four primary evidential interpretations.

Approaches to statistical inference, such as hypothesis testing and estimate theories, are linked to certain interpretations of probability. For instance, proponents of "frequentist" statistical approaches like R.A. Fischer, J. Neyman, and E. Pearson take the physical interpretation.

While the Bayesian school of statistics generally acknowledges the existence and significance of physical probabilities, they also believe that the computation of evidentiary probabilities is both necessary and legitimate. However, the interpretations of probability, not statistical inference theories, are the main subject of this article.

Since probabilities are examined in many different academic domains, the language used in this topic is a little perplexing. "Frequentist" is a very hard work. It alludes to a specific, largely abandoned theory of physical probability in the minds of philosophers. For scientists, however, "frequentist probability" is simply another term for objective or physical probability. For proponents of the Bayesian inference perspective, "frequentist statistics" refers to a method of statistical inference that just acknowledges physical probabilities. In addition, the term "objective" in probability can refer to evidentiary probabilities that are constrained by rational principles, such as logical and epistemic probabilities, but it can also refer to the exact meaning of "physical" in this context.

4. TOSING A COIN OR THROWING DICES

Since it is assumed that the dice (or coin) are "fair" (perfect) and that there is no interaction effect from the way the throw is executed, the natural world of the dice, or its impact with a surface, the required approach is strictly mathematical and has nothing to do with the science of the process. A toss's result is independent of time and is not influenced by any other factors. The throws are therefore thought to constitute a rigorously repeated ideal procedure for an infinitely long time, from which the equal and certain probability of receiving each face is derived. Furthermore, every throw is independent of every throw in the past or future, and every face is mutually exclusive. "Discrete probability" is another name for this kind.

According to this theory, uncertainty refers to the fact that the throwing process is strictly stochastic, making it difficult to forecast the outcome of the succeeding throws. Under ideal circumstances, there is no definitional ambiguity. The long-term repercussions of departures from the rigorously ideal assumption are also the subject of speculation, but this is the domain of experimental science.

There are numerous different frames that can be assimilated to the earlier ones (cards, etc.) in place of dice or coins.

5. THE GUIDE FOR EXPERIMENTATION

At least two characteristics define the scientific experimental framework:

a. the idea of sample repeatability is constrained, specifically in time, which thus limits the rationale for consistently viewing the process as (only) stochastic;

b. The idea of uncertainty encompasses not only the unpredictability of the subsequent measurement result—within specified, mostly always finite, bounds—but also the potential for systematic effects, such as those resulting from epistemic considerations.

The VIM3 defines a "repeatability condition" as follows: "condition of measurement, out of a set of situations that includes the same measurement method, same operators, same measuring system, same operating conditions, and same location, and replicate measures on the same or similar objects over a short period of time" (emphasis added). The phrase "short period of time" actually means "so short that the conditions for repeatability hold," hence it is very much a tautology. When feasible, it is considered required to conduct some form of independent verification of the condition's truth regarding two features:

(i) the measurand's (same or comparable objects) stability over time; and

(ii) the repeatability of the operational conditions (all other conditions). Fair dice or coins, on the other hand, are presumed to meet both of these requirements.

Regarding b., systematic effects create elements of uncertainty that aren't there when fair dice or coins are used. They are also referred to as "bias," signifying the variations between the measured values and a collection of reference values that constitute what is known as a "reference condition." An example of this would be Xi = ^{9}Xi + Bi for a single additive bias Bi influencing a quantity Xi (the "standard state" sign ^{9} is taken from physical chemistry for analogy's sake to represent the reference condition for which E(Bi) =: 0), and ^{9}Xi + Bi = ^{9}Xi – Ci, where Ci is the so-called "correction."

The so-called epistemic uncertainty, or the uncertainty resulting from inadequate understanding of known effects, is caused by a separate category of systematic impacts and manifests itself in an imperfect model of the conditions of the experiment.

This category is actually not all-inclusive and would be better divided into two different types of uncertainty: ontological (ignorance about (certain aspects of) the phenomenon being studied) and epistemic (imperfect understanding, specifically in science and technique).

When a impact quantity is mis-estimated or mis-modelled, the former happens. The latter includes, for example, the situation in which an influence quantity was left out of the model due to the fact that its effect was not recognised. It leads to a flawed model in both situations.

The "definitional uncertainty," which is not to be mistaken with any of the other categories of uncertainty, is still a separate one. It is about the fact that there are various (known) definitions of the measurand, which means that they are not unique. When it comes to recognised problems, it cannot be regarded as epistemic; nonetheless, whether or not to include any of the situations in the model depends on the experimenter's judgement; this is a feature of the model rather than a flaw in it.

Uncertainties in ontology, epistemology, and definition are not stochastic. The various elements of uncertainty in scientific experimentation are summarised in Figure 1.

Another aspect of the experiment frame that is helpful to take into account independently is the potential temporal dependence of the procedure being studied.

Time series variations or the non-repeatability of compounding data series over time could be the cause.

In the first instance, the terminology used to denote invariance with respect to time (static, stationary, etc.) or dependency on time (drifting, dynamic, etc.) may differ semantically or in various scientific and technological contexts. The reader is referred to the review conducted regarding the possibility of resulting confusion.

The two-sample variance (Allen variance) studies of frequencies standards provide an example of how the time scale itself may contribute to the appearance of various random effects. One cannot, however, infer from it that all variations over time are random; a mixed effect may accumulate, as is the case with the frequently seen "drifting" of an instrument's features from its first calibrated condition.

Figure 1 shows the four potential combinations of both known and unidentified knowledge, representing the many elements of uncertainty in the field of experimentation.

6. DIFFERENCES AND SOME IMPLICATIONS BETWEEN THE TWO FRAMES

There is an absolute limit to aleatory uncertainty. For example, after throwing the coin a thousand times, we could confidently state the likelihood that a head would land, but that's all we have to say about the next time we toss the coin (emphasis added).

Without the aleatoric components of uncertainty, probability cannot have the same significance in experimental research.

Three categories of uncertainty components are helpfully identified:

(i) Uncertainty in model parameters,

(ii) Uncertainty in influence quantities, and

(iii) Uncertainty in the model itself.

If referenced to the qualities of influence quantities, components (i) may be epistemic; if referred to their values that were measured, they may be random; or both may be considered. The influence quantities are typically separated into two categories: the "basic" and the "derived" ones, which are the ones that cause "bias", as well as those that require "corrections". While the latter can either be measured or have their values derived or inferred from prior knowledge, the former are measured indirectly (meaning they have at least one Type A uncertainty component) or only have Type B components.

Component (ii) may be ontological if it refers to missed quantities or epistemic if it refers to model imperfection. The definitional uncertainty is also included.

Component (iii) may be epistemic if the values are calculated or inferred, or stochastic if the values are derived from measurements.

Definitional uncertainty is typically resolved by defining the appropriate sort of definition; otherwise, it becomes an ontological risk component that is not addressed in the new GUM. Ontological uncertainty is typically not included in experimental science budgets.

It is common practice to assume ignorance about a position parameter, such as assuming a null mean and predicting a range for the resultant uncertainty component, in order to "randomise," or turn, epistemic uncertainty into a stochastic component. The latter can be set as the standard deviation, or a multiple of it, set by the selected confidence interval (or degree of belief), or as an interval (such as the "Maximum Permissible Error" (MPE, term 4.26 in 2013) or "Worst Case Uncertainty" (WCU), or a non-probabilistic interval).

7. FINAL COMMENTS AND AN EXAMPLE

According to the previous review, dice and related cases represent an oversimplification of a considerably more complex framework for uncertainty in experimental science, with the former focussing primarily on its intrinsic stochastic component.

In contrast, the main worry in the latter instance is generally the systemic consequences. Despite the possibility of a stochastic part, they ultimately include the necessity of an evaluation that essentially entails a subjective judgement that necessitates a decision. The seminal case (though a larger range of comparable situations in measurement apply) of the treatment for uncertainty in the evaluation of the values of fundamental constants is discussed from the perspective of measurement psychology and decision theory. This is an intriguing perspective that is essentially "external" to the metrology field.

Various explanations for the judgment's "bias" are taken into account, leading to what is referred to as their "overconfidence": We have consistently replicated a strong conclusion from laboratory studies of human judgement in multiple sets of examined measurements of physical constants: reported uncertainties are too small. How did this seeming arrogance come about? Such biases might develop rather accidentally from cognitive methods used in processing unclear information, according to experimental investigations of human judgement.

This attitude isn't always inadvertent, though; in one instance, it "relates to the methods selected to evaluate the uncertainty." When reporting results, physicists are advised to take into account all potential sources of systematic uncertainty. However, it is impossible to know how thoroughly individual scientists have investigated the uncertainty regarding their own experiments without clear rules about what to take into account and an explicit acknowledgement of the subjective aspects of uncertainty assessment. A second potential source of bias is that, in contrast to laboratory experiments on judgement, which can take great care to make sure that subjects are motivated to express their doubts candidly, real-world settings create other pressures. It is conceivable that some of the apparent overconfidence reflects a conscious decision to disregard the herder-to-assess sources of uncertainty (emphasis added).

In instance, "having a pre-existing recommended value may especially encourage researchers to discard or adjust unforeseen outcomes, and so induce correlated errors in apparently independent experiments" in the latter sense. According to these authors, this has been seen in a number of situations, specifically for the speed of light in vacuum (c_0), which serves as the basis for the unit of length, and the inverse of the fine structure constant (–1), which serves as the basis for the value of the electron change (e), or the unit of electrical current.

While the goal of the actual CODATA Task Group is also to adjust the value of the constants using the Least Squares Analysis (LSA) method, rather than using the mean—or a different one strictly statistical parameter—of the experimental values (based on probability), the analysis in focusses solely on the uncertainties associated with the recommended values.

Although the focus on the subjective aspect of uncertainty analysis may appear to exclude "frequentist" methods, this does not automatically imply that the Bayesian approach is always preferable. The authors state that the Bayes theorem is an uncontroversial aspect of probability theory in Chapter 1 of 22, pp. 30-31. Because it views probabilities as subjective, Bayesian inference is more contentious because it permits judgements that use a variety of evidence types. Evidence for frequentist probability must be of a single type, such as coin flips. If subjective perceptions satisfy coherence tests, they are merely probability. As a result, probabilities are more than merely a belief statement (emphasis added).

In reality, CODATA's stance fluctuated throughout the years, beginning with a stance where its primary strength was the initial critical evaluation of the data and its screening, rather than the application of the particular analytical treatment (the LSA). They later moved to a position that made use of all the data that was available. This was the 2010 role, which is currently the final one. Since a few outlier data would no longer significantly affect the final value and related uncertainty, it is possible that the plethora of data, in addition to the criticism of the subjective character of the screening, led to the final decision.

The dependability of the estimations of the constant values, specifically the suggested values for the years 1928–1973, is statistically analysed. A remarkable 57% is the "surprise index," which is calculated as the percentage of the population that falls outside the evaluated 98% confidence interval, or outside 2.33s. The experimental uncertainty has been significantly reduced in more recent experiments; therefore, the earlier statistics may have been much better. Nonetheless, the previously mentioned "bandwagon" effect of the suggested values may have grown in significance, warranting an immediate revision of these kinds of evaluations.

The following are the conclusions regarding that particular case: "It appears that there is widespread underestimate of uncertainty in measurements of physics constants and compilation of recommended values. This data expands on earlier findings of overconfidence in human judgement laboratory studies to a task domain of significant practical relevance. The utility of those measurements is greatly reduced if stated uncertainties do not accurately reflect the size of actual mistakes, whether as a result of judgement biases or insufficient analysis. For instance, this is essential for use in the "New SI."

Generally speaking, it turns out that using examples involving dice, coins, and the like is not particularly useful for learning how to handle experimental uncertainty and can even be misleading. As a result, it should be avoided.

8. APPENDIX: EXPERIMENTAL SCIENCE FALSIFICATION

With the requirement for judgement and decision-making, the previously stated subjectivism in experimental sciences necessitates "verifications" (from Wittgenstein onwards) or "falsification" (from Popper forward) criteria.

Due to the challenges posed by inevitable epistemological limits and the absence of a uniform criterion about the appropriate number of verifications, the first technique was later widely regarded as an unachievable aim.

The second approach was essentially developed in an environment devoid of ambiguity; yet, we discovered that, in the context of uncertain knowledge, a single instance of falsification in measurement, even if it is credibly demonstrated, cannot be deemed adequate. "Falsification is not possible without some threshold deviation which would be considered enough unlikely to reject the theory," in addition to the requirement for repeated occurrences.

"One of the defences of obscurantism which bar the way of scientific advance is the idol of certainty (including that of degrees of imperfect certainty or probability)," according to Popper (1936), and "the relations between probability and experience are also still in need of clarification." Investigating this issue will reveal what initially appears to be an insurmountable challenge to my methodological beliefs. Because, despite their crucial significance in empirical science, probability assertions are, in theory, immune to rigorous verification (emphasis added). Later, by initially putting out the theory of "propensity," which is an additional interpretation of probability, he was able to reconcile the concept of probability.

***Department of Mathematics**
Kamla Modi Govt. Girls College
Neem Ka Thana

REFERENCES

[1] B.N. Taylor, NBSIR 81-2426, 1982, National Bureau of Standards (today NIST), Gaithersburg USA.

[2] F. Pavese, Some problems concerning the use of the CODATA fundamental constants in the definition of measurement units, Metrologia 51 (2014) L1–L4, with online Supplementary Information.

[3] K.H. Ruhm, Are fluctuating measurement results instable, drifting or non-stationary? - A survey on related terms in Metrology, Journal of Physics Conference Series 459 (2013): 2043, doi: 10.1088/1742-6596/459/1/012043.

[4] Wikipedia, term "Probability interpretations", as consulted on February 24, 2015, and references therein.

[5] G.L.S. Shackle, A Non-Additive Measure of Uncertainty, The Review of Economic Studies 17 (1949-1950) 70-74.

[6] JGCM, Guide for the expression of uncertainty in measurement, Draft of 14 December 2014 delivered for comments to ISO TC69.

[7] M. Henrion, B. Fischhoff, Assessing Uncertainty in Physical Constants, Ch. 9 in "Judgement and Decision Making" (B. Fischhoff, Ed.) Routhledge (2013) pp. 172-187, 113649734X, 9781136497346, taken from American Journal of Physics 54, 791 (1986); doi: 10.1119/1.14447.

[8] K. Popper, The Logic of Scientific Discovery (Taylor & Francis e-Library ed.). London and New York: Routledge / Taylor & Francis e-Library (2005).

[9] L.J. Savage, The foundations of statistics. New York: John Wiley & Sons, Inc. (1954). ISBN 0-486-62349-1.

[10] Wikipedia, term "imprecise probability", as consulted on June 30, 2015, and references therein.

[11] J.R. Grozier, Falsificationism, Science and Uncertainty, MSc Dissertation, London Centre for the History of Sciences, Technology and Medicine, 2013, pp. 1–29.

[12] P.J. Mohr, B.N. Taylor and D.B. Newell, CODATA Recommended Values of the Fundamental Physical Constants: 2010, Rev. Modern Phys. 84 (2012) 1–94.

[13] T. O'Hagan, Dicing with the unknown, Significance 2005, 132–133.

[14] Enric Trillas, Some uncertain reflections on uncertainty, Archives for the Philosophy and History of Soft Computing, International Online Journal, Issue 1 (2013).

[15] J.R. Grozier, Can Theories be Falsified by Experiment?, MSc essay (UCL, London, 2012).

[16] Wikipedia, term "uncertainty quantification", as consulted on June 30, 2015, and references therein.

[17] F. Pavese: "Mathematical and statistical tools in metrological measurement", 2013, Chapter in Physical Methods, Instruments and Measurements, [Ed. UNESCO-EOLSS Joint Committee], in Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford, UK, http://www.eolss.net.

[18] Zadeh, L. A., Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1 (1978) 3–28. doi:10.1016/0165-0114(78)90029-5.

[19] F. Pavese, Key comparisons: the chance for discrepant results and some consequences, Acta IMEKO, 2015, in press.

[20] Dubois, Didier; Henri Prade (1985). Théorie des possibilité. Masson, Paris.

[21] F. Pavese, Subjectively vs. objectively-based uncertainty evaluation in metrology and testing, IMEKO-TC1-TC7-2008-030, available on http://www.imeko.org/index.php/proceedings# under TC7 (2008); F. Pavese, On the degree of objectivity of uncertainty evaluation in metrology and testing" Measurement 2009 42 1297–1303.

[22] da Silva R.J., Santos J.R., Camões M.F., Worst case uncertainty estimates for routine instrumental analysis, Observations on the worst case uncertainty, Analyst 127 (2002) 957–963; L. Fabbiano, N. Giaquinto, M. Savino, G. Vacca, Observations on the worst case uncertainty, Journal of Physics: Conference Series 459 (2013): 2043, doi:10.1088/1742-6596/459/1/012038.

[23] BIPM, Guide for the expression of uncertainty in measurement; JCGM 00:2008, ISO Geneva, at http://www.bipm.org/en/publications/guides/gum.html.

[24] A. Der Kiureghian, O. Ditlevsen, Aleatory or epistemic? Does it matter?, Special Workshop on Risk Acceptance and Risk Communication, March 26, 2007, Stanford University, pp. 1–13.

[25] D. Skulj, Non-additive probability, 2002. https://www.stat.aau.at/Tagungen/Ossiach/Skulj.pdf.

[26]CODATAhttp://physics.nist.gov/cuu/Constants/index.html,

http://www.bipm.org/extra/codata/.

[27] L.P. Swiler, T.L. Paez, R. L. Mayes, Epistemic Uncertainty Quantification Tutorial, Proceedings of the IMAC-XXVII, 2009 Orlando, Florida USA, 2009 Society for Experimental Mechanics Inc.

[28] I. Gilboa, D. Schmeidler, Additive representations of non-additive measures and the Choquet integral, Annals of Operations Research 52 (1994) 43-65.

[29] Hájek, Alan, Zalta, Edward N., ed., Interpretations of Probability, The Stanford Encyclopedia of Philosophy, online.

[30] F. Pavese and P. De Bièvre, Fostering diversity of thought in measurement science, in Advanced Mathematical and Computational Tools in Metrology and Testing X (F. Pavese, W. Bremser, A. Chunovkina, N. Fischer, A.B. Forbes, eds.), Series on Advances in Mathematics for Applied Sciences vol. 86, World Scientific, Singapore, 2015, pp. 1–8.

[31]M. Squair, Epistemic, ontological and aleatory risk, 2009, blog http://criticaluncertainties.com/2009/10/11/epistemic-and-aleatory-risk/.

[32] K. Popper, The Propensity Interpretation of the Calculus of Probability and of the Quantum Theory. In Observation and Interpretation, Korner & Price (eds.), Buttersworth Scientific Publications, 1957, pp. 65–70.

[33] de Elía, Ramón; Laprise, René, Diversity in interpretations of probability: implications for weather forecasting, Monthly Weather Review 133 (2005) 1129–1143.

[34] BIPM (2012) International vocabulary of basic and general terms in metrology (VIM), 3rd edn. BIPM/ISO, JGCM 200:2012. 2008 version with minor corrections. http://www.bipm.org/en/publications/guides/vim.html.