The Wine/Water Paradox:
background, provenance and proposed resolutions

Michael A.B. Deakin

Abstract

The so-called Wine/Water Paradox is one of a number of puzzles associated with elementary probability theory. It has recently been the subject of renewed interest with the proposal of two competing attempts at resolution. The origins of the paradox will be explored, as will some of its consequences for the understanding of the nature of probability. It will then be argued that the attempted resolutions do not carry the force their proposers attach to them.

1 Statement of the Paradox

The Wine/Water Paradox has been given in a number of forms, differing mostly in the numerical values of the various parameters involved. For convenience, I adopt those given in a recent re-examination of it [10].

> A mixture is known to contain a mix of wine and water in proportions such that the amount of wine divided by the amount of water (both measured in terms of some agreed unit) is a ratio x lying in the interval $1/3 \leq x \leq 3$. We seek the probability, P^* say, that $x \leq 2$.

The problem is, in one sense, insoluble, because insufficient information is supplied to enable a rigorous solution. However, some would argue, according to the so-called Principle of Indifference, that, failing such information, we should suppose that x is uniformly distributed. That is to say that for any number a in the given interval,

$$\text{Prob}\{x \leq a\} = \frac{1}{8}(3a - 1).$$

From this it follows that

$$P^* = \text{Prob}\{x \leq 2\} = \frac{5}{8}.$$

However it is perfectly possible to argue that the amount of wine divided by the amount of water (both measured in terms of the same unit as before) is a ratio y also lying in the interval $1/3 \leq y \leq 3$. If we now suppose (with the same validity as before) that y is uniformly distributed over this interval, we have for $1/3 \leq a \leq 3$:

$$\text{Prob}\{y \geq a\} = \frac{3}{8}(3 - a).$$

From this it follows that

$$P^* = \text{Prob}\{x \leq 2\} = \text{Prob}\left\{y \geq \frac{1}{2}\right\} = \frac{3}{8}\left(3 - \frac{1}{2}\right) = \frac{15}{16}.$$

The paradox lies in the discrepancy between the two values of P^*.

2   Provenance of the Paradox

The origins of the paradox are somewhat obscure. Mikkelson [10] attributes it to von Mises, but gives no details. However, von Mises himself [11] claims to follow Poincaré [12] in attributing it to Bertrand [1]. [I include the Poincaré reference for completeness, although it has not been available to me.] Certainly Bertrand produced a number of paradoxes in his exposition of the laws of probability. The best-known is his Chord Paradox, which is briefly discussed in an appendix to the present paper. I have been unable to find in Bertrand's book any discussion of the Wine/Water Paradox or any equivalent to it. (However, see the appendix.) It is therefore likely that the attribution of the paradox to Bertrand is in fact a misattribution.

Gillies [6] also presents the paradox, and he attributes it to Keynes [8], who did in fact espouse a form of the Principle of Indifference, which he termed the Principle of Irrelevance. Gillies thus finds it admirable that Keynes devotes time to a careful examination of the objections to the principle. Keynes, however, in his turn, attributes the paradox to von Kreis [9] (whose work is also here referenced for completeness, although it too has not been available to me). Keynes, in his account, regards the paradox as saying that there are limits to the extent to which we can precisely quantify probabilities of this type.

Something of Keynes' approach is apparent in the work of Schlesinger [13]. Indeed, Schlesinger goes to some lengths to make more explicit what Keynes merely indicates. In particular, and for purposes of illustration, he uniformly distributes neither x nor y but another quantity (see Equation (2) below), and so reaches the value

$$P^* = \frac{5}{6}.$$

Mikkelson [10] describes Schlesinger as "[one of the] lonely defenders of PI [the Principle of Indifference]". The other is Jaynes [7], whose discussion of the Wine/Water Paradox is peripheral to an attempted resolution of the Bertrand Chord Paradox. Mikkelson describes both as "[throwing] up their hands at [the Wine/Water] Paradox", and certainly this is a fair description of the discussion by Jaynes.

Mikkelson presents his own attempted resolution of the paradox, but there has also recently been another. This is by Burock [3], and it produces yet another value for P^*. These two "solutions" will both be examined below, but before this we need to consider some basics.

3   A Simple Resolution

It is important to note not only what we *are* told about the situation, but also what we are *not*. We have no information at all in the statement of the problem as to the frequency distribution of x. The cumulative distribution $\text{Prob}\{x \le a\} = (3a - 1)/8$ of Section 1 is supplied as an assumption to substitute for information that is simply not there. The same can be said of the assumed distribution of y. The assumption of uniformity means that these cumulative distributions are linear. As $y = 1/x$ is not itself a linear function, the two cumulative distributions cannot simultaneously be linear and hence x and y cannot both be uniformly distributed. This is what Jaynes [7] refers to when he describes the problem as "overdetermined".

Any cumulative probability distribution could have been used for $\text{Prob}\{x \le a\}$. That is to say, we could set

$$\text{Prob}\{x \le a\} = P(a),$$

where the only restrictions on $P(a)$ are that it should be an increasing function of a for which $P(1/3) = 0$ and $P(3) = 1$. It follows that the required probability P^* could be any number at all between 0 and 1.

However, those who, like Mikkelson [10] and Schlesinger [13], seek to defend the use of uniform distributions do so from the point of view of either subjectivist or of "logical" theory. For the subjectivists, probability seeks to quantify a degree of personal belief, rather than to represent a measurable underlying reality. But if we consider the ramifications of this position, we see no paradox at all. All that has happened is that Subject 1, using a uniform distribution on x, comes up with one subjective probability, while Subject 2, using a uniform distribution on y, comes up with another. There is no reason whatsoever for their subjective probabilities to agree! What is true, of course, is that there is no "objective reason" to prefer either subjective probability over the other. This might be seen by some as an argument against the underlying theory.

Now consider a variation: the "logical theory". This attempts to replace the concept of personal belief with that of "rational belief". This concept advances the view that the probability to be assigned to an event is the probability that a rational mind would assign to it, given the available data. If this concept can be shown to be viable, then clearly we do indeed have a paradox that needs to be resolved, or as Mikkelson has it, "dissolved".

However, the mere naming of such a position is not enough to demonstrate that it is consistent. The burden of proof lies with the advocates of such a view. To say this is essentially to state in starker form Ramsey's objection to the account espoused by Keynes [8, Chapter IV], who used not quite the Principle of Indifference but a closely related Principle of Irrelevance. There is a good analysis of this debate by Cottrell [4, pp. 30–35] and a summary and critique of both points of view by Gillies [6, Chapters 3 and 4]

It could indeed be urged that the Wine/Water Paradox constitutes a *reductio ad absurdum* of the logical theory. In this connection, it is of interest to note that Jaynes [7, p. 490], who has considerable sympathy with the "logical theory", does not accept that the Wine/Water Paradox is susceptible of sensible resolution.

4   Mikkelson's Proposed Resolution

In order to consider the alternative resolution favored by Mikkelson [10] and in fact mooted, before him, by Schlesinger [13, p. 191], introduce some further notation. Let:

      W_1 be the quantity of water, and
      W_2 be the quantity of wine,

both in the same agreed unit. In terms of the earlier notation, $x = W_1/W_2$ and $y = 1/x$.

Mikkelson imposes the condition

$$W_1 + W_2 = 1 \tag{1}$$

(which may be done without any loss of generality) and then considers the distribution of a quantity h defined by

$$h \ (= W_1) \ = \ \frac{x}{1+x} \ . \tag{2}$$

This variable is then supposed to be uniformly distributed over the relevant interval, which is easily seen to be $1/4 \leq h \leq 3/4$, and it is shown that on *this* hypothesis, the required probability $P^* = 5/6$.

He further shows that if h is expressed in terms of y rather than x the same result obtains, and that the variable h has an attractive intuitive interpretation in terms of the height of a virtual separation surface in the imagined event that the two liquids do not in fact mix.

It will be shown in Section 6 that this initially attractive resolution of the paradox is not unique, and that a fuller analysis leads us to conclude that the required probability P^* could be any number between $1/2$ and 1.

5 Burock's Proposed Resolution

Burock [3], noting that the condition $xy = 1$ is imposed on any proposed joint probability distribution over (x, y), champions the use of arclength along this curve between the end-points $(1/3, 3)$ and $(3, 1/3)$ as the quantity to be uniformly distributed. Thus Burock has

$$P(a) = \frac{\int_{1/3}^{a} x^{-2}\sqrt{1 + x^4}\,\mathrm{d}x}{\int_{1/3}^{3} x^{-2}\sqrt{1 + x^4}\,\mathrm{d}x}. \tag{3}$$

This expression can be put into a (very cumbersome) closed form involving elliptic integrals. However, this course will not be pursued here. Rather, as with Burock's analysis, mine will proceed in terms of numerical computation. Put

$$F(a) - F(1/3) = \int_{1/3}^{a} x^{-2}\sqrt{1 + x^4}\,\mathrm{d}x. \tag{4}$$

Then Burock's solution is

$$P^* = \frac{F(2) - F(1/3)}{F(3) - F(1/3)}. \tag{5}$$

This evaluates as $3.2787\ldots/4.2932\ldots = 0.7637\ldots$.

6 A Generalized Analysis

Mikkelson's function h is by no means the only one giving the same result whether we proceed in terms of x or of y. In order to generalize, replace Equation (2) by the more general

$$h = f(x), \tag{6}$$

where, to impose symmetry between x and y and to preserve Equation (1), we require

$$f(x) + f(1/x) = 1. \tag{7}$$

Equation (7) is a functional equation, one possible solution of which is given by Equation (2), as the reader may verify. Because of the imposed symmetry, a uniform distribution over $f(x)$ will also force a uniform distribution over $f(y)$ and *vice versa*.

We now proceed to the general solution of Equation (7). To this end, set

$$\frac{f(x)}{f(1/x)} = \frac{f(x)}{1 - f(x)} = \varphi(x) \tag{8}$$

from which it follows that

$$f(x) = \frac{\varphi(x)}{1 + \varphi(x)} \tag{9}$$

Also, from (8), $\varphi(1/x) = f(1/x)/f(x)$ and hence

$$\varphi(x)\varphi(1/x) = 1. \tag{10}$$

Mikkelson's function h, Equation (2), corresponds to the simple solution $\varphi(x) = x$, but this solution is by no means forced upon us. An obvious generalization is $\varphi(x) = x^n$, but solutions even more general than this are clearly possible.

In order to set up a uniform distribution and to derive meaningful results from it, we require the correspondence between x and $f(x)$ to be 1–1. The relationship between f and

ϕ is 1–1 by virtue of Equations (8) and (9). It follows that the relation between x and ϕ is 1–1 and thus that the function $\varphi(x)$ is monotonic. Without loss of generality, suppose it to be increasing with x. Directly from Equation (10), $\varphi(1) = 1$. Otherwise $\varphi(x)$ is arbitrary on the interval $1 \leq x \leq 3$ and we may define it there (under these very broad constraints) *ad lib*. Then on the interval $1/3 \leq x \leq 1$, define $\varphi(x)$ by direct appeal to Equation (10).

Thus $\varphi(x)$, and hence $f(x)$, and so h, may be defined in (infinitely) many different ways. For a general $f(x)$ satisfying Equation (7),

$$P^* = \frac{f(2) - f(1/3)}{f(3) - f(1/3)} \tag{11}$$

which, by Equation (9) and following some simplification, yields

$$P^* = \frac{\varphi(3)\varphi(2) - 1}{[1 + \varphi(2)] \cdot [\varphi(3) - 1]} \ . \tag{12}$$

The only constraint on the value of $\varphi(2)$ is provided by the condition noted above, namely $1 = \varphi(1) < \varphi(2) < \varphi(3)$. The value of P^* varies monotonically with the value of $\varphi(2)$ from a minimum of $1/2$, the limit as $\varphi(2) \to 1$, to 1, the limit as $\varphi(2) \to \varphi(3)$. This reduces the range of possible values of P^* to one half of the general value discussed in Section 2 above. This reduction is directly attributable to the Symmetry condition imposed by Equation (7).

Comparison between Equations (5) and (11) shows that Burock's proposed resolution amounts to the identification $f(x) = F(x)$. The fact that $F(x)$ satisfies Equation (7) is capable of direct demonstration, which I leave to the reader.

7   A Sociological Remark

The two recent attempts at resolution effectively cancel one another out. Both Mikkelson and Burock claim to have produced a uniquely correct value for P^*. This means that both cannot be "right" unless there is no agreed definition of "uniquely right". Mikkelson's resolution certainly selects the simplest solution of Equation (10), but Burock's geometric argument has its own intrinsic charm. Which we regard as "simplest" is a matter of personal choice. But this is to wander outside Mathematics and into sociology!

Appendix

Although Bertrand [1] does not deal with the Wine/Water Paradox directly, nor with any direct equivalent of it, he does consider a close relative (certainly not a clone but perhaps a sibling, or at most a first cousin). This is the so-called Horizon Paradox [1, p. 5].

A plane is chosen "at random" in space. What is the probability that it makes an angle of less than $\pi/4$ with the horizon [i.e. a given plane]?

If we term the required angle θ, and if θ is regarded as uniformly distributed over the interval $0 \leq \theta \leq \pi/2$, then the required probability is $1/2$. On the other hand, if we regard $\cos\theta$ as uniformly distributed over the interval $1 \geq \cos\theta \geq 0$, then the required probability is $1 - 1/\sqrt{2} \approx 0.29$. [The choice of $\cos\theta$ as the required function is motivated by a geometric argument.]

The Horizon Paradox seems not to have attracted the same attention as either the Wine/Water Paradox or the Chord Paradox. This latter is introduced immediately prior to the discussion of the Horizon Paradox. Bertrand comments: "This question [the Horizon Paradox], like the previous one [i.e. the Chord Paradox] is ill-posed, and the two conflicting answers constitute the proof [of this]" (my translation).

This has indeed been the almost universal judgment on the Chord Paradox. This paradox requests the probability that a chord drawn "at random" in a circle has a length greater than the side of an inscribed equilateral triangle. Bertrand himself used three different interpretations of "at random" to produce answers of 1/3, 1/2 *and* 1/4. Subsequently, Czuber [5] gave plausible arguments leading to three more possible answers, and Bower [2] later extended the number of plausible answers to ∞.

In a more recent discussion, Jaynes [7] attempted to argue that the answer 1/2 was the "correct" one. He lists various earlier discussions, all at variance with his conclusion, except for a somewhat vague statement by Borel. His otherwise full list overlooks a contribution from Keynes [8, p. 63], who rather imaginatively views the chord as a limiting case of an inscribed plane figure. He also does less than full credit to Uspensky [14, p. 251]. This is a more serious oversight, for Jaynes claims to have verified his result experimentally. This is not really the knockdown argument he supposes, for his experimental design, in which "long" straws were dropped onto a circular outline, almost guarantees this result. Uspensky envisaged a similar *gedankenexperiment* in which a disk was dropped onto a ruled grid to the same effect. In essence, Jaynes imposes a further condition on the problem in order to reach his uniqueness result. Recently, however, the Jaynes solution has been influentially endorsed by Weisstein [15].

Acknowledgement

References

[1] J. Bertrand, *Calcul des probabilités*, (Gauthier Paris 1889).

[2] O.K. Bower, *Note concerning two problems of geometric probability*, Am. Math Monthly **41** (1934), 506–510.

[3] M. Burock, *Indifference, sample space, and the wine/water paradox*, http://philsci-archive.pitt.edu/archive/00002487/01/Indifference_new_...Burock_2005.pdf (2005).

[4] A. Cottrell, *Keynes's theory of probability and its relevance to economics. Three theses*, Economics and Probability **9** (1993), 25–51.

[5] E. Czuber, *Geometrische Wahrscheindlichkeiten und Mittelwerte*, (Teubner Leibzig 1908).

[6] D. Gillies, *Philosophical Theories of Probabilitya,* (Routledge London 2000).

[7] E.T. Jaynes, *The well-posed problem*, Foundations of Physics **3** (1973), 477–492.

[8] J.M. Keynes, *Treatise on Probability*, (Macmillan London 1921).

[9] J. von Kreis, *Die Principien der Wahrscheinlichkeitsrechnung*, (Freiburg 1886).

[10] J.M. Mikkelson, *Dissolving the wine/water paradox*, Brit. J. Phil. Sci. **55** (2004), 137–145.

[11] R. von Mises, *Probability, Statistics and Truth*, Second revised English Edition (Allen and Unwin London 1957) (Translation of a 1928 German original).

[12] H. Poincaré, *Calcul des probabilités*, 2nd ed. (Gauthier–Villars Paris 1912), see pages 118–152.

[13] G.N. Schlesinger, *The Sweep of Probability*, (University of Notre Dame Press 1991).

[14] J.V. Uspensky, *Introduction to Mathematical Probability*, (McGraw–Hill New York 1937), see page 251.

[15] E. Weisstein, *Bertrand's problem*, *MathWorld*–A Wolfram Web Resource, http://mathworld.wolfram.com/BertrandsProblem.html

School of Mathematical Sciences, Monash University, VIC 3800
E-mail: Michael.Deakin@sci.monash.edu.au