Decision Making Under Uncertainty in the *Treatise on Probability*

Keynes' Mathematical Solution of the 1961 Ellsberg Two Color Ambiguous Urn Ball Problem in 1921

Michael Brady*

J.M. Keynes, in Chapter 30 of his 1921 book, A Treatise on Probability (TP), presented a mathematical analysis of the two color Ellsberg urn ball problem. In light of this, the two-color problem should be renamed the "Keynes Ambiguity Problem".

1. Introduction

Although Ellsberg presented two different urn ball problems in his 1961 article on ambiguity (Ellsberg, 1961, pp. 649-659), a three color problem and a two color problem, it is the two color problem which has captured by far the greater amount of attention in the academic literature.

The two color problem consists of two urns. Urn one has 50 red and 50 black balls, while the other urn, urn two, has 100 red and black balls in any of 101 possible distributions. Ellsberg, and all researchers after him, found that many experimental subjects prefer to bet on a red (black) ball from urn 1, but are indifferent between red and black when betting on only one of the two urns. This response pattern falsified the predictions based on Subjective Expected Utility (SEU) Theory, that decision makers should be indifferent in their choices between the two urns.

Brady (1987b) and Brady and Lee (1989a,b) pointed out that Keynes' urn ball example on page 75 of the TP was practically identical to Ellsberg's two color problem. Given this fact, it was suggested that what Keynes called the "weight of argument" or "evidence" was describing essentially the same phenomena that Ellsberg was calling by the term "ambiguity".

The purpose of this paper is to demonstrate that Keynes solved the problem of estimating a probability value from urn 2, using his principle of indifference, and then generalized this solution technique by developing a general formula for any number of draws from a type two urn in Chapter 30 of the TP, as well as covering the mathematical theory underlying this problem at a very advanced level.

2. Keynes' Preliminary Analysis in Chapters 4 and 6 of the TP of the Two Color Urn Problem

After discussing his two color Ellsberg-type urn ball model, where urn 1 contains black and white balls in equal proportion and urn 2 contains an unknown proportion of each color, Keynes states:

"It is evident that in either case the probability of drawing a white ball is ½, but the weight of the argument in favor of this conclusion is greater in the first case." (Keynes, 1921, p. 75).

In Ellsberg's terminology, the ambiguity would be less in the first case. Increasing ambiguity for Ellsberg would be a decrease in the weight of argument, or evidence, for Keynes, and vice-versa.

Keynes devoted an entire chapter (Chapter 4) of the TP to an attempted rehabilitation of the principle of non-sufficient reason. The sound aspects he renamed the principle of indifference. On page 50 of the TP, in footnotes 1 and 3, Keynes points out that Boole, in 1854 (!) and Stumpf, in 1892 (!), had already demonstrated that, for *both* Ellsberg urns, the probability of a red (black) is ½. We will discuss Keynes' generalized formula for calculating probabilities for multiple draws from an Ellsberg type urn 2 in section 3.

In urn 2, there are 100 red and black balls, in any possible proportion, ranging from 0 red, 100 black to 100 red, 0 black. Keynes discusses this example as a "typical" example (Keynes, p. 49-51, 56-57) and notes that Boole was the first to analyze this question on pages 369-375 of his 1854 book, *The Laws of Thought*. Boole arrives at a probability value of ½ for the black and red balls in urn 2.

The solution approach is to assume that each of the possible 101 different probability distributions is equally likely a priori. The probability of either red or black in urn 2 is

$$(1/101)[0/100+1/100+2/100+3/100+\cdots +97/100+98/100+99/100+100/100] = 5{,}050/10{,}100 = ½$$

or, in summation notation, à la Boole,

$$(1/101) \sum_{0}^{100} (N/100) = ½$$ (Boole, 1854, p. 372, lines 14-22)[1]

The variance can also easily be calculated. It is

$$(1/101) \sum_{0}^{100} [(N/100)-(½)]^2 = (1/101) \sum_{0}^{100} [N/100(1-(N/100)^2 + (100-N)/100(0-(N/100)^2] = ¼.$$

The literature in this area is not satisfactory. Neither Segal, 1987; Sinn, 1980[2]; Sherman, 1974; Smith, 1969; Ellsberg, 1961; Raiffa, 1961; de Finetti, 1977; or Camerer and Weber, 1992[3], ever mention the analysis of Boole, Stumpf, or Keynes. The result is a number of errors in the literature. For instance, Smith claims,

"Indifference between red and black in urn I (our Urn 2) implies that red and black are equally probable (but not necessarily = ½)" (Smith, 1969, p. 327).

This is simply mathematically false, as demonstrated by Boole, Stumpf, and Keynes. Smith only goes half-way toward a solution, giving

"$p_1+q_1 = N/100+(100-N)/100 = 1$." (Smith, 1969, p. 328).

Sherman gives the wrong answers on page 168 of his 1974 article in his footnote 2. See footnote 4. The correct answers are:

$$(X/101) \sum_{0}^{100} (N/100) = X/2$$ is the mean payoff from Ellsberg's urn 2, where X=$100 if red and $0 if black, while the variance equals

$$(1/101) \sum_{0}^{100} [(NX/100)-(X/2)]^2+(1/101) \sum_{0}^{100} [N/100(X-(NX/100))^2+ (100-N)/100(0-(NX/100))^2] = X^2/4.$$

Sinn[2] (1980) is unaware of Keynes' handling of the principle of insufficient reason and reformulation of it as the principle of indifference. His "rehabilitation" of the principle of non-sufficient reason, as it pertains to the Ellsberg paradox, merely duplicates the work of Boole, Stumpf, and Keynes (from Chapter 30 of the TP). He adds nothing new to Keynes' discussion.

Similarly, Segal (1987) also duplicates the arithmetic of Boole, Stumpf, and Keynes. Again, he is unaware that his result, that the probabilities for both Ellsberg urns is ½, was worked out 133 years before he published his article! Note also that, due to the very small amounts of money involved in the Ellsberg problems, i.e., $100.00, the utility functions are *linear*, which means that they give the same answer as expected *value* analysis. Segal's resort to utility functions, which are linear, adds nothing new to an expected value analysis.

3. Keynes' Analysis in Chapter 30 of the TP.

Making use of the Beta and Gamma functions, Keynes demonstrates that underlying Ellsberg's second urn ball model is Laplace's rule of succession. Keynes states:

> "If X stands for the a priori probability of an event in given conditions, then the probability that the event will occur m times and fail n times ...is $X^m(1-X)^n$. If, however, X is unknown, all values of it between 0 and 1 are a priori equally probable." (Keynes, 1921, pp. 375-376).

The analysis then leads, where A is a constant, to A being determined by

$$\int_0^1 AX^m(1-X)^n dX = 1$$

so that $A = \Gamma(m+n+2)/\Gamma(m+1)\Gamma(n+1)$.

"...the probability that the event will occur at the (m+n+1) th trial, when we know it has occurred m times in (m+n) trials, is

$$A\int_0^1 X^{m+1}(1-X)^n dX" \qquad \text{(Keynes, p. 376).}$$

Substituting in the value of A given above gives us $(m+1)/(m+n+2)$.

The basis of Keynes' analysis is the gamma function, whose integral is

$$\Gamma(\propto) = \int_0^\infty y^{\propto-1} e^{-y} dy.$$

For $\propto = 1, \Gamma(1) = \int_0^\infty e^{-y} dy = 1 = -e^{-y} \mid^\infty_0$.

For $\propto > 1, \Gamma(\propto) = (\propto-1) \int_0^\infty y^{\propto-2} e^{-y} dy = (\propto-1)\Gamma(\propto-1)$. A further result is $\Gamma(\propto+1) = \propto\Gamma(\propto)$.

Now $\int_0^1 x^{m-1}(1-X)^{n-1} dX = B(m,n)$ = the Beta Function. Then $B(m,n) = B(n,m) = \Gamma(m)\Gamma(n)/\Gamma(m+n)$. (See formulas 479-481, p. A-192, Handbook of Chemistry and Physics).

Thus, $\int_0^1 X^m(1-X)^n = \Gamma(m+1)\Gamma(n+1)/\Gamma(m+n+2)$. If $\int A X^m(1-X)^n dX = 1$, then $A = \Gamma(m+n+2)/\Gamma(m+1)\Gamma(n+1)$.

If "the event has invariably occurred", n = o and we obtain the result $(m+1)/(m+n+2) = (m+1)/(m+2)$. A priori, before any draws, m = o, and we obtain ½.

Keynes' summarizes:

"We begin with the assumption that the a priori probability of an event, about which we have no information and no experience, is unknown, and that all values between 0 and 1 are equally probable. We end with the conclusion that the a priori probability of such an event is ½. It has been pointed out...that this contradiction was latent, as soon as the Principle of Indifference was superimposed on the principle of unknown probabilities." (Keynes, p. 377).

This result is purely analytic. In Section 11, Chapter 30, Keynes points out that Laplace's Rule of Succession "was first suggested by the problem of the urn which contains black and white balls in unknown proportions. It is supposed that all compositions of the urn are equally probable and the proof proceeds precisely as in the case of the more general rule of succession" (Keynes, p. 378). The general case is given above.

Of course, this urn is identical to the ambiguous urn 2 of Ellsberg (1961). Keynes continues:

"On the hypothesis that all compositions of the urn are equally probable...and on the further hypothesis that the number of balls is infinite, this solution is correct" (Keynes, p. 378).

However, this solution is not correct if the number of balls is finite and we are dealing with more than one draw. Then, the correct solution is:

"...more generally, if p black balls and q white balls have been drawn and replaced, the chance that the next ball will be black is

$$(1/n)\sum_{r=0}^{r=n} r^{p+1}(n-r)^q \Big| \sum_{r=0}^{r=n} r^p (n-r)^q ."$$

In conclusion, Keynes, following Boole and Peirce, made an exhaustive and complete mathematical study of the two color "Ellsberg" urn ball problem. Keynes solved this problem mathematically in 1921. Keynes didn't just "wonder" about ambiguity. Nor did he just write about his "intuitions" about ambiguity. The same could also be said about Boole and Peirce. In light of the historical evidence presented in this paper, I suggest that one should refer to the "two-color Keynes problem" or "two-color Boole-Keynes-Peirce problem", while reserving the "Ellsberg problem" for the three color urn ball problem only.

---

* 9426 Flower Street, Bellflower, California 90706-5706, U.S.A.

Notes

1. The reader should note that C.S. Peirce's 1878 bean bag model is practically the same as Keynes' and Ellsberg's. Suppose you have a bean bag composed of 1000 black and white beans in unknown proportions. *Before* you bet on the color of a bean hidden under a thimble drawn from the bag, you are allowed four different options. Option 1 allows you to sample the bag 1000 times before betting. Option 2 allows you to sample 2000 times. Option 3 allows you to sample 10 times. Option 4 allows you to sample 0 times. Suppose that four individuals chose each of the four options and obtained these results. Option 1 yielded 505 black beans and 495 white. Option 2 yielded 1,010 black beans and 990 white. Option 3 yielded 6 black beans and 4 white beans. In option 4, in complete ignorance, the individual applies the principle of indifference to obtain:

1/10001[0/1000+1/1000+2/1000+3/1000+--997/1000+998/1000+999/1000+1000/1000] = 500,500/1,001,000 = ½. The variance is also ¼.

Peirce argues that in options 1 and 2 we can have great confidence in the probabilities. In option 3 we can have very little confidence. In option 4, we can have no confidence and what we should say is "that the chance is entirely indefinite". Sinn, Raiffa, and de Finetti argue the *opposite* case, i.e., that there *is* sense in saying that the chance of a totally unknown event is even, provided, of course, that you *assume* that all events are equally likely.

Peirce, of course, like Boole, Keynes, and the 1950's Carnap, would have none of this. According to Peirce, options 1 and 2 inspire great confidence while options 3 and 4 are greatly uncertain and the probability indefinite. Boole, like Peirce, would describe any probability assessment as being "arbitrary".

Similarly, von Mises understood the "idea" of weight of evidence or ambiguity. Unfortunately, he never incorporated a distinct analysis within his limiting-relative frequency approach to probability:

> "Suppose we draw ten times from an urn which contains black and white balls...it is clear that the numerical ratio 7/10 as such cannot be the only decisive instance for justifying the validity of our hypothetical assumption (that the probability p of drawing a black ball lies between 0.6 and 0.8, given an observed frequency of 0.7). If in 1000 experiments we obtain 700 black balls, the ratio is still 0.7, yet the hypothesis that the original probability p lies between 0.6 and 0.8 now has much better 'backing'" (Mises, 1957, pp. 155-156).

or

> "...if an event has occurred twice in three trials, we cannot conclude anything from this fact; if it occurs 2000 times in 3000 observations, we can draw fairly precise conclusions...," (Mises, 1957, pp. 156).

Given Mises attack on the use of small sample statistical methods (Mises, 1957, pp. 158-159) which follows directly from his discussion above, he concludes that without some a priori information, gained independently from the probability analysis or statistics, "we cannot draw any conclusions" (Mises, p. 159) and "we have no other recourse but to extend our sequence of observations to many hundreds or thousands of cases" (Mises, p. 159).

Unfortunately, Von Mises never provides the reader any hint of what exact number of observations one must have in order to obtain "much better 'backing'". He merely states that n, the number of observations, must be "sufficiently large". A good question to have asked Mises is "what number of observations *is* sufficiently large?" Obviously, if you *assume* unlimited, indefinite or infinite amounts of observations, then there is no problem. It appears that Mises overlooked both Peirce and Keynes on this topic of "better backing".

2. Hey implicitly agrees with Sinn that SEU Theory uses the principle of indifference combined with Laplace's rule of succession. Consider his discussion of the usual tree diagram analysis for tossing a coin. Then,

> "a head will send you on the upper path, a tail on the lower path. Suppose further that you do not know with certainty P(H) = p, the probability of the coin landing heads. After the first chance node (i.e., branch), you will have an observation from which you can learn. The information will be incorporated using Bayes Theorem. More precisely, suppose your prior on p is beta with parameters a and b; after the observation, the posterior will be beta with parameters a +1 and b or a and b+ 1 depending on whether a head or tail occurred. Using the fact that the mean of a beta with parameters a and b is a/(a+b), it follows that we must have:
>
> $p_1 = p_2 = a/(a+b)$.
>
> $q_1 = q_2 = q_5 = q_6 = (a+1)/(a+b+1)$.
>
> $q_3 = q_4 = q_7 = q_8 = a/(a+b+1)$.

In other words, all the relevant p's and q's can be calculated before *any* choice is made...

> Of course, crucial to this is the notion that information is *endogenous*, that is, generated within the choice problem itself. If, in contrast, information was exogenous...then the q's could not be calculated ab initio..." (Hey, 1981, pp. 134-135).

Of course, Keynes already provided a much more detailed analysis of this exact question in Chapter 30 of the TP, sixty years (!) before Hey (Sinn) implicitly (explicitly) recognized that the independence axiom or sure thing principle of SEU Theory is nothing other than Laplace's rule of succession combined with the principle of indifference.

Examples of other authors who rely on Laplace's rule of succession, combined with the principle of indifference, are Dobbs (1991, p. 421) and Viscusi, Magat, et al. (1987, p. 62).

3. Camerer and Weber ignore Keynes' mathematical solution. They state that

> "Keynes (1921) drew the distinction between the *implication* of evidence...and the *weight* of evidence...Keynes wondered whether a single probability could express both dimensions of evidence" (Camerer and Weber, p. 327, see also pp. 331 and 347-348).

They then discuss Ellsberg's two and three color urn ball problems; Knight is also mentioned as having discussed the two color problem. Keynes' urn ball model analysis is not even mentioned. They forget to inform the reader that Knight's analysis led him to recommend indifference in the two color urn ball problem, i.e., the SEU answer. In their conclusion, they claim that, "But as Ellsberg (1961) showed (following the *intuitions* (author's underscore) of Keynes and Knight)..." (Camerer and Weber, p. 360). Camerer and Weber are ignorant of basic literature going back to the early 1850's as I note in my concluding comment.

4. Sherman gives formulas, for the calculation of the expected value and the variance of Ellsberg's ambiguous urn, which are incorrect.

His incorrect formula for the mean is

$(X/100) \sum_{N=1}^{100} (N/100) = X/2$, while his incorrect formula for the variance equals

$$(1/100) \sum_{N=1}^{100} NX/100 - X/2)^2 + (1/100) \sum_{N=1}^{100} [N/100(X-NX/100)^2 + 100\text{-}N/100\ (0\text{-}NX/100)^2] = X^2/4.$$

The correct answers are given above.

References

Boole, (G. (1854), The Laws of Thought. Dover Publications: New York.

Brady, M.E. (1987b), "From Keynes to Ellsberg and Back. The concept of weight of evidence rediscovered." Unpublished manuscript.

Brady, M.E. and Lee, H.B. (1989a), "Dynamics of choice behavior: the logical relation between linear objective probability and nonlinear subjective probability". *Psychological Reports*, 64, pp. 91-97.

Brady, M.E. and Lee, H.B. (1989b), "Is there an Ellsberg-Fellner paradox? A note on its resolution. *Psychological Reports*, 64, pp. 1087-1090.

Camerer, C. and Weber, M. (1992), "Recent developments in modeling preferences: uncertainty and ambiguity". *Journal of Risk and Uncertainty*, 5, pp. 325-370.

Carnap, R. (1950), *Logical Foundations of Probability*. Chicago: University of Chicago Press.

deFinetti, B. (1977), "Probabilities of probabilities". In A. Aykal and C. Brumet (eds.). *New Dimensions in the application of Bayesian methods*. Amsterdam: N. Holland, pp. 1-10.

Dobbs, I.M. (1991), "A Bayesian approach to decision-making under ambiguity", *Economica*, 58, pp. 417-40.

Ellsberg, D. (1961), "Risk, ambiguity and the Savage axioms". *Quarterly Journal of Economics*, 75, pp. 643-669.

Handbook of Chemistry and Physics, 1969. Chemical Rubber Company.

Hey, J.D. (1983), "Whither uncertainty". *Economic Journal*, 91, pp. 130-139.

Madan, D.B. and Owings, J.C. (1988), "Decision theory with complex uncertainties". *Synthese*, 75, pp. 25-44.

Mises, R. Von (1957), *Probability, Statistics, and Truth*. Dover: New York.

Peirce, C.S. (1878), "The probability of induction". In Hartshorne and Weiss, eds., *Collected Works of C.S. Peirce*, Vol. II, Cambridge University Press, 1930-35, pp. 419-421.

Raiffa, H. (1961), "Risk, ambiguity, and the Savage Axioms: Comment". *Quarterly Journal of Economics*, 75, pp. 690-694.

Segal, U. (1987), "The Ellsberg paradox and risk aversion: An anticipated utility approach". *International Economic Review*, 28, pp. 175-202.

Sinn, H.W. (1980), "A rehabilitation of the principle of insufficient reason. *Quarterly Journal of Economics*, 94, pp. 493-506.

Sherman, R. (1974), "The psychological differences between ambiguity and risk". *Quarterly Journal of Economics*, 88, pp. 166-169.

Smith, V.L. (1969), "Measuring nonmonetary utilities in uncertain choices: The Ellsberg urn". *Quarterly Journal of Economics*, 81,pp. 324-29.

Viscusi, W.K., Magat, W., et al., (1987), *Learning about Risk*, (Cambridge, Mass.: Harvard University Press).