

Subjective Probability with Black Swans

Graciela Chichilnisky*

Columbia University, New York 10027 USA

November 23, 2009

Abstract

We extend the foundation of subjective probability to integrate rare events that are potentially catastrophic, called *black swans*, such as natural hazards, market crashes, catastrophic climate change and major episodes of species extinction. Such events are generally treated as ‘outliers’ and often disregarded. We propose a new axiomatization of subjective probability requiring equal treatment for rare and frequent events - the Swan Axiom - and characterize the subjective probabilities that the axioms imply: these are neither finitely additive nor countably additive but a combination of both. They exclude countably additive probabilities as in De Groot (1970) and Arrow (1971), and are a strict subset of Savage (1954) probabilities that are finitely additive measures. Our subjective probabilities are standard distributions when the sample has no black swans. The finitely additive part assigns however more weight to rare events than do standard distributions and in that sense explains the persistent observation of ‘power laws’ and ‘heavy tails’ that eludes classic theory. The axioms extend earlier work in Chichilnisky (2000, 2002) to encompass the foundation of subjective probability, and axiomatic treatments of subjective probability by Savage (1954), Villegas (1964), De Groot (1963) and the foundations of choice under uncertainty in Arrow (1971).

*Director, Columbia Consortium for Risk Management, Columbia University, and Professor Departments of Economics and Mathematical Statistics, Columbia University, New York 10027, 335 Riverside Drive, NY NY 10025, tel. 212 678 1148, fax 212 678 0405, email: gc9@columbia.edu; website: www.chichilnisky.com,

1 Introduction

Black swans are rare events with major consequences, such as market crashes, natural hazards, global warming and major episodes of extinction. This article is about the foundations of probability when catastrophic events are at stake. It provides a new axiomatic foundation for subjective probability requiring sensitivity to rare and frequent events. The study culminates in Theorem 2, that proves existence and representation of a subjective probability satisfying three axioms. The last of these axioms requires that the subjective probability be sensitive to rare events, a property that is desirable but not respected by standard probabilities. The article discusses the connection between those axioms and the Axiom of Choice at the foundation of Mathematics. It defines a new type of subjective probabilities that coincide with standard distributions when the sample is populated by frequent events. Generally, however, they are a mixture of countable and finitely additive measures, assigning more weight to black swans than do normal distributions, and predicting more realistically the incidence of ‘outliers,’ ‘power laws’ and ‘heavy tails.’

The article refines and extends the formulation of subjective probability as a decision characteristic of an agent choosing actions ("acts") in an uncertain world. It provides an argument, and formalization, that subjective probabilities must be additive functionals on $L_\infty(\mathcal{U})$, where \mathcal{U} is a σ -field of "events" represented by their indicator (bounded, real valued) functions, that are neither countably additive nor finitely additive. The contribution is to provide an axiomatization showing that subjective probabilities must lie in the full space L_∞^* rather than L_1 as the usual formalization (Arrow, 1971) forcing countable additivity, implies. The new axioms incorporate both Savage’s (1954) axiomatization of finitely additive measures, as well as Villegas’ (1964) and Arrow’s (1971) that are based on countably additive measures, and extend both to deal more realistically with catastrophic events.

Savage (1954) axiomatized subjective probabilities as finitely additive measures representing the decision makers’ beliefs, an approach that can ignore frequent events as shown in the Appendix. To overcome this, Villegas (1964) and Arrow (1971) introduced an additional continuity axiom (called ‘monotone continuity’) that yields countably additivity of the measures. However monotone continuity has unusual implications when the subject is confronted with rare events, for example it predicts that in exchange for a couple of cents, one should be willing to accept a small risk of death, a possibility that Arrow called ‘outrageous’ (1971, P. 48 and 49). This article defines a realistic solution: for some payoffs and in certain situations, one may be willing to accept a small risk of death - but not in others. This means that monotone continuity holds in some cases but not in others, a possibility that leads to the axiomatization proposed in this article and is consistent with the experimental observations reported in Chanel and Chichilnisky (2009a, b). The results are as follows. First we show that countably additive measures are insensitive to *black swans*: they assign negligible weight to rare events, no matter how important these may be, treating catastrophes as outliers, Chichilnisky (2009 a,b). Finitely additive measures,

on the other hand, may assign no weight to frequent events, which is equally troubling. Our new axiomatization balances the two approaches and extends both, requiring sensitivity to rare events as well as frequent events. We provide an existence theorem for probabilities that satisfy our axioms, and a characterization of all that do.

The results start from an axiomatic approach to choice under uncertainty and sustainable development introduced by the author Chichilnisky (1996a, 2000, 2009), and illuminate the classic issue of continuity that has always been at the core of ‘subjective probability’ axioms (Villegas, 1964, Arrow 1971). To define continuity, we use a topology that tallies with the experimental evidence of how people react to rare events that cause fear (Le Doux (1996), Chichilnisky (2009)), previously used by Debreu (1953) to formalize a market’s Invisible Hand, and in Chichilnisky (2000, 2002) to axiomatize choice under uncertainty with rare events. The new results provided here show that a standard axiom of decision theory, Monotone Continuity, is equivalent to De Groot’s Axiom SP_4 that lies at the foundation of classic likelihood theory (Proposition 1) and that both of these axioms underestimate rare events no matter how catastrophic they may be. We introduce here a new Swan Axiom (Section 3) that negates them both, and show it is a combination of two axioms defined in Chichilnisky (2000, 2002) and prove that any subjective probability satisfying the Swan Axiom is neither countably additive nor finitely additive: it is both (Theorem 1). Theorem 2 provides a complete characterization of all subjective probabilities that satisfy linearity and the Swan Axiom, thus extending earlier results of Chichilnisky (2000, 2002, 2009).

There are other approaches to subjective probability such as Choquet Expected Utility Model (CEU, Schmeidler, 1989) and Prospect Theory (Kahneman and Tversky, 1975 and 1992). They use a non-linear treatment of probabilities of likelihoods (see e.g. Dreze, 1987, or Bernstein, 1996), while we retain linear probabilities. Both have a tendency to give higher weight to small probabilities, and are theoretical answers to experimental paradoxes (Allais, 1953, Ellsberg, 1961) refuting the *Independence Axiom* of the Subjective Expected Utility (SEU) model. Our work focuses instead directly on the foundations of probability by taking the logical negation of the *Monotone Continuity Axiom*. It is striking that weakening or rejecting this axiom - respectively in decision theory and in probability theory - ends up in probability models that are more in tune with observed attitudes when facing catastrophic events. Presumably each approach has advantages and shortcomings. It seems that the approach offered here may be superior on four counts: (i) It retains linearity of probabilities, (ii) It identifies Monotone Continuity as the reason for underestimating catastrophic events, an axiom that depends on a technical definition of continuity and has no other compelling feature, (iii) it seems easier to explain and to grasp, and therefore (iv) it may be easier to use in applications.

2 The Mathematics of Uncertainty

Uncertainty is described by a set of distinctive and exhaustive possible **events** represented by a family of sets $\{U_\alpha\}$, $\alpha \in N$, whose union describes a universe $\mathcal{U} = \cup_\alpha U_\alpha$. An event $U \in \mathcal{U}$ is identified with its **characteristic function** $\phi_U : \mathcal{U} \rightarrow R$ where $\phi_U(x) = 1$ when $x \in U$ and $\phi_U(x) = 0$ when $x \notin U$, and therefore can be identified as a function $f(x) = \phi_U(x)$ for some U . The subjective probability of an event U is a real number $W(U)$ that measures how likely it is to occur according to the subject. Generally we assume that the probability of the universe is 1 and that of the empty set is zero $W(\emptyset) = 0$. In this article we make no difference between subjective probabilities and likelihoods, using both terms interchangeably. Classic axioms for subjective probability (respectively likelihoods) are in Savage (1954) and De Groot (1970). The likelihood of two disjoint events is the sum of their likelihoods: $W(U_1 \cup U_2) = W(U_1) + W(U_2)$ when $U_1 \cap U_2 = \emptyset$. These properties correspond to the definition of a subjective probability or likelihood as a *finite additive measure* on a family (σ -algebra) of measurable sets of \mathcal{U} , which is Savage's (1954) definition of subjective probability. A *purely finitely additive probability* is one that is additive but not countably additive. Savage's subjective probabilities can be purely finitely additive or countably additive. In that sense they include all the probabilities in this article. However as seen below, this article excludes subjective probabilities that are either purely finitely additive, or countably additive, and therefore our characterization of a subjective probability is strictly finer than that Savage's (1954), and different from the view of a measure as a countably additive set function (De Groot, 1979). The following two axioms were introduced in Villegas (1964), see Arrow (1971) and DeGroot (1970) for the purpose of obtaining countable additivity:

Monotone Continuity Axiom (MC), (Arrow, 1971): for every two events f and g with $W(f) > W(g)$, and every **vanishing** sequence of events $\{E_\alpha\}_{\alpha=1,2,\dots}$ (defined as $\forall \alpha, E_{\alpha+1} \subset E_\alpha$ and $\cap_{\alpha=1}^\infty E_\alpha = \emptyset$) there exists N such that altering arbitrarily the events f and g on the set E_i , where $i > N$, does not alter the subjective probability ranking of the events, namely $W(f') > W(g')$, where f' and g' are the altered events.

This axiom is equivalent to requiring that the probability of the sets along a vanishing sequence goes to zero. The decreasing sequence could consist of infinite intervals of the form (n, ∞) for $n = 1, 2, \dots$. Monotone continuity therefore implies that the likelihood of this sequence of events goes to zero, even though all its sets are unbounded. A similar example can be constructed with a decreasing sequence of bounded sets, $(-1/n, 1/n)$ for $n = 1, 2, \dots$, which is also a vanishing sequence as it is a decreasing and their intersection is empty.

De Groot's Axiom SP_4 (De Groot, 1970), Chapter 6, p. 71): if $A_1 \supset A_2 \supset \dots$ is a decreasing sequence of events and B is some fixed event that is less likely than A_i for all i , then the probability of the intersection $\cap_i^\infty A_i$ is larger than that of B .

The following proposition establishes that these two axioms are one and the same; both imply countable additivity:

Proposition 1: *A relative likelihood satisfies the **Monotone Continuity Axiom** if and only if it satisfies Axiom SP_4 . Each of the two axioms implies countable additivity.*

Proof:

Assume that De Groot's axiom SP_4 is satisfied. When the intersection of a decreasing sequence of events is empty $\cap_i A_i = \emptyset$ and the set B is less likely to occur than every set A_i , then the subset B must be as likely as the empty set, namely its probability must be zero. In other words, if B is more likely than the empty set, then regardless of how small is the set B , it is impossible for every set A_i to be as likely as B . Equivalently, the probability of the sets that are far away in the vanishing sequence must go to zero. Therefore SP_4 implies Monotone Continuity. Reciprocally, assume MC is satisfied. Consider a decreasing sequence of events A_i and define a new sequence by subtracting from each set the intersection of the family, namely $A_1 - \cap_i^\infty A_i, A_2 - \cap_i^\infty A_i, \dots$. Let B be a set that is more likely than the empty set but less likely than every A_i . Observe that the intersection of the new sequence is empty, $\cap_i (A_i - \cap_i^\infty A_i) = \emptyset$ and since $A_i \supset A_{i+1}$ the new sequence is, by definition, a vanishing sequence. Therefore by MC $\lim_i W(A_i - \cap_i^\infty A_i) = 0$. Since $W(B) > 0$, B must be more likely than $A_i - \cap_i^\infty A_i$ for some i onwards. Furthermore, $A_i = (A_i - \cap_i^\infty A_i) \cup (\cap_i^\infty A_i)$ and $(A_i - \cap_i^\infty A_i) \cap (\cap_i^\infty A_i) = \emptyset$, so that $W(A_i) > W(B)$ is equivalent to $W(A_i - \cap_i^\infty A_i) + W((\cap_i^\infty A_i)) > W(B)$. Observe that $W(\cap_i^\infty A_i) < W(B)$ would contradict the inequality $W(A_i) = W(A_i - \cap_i^\infty A_i) + W((\cap_i^\infty A_i)) > W(B)$, since as we saw above, by MC, $\lim_i W(A_i - \cap_i^\infty A_i) = 0$, and $W(A_i - \cap_i^\infty A_i) + W((\cap_i^\infty A_i)) > W(B)$. It follows that $W(\cap_i^\infty A_i) > W(B)$, which establishes De Groot's Axiom SP_4 . Therefore Monotone Continuity is equivalent to De Groot's Axiom SP_4 . A proof that each of the axioms implies countable additivity is in Villegas (1964), Arrow (1971) and De Groot (1970).■

The next section shows that the two axioms, monotone continuity and SP_4 are biased against rare events no matter how important these may be.

3 The Value of Life

The best way to explain the role of *monotone continuity* is by means of an example provided by Arrow (1971), p. 48 and 49. He explains that if a is an action that involves receiving one cent, b is another that involves receiving zero cents, and c is a third action involving receiving one cent and facing a small probability of death, then *Monotone Continuity* requires that the third action involving death and one cent should be preferred to the action with zero cents when the probability of death is small enough. Even Arrow says of his requirement 'this may sound outrageous at first blush...' Arrow (1971) p. 48 and 49. Outrageous or not, Monotone Continuity (MC) leads to neglect rare events with major consequences, like death. Death is a black swan.

To overcome the bias we introduce an axiom that is the logical negation of MC: this means that sometimes MC holds and others it does not. We call this

the **Swan Axiom**, and it is stated formally below. To illustrate this, consider an experiment where subjects are offered a certain amount of money to choose a pill at random from a pile, which is known to contain one pill that causes death (Chanel and Chichilnisky 2009b). It was shown experimentally (Chanel and Chichilnisky (2009b)) that in some cases people accept a sum of money and choose a pill provided the pile is large enough - namely when the probability of death is small enough - thus satisfying the monotone continuity axiom and determining the statistical value of their lives. But there are also cases where the subjects will not accept to choose any pill, no matter how large is the pile. Some people refuse the payment if it involves a small probability of death, no matter how small the probability may be (Chanel and Chichilnisky, 2009a,b). This conflicts with the Monotone Continuity axiom, as explicitly presented by Arrow (1971).

Our Axiom provides a reasonable resolution to this dilemma that is realistic and consistent with the experimental evidence. It implies that there exist catastrophic outcomes such as the risk of death, so terrible that one is unwilling to face a small probability of death to obtain one cent versus half a cent, no matter how small the probability may be. According to our Axiom, no probability of death may be acceptable when one cent and half a cent are involved. Our Axiom also implies that in other cases there may be a small enough probability that the lottery involving death may be acceptable, or the payoff is large enough to justify the small risk. This is a possibility discussed by Arrow (1971). In other words: sometimes one is willing to take a risk with a small enough probability of a catastrophe, in other cases one is not. This is the content of our Axiom, which is formally stated below:

The Swan Axiom: There exist events f and g with $W(f) > W(g)$, and for every vanishing sequence of events $\{E_\alpha\}_{\alpha=1,2,\dots}$ an $N > 0$ such that altering arbitrarily the events f and g on the set E_i , where $i > N$, does not alter the probability ranking of the events, namely $W(f') > W(g')$, where f' and g' are the altered events. For other events f and g with $W(f) > W(g)$, there exist vanishing sequence of events $\{E_\alpha\}_{\alpha=1,2,\dots}$ where for every N , altering arbitrarily the events f and g on the set E_i , where $i > N$, does alter the probability ranking of the events, namely $W(f') < W(g')$, where f' and g' are the altered events.

Definition: A probability W is said to be **biased against rare events** or **insensitive to rare events** when it neglects events that are small according to Villegas and Arrow; as stated in Arrow (1971) p. 48: "An event that is far out on a *vanishing sequence* is 'small' by any reasonable standards" Arrow (1971) p. 48. Formally, a probability is insensitive to rare events when given two events f and g and any vanishing sequence of events (E_j) , $\exists N = N(f, g) > 0$, such that $W(f) > W(g) \Leftrightarrow W(f') > W(g') \forall f', g'$ satisfying $f' = f$ and $g' = g$ a.e. on $E_j^c \subset R$ when $j > N$, where E^c denotes the complement of the set E .

Proposition 2. A subjective probability satisfies Monotone Continuity if and only if it is biased against rare events.

Proof: This is immediate from the definitions of both. ■

Corollary 1: Countably additive probabilities are biased against rare events.

Proof: It follows from Propositions 1 and 2. ■

Proposition 3. Purely finitely additive probabilities can be biased against frequent events.

Proof: See example in the Appendix. ■

Proposition 4: A subjective probability that satisfies the Swan Axiom is neither biased against rare events, nor biased against frequent events.

Proof: This is immediate from the definition. ■

4 An Axiomatic Approach to Unbiased Statistics

This section proposes an axiomatic foundation for subjective probability that is unbiased against rare and frequent events. The axioms are as follows:

Axiom 1: Subjective probabilities are continuous and additive

Axiom 2: Subjective probabilities are unbiased against rare events

Axiom 3: Subjective probabilities are unbiased against frequent events.

Axioms 2 and 3 together are equivalent to the Swan Axiom defined in the previous section, which is required to avoid bias against rare and frequent events as shown in Section 3. Additivity is a natural condition and **continuity** captures the notion that ‘nearby’ events are thought as being similarly likely to occur; this property is important to ensure that ‘sufficient statistics’ exist. ‘Nearby’ has been defined by Villegas (1964) and Arrow (1971) as follows: two events are **close** or **nearby** when they differ on a **small set** as defined in Arrow (1971), see previous section. We saw in Proposition 2 that the notion of continuity defined by Villegas and Arrow - namely monotone continuity - conflicts with the Swan Axiom. Indeed Proposition 2 shows that countably additive measures are biased against rare events. On the other hand, Proposition 3 and the Example in the Appendix show that purely finitely additive measures can be biased against frequent events. A natural question is whether after one eliminates both biases there is anything left. The following proposition addresses this issue:

Theorem 1: A subjective probability that satisfies the Swan Axiom is neither finitely additive nor countably additive; it is a strict convex combination of both.

Proof: The result is immediate. The next Section develops the proofs and provides examples when the events are Borel sets in R or an interval (a, b) . ■

Theorem 1 establishes that neither Savage’s approach, nor Villegas’ and Arrow’s, satisfy the three axioms stated above. These three axioms require more than the additive subjective probabilities of Savage, since purely finitely additive probabilities are finitely additive and yet they must be excluded here. At the same time the axioms require less than the countably subjective additivity of Villegas and Arrow, since countably additive probabilities are biased against rare events. Theorem 1 above shows that a strict combination of both does the

job.

Theorem 1 does not however prove the existence of likelihoods that satisfy all three axioms. What is missing is an appropriate definition of continuity that does not conflict with the Swan Axiom. The following Section shows that this can be achieved by identifying an event with its characteristic function, so that events are contained in the space of bounded real valued functions on the universe space \mathcal{U} , $L_\infty(\mathcal{U})$, and endowing this space with the sup norm. In this case the likelihood $W : L_\infty(\mathcal{U}) \rightarrow R$ is taken to be continuous with respect to the sup norm, a topology used in Debreu (1953).

5 Axiomatic Statistics on R or (a, b)

From now on events are taken to be the Borel sets of the real line R or the interval (a, b) , a widely used case that make the results concrete and compare the results with the earlier axioms on choice under uncertainty of Chichilnisky (2000, 2002, 2009). We use a concept of ‘continuity’ based on a topology that was used earlier in Debreu (1953) and in Chichilnisky (2000, 2002, 2009 a,b): observable events are in the space of measurable and essentially bounded functions $L = L_\infty(R)$ with the sup norm $\|f\| = \text{ess sup}_{x \in R} |f(x)|$. This is a sharper and more stringent

definition of closeness than the one used by Villegas and Arrow, since an event can be small under the Villegas-Arrow definition but not under ours, see the Appendix. The difference as shown below determines sensitivity to rare events.

A subjective probability satisfying the classic axioms in De Groot (1970) is called a **standard probability**, and is therefore countably additive. A classic result is that for any event $f \in L_\infty$ a standard probability has the form $W(f) = \int_R f(x) \cdot \phi(x) d\mu$, where $\phi \in L_1(R)$ is an integrable function in R .

The next step is to introduce the new axioms, show existence and characterize all the distributions that satisfy the axioms. We need more definitions. A subjective probability $W : L_\infty \rightarrow R$ is called **biased against rare events**, or **insensitive to rare events** when it neglects events that are small according to a probability measure μ on R that is absolutely continuous with respect to the Lebesgue measure. Formally, a probability is insensitive to rare events when given two events f and $g \exists \varepsilon = \varepsilon(f, g) > 0$, such that $W(f) > W(g) \Leftrightarrow W(f') > W(g') \forall f', g'$ satisfying $f' = f$ and $g' = g$ a.e. on $A \subset R$ and $\mu(A^c) < \varepsilon$. Here A^c denotes the complement of the set A . $W : L \rightarrow R$ is said to be **insensitive to frequent events** when given any two events $f, g \exists \varepsilon(f, g) > 0$ that $W(f) > W(g) \Leftrightarrow W(f') > W(g') \forall f', g'$ satisfying $f' = f$ and $g' = g$ a.e. on $A \subset R$ and $\mu(A^c) > 1 - \varepsilon$. W is called **sensitive** to rare (or frequent) events when it is **not insensitive** to rare (or frequent) events.

The following three axioms are identical to the axioms in last section, specialized to the case at hand:

Axiom 1: $W : L_\infty \rightarrow R$ is linear and continuous

Axiom 2: $W : L_\infty \rightarrow R$ is sensitive to frequent events

Axiom 3: $W : L_\infty \rightarrow R$ is sensitive to rare events

The first and the second axiom agree with classic theory and standard likelihoods satisfy them. The third axiom is new.

Lemma 1: *A standard probability satisfies Axioms 1 and 2, but it is biased against rare events and therefore does not satisfy Axiom 3.*

Proof:

Consider $W(f) = \int_{\mathcal{R}} f(x)\phi(x)dx$, $\int_{\mathcal{R}} \phi(x)dx = K < \infty$. Then

$$W(f)+W(g) = \int_{\mathcal{R}} f(x)\phi(x)dx + \int_{\mathcal{R}} g(x)\phi(x)dx = \int_{\mathcal{R}} |f(x)+g(x)| \phi(x)dx = W(f+g),$$

and therefore W is linear. It is **continuous** with respect to the L_1 norm $\|f\|_1 = \int_{\mathcal{R}} |f(x)| \phi(x)d\mu$ (19) because $\|f\|_{\infty} < \varepsilon$ implies

$$W(f) = \int_{\mathcal{R}} f(x).\phi(x)dx \leq \int_{\mathcal{R}} |f(x)| .\phi(x)dx \leq \varepsilon \int_{\mathcal{R}} \phi(x)dx = \varepsilon K.$$

Since the sup norm is finer than the L_1 norm, continuity in L_1 implies continuity with respect to the sup norm (Dunford and Schwartz, 1958). Thus a standard subjective probability satisfies Axiom 1. It is obvious that for every two events f, g , with $W(f) > W(g)$, the inequality is reversed namely $W(g') > W(f')$ when f' and g' are appropriate variations of f and g that differ from f and g on sets of sufficiently large Lebesgue measure. Therefore Axiom 2 is satisfied. A standard subjective probability is however not sensitive to rare events, as shown in Chichilnisky (2000, 2002, 2006, 2008, 2009a,b).. ■

6 Existence and Representation

Theorem 2: *There exists a subjective probability $W : L_{\infty} \rightarrow \mathcal{R}$ satisfying Axioms 1, 2, and 3. A probability satisfies Axioms 1,2 and 3 if and only if there exist two continuous linear functions on L_{∞} , denoted ϕ_1 and ϕ_2 and a real number λ , $0 < \lambda < 1$, such that for any observable event $f \in L_{\infty}$*

$$W(f) = \lambda \int_{x \in \mathcal{R}} f(x)\phi_1(x)dx + (1 - \lambda)\phi_2(f) \quad [1]$$

where $\phi_1 \in L_1(\mathcal{R}, \mu)$ defines a countably additive measure on \mathcal{R} and ϕ_2 is a purely finitely additive measure.

Proof:

This result follows from the representation Theorem in Chichilnisky (2000, 2009).■

Example 1 ‘Heavy’ Tails.

The following illustrates the additional weight that the new axioms assign to rare events; in this example in the form of ‘heavy tails’. The finitely additive measure ϕ_2 appearing in the second term in [1] can be illustrated as follows. On the subspace of events with limiting values at infinity, $L'_{\infty} = \{f \in L_{\infty} :$

$\lim_{x \rightarrow \infty} f(x) < \infty$ }, define $\phi_2(f) = \lim_{x \rightarrow \infty} f(x)$ and extend this to a function on all of L_∞ using Hahn Banach's theorem. The difference between a standard probability and the likelihood defined in [1] is the second term ϕ_2 , which focuses all the weight at infinity. This can be interpreted as a 'heavy tail' namely a part of the distribution that is not part of the standard density function ϕ_1 and gives more weight to the sets that contain *terminal* events namely sets of the form (x, ∞) . ■

Corollary 1: Absent rare events, a subjective probability that satisfies Axioms 1, 2, and 3 is consistent with classic axioms and yields a countably additive measure.

Proof: Axiom 3 is an empty requirement when there are no rare events while, as shown above, Axioms 1 and 2 are consistent with standard relative likelihood. ■

7 The Axiom of Choice

There is a connection between the axioms presented here and the Axiom of Choice in the foundation of mathematics (Godel, 1940), which postulates that there exists a universal and consistent fashion to select an element from every set. The best way to describe the situation is by means of an example, see also Dunford and Schwartz (1958), Yosida (1952, 1974), Chichilnisky and Heal (1997) and Kadane and O'Hagan (1995).

Example 2: Illustration of a Purely Finitely Additive Measure

Cosnider a measure ρ that satisfies the following: for every interval $A \subset R$, $\rho(A) = 1$ if $A \supset \{x : x > a, \text{ for some } a \in R\}$, and otherwise $\rho(A) = 0$. Then ρ is not countably additive, because the family of countably many disjoint sets $\{V_i\}_{i=0,1,\dots}$ defined as $V_i = (i, i + 1] \cup (-i - 1, -i]$, satisfy $V_i \cap V_j = \emptyset$ when $i \neq j$, and $\bigcup_{i=0}^{\infty} V_i = \bigcup_{i=0}^{\infty} (i, i + 1] \cup (-i - 1, -i] = R$, so that $\rho(\bigcup_{i=0}^{\infty} V_i) = 1$, while $\sum_{i=0}^{\infty} \rho(V_i) = 0$, which contradicts countable additivity. Since the contradiction arises from assuming that ρ is countably additive, ρ must be purely finitely additive. Observe that ρ assigns zero measure to any bounded set, and a positive measure only to unbounded sets that contain a 'terminal set' of the form

$$\{x \in R : x > a \text{ for some } a \in R\}.$$

One can define a function on L_∞ that represents this purely finitely additive measure ρ if we restrict our attention to the closed subspace L'_∞ of L_∞ consisting of those functions $f(x)$ in L_∞ that have a limit when $x \rightarrow \infty$, by the formula $\rho(f) = \lim_{x \rightarrow \infty} f(x)$, as in Example 1 of the previous section. The function $\rho(\cdot)$ can be illustrated as a limit of a sequence of delta functions whose support increases without bound. The problem is now to extend the function ρ to another defined on the entire space L_∞ . This could be achieved in various ways but as we will see, each of them requires the Axiom of Choice.

One can use Hahn - Banach's theorem (Dunford Schwartz, 1958) to extend the function ρ from the closed subspace $L'_\infty \subset L_\infty$ to the entire space L_∞ preserving its norm. However, in its general form Hahn - Banach's theorem requires the Axiom of Choice (Dunford Schwartz, 1958). Alternatively, one can extend the notion of a *limit* to encompass all functions in L_∞ including those with no standard limit. This can be achieved by using the notion of convergence along a *free ultrafilter* arising from compactifying the real line R as in Chichilnisky and Heal (1997). However the existence of a *free ultrafilter* also requires the Axiom of Choice.

This illustrates why the attempts to construct *purely finitely additive measures* that are representable as functions on L_∞ , require the Axiom of Choice. Since our criteria require purely finitely additive measures, this provides a connection between the Axiom of Choice and our axioms for relative likelihood. It is somewhat surprising that the consideration of rare events that are neglected in standard statistical theory conjures up the Axiom of Choice, which is independent from the rest of mathematics (Godel, 1940).

8 Appendix

Example 3: A Probability that is Biased Against Frequent Events

Consider $W(f) = \liminf_{x \in R} (f(x))$. This is insensitive to frequent events of arbitrarily large Lebesgue measure (Dunford and Schwartz, 1958) and therefore does not satisfy Axiom 2. In addition it is not linear, failing Axiom 1.

Example 4: Two Approaches to 'Closeness'

Consider the family $\{E_i\}$ where $E_i = [i, \infty)$, $i = 1, 2, \dots$. This is a vanishing family because $\forall i E_i \supset E_{i+1}$ and $\bigcap_{i=1}^{\infty} E_i = \emptyset$. Consider now the events $f^i(t) = K$ when $t \in E_i$ and $f^i(t) = 0$ otherwise, and $g^i(t) = 2K$ when $t \in E_i$ and $g^i(t) = 0$ otherwise. Then for all i , $\sup_{E^i} |f^i(t) - g^i(t)| = K$. In the sup norm topology this implies that f^i and g^i are **not** 'close' to each other, as the difference $f^i - g^i$ does not converge to zero. No matter how far along we are along the vanishing sequence E^i the two events f^i, g^i differ by K . Yet since the events f^i, g^i differ from $f \equiv 0$ and $g \equiv 0$ respectively only in the set E_i , and $\{E_i\}$ is a vanishing sequence, for large enough i they are as 'close' as desired according to Villegas-Arrow's definition of 'nearby' events.

The dual space L_∞^* : countably additive and finitely additive measures

The space of continuous linear functions on L_∞ is the 'dual' of L_∞ , and is denoted L_∞^* . It has been characterized e.g. in Yosida (1952, 1974). L_∞^* consists of the sum of two subspaces (i) L_1 functions g that define countably additive measures ν on R by the rule $\nu(A) = \int_A g(x) dx$ where $\int_R |g(x)| dx < \infty$ so that ν is *absolutely continuous* with respect to the Lebesgue measure, and (ii) a subspace consisting of purely finitely additive measure. A countable measure can be identified with an L_1 function, called its 'density,' but purely finitely additive measures cannot be identified by such functions.

Example 5: A Finitely Additive Measure that is not Countably Additive

See Example 2 in Section 7.

9 Acknowledgements

This research was conducted at Columbia University's Program on Information and Resources and its Columbia Consortium for Risk Management (CCRM). We acknowledge support from Grant No 5222-72 of the US Air Force Office of Research directed by Professor Jun Zheng, Washington DC. The initial results (Chichilnisku (1996a)) were presented at Stanford University's 1993 Seminar on Reconsideration of Values, organized by Professor Kenneth Arrow, at a 1996 Workshop on Catastrophic Risks organized at the Fields Institute for Mathematical Sciences of the University of Toronto, Canada, at the NBER Conference *Mathematical Economics: The Legacy of Gerard Debreu* at UC Berkeley, October 21, 2005, the Department of Economics of the University of Kansas National Bureau of Economic Research General Equilibrium Conference, September 2006, at the Departments of Statistics of the University of Oslo, Norway, Fall 2007, at a seminar organized by the late Professor Chris Heyde at the Department of Statistics of Columbia University, Fall 2007, and at seminars organized by Drs. Charles Figuières and Mabel Tidball at LAMETA Université de Montpellier, France December 19 and 20, 2008, and by Professor Alan Kirman at GREQAM Université de Marseille, December 18 2008. We are grateful to the above institutions and individuals for supporting the research, and for helpful comments and suggestions. An anonymous referee provided insightful comments that improved this article.

10 References

- Anscombe F.J. and R.J. Aumann (1963) A Definition of Subjective Probability *Annals of Mathematical Statistics* Vol 43, p. 199-295.
- Arrow K (1971) *Essays in the Theory of Risk Bearing*. (North Holland, Amsterdam)
- Bernstein, P.L. (1996) *Against the Gods: A Remarkable Story of Risk* Wiley, New York.
- Chanel, O. and G. Chichilnisky (2009a) The Influence of Fear in Decisions: Experimental Evidence *Journal of Risk and Uncertainty*, December 2009, Vol 39, No. 3.
- Chanel, O. and G. Chichilnisky (2009b) *The Value of Life: Theory and Experiments* Working Paper, GREQE, Université de Marseille and Columbia University, New York.
- Chichilnisky G (1996a) An Axiomatic Approach to Sustainable Development. *Soc Choice Welf* 13:321-257.

- Chichilnisky G (1996b) Updating Von Neumann Morgenstern Axioms for Choice under Uncertainty. Proceedings of a Conference on Catastrophic Risks, The Fields Institute for Mathematical Sciences, Toronto Canada.
- Chichilnisky G and Heal G M (1997) Social Choice with Infinite Populations. *Soc Choice Welf* 14:303-319.
- Chichilnisky G (2000) An Axiomatic Approach to Choice under Uncertainty with Catastrophic Risks. *Resource Energ Econ* 22:221-231.
- Chichilnisky G (2002) Catastrophic Risk, *Encyclopedia of Environmetrics*, eds El-Shaarawi AH and Piegorsch WW (John Wiley & Sons, Chichester). Vol 1 pp 274-279
- Chichilnisky G and Wu H-M (2006) General Equilibrium with Endogenous Uncertainty and Default. *J Math Econ* 42: 499–524.
- Chichilnisky G (2009) The Limits of Econometrics: Non Parametric Estimation in Hilbert Spaces. *Econ Theor*, 2009
- Chichilnisky G (2009) Choice under Uncertainty: The Work and Legacy of Kenneth Arrow *Encyclopedia of Quantitative Finance*, ed Cont R. (John Wiley and Sons, Chichester), in press.
- Chichilnisky G (2009) The Topology of Fear. *J Math Econ*, December 2009.
- Debreu G (1953) Valuation Equilibrium and Pareto Optimum. *Proceedings of the National Academy of Sciences* 40:588-592.
- De Groot MH. (1970 and 2004) *Optimal Statistical Decisions*. (John Wiley and Sons, Hoboken, New Jersey)
- Dreze, J.H. (1987) *Essays on Economic Decisions Under Uncertainty*, Cambridge, Cambridge University Press.
- Dubins L and Savage L (1965) *How to Gamble if you Must* (McGraw Hill, New York)
- Dubins L (1975) Finitely Additive Conditional Probabilities, Conglomerability and Disintegration. *The Annals of Probability* 3:89-99.
- Dunford N and Schwartz JT (1958) *Linear Operators, Part I* (Interscience, New York)
- Godel K (1940) The Consistency of the Continuum Hypothesis. *Annals of Mathematical Studies* 3 (Princeton University Press, Princeton)
- Kadane JB and O'Hagan A (1995) Using finitely additive probability: uniform distributions of the natural numbers. *J Am Stat Assoc* 90:525-631.
- Kahneman, D. and A. Tversky (1979) Prospect Theory: An analysis of Decisions Under Risk, *Econometrica*, 47, 263-291.
- Le Doux J (1996) *The Emotional Brain* (Simon and Schuster, New York)
- Purves RW and Sudderth (1976) Some Finitely Additive Probability. *The Annals of Probability* 4:259-276.
- Savage, L. *The Foundations of Statistics*, John Wiley and Sons New York 1054, revised edition 1972.
- Schmeidler D. (1989) Subjective Probability and Expected Utility Without Additivity, *Econometrica*, 47, 263-291.
- Tversky A. and Kahneman D. (1992) Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5, 297-323.

Yosida K and Hewitt E (1952) Finitely Level Independent Measures. Trans Amer Math Soc 72:46-66.

Yosida K (1974) Functional Analysis, 4th Edition (Springer Verlag, New York and Heidelberg)

Villegas C (1964) On Quantitative Probability σ -Algebras. Ann Math Stat 35:1789 - 1800.