Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/econbase

The foundations of statistics with black swans

Graciela Chichilnisky [a,b,c,∗]

[a] Columbia Consortium for Risk Management, Columbia University, NY 10027, USA

[b] Department of Economics, Columbia University, NY 10027, USA

[c] Department of Mathematical Statistics, Columbia University, NY 10027, USA

ARTICLE INFO

Article history:
Received 21 February 2009
Received in revised form
11 September 2009
Accepted 18 September 2009
Available online xxxx

JEL classification:
C00
C02
C40
G44
D81
G01
G32
Q54

Keywords:
Foundations of statistics
Black swans
Rare events
Catastrophic risks
Likelihood
Subjective probability
Finitely additive measures

ABSTRACT

We extend the foundation of statistics to integrate rare events that are potentially catastrophic, called *black swans*.These include natural hazards, regime change in complex systems, market crashes, catastrophic climate change and major episodes of species extinction. Classic statistics and physics treat such events as 'outliers' and often disregard them. We propose a new axiomatization of subjective probability requiring equal treatment for rare and frequent events, and characterize the likelihoods or subjective probabilities that the axioms imply. These coincide with countably additive measures and yield normal distributions when the sample has no black swans. When the sample includes black swans, the new likelihoods are represented by a combination of countable and finitely additive measures with both parts present. The axioms were introduced in Chichilnisky (2000, 2002); they extend the axiomatic foundations of Savage (1954), Villegas (1964) and Arrow (1971) and they are valid for bounded and unbounded samples (Chichilnisky, 1996b). The finitely additive measures assign more weight to rare events than do standard distributions and in that sense explain the persistent observation of power laws and 'heavy tails' that eludes classic theory.

1. Introduction

Black swans are rare events with major consequences, such as market crashes, natural hazards, global warming and major episodes of extinction. For such events one has few or no prior observations,

∗ Corresponding address: 335 Riverside Drive, NY 10025, USA. Tel.: +1 212 678 1148; fax: +1 212 678 0405.
E-mail addresses: gc9@columbia.edu, chichilnisky1@gmail.com.
URL: http://www.chichilnisky.com.

0165-4896/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.mathsocsci.2009.09.007

challenging standard statistics. Statistical measurements that rely on standard observations discard 'outliers', underestimate the occurrence of rare events no matter how important these may be. To correct this bias, the natural hazards literature uses power laws to predict such risks, without however an axiomatic foundation.

This article provides a new axiomatic foundation for statistics requiring sensitivity to rare and frequent events, and characterizes the likelihoods or subjective probability distributions that these axioms imply. The new distributions coincide with standard distributions when the sample is populated by frequent events. Generally 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 new axioms incorporate and extend both Savage's (1954) axiomatization based on finitely additive measures, as well as Villegas' (1964) and Arrow's (1971) that are based on countable measures. Savage (1954) axiomatized subjective expected utility with a finitely additive measure representing the decision makers' beliefs, an approach that can ignore frequent events as shown in the Appendix. Villegas (1964) and Arrow (1971) introduced an extra continuity axiom (monotone continuity) that yields countable additivity. However monotone continuity has unusual implications: it predicts that in exchange for a couple of cents, one should be willing to accept a small risk of death, a possibility that Kenneth 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 for 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 work 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, Chanel and Chichilnisky (2009a,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 likelihoods that satisfy our axioms, and a characterization of all the likelihoods that do.

The results are based on an axiomatic approach to choice under uncertainty and sustainable development introduced by the author Chichilnisky (2000, 2002), and illuminate the issue of continuity that is at the core of 'sufficient statistics'. Continuity depends on the topology that is used to measure statistical proximity. Villegas and Arrow's 'monotone continuity' is implicitly based on a topology that neglects rare events. We use instead a topological approach that reflects recent neurological findings about catastrophes and the special role of fear in perception Le Doux (1996) and Chichilnisky (2000, 2002), yielding 'extremal' responses to extremal risks Accordingly, we define 'nearby' observations in a way that is sensitive both to extremal as well as average events. The proximity notion used here has been called the 'topology of fear' (2009a) and is sharply sensitive to extremal or catastrophic events even when these are infrequent. A similar topology was used by Debreu (1953) in the initial formulation of Adam Smith's Invisible Hand theorem. The new axioms give rise to new types of distribution, a mix of 'countably additive' and 'finitely additive' measures, with both parts present. Finitely additive measures were studied in Savage (1954), Dubins and Savage (1965), Dubins (1975) and Purves and Sudderth (1976) among others. The new combined distributions were introduced and studied in Chichilnisky (2000, 2002), Chichilnisky (1996a,b), Chichilnisky (2009a,b, in press) and could play an important role in the measurement and management of risks such as climate change, market crashes (Chichilnisky and Wu, 2006), extinction events, and the statistics of change.

2. The mathematics of uncertainty

Uncertainty is described by a set of distinctive and exhaustive possible **events** that are represented by a family of sets $\{U_\alpha\}$ whose union describes a universe $\mathcal{U} = \cup_\alpha U_\alpha$. An event $U \in \mathcal{U}$ can be identified with its **characteristic function** $\phi_U : \mathcal{U} \to R$ where $\phi_U(x) = 1$ when $x \in U$ and $\phi_U(x) = 0$ when $x \notin U$.

Please cite this article in press as: Chichilnisky, G., The foundations of statistics with black swans. Mathematical Social Sciences (2009), doi:10.1016/j.mathsocsci.2009.09.007

The (relative) **likelihood** of an event measures how likely it is to occur; it is also called a subjective probability (Anscombe and Aumann, 1963) and in this article we make no distinction between the two concepts. **Axioms** for relative likelihoods are natural and self evident properties; classic axioms can be found in De Groot (1970) Chapter 6.

It is generally assumed that 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 $U_1 \cap U_2 = \emptyset$, and the likelihood of the empty set is zero $W(\emptyset) = 0$. These properties correspond to the definition of a finite additive measure on the measurable sets of $\mathcal{U}$, which is Savage's (1954) definition of a subjective probability. C**ountable additivity** means $W(\cup_{h=1^\infty} U_h) = \sum_{h=1}^{\infty} W(U_h)$ if $\forall i, j, U_i \cap U_j = \emptyset$. A **purely finitely additive likelihood** is additive but not countably additive. The following two axioms appear in Villegas (1964), Arrow (1971) and De Groot (1970); their role is to yield 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 $\{E_\alpha\}_{=1,2\ldots}$ (i.e.$\forall \alpha, E_{\alpha+1} \subset E_\alpha$ 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 likelihood 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 likelihood of sets along a vanishing sequence goes to zero. The decreasing sequence could consist of infinite intervals of the form (n, ∞) for $n = 1, 2 \ldots$. 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 $n = 1, 2 \ldots$, 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 $A_1 \supset A_2 \supset \cdots$ is a decreasing sequence of events and B is some fixed event that is less likely than A_i for all i, then the intersection $\cap_i^\infty A_i$ is more likely than 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 likelihood 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 likelihood 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 $A_1 - \cap_i^\infty A_i, A_2 - \cap_i^\infty A_i, \ldots$. 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 Groots'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 Section 3 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 if the probability of death is small enough. Even Kenneth 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 how this works. consider an experiment (Chanel and Chichilnisky, 2009a,b) where the 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. 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 in the process determining a 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. They 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. For example, if instead of one cent and half a cent one considers one billion dollars and half a cent — as Arrow (1971) suggests in p. 48 and 49: then under certain conditions one may be willing to take the lottery that involves a small probability of death and one billion dollars over the one that offers half a cent. In another example, a person who has a short term to live – for example an incurable cancer – may be inclined to choose a pill if the payment and the size of the pile are large enough. This seems a reasonable solution to the dilemma that Arrow raises. 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 $\{E_\alpha\}_{=1,2...}$ an $N > 0$ such that altering arbitrarily the events f and g on the set E^i, where $i > N$, does not alter the likelihood ranking of the events, namely $W(f') > W(g')$, where f' and g' are the altered events. Observe that for other events f and g with $W(f) > W(g)$, there exist vanishing sequence of events $\{E_\alpha\}_{=1,2...}$ where for every N, altering arbitrarily the events f and g on the set E^i, where $i > N$, does alter the likelihood ranking of the events, namely $W(f') < W(g')$, where f' and g' are the altered events.

Definition: A likelihood 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. Formally, a likelihood 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 likelihood 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 likelihoods are biased against rare events.*

Proof. It follows from Propositions 1 and 2. ■

Please cite this article in press as: Chichilnisky, G., The foundations of statistics with black swans. Mathematical Social Sciences (2009), doi:10.1016/j.mathsocsci.2009.09.007

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

Proof. See example in the Appendix. ■

Proposition 4. *A likelihood 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 statistics that is unbiased against rare and frequent events. The axioms are as follows:

Axiom 1: *Likelihoods are continuous and additive.*

Axiom 2: *Likelihoods are unbiased against rare events.*

Axiom 3: *Likelihoods 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** of the likelihood captures the notion that 'nearby' events are 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 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. 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 likelihood 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. The three axioms above require more than the additive likelihoods of Savage, since purely finitely additive likelihoods are finitely additive and yet they must be excluded here; at the same time the axioms require less than the countably additivity of Villegas and Arrow, since countably additive likelihoods 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 $W : L_\infty(\mathcal{U}) \to R$ is taken to be continuous with respect to the sup norm.

5. Axiomatic statistics on R or (a, b)

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

Please cite this article in press as: Chichilnisky, G., The foundations of statistics with black swans. Mathematical Social Sciences (2009), doi:10.1016/j.mathsocsci.2009.09.007

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 likelihood that satisfies the classic axioms in De Groot (1970) is called a **standard likelihood**. A classic result is that for any event $f \in L_\infty$ a standard likelihood has the form $W(f) = \int_R f(x).\phi(x)\mathrm{d}\mu$, where $\phi \in L_1(R)$ is an integrable 'density' function on R.

The next step is to introduce the new axioms, show the existence of likelihoods that satisfy them and characterize all the likelihoods that satisfy the axioms. We need more definitions. A likelihood $W : L_\infty \to R$ 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 likelihood 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. A likelihood $W : L \to R$ 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: *$W : L_\infty \to R$ is linear and continuous.*

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

Axiom 3: *$W : L_\infty \to R$ 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 likelihood function satisfies Axioms 1 and 2 , but it is biased against rare events and therefore does not satisfy Axiom 3.*

Proof. Consider a standard likelihood $W(f) = \int_R f(x)\phi(x)\mathrm{d}x$, $\int_R \phi(x)\mathrm{d}x = K < \infty$. Then

$$W(f) + W(g) = \int_R f(x)\phi(x)\mathrm{d}x + \int_R g(x)\phi(x)\mathrm{d}x = \int_R f(x) + g(x).\phi(x)\mathrm{d}x = W(f+g),$$

and therefore W is **linear**. A standard likelihood is **continuous** with respect to the L_1 norm $\| f \|_1 = \int_R | f(x) | \phi(x)d\mu$ (19) because $\| f \|_\infty < \varepsilon$ implies

$$W(f) = \int_R f(x).\phi(x)\mathrm{d}x \leq \int_R | f(x) | .\phi(x)\mathrm{d}x \leq \varepsilon \int \phi(x)\mathrm{d}x = \varepsilon K.$$

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 likelihood 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 relative likelihood is however not sensitive to rare events, as shown in Chichilnisky (2000, 2002), Chichilnisky (1996b), Chichilnisky (2009a,b). ■

6. Existence and representation

Theorem 2. *There exists a relative likelihood $W : L_\infty \to R$ satisfying Axioms 1, 2, and 3. A likelihood 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 $f \in L_\infty$*

$$W(f) = \lambda \int_{x \epsilon R} f(x)\phi_1(x)\mathrm{d}x + (1-\lambda)\phi_2(f) \tag{1}$$

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

Please cite this article in press as: Chichilnisky, G., The foundations of statistics with black swans. Mathematical Social Sciences (2009), doi:10.1016/j.mathsocsci.2009.09.007

Proof. This result follows from the representation theorem in Chichilnisky (2000, 2002). ■

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, $L'_\infty = \{f \epsilon L_\infty : \lim_{x\to\infty} f(x) < \infty\}$, define $\phi_2(f) = \lim_{x\to\infty} f(x)$ and extend this to a function on all of L_∞ using Hahn Banach's theorem. The difference between a standard likelihood 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 2. *Absent rare events, a likelihood that satisfies Axioms* 1, 2, *and* 3 *is consistent with classic statistical axioms and yields a standard likelihood.*

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 and Hewitt (1952), Yosida (1974), Chichilnisky and Heal (1997) and Kadane and O'Hagan (1995).

Example 2 (*Representing a Purely Finitely Additive Measure*). Define a measure ρ as follows: for every Borel measurable set $A \subset R$, $\rho(A) = 1$ if $A \supset \{x : x > a$, 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 $V_i = (i, i+1] \bigcup (-i-1, -i]$, satisfy $V_i \bigcap V_i = \varnothing$ when $i \neq j$, and $\bigcup_{i=0}^{\infty} V_i = \bigcup_{i=0}^{\infty} (i, i+1] \bigcup (-i-1, -i] = R$, 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

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

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 $x \to \infty$, by the formula $\rho(f) = \lim_{x\to\infty} f(x)$, as in Example 1 of the previous section. The measure $\rho(.)$ can be seen 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 and 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 and 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).

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 were presented at Stanford University's 1993 Seminar on Reconsideration of Values, organized by 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 Figuieres and Mabel Tidball at LAMETA Universite de Montpellier, France December 19 and 20, 2008, and by Professor Alan Kirman at GREQAM Universite 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.

Appendix

Example 3 (*A Likelihood that is Biased Against Frequent Events*). Consider $W(f) = \liminf_{x\epsilon R}(f(x))$. 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 $E^i = [i, \infty)$, $i = 1, 2, \ldots$. This is a vanishing family because $\forall i\ E^i \supset E^{i+1}$ and $\bigcap_{i=1}^{\infty} E^i = \varnothing$. 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 and Hewitt (1952), Yosida (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 υ is *absolutely continuous*with respect to the Lebesgue measure, and (ii) a subspace consisting of purely finitely additive measures. 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.

References

Anscombe, F.J., Aumann, R.J., 1963. A definition of subjective probability. Ann. Math. Statist. 43, 199–295.
Arrow, K., 1971. Essays in the Theory of Risk Bearing. North Holland, Amsterdam.
Chanel, O., Chichilnisky, G., 2009a. The influence of fear in decisions: Experimental evidence. J. Risk Uncertainty 39 (3).
Chanel, O., Chichilnisky, G., 2009b, The Value of Life: Theory and Experiments Working Paper. GREQE. Universite de Marseille and Columbia University, New York.

Chichilnisky, G., 2000. An axiomatic approach to choice under uncertainty with catastrophic risks. Resour. Energ. Econ. 22, 221–231.

Chichilnisky, G., 2002. In: El-Shaarawi, A.H., Piegorsch, W.W. (Eds.), Catastrophic Risk. In: Encyclopedia of Environmetrics, vol. 1. John Wiley & Sons, Chichester, pp. 274–279.

Chichilnisky, G., 1996a. An Axiomatic Approach to Sustainable Development. Soc. Choice Welf. 13, 321–357.

Chichilnisky, G., 1996b. Updating Von Neumann Morgenstern axioms for choice under uncertainty. In: Proceedings of a Conference on Catastrophic Risks. The Fields Institute for Mathematical Sciences, Toronto Canada.

Chichilnisky, G., 2009a. The limits of econometrics: Non parametric estimation in hilbert spaces. Econ. Theor. 25 (4), 1070–1086.

Chichilnisky, G., 2009b. The topology of fear. J. Math. Econ. Special Issue NBER General Equilibrium Conference in Honor of Gerard Debreu, University of Kansas, volume 45, issue 11–12, December 2009.

Chichilnisky, G., 2009. Choice under Uncertainty: The Work and Legacy of Kenneth Arrow Encyclopedia of Quantitative Finance, Cont, R. (Ed.), John Wiley and Sons, Chichester, (in press).

Chichilnisky, G., Heal, G.M., 1997. Social choice with infinite populations. Soc Choice Welf 14, 303–319.

Chichilnisky, G., W.u., H.-M., 2006. General equilibrium with endogenous uncertainty and default. J. Math. Econ. 42, 499–524.

Debreu, G., 1953. Valuation equilibrium and pareto optimum. Proceedings of the National Academy of Sciences 40, 588–592.

De Groot, M.H., 1970. Optimal Statistical Decisions. John Wiley and Sons, Hoboken, New Jersey.

Dubins, L., 1975. Finitely additive conditional probabilities, conglomerability and disintegration. Ann. Probab. 3, 89–99.

Dubins, L., Savage, L., 1965. How to Gamble if You Must. McGraw Hill, New York.

Dunford, N., Schwartz, J.T., 1958. Linear Operators, Part I. Interscience, New York.

Godel, K., 1940. The consistency of the continuum hypothesis. In: Annals of Mathematical Studies 3. Princeton University Press, Princeton.

Kadane, J.B., O'Hagan, A., 1995. Using finitely additive probability: Uniform distributions of the natural numbers. J. Amer. Statist. Assoc. 90, 525–631.

Le Doux, J., 1996. The Emotional Brain. Simon and Schuster, New York.

Purves, R.W., Sudderth, 1976. Some finitely additive probability. The Annals of Probability 4, 259–276.

Savage, Leonard J., 1954. The Foundations of Statistics. John Wiley and Sons, New York, revised edition 1972.

Villegas, C., 1964. On quantitative probability σ-algebras. Ann. Math. Stat. 35, 1789–1800.

Yosida, K., 1974. Functional Analysis, 4th edition. Springer Verlag, New York and Heidelberg.

Yosida, K., Hewitt, E., 1952. Finitely level independent measures. Trans. Amer. Math. Soc. 72, 46–66.

Please cite this article in press as: Chichilnisky, G., The foundations of statistics with black swans. Mathematical Social Sciences (2009), doi:10.1016/j.mathsocsci.2009.09.007