When coherent preferences may not preserve indifference between equivalent random variables: A price for unbounded utilities.*

Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane

Abstract

We extend de Finetti's (1974) theory of coherence to apply also to unbounded random variables. We show that for random variables with mandated infinite prevision, such as for the St. Petersburg gamble, coherence precludes indifference between equivalent random quantities. That is, we demonstrate when the prevision of the difference between two such equivalent random variables must be positive. This result conflicts with the usual approach to theories of Subjective Expected Utility, where preference is defined over lotteries. In addition, we explore similar results for unbounded variables when their previsions, though finite, exceed their expected values, as is permitted within de Finetti's theory. In such cases, the decision maker's coherent preferences over random quantities is not even a function of probability and utility. One upshot of these findings is to explain further the differences between Savage's theory (1954), which requires bounded utility for non-simple acts, and de Finetti's theory, which does not. And it raises a question whether there is a theory that fits between these two.

* We thank Fabio Cozman and Matthias Troffaes for helpful discussions. Seidenfeld gratefully acknowledges support by the State of Sao Paulo, Brazil, grant FAPESP #05/59541-4, during his stay at the University of Sao Paulo, Spring term 2006.

1. **Introduction**. In this paper we examine coherent preferences over unbounded variables in order to demonstrate general circumstances when rational preference cannot be a function of probability and utility, let alone an ordering represented by subjective expected utility. We use de Finetti's (1974) theory of coherent previsions, which we summarize and extend to unbounded values, as follows.

For each real-valued random variable, X, defined on a common state space the decision maker has a prevision $\mathbf{Prev}(X)$, which may be real-valued or, when X is unbounded, possibly infinite, negative or positive.

When the prevision for X is real-valued, it is subject to a two-sided, real-valued payoff $\boldsymbol{c_X}(X - \mathbf{Prev}(X))$, where $\boldsymbol{c}_{\mathrm{X}}$ is a real number. The prevision is said to be two-sided, as $\boldsymbol{c}_{\mathrm{X}}$ may be either positive or negative, corresponding informally to the decision maker "buying" or "selling" the random payoff X for the amount $\mathbf{Prev}(X)$, scaled by the magnitude $|\boldsymbol{c_X}|$. That is, the decision maker is committed to using $\mathbf{Prev}(X)$ as the "fair price" when buying or selling the $|\boldsymbol{c_X}|$-multiples of the random quantity X.

When the prevision for X is infinite-positive, i.e., when X has a value to the decision maker greater than any finite amount, then for each real constant $\boldsymbol{k}$ and for each $\boldsymbol{c_X} > 0$, we require that the decision maker is willing to accept (i.e., is committed to "buy") a one-sided payoff $\boldsymbol{c_X}(X - \boldsymbol{k})$. Likewise, when the prevision for X is infinite-negative, with value less than any finite amount, then for each real constant $\boldsymbol{k}$ and for $\boldsymbol{c_X} < 0$, we require the decision maker is willing to accept (i.e., is committed to "sell") a one-sided payoff $\boldsymbol{c_X}(X - \boldsymbol{k})$. Moreover, in accord with de Finetti's theory, the decision maker is required to accept an arbitrary, finite sum of such real-valued payoffs across (finitely many) random variables.

Definition: Previsions are ***coherent*** if there is no finite selection of non-zero constants, $\boldsymbol{c}_{\mathrm{X}}$, with the sum of the payoffs *uniformly* dominated by 0. The previsions are ***incoherent*** otherwise.

Theorem (de Finetti, 1974, 3.10 & 3.12): Previsions over the set of *bounded* random variables are coherent if and only if they are the expectations of some finitely additive probability.

Note that when the random variables in question are the indicator functions for events in the space, then their coherent previsions are their probabilities under the finitely additive measure that satisfies the theorem above.

1.1 **Coherent previsions for unbounded random variables**. The theory of coherent previsions is not confined to bounded random variables, as de Finetti notes (1974, Sections 3.12.4, 6.5.4-6.5.9). We extend his theory, as follows

Define a (weak) order, «, over random variables according to their coherent previsions by:

Definition: $X \ll Y$ if and only if $\mathbf{Prev}(Y - X) > 0$,

with $X \equiv Y$ if and only if $\mathbf{Prev}(Y - X) = \mathbf{Prev}(X - Y) = 0$.

However, in such cases, as previsions are not necessarily real-valued, they may induce an ordering of random variables, «, that violates standard Archimedean and Continuity conditions, as explained in section 4.

Random variables can be unbounded either above, below, or in both directions. In the results that follow, when we refer to an unbounded random variable, we will mean unbounded above only. Everything that we prove for such random variables carries over in an obvious way to random variables unbounded below only. When we need to refer to the prevision of a random variable unbounded in both directions, for example Y-X when both are unbounded above, we appeal to the Fundamental Theorem of Prevision, essentially proven as Theorem 3.10 of (de Finetti, 1974). This theorem guarantees the existence of a coherent prevision for each such random variable, but generally provides only an interval each of whose values is a possible coherent prevision.

Also in what follows we use only countably additive probabilities, for two reasons. It simplifies our analysis and it shows that the difficulties demonstrated here with de

Finetti's theory of coherence applied to unbounded random variables are not avoided merely by restricting previsions to countably additive probabilities.

Our central goal in this paper is to explore the class of coherent weak-orders as defined above, and establish general conditions under which such coherent previsions for unbounded random variables must distinguish by strict preference among a (finite) set of ***equivalent*** ($\approx$) random variables. That is let $\{Y_1, \ldots., Y_m\}$ be a finite set of equivalent but unbounded discrete random variables. In other words, for each real number r, **Prob**($Y_i = r$) is the same for each Y_i, $i = 1, \ldots m$. We provide general conditions under which coherence **precludes** $Y_i \equiv Y_j$ though $Y_i \approx Y_j$.

There are two cases of previsions for unbounded variables relevant to our analysis of equivalent random variables.

> **Case 1**: **When a coherent prevision differs from its expected value**: Let Z be an unbounded, discrete random variable ($Z = 1, 2, \ldots$), e.g., a Geometric(**p**) distribution, **Prob**($Z = n$) = $\mathbf{p}(1\text{-}\mathbf{p})^{n-1}$ ($n = 1, 2, \ldots$). Suppose the prevision for Z, **Prev**(Z), is greater than its expectation, **E**[Z], where for the Geometric(**p**), **E**[Z] = $\mathbf{p}^{-1}$. Then, generally, write **Prev**(Z) = **E**[Z] + $\boldsymbol{b}$, with $\boldsymbol{b} > 0$. We let '$\boldsymbol{b}$' denote the ***boost*** that the prevision of Z receives in excess of its expected value. (Possibly, $\boldsymbol{b}$ is positive-infinity in the example.) Surprisingly, a finite prevision for a non-negative, unbounded random variable that includes a finite, positive boost $\boldsymbol{b}$ is a coherent 2-sided prevision in de Finetti's sense. This is so because there can be no sure loss when this prevision is combined with coherent previsions for bounded variables.

> **Case 2: When a coherent prevision must be infinite**: Let Z be an unbounded random variable such that for each coherent prevision, and for each real number $r > 0$, there is a linear combination W of random quantities such that either $Z > W$ and **Prev**(W) $\geq r$, or $W > Z$ and **Prev**(W) $\leq r$. That is, in this case each coherent prevision for Z must be infinite. An example of this is the familiar St. Petersburg

variable, Z, where with probability 2^{-n}, $Z = 2^n$. Flip a fair coin until it lands heads for the first time at flip n, when Z equals 2^n.

2. Strict preference among equivalent random variables having finite previsions greater than their expectations.

2.1 We begin with previsions under Case 1.

***Theorem* 1**: Let X be a Geometric($\mathbf{p}$) random variable. If the prevision for X, $\mathbf{Prev}(X)$, is finite but greater than its expectation ($\mathbf{E}[X] = \mathbf{p}^{-1}$), though this is a coherent prevision in de Finetti's sense, then there exist three equivalent random variables, W_1,W_2, and W_3, each with finite previsions such that some two cannot have the same prevision.

Example 2.1: Before we demonstrate this result, we offer an illustration for the special case of fair-coin flipping. Let X be a Geometric(1/2) random variable. For convenience, denote by $\mathbf{\Omega}$ the partition of the state space into events $\{\omega_n: n = 1, 2, ...\}$ such that $X(\omega_n) = n$, i.e., where ω_n denotes the event where the fair coin lands heads first on the n^{th} flip. Hence, $\mathbf{Prob}(X = n) = \mathbf{Prob}(\omega_n) = 2^{-n}$, $n = 1, 2, \ldots$, and $\mathbf{E}[X] = 2$.

Let $\{B, B^{\mathbf{c}}\}$ be the outcome of an independent flip of another fair coin, so that $\mathbf{Prob}(B, \omega_n) = 2^{-(n+1)}$, for $n = 1, 2, \ldots$.

With the state space $\{B, B^{\mathbf{c}}\} \times \mathbf{\Omega}$, define two other random variables W_1, and W_2 as follows:

$W_1(B, \omega_n) = n+1$; $W_1(B^c, \omega_n) = 1$ $(n = 1, 2, ...)$

and $W_2(B^{\mathbf{c}}, \omega_n) = n+1$; $W_1(B, \omega_n) = 1$ $(n = 1, 2, ...)$.

Table 1, below, displays the three equivalent variables, X, W_1, and W_2, defined over the space $\{B, B^{\mathbf{c}}\} \times \mathbf{\Omega}$.

TABLE 1

	ω_1	ω_2	…	ω_n	…
	$X = 1$	$X = 2$		$X = n$	
B	$W_1 = 2$	$W_1 = 3$		$W_1 = n+1$	
	$W_2 = 1$	$W_2 = 1$		$W_2 = 1$	
	$X = 1$	$X = 2$		$X = n$	
$\mathbf{B}^c$	$W_1 = 1$	$W_1 = 1$		$W_1 = 1$	
	$W_2 = 2$	$W_2 = 3$		$W_2 = n+1$	

Obviously, W_1 and W_2 are equivalent. Moreover, each has a Geometric(½) distribution; hence, $X \approx W_1 \approx W_2$. However, for each state (b,ω) in $\{B, B^c\} \times \Omega$, $W_1(b,\omega) + W_2(b,\omega) - X(b,\omega) = 2$. Thus, $\mathbf{Prev}(W_1\text{-}X) + \mathbf{Prev}(W_2\text{-}X) = 0$ if and only if $\mathbf{Prev}(W_1) = \mathbf{Prev}(W_2) = \mathbf{Prev}(X) = 2$, when the prevision for a Geometric(½) variable is its expectation, and then $\boldsymbol{b} = 0$.

***Proof of Theorem* 1**: We offer an indirect proof, assuming for the *reductio* the hypothesis that equivalent random variables with finite expectations carry equal prevision. The argument is presented in 3 parts, in Appendix 1. Part 1 of the proof defines the equivalent random variables whose previsions, in the end, cannot all be equal. Part 2 develops general results how previsions for independent random variables relate to their expected values. Part 3 puts the pieces together.

2.2 *An elemental property of coherent boost.*

For a random variables X with finite expectation, define the *boost function*

Definition: $\beta(X) = \mathbf{Prev}(X) - \mathbf{E}(X)$.

It is straightforward to show that when $-\infty < \mathbf{E}(X) < \infty$, the boost function $\beta(\bullet)$ is a finitely additive linear operator (Dunford and Schwartz, 1988, p. 36) that has the value 0 on all bounded random variables. That is, $\beta(X+Y) = \beta(X) + \beta(Y)$, $\beta(\boldsymbol{a}X) = \boldsymbol{a}\beta(X)$, and $\beta(X) = 0$ when X is a bounded random variable.

From this observation we obtain the following elementary result about non-negative boosts.

Proposition: Let $0 < \beta(X) < \infty$. Then, $\beta(X^k) = \infty$ if $k > 1$, and $\beta(X^k) = 0$ if $k < 1$.

Thus, the family of random variables with finite expectations and finite boost have a common tail distribution, up to scalar multiples of one another. This leads to the following generalization of Theorem 1.

Corollary: Let **Q** be a probability distribution whose tail is stochastically dominated by the tail of some Geometric(**p**) distribution, i.e., there exists k and **p** such that for all $n \geq k$, $\mathbf{Q}(Y = n) = q_n \leq \mathbf{p}(1-\mathbf{p})^{n-1} = p_n$. If all Geometric(**p**) distributions have 0 boost, then by coherence the **Q**-distribution also has 0 boost.

3. Strict preferences among generalized St. Petersburg random variables.

We turn next to previsions that result from random variables whose coherent prevision is mandated to be infinite, Case 2. For each Geometric($1-2^{-m}$) distribution, ($m = 2, 3, \ldots$) we define a generalized St. Petersburg gamble, Z_m, and construct a set of 2^{m-1} equivalent random variables $X_1 \approx X_2 \approx \ldots \approx X_{2^{m-1}}$ ($\approx Z_m$), such that if previsions are coherent, though infinite, then

$$\sum_{i=1}^{2^{m-1}} \mathbf{Prev}(X_i - Z_m) = 2^{m-1}. \qquad (*)$$

That is, though the X_i ($i = 1, \ldots, 2^{m-1}$) are pairwise equivalent random variables (and equivalent to Z_m), the coherent prevision of their pairwise differences with Z_m cannot all be 0. Then we extend this argument to include random variables that are equivalent to the “tail” of a Geometric($1-2^{-m}$) distribution.

Example 3.1: First, however, we illustrate these two results for the Geometric(½) distribution. Let ω_n ($n = 1, \ldots$) have a Geometric(½) distribution, $\mathbf{Prob}(\omega_n) = 2^{-n}$ ($n = 1, 2, \ldots$). Let $\{B, B^c\}$ be the outcome of a fair-coin flip, independent of the states, ω_n. Partition each ω_n into two equi-probable cells using independent, probability 1/2 events $\{B, B^c\}$, and define three equivalent random variables, X_1, X_2, and Z, on the product space $\{B, B^c\} \times \Omega$ as follows:

- the traditional St. Petersburg random variable: $Z(\omega_n) = 2^n$, independent of B.
- the random variable $X_1(B \cap \omega_n) = 2^{n+1}$, and $X_1(B^c \cap \omega_n) = 2$.
- the random variable $X_2(B \cap \omega_n) = 2$, and $X_2(B^c \cap \omega_n)= 2^{n+1}$.

Table 2 displays these variables, defined on the space $\{B, B^c\} \times \Omega$, where ω_k denotes the state where the fair coin lands heads first on the k^{th} flip.

TABLE 2

	ω_1	ω_2	....	ω_n	....
	$Z = 2$	$Z = 4$		$Z = 2^n$	
B	$X_1 = 4$	$X_1 = 8$		$X_1 = 2^{n+1}$	
	$X_2 = 2$	$X_2 = 2$		$X_2 = 2$	
	$Z = 2$	$Z = 4$		$Z = 2^n$	
B^c	$X_1 = 2$	$X_1 = 2$		$X_1 = 2$	
	$X_2 = 4$	$X_2 = 8$		$X_2 = 2^{n+1}$	

Though $X_2 \approx X_2 \approx Z$, if previsions are coherent, for each state (b,ω) in $\{B, B^c\} \times \Omega$, $X_1(b,\omega) + X_2(b,\omega) - 2Z(b,\omega) = 2$. Hence, **Prev**$(X_1+X_2-2Z) = 2$. This is in contradiction with the hypothesis that the prevision of the difference between equivalent random variables is 0, as then **Prev**(X_1-Z) + **Prev**(X_2-Z) = **Prev**$(X_1+X_2-2Z) = 0$.

Next, consider a random variable Z_k that agrees with Z on a "tail," and is 0 elsewhere, i.e., $Z_k(\omega_n) = Z(\omega)$ for all $n \geq k$ and $Z_k(\omega_n) = 0$ otherwise. Define the two other equivalent random variables, X_{k1} and X_{k2}, as follows. These definitions generalize the previous construction. That is, in what comes next, by letting $k = 1$ we obtain the construction above.

- a called-off St. Petersburg variable: $Z_k(\omega_n) = 2^n$ for $n \geq k$, and $Z_k(\omega_n) = 0$ for $n < k$.
- a called-off variable $X_{k1}(\omega_n \cap B) = 2^{n+1}$ and $X_{k1}(\omega_n \cap B^c) = 2^k$ for $n \geq k$, $X_{k1}(\omega_n) = 0$ for $n < k$.
- a called-off variable $X_{k2}(\omega_n \cap B^c) = 2^{n+1}$ and $X_{k2}(\omega_n \cap B) = 2^k$ for $n \geq k$, $X_{k2}(\omega_n) = 0$ for $n < k$.

Table 3 displays these equivalent random variables which are again defined on the product space, $\{B, B^c\} \times \Omega$.

TABLE 3

	ω_1	…	ω_{k-1}	…	ω_k	…	ω_n	…
	$Z = 0$		$Z = 0$		$Z = 2^k$		$Z = 2^n$	
B	$X_{k1} = 0$		$X_{k1} = 0$		$X_{k1} = 2^{k+1}$		$X_{k1} = 2^{n+1}$	
	$X_{k2} = 0$		$X_{k2} = 0$		$X_{k2} = 2^k$		$X_{k2} = 2^k$	
	$Z = 0$		$Z = 0$		$Z = 2^k$		$Z = 2^n$	
B^c	$X_{k1} = 0$		$X_{k1} = 0$		$X_{k1} = 2^k$		$X_{k1} = 2^k$	
	$X_{k2} = 0$		$X_{k2} = 0$		$X_{k2} = 2^{k+1}$		$X_{k2} = 2^{n+1}$	

Though these are three equivalent random variables then, by similar reasoning as above, $\mathbf{Prev}(X_{k1} + X_{k2} - 2Z_k) = 2$, and it cannot be that $X_{k1} \approx X_{k2} \approx Z_k$.

Generalized St. Petersburg variables.

Theorem 2: For each Geometric($1\text{-}2^{-m}$) distribution, ($m = 2, 3, \ldots$), we define a generalized St. Petersburg gamble, Z_m, and construct a set of another $2^{m\text{-}1}$ equivalent random variables $X_1 \approx X_2 \approx \ldots \approx X_{2^{m-1}}$ ($\approx Z_m$), such that if previsions are coherent, though infinite, then

$$\sum_{i=1}^{2^{m-1}} \mathbf{Prev}(X_i - Z_m) = 2^{m\text{-}1}. \qquad (*)$$

The details are provided in Appendix 2.

*Equivalent variables for a tail of the Geometric(**p**),* $\mathbf{p} = 1 - 2^m$.

In Appendix 3, we modify the construction used for Theorem 2 in order to address the case of agreement with a tail of the Geometric(**p**), resulting in the following.

Theorem 3: Assume that distribution $\boldsymbol{Q}$ agrees with the Geometric(**p**) for all (coarse) states ω_n, $n \geq k$. That is, $\boldsymbol{Q}(\omega_n) = \mathbf{p}(1\text{-}\mathbf{p})^n$, for all $n \geq k$. We define a called-off generalized St. Petersburg gamble, Z_m, and construct a set of $2^{m\text{-}1}$ equivalent called-off random variables $X_1 \approx X_2 \approx \ldots \approx X_{2^{m-1}} \approx Z_m$, such that if previsions are coherent, though infinite, then

$$\sum_{i=1}^{2^{m-1}} \mathbf{Prev}(X_i - Z_m) = 2^{m-1}. \qquad (*)$$

Our primary reason for the added generality of Theorem 3 over Theorem 2 is the following corollary.

Corollary: Let **Q** be a probability distribution whose tail stochastically dominates the tail of some Geometric(**p**) distribution, i.e., there exists $\boldsymbol{k}$ and **p** such that for all $n \geq \boldsymbol{k}$, $\mathbf{Q}(Y = n) = q_n \geq \mathbf{p}(1-\mathbf{p})^{n-1} = p_n$. Then coherent previsions under the **Q**-distribution also must distinguish between some equivalent random variables.

Proof: Thin out the tail of the **Q-**distribution using independent Benouilli trials in order to define the random variable X that agrees with the tail of this Geometric(**p**) distribution. Specifically, for each $n \geq \boldsymbol{k}$ define an event E_n with independent Bernoulli probability $\mathbf{Prob}(E_n) = e_n = 1 - (q_n - p_n)$, and where $\mathbf{Prob}(E_n, Y = n) = e_n q_n$. Define the random variable X so that $X(\omega_n) = 0$ if $n < \boldsymbol{k}$, and for $n \geq \boldsymbol{k}$, $X(\omega_n) = Y(\omega_n)$ when E_n does not obtain, otherwise, i.e., when E_n obtains, $Y(\omega_n) = 0$. Then, under the **Q**-distribution, X is equivalent to the $\boldsymbol{k}$-tail of the Geometric(**p**) distribution. The corollary follows from Theorem 3.

4. Archimedean and Continuity principles applied to Cases 1 and 2.

In order better to understand the differences between Cases 1 and 2, we consider two principles that relate coherent previsions to expected values for random variables.

> *Definition*: Given random variables X and Y, and real number $0 \leq \alpha \leq 1$, let $\alpha X \oplus (1-\alpha)Y$ denote their convex mixture.

First, recall the familiar Archimedean axiom of von Neumann-Morgenstern Expected Utility theory.

- *Archimedean principle*: Let X, Y, and Z be random variables strictly ordered as $X \ll Y \ll Z$. Then there exist $0 < \alpha, \beta < 1$ satisfying

$$\alpha X \oplus (1\text{-}\alpha)Z \ll Y \ll \beta X \oplus (1\text{-}\beta)Z.$$

This principle is violated by infinite coherent previsions, such as the St. Petersburg lottery of Case 2. The argument is elementary. Let X and Y be constant random variables, $X = 1$ and $Y = 2$. And let Z be a variable with infinite prevision, e.g., the St. Petersburg variable. Then though $X \ll Y \ll Z$ there is no $0 < \alpha < 1$ satisfying $\alpha X \oplus (1\text{-}\alpha)Z \ll Y$. At the same time, this principle does not preclude a coherent prevision that uses a positive boost, as in examples from Case 1. That is, this Archimedean principle blocks only coherent previsions falling under Case 2, but does not prevent those from Case 1. But previsions from Case 1 cannot be represented as a function of probability and utility.

Together with the Ordering and Independence axioms, the Archimedean principle is adequate to insure a real-valued expected utility representation for coherent preference over *simple* random variables. But this principle fails to produce that same result when random variables are non-simple, even though utility is bounded. [See Fishburn (1979, section 10) and (1982, section 11.3) for helpful discussion of this point.] And as just noted, since it allows previsions with finite, positive boost (Case 1), it is not adequate either to insure that previsions for unbounded random variables are a function of probability and utility.

A different approach, used in the definition of the Lebesgue integral for unbounded functions – see Royden (1968, p. 226), is this.

- *Continuity principle*: Let X be a non-negative variable and let $\{X_n\}$ be a sequence of non-negative random variables converging "from below" to the random variable X. That is, for each state ω, $X_n(\omega) \leq X(\omega)$ ($n = 1, \ldots$) and $lim_{\text{n}}\, X_n = X$. Then lim_{n} $\mathbf{Prev}(X_n) = \mathbf{Prev}(X)$.

Evidently, this Continuity principle blocks previsions from Case 1 – positive ***boost*** creates discontinuous previsions. The reasoning again is elementary. Let X be an unbounded non-negative random variable with finite expectation $\mathbf{E}[X] = \mu$, but whose prevision includes a positive boost, $\boldsymbol{b}$, $\mathbf{Prev}[X] = \mu + \boldsymbol{b}$. Define X_n as the bounded random variable which is a truncation of X at state ω_n: $X_n(\omega_i) = X(\omega_i)$, for $i \leq n$, and $X_n(\omega_i) = 0$ for $i > n$. Then the sequence of random variables $\{X_n\}$ converges to X (from below). But since each X_n is simple $\mathbf{Prev}[X_n) = \mathbf{E}[X_n]$ and then $lim_n\, \mathbf{Prev}[X_n] < \mathbf{Prev}[X]$. By contrast, this Continuity principle does not preclude previsions in Case 2. Hence, as with the Archimedean principle, which blocks Case 2 but not Case 1, it too is insufficient to insure that coherent previsions are represented by some function of probability and utility.

If we entertain the combination of these two principles, in order to block the problematic previsions from Case 1 and Case 2, the results are troubling because also we add other deficiencies that each principle brings in its wake.

We have already indicated that the Archimedean principle "solves" the problems posed in Case 2 by precluding infinite previsions. But, if utility is unbounded, then the construction of a St. Petersburg-style variable follows from the presence of a geometric distribution. Hence, adopting the Archimedean principle combined with allowing non-simple, discrete distributions will compel a bounded utility function.

De Finetti (1974, section 3.13) shows that when previsions for the variables X_n are coherent, and X is their (pointwise) limit, then using the Continuity principle to fix the prevision for X preserves coherence. Moreover, for a random variable with finite prevision then its prevision equals its expected utility. However, he rejects the Continuity principle as mandatory for coherence, since it precludes finitely but not countably additive probabilities. Thus, if it is important to allow merely finitely additive probability as coherent, the Continuity principle is too strong.

5. Conclusions and further questions.

The results in this paper extend de Finetti's (1974) theory of coherence to unbounded random variables without assuming either Continuity of preferences or the usual Archimedean axiom. We show that for random variables with infinite previsions, such as the St. Petersburg gamble, coherence precludes indifference between some equivalent random quantities. That is, according to Theorems 2 and 3, and their Corollary, the prevision of the difference between two such equivalent random variables must sometimes be positive. This result follows from a very liberal standard of coherence, i.e., we do no see how further to weaken de Finetti's theory of coherence beyond what we propose with 1-sided previsions in order to avoid this problem. Nor is it an issue where the debate over coherence of merely finitely additive prevision is relevant.

In addition to problems arising with St. Petersburg-styled variables, we explore similar results for unbounded variables when their previsions, though finite, exceed their countably additive expected values, as is permitted within de Finetti's theory. In such cases, the decision maker's coherent preferences over random quantities, though real valued, are not a function of probability and utility.

These results conflict with the usual approach to theories of Subjective Expected Utility, such as Savage's (1954) theory, where preference is defined over (equivalent) lotteries. The contrast is a subtle one, however. Like de Finetti's theory, Savage's theory permits merely finitely additive personal probability, i.e., preference in Savage's theory is not required to be continuous in the sense that we use here. Nor do we think it reasonable to mandate full Continuity of previsions in general as a way to avoid these problems.

In Savage's theory, the problem with unbounded random variables is sidestepped entirely because, in the presence of his postulate P7 (which is needed to control the preference relation over non-simple acts), utility is bounded (Savage, Section 5.4). Savage's postulates P1-P6 constitute a theory for simple acts that does not require a bounded utility. However, in that setting there are no unbounded random variables either, and so the difficulties with SEU theory that are the subject of this paper cannot occur.

We see the results of this paper as pointing to the need for developing a normative theory that stands between Savage's P1-P7, where utility is bounded, and de Finetti's theory of coherence, which allows finite but discontinuous previsions for unbounded random quantities, even when all bounded random quantities have continuous previsions and probability is countably additive. We hope to find a theory that navigates satisfactorily between these two landmarks.

One example of a theory taking the middle ground, and with which we are not completely satisfied, is to restrict coherence to random variables with finite previsions and to require that when previsions for indicator functions are continuous, then all random variables with finite expectations have previsions that are continuous. This combination results in a theory that permits:

1) merely finitely additive, discontinuous previsions for non-simple, bounded random variables,
2) countably additive, continuous previsions for unbounded random variables with finite expectations,

and 3) finitely additive previsions for unbounded random variables that are a convex combination of these first two.

But such a theory is not satisfactory, we think, since it requires that merely finitely additive probabilities live on compact sets. This will not do even to reconstruct text-book inference with so-called "improper" priors.

Appendix 1

***Theorem* 1**: Let X be a Geometric($\mathbf{p}$) random variable. If the prevision for X, $\mathbf{Prev}(X)$, is finite but greater than its expectation ($\mathbf{E}[X] = \mathbf{p}^{-1}$), though this is a coherent prevision in de Finetti's sense, then there exist three equivalent random variables, W_1,W_2, and W_3, each with finite previsions such that some two cannot have the same prevision.

***Proof of Theorem* 1**: We offer an indirect proof, assuming for the *reductio* the hypothesis that equivalent random variables with finite expectations carry equal prevision. The argument is presented in 3 parts, in the appendix: Part 1 of the proof defines the equivalent random variables whose previsions, in the end, cannot all be equal. Part 2 develops general results how previsions for independent random variables relate to their expected values. Part 3 puts the pieces together.

Part 1 of the proof: Let $\mathbf{Prev}(X) = \mathbf{E}[X] + \boldsymbol{b} = \boldsymbol{t} > \mathbf{p}^{-1}$. Consider two, *iid* draws from this Geometric($\mathbf{p}$) distribution, X_1 and X_2. By the hypothesis $\mathbf{Prev}(X_i) = \boldsymbol{t}$ (i = 1, 2).

Define the random variable $W = X_1+X_2$, which has a NegBin(2,$\mathbf{p}$) distribution. By coherence, then $\mathbf{Prev}(W) = 2\boldsymbol{t}$.

Note that the conditional distribution $\mathbf{Prob}(X_1 \mid W = n) = (n\text{-}1)^{-1}$ for $(1 \leq X_1 \leq n\text{-}1)$ is uniform, because $\mathbf{Prob}(X_1 = k \mid W = n) = \mathbf{Prob}(X_1 = k, X_2 = n\text{-}k, W = n) / \mathbf{Prob}(W = n)$.
$\mathbf{Prob}(X_1 = k, X_2 = n\text{-}k, W = n) = \mathbf{Prob}(X_1 = k, X_2 = n\text{-}k)$

$$= \mathbf{p}(1\text{-}\mathbf{p})^{k-1}\mathbf{p}(1\text{-}\mathbf{p})^{n-k-1}$$

$$= \mathbf{p}^2(1\text{-}\mathbf{p})^{n-2}$$

which is constant (and positive) for $1 \leq k \leq n\text{-}1$. Hence, $\mathbf{Prob}(X_1 = i \mid W = n) / \mathbf{Prob}(X_1 = j \mid W = n) = 1$ for $1 \leq i, j \leq n\text{-}1$.

Write W as a sum of four random variables: $W = W_1+W_2+\mathrm{W}_3+W_4$, with the first three equivalent random variables, defined as follows. The set of points where $W_i = n$ is defined according to 3 cases, as n = 1, 2, or 0, *mod* 3. The idea is that the n-1 many equally probable points ($X_1 = k$, $X_2 = n\text{-}k$) that comprise the event $W = n$ are divided equally among the three events $W_i = n$, (i = 1, 2, 3), with the remaining one or two points relegated to $W_4 = n$, which is a non-empty set whenever $n \neq 1$ *mod* 3.

For $n = 1$ *mod* 3:

$W_1 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 1 \leq X_1 \leq (n\text{-}1)/3\}$

$W_2 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } (n\text{-}1)/3 < X_1 \leq 2(n\text{-}1)/3\}$

$W_3 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 2(n\text{-}1)/3 < X_1 \leq (n\text{-}1)\}$

and $\{W_4 = n\} = \varnothing$.

For $n = 2$, *mod* 3:

$W_1 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 1 < X_1 \leq (n+1)/3\}$

$W_2 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } (n+1)/3 < X_1 < 2(n+1)/3\}$

$W_3 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 2(n+1)/3 \leq X_1 \leq (n\text{-}1)\}$

and $W_4 = n$ for the event $\{(X_1 = 1, X_2 = n\text{-}1)\}$

For $n = 0$, *mod* 3:

$W_1 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 1 < X_1 \leq n/3\}$

$W_2 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } n/3 < X_1 < 2n/3\}$

$W_3 = n$ for states satisfying $\{X_1 + X_2 = n \text{ and } 2n/3 \leq X_1 < (n\text{-}1)\}$

and $W_4 = n$ for either of the two event $\{(X_1 = 1, X_2 = \text{n-1}), (X_1 = \text{n-1}, X_2 = 1)\}$.

$W_i = 0$ (i = 1, 2, 3, 4) for all other states.

It is evident that $W = W_1 + W_2 + W_3 + W_4$, so **Prev**($W$) = $\sum_i$ **Prev**(W_i). Also evident is that $W_1 \approx W_2 \approx W_3$, as they are constructed so that Prob($W_i = n$) is, for each n = 1, 2, …, the same value for i = 1, 2, 3.

Part 2 of the proof: Next, we develop two general claims about previsions for independent variables, Lemmas 1 and 2, from which, in Part 3, we derive that **Prev**(W_2) < **Prev**(W_1) < **Prev**(W_3), in contradiction with the hypothesis that equivalent random variables are equally preferred.

Lemma 1: Let Y be an integer random variable with finite mean, $\mathbf{E}(Y) = \mu < \infty$, and finite prevision $\mathbf{Prev}(Y) = \pi < \infty$. Coherence assures that $\mu \leq \pi$. Let F be the indicator for an event, independent of Y, with $\mathbf{Prob}(F) = \alpha$. Then $\mathbf{Prev}(FY) = \alpha\pi$.

Proof: If α is a rational fraction, $\alpha = k/m$, the lemma follows using the *reductio* hypothesis applied to the m-many equivalent random variables F_iX , where $\{F_1, \ldots, F_m\}$ is a partition into equiprobable events F_i. That is, from the hypothesis, $\mathbf{Prev}(F_iY) = c$ ($i = 1, \ldots, m$), and by finite additivity of previsions, then $c = \pi/m$, so that $\mathbf{Prev}(FY) = k\pi/m = \alpha\pi$. If α is an irrational fraction, the lemma follows by dominance applied to two sequences of finite partitions of equally probable events. One sequence provides bounds on $\mathbf{Prev}(FY)$ from below, and the other sequence provides bounds from above.

Next, let X and Y be independent random variables defined on the positive intergers $\boldsymbol{N}$. Consider a function $\boldsymbol{g}(i) = j$, $\boldsymbol{g}:N\rightarrow N$, with the sole restriction that for each value j, $\boldsymbol{g}^{-1}(j)$ is a finite (and possibly empty) set. The graph of the function $\boldsymbol{g}$ forms a binary partition of the positive quadrant of the (X, Y)-plane into events G and G^c, with G defined as: $G = \{(x, y): \boldsymbol{g}(x) \leq y\}$. G is the region at or above the graph of $\boldsymbol{g}$. Then, on each horizontal line of points in the positive quadrant of the (X, Y)-plane, on a line satisfying $\{Y = j\}$, only finitely many points belong to the event G.

Let GX denote the random variable that equals X on G and 0 otherwise, and likewise for the random variable G^cX. So, $X = GX + G^cX$. The next lemma shows how the *boost* $\boldsymbol{b}$ for the random variable X divides over the binary partition formed by the event G.

Lemma 2. With X, Y, and G defined above,

$$\mathbf{Prev}(GX) = \mathbf{E}(GX), \text{ whereas } \mathbf{Prev}(G^cX) = \mathbf{E}(G^cX) + \boldsymbol{b}.$$

That is, ***all*** of the *boost* associated with the prevision of X attaches to the event G^c, regardless the probability of G^c.

Proof: For each value of $j = 1, 2, \ldots$, write the random variable $[Y = j]X$ as a sum of two random variables, using G (respectively G^c) also as its indicator function:

$$[Y = j]X = [Y = j]GX + [Y = j]G^cX.$$

So, $\quad$ **Prev**$([Y = j]G^cX)$ = **Prev**$([Y = j]X)$ - **Prev**$([Y= j]GX)$

and $\quad$ $\mathbf{E}([Y = j]G^cX)$ = $\mathbf{E}([Y = j]X) - \mathbf{E}([Y= j]GX)$.

But $[Y = j]GX$ is a simple random variable, as the event G contains only finitely many points along the strip $[Y = j]$. Thus, **Prev**$([Y= j]GX)$ = $\mathbf{E}([Y = j]GX)$.

Since X and Y are independent, by Lemma 1,

Prev$([Y= j]X)$ = **Prob**$(Y = j)(\mathbf{E}[X] + \boldsymbol{b})$

So, $\quad$ **Prev**$([Y = j]G^cX)$ = **Prob**$(Y = j)(\mathbf{E}[X] + \boldsymbol{b})$ - $\mathbf{E}([Y = j]GX)$

= **Prob**$(Y = j)\boldsymbol{b} + \mathbf{E}([Y= j]X) - \mathbf{E}([Y = j]GX)$

= **Prob**$(Y = j)\boldsymbol{b} + \mathbf{E}([Y = j]G^cX)$.

Thus, the prevision for $[Y = j]G^cX$ contains a boost equal to **Prob**$(Y = j)\boldsymbol{b}$. But as $\sum_j$ **Prob**$(Y = j)\boldsymbol{b} = \boldsymbol{b}$, we have **Prev**$(G^cX) = \sum_j (\mathbf{E}([Y = j]G^cX) +$ **Prob**$(Y = j)\boldsymbol{b}) = \mathbf{E}[G^cX] + \boldsymbol{b}$ and there is no boost associated with GX, **Prev**$(GX) = \mathbf{E}[GX]$.

Part 3 of the proof: Observe that, in accord with Lemma 2,

the random variable W_2 contains none of the boost associated with either X_1 or X_2,

the random variable W_1 contains none of the boost associated with X_1

and the random variable W_3 contains none of the boost associated with X_2.

By application of Lemma 1, we see that:

the boost $\boldsymbol{b}$ associated with X_2 divides between W_1 and W_4 in the proportion α:(1-α) where $\alpha = (1-\mathbf{p})(1+ \mathbf{p}(1-\mathbf{p})/[3(1-\mathbf{p}) + \mathbf{p}^2])$

and the boost $\boldsymbol{b}$ associated with X_1 divides between W_3 and W_4 in the proportion (1-γ):γ where $\gamma = \mathbf{p}(1-\mathbf{p})/[3(1-\mathbf{p}) + \mathbf{p}^2]$.

Since $W_1 \approx W_2, \approx W_3$, then $\mathbf{E}[W_1] = \mathbf{E}[W_2] = \mathbf{E}[W_3] = (2/3\mathbf{p}) - \mathbf{E}[W_4]$. Therefore, by adding the respective boosts, **Prev**(W_2) < **Prev**(W_1) < **Prev**(W_3), which establishes the theorem, and somewhat more.

Appendix 2

Theorem 2: For each Geometric($1\text{-}2^{-m}$) distribution, (m = 2, 3, …) we define a generalized St. Petersburg gamble, Z_m, and construct a set of another 2^{m-1} equivalent random variables $X_1 \approx X_2 \approx \ldots \approx X_{2^{m-1}}$ ($\approx Z_m$), such that if previsions are coherent, though infinite, then

$$\sum_{i=1}^{2^{m-1}} \mathbf{Prev}(X_i - Z_m) = 2^{m-1}. \qquad (*)$$

***Proof of Theorem* 2**: Fix $m \geq 2$. Begin with the coarse states ω_n (n = 1, ….) having Geometric(**p**) probability,

$$\text{Prob}(\omega_n) = \mathbf{p}(1\text{-}\mathbf{p})^{n-1},$$

where $\mathbf{p} = 1 - 2^{-m}$.

Partition each ω_n into $[2^{m-1} + 1] \times 2$ many cells, as follows:

The rows of the partition are comprised by a 2^{m-1}+1 fold event, with disjoint outcomes: $B_1, \ldots, B_{2^{m-1}}$, and B^c. Call $B = B_1 \cup \ldots \cup B_{2^{m-1}}$, for reasons that will become clear shortly.

Denote the two columns of the partition of ω by (t_{n1}, t_{n2}), where $t_{n1} \cup t_{n2} = \omega_n$.

The marginal probabilities for the two column events that partition the coarse state ω_n satisfy: $\mathbf{Prob}(t_{n1}) = (1\text{-}\mathbf{p})^n$ and so $\mathbf{Prob}(t_{n2}) = (2\mathbf{p}\text{-}1)(1\text{-}\mathbf{p})^{n-1}$.

The marginal probabilities for the 2^{m-1}+1 row events satisfy: $\mathbf{Prob}(B_i) = 1\text{-}\mathbf{p}$ (i = 1, …, 2^{m-1}), so that $\mathbf{Prob}(B^c) = ½$.

Let the rows and columns be independent, so that for i = 1, .., 2^{m-1},

$$\mathbf{Prob}(B_i \cap t_{n1}) = (1\text{-p})^{n+1}.$$

Next, define Z_m, the generalized St. Petersburg variable as follows:

$$Z_m(t_{n1}) = (1\text{-}\mathbf{p})^{-n} \text{ and } Z_m(t_{n2}) = 0.$$

Note that Z_m does not depend on B_i or B^c, and Z_m has infinite prevision.

Define the random variables X_i so that for i = 1, …2^{m-1}-1,

$$X_i(B_i \cap t_{n1}) = (1\text{-}\mathbf{p})^{-(n+1)}$$

$$X_i(B_i \cap t_{n2}) = 0$$

$$X_i(B_{i+1} \cap t_{n1}) = X_i(B_{i+1} \cap t_{n2}) = (1\text{-}\mathbf{p})^{-1}$$

for other states, $(j \neq i, i+1)$ $\quad X_i(B_j \cap t_{n1}) = X_i(B_j \cap t_{n2}) = 0$

and $\quad X_i(B^c \cap t_{n1}) = X_i(B^c \cap t_{n2}) = 0.$

For $X_{2m\text{-}1}$, modify only the third line of this definition, as follows:

$$X_{2m\text{-}1}(B_{2m\text{-}1} \cap t_{n1}) = (1\text{-}\mathbf{p})^{-(n+1)}$$

$$X_{2m\text{-}1}(B_{2m\text{-}1} \cap t_{n2}) = 0$$

$$X_{2m\text{-}1}(B_1 \cap t_{n1}) = X_{2m\text{-}1}(B_1 \cap t_{n2}) = (1\text{-}\mathbf{p})^{-1}$$

for other states, $(j \neq 2^{m\text{-}1}, 1)$ $\quad X_{2m\text{-}1}(B_j \cap t_{n1}) = X_{2m\text{-}1}(B_j \cap t_{n2}) = 0$

and $\quad X_{2m\text{-}1}(B^c \cap t_{n1}) = X_{2m\text{-}1}(B^c \cap t_{n2}) = 0.$

The X_i are pairwise equivalent random variables as is evident from the symmetry of their definitions and the fact that the first $2^{m\text{-}1}$ rows have equal probability. Each X_i is equivalent to Z_m as well, since the probability that each assumes the value $(1\text{-}\mathbf{p})^{-n}$ is $(1\text{-}\mathbf{p})^n$ for $n = 1, 2, \ldots$. Table 4, below, displays these $2^{m\text{-}1}+1$ random variables defined over the $2^{m\text{-}1}+1 \times 2$ partition of the one coarse state ω_n.

> *Technical Aside*: The equivalence among the $2^{m\text{-}1}+1$ variables $Z_m, X_1, X_2, \ldots, X_{2m\text{-}1}$, obtains over all values of $\mathbf{p}$ for which the construction above is well defined, i.e., the equivalence obtains for all $1 > \mathbf{p} \geq 1 - 2^{-(m-1)}$. However, in order to avoid appeal to the following extra assumption, we apply the construction solely to the case where $\mathbf{p} = 1\text{–}2^{-m}$, when the proof of the theorem does not require an extra assumption. The additional assumption needed to apply the construction to the other values of $\mathbf{p}$ is that, if (*i*) X is simple with $\mathbf{Prev}(X) = 0$, and (*ii*) X and Y are independent, then $\mathbf{Prev}(XY) = 0$.

TABLE 4 – The partition of coarse state ω_n into $2^{m-1}+1$ rows and 2 columns, with the values of the $2^{m-1}+1$ equivalent variables displayed within the table.

	t_{n1}	t_{n2}
B_1	$Z_m = (1-p)^{-n}$ $X_1 = (1-p)^{-(n+1)}$ $X_2 = 0$ … $X_{2m-1-1} = 0$ $X_{2m-1} = (1-p)^{-1}$	$Z_m = 0$ $X_1 = 0$ $X_2 = 0$ … $X_{2m-1-1} = 0$ $X_{2m-1} = (1-p)^{-1}$
B_2	$Z_m = (1-p)^{-n}$ $X_1 = (1-p)^{-1}$ $X_2 = (1-p)^{-(n+1)}$ $X_3 = 0$ … $X_{2m-1} = 0$	$Z_m = 0$ $X_1 = (1-p)^{-1}$ $X_2 = 0$ $X_3 = 0$ … $X_{2m-1} = 0$
. . .		
B_i	$Z_m = (1-p)^{-n}$ $X_1 = 0$ … $X_{i-2} = 0$ $X_{i-1} = (1-p)^{-1}$ $X_i = (1-p)^{-(n+1)}$ $X_{i+1} = 0$ … $X_{2m-1} = 0$	$Z_m = 0$ $X_1 = 0$ … $X_{i-2} = 0$ $X_{i-1} = (1-p)^{-1}$ $X_i = 0$ $X_{i+1} = 0$ … $X_{2m-1} = 0$
. . .		
B_{2m-1}	$Z_m = (1-p)^{-n}$ $X_1 = 0$ … $X_{2m-1-2} = 0$ $X_{2m-1-1} = (1-p)^{-1}$ $X_{2m-1} = (1-p)^{-(n+1)}$	$Z_m = 0$ $X_1 = 0$ … $X_{2m-1-2} = 0$ $X_{2m-1-1} = (1-p)^{-1}$ $X_{2m-1} = 0$
B^c	$Z_m = (1-p)^{-n}$ $X_1 = 0$ … $X_{2m-1} = 0$	$Z_m = 0$ $X_1 = 0$ … $X_{2m-1} = 0$

We establish a contradiction with the hypothesis that the prevision for a difference in equivalent random variables is 0 as follows. Consider the random variable obtained by the finite sum

$$W_m = \sum_{i=1}^{2^{m-1}} (X_i - Z_m).$$

Then

$$W_m(B_i \cap t_{n1}) = (1\text{-}\mathbf{p})^{-n}/2 \; + \; (1\text{-}\mathbf{p})^{-1}$$

$$W_m(B_i \cap t_{n2}) = (1\text{-}\mathbf{p})^{-1}$$

$$W_m(B^c \cap t_{n1}) = -2^{m-1}(1\text{-}\mathbf{p})^{-n} = -(1\text{-}\mathbf{p})^{-n}/2$$

$$W_m(B^c \cap t_{n2}) = 0.$$

Note that W_m does not distinguish among the B_i, which we may now collapse into a single row of cells, denoted B, with combined probability ½.

Write W_m as a sum of three random variables, T_m, U_m, and V_m, defined as follows on 4 cells per coarse state $\omega_n = \{B, B^c\} \times \{t_{n1}, t_{n2}\}$

$$T_m(B \cap t_{n1}) = -U_m(B^c \cap t_{n1}) = (1\text{-}\mathbf{p})^{-n}/2$$

$$T_m(B \cap t_{n2}) = T_m(B^c \cap t_{n1}) = T_m(B^c \cap t_{n2}) = 0$$

$$U_m(B \cap t_{n1}) = U_m(B \cap t_{n2}) = U_m(B^c \cap t_{n2}) = 0$$

$$V_m(B \cap t_{n1}) = V_m(B \cap t_{n2}) = (1\text{-}\mathbf{p})^{-1}$$

$$V_m(B^c \cap t_{n1}) = V_m(B^c \cap t_{n2}) = 0.$$

Note that as **Prob**(B) = ½, **Prev**(V_m) = 2^{m-1}. Observe also that T_m and $-U_m$ are equivalent random variables. Then, by the *reductio* hypothesis **Prev**($T_m + U_m$) = 0. Adding these two previsions, we obtain the desired equation (*), which contradicts the hypothesis that the difference in equivalent random variables carries prevision 0, and proving the theorem.

Appendix 3

Theorem 3: Assume that distribution $\boldsymbol{Q}$ agrees with the Geometric($\mathbf{p}$) for all (coarse) states ω_n, $n \geq k$. That is, $\boldsymbol{Q}(\omega_n) = \mathbf{p}(1\text{-}\mathbf{p})^n$, for all $n \geq k$. We define a called-off generalized St. Petersburg gamble, Z_m, and construct a set of 2^{m-1} equivalent called-off random variables $X_1 \approx X_2 \approx \ldots \approx X_{2m-1} \approx Z_m$, such that if previsions are coherent, though infinite, then

$$\sum_{i=1}^{2^{m-1}} \mathbf{Prev}(X_i - Z_m) \;=\; 2^{m-1}. \qquad (*)$$

Proof of Theorem 3. For notational convenience, we suppress the subscripts k and m in what follows. Define Z, the called-off St. Petersburg variable so that:

$$Z(\omega_n) = 0 \text{ for } n < k.$$

$$Z(t_{n1}) = (1\text{-}\mathbf{p})^{-n} \text{ and } Z(t_{n2}) = 0 \text{ for } n \geq k.$$

Next define the random variables X_i so that for $i = 1, \ldots 2^{m-1}\text{-}1$,

$$X_i(\omega_n) = 0 \text{ for all } n < k$$

$$X_i(B_i \cap t_{n1}) = (1\text{-}\mathbf{p})^{-(n+1)} \text{ for } n \geq k$$

$$X_i(B_i \cap t_{n2}) = 0$$

$$X_i(B_{i+1} \cap t_{n1}) = X_i(B_{i+1} \cap t_{n2}) = (1\text{-}\mathbf{p})^{-k}$$

for other states, $(j \neq i, i+1)$ $\quad X_i(B_j \cap t_{n1}) = X_i(B_j \cap t_{n2}) = 0$

and $\quad X_i(B^c \cap t_{n1}) = X_i(B^c \cap t_{n2}) = 0.$

For $X_{2m\text{-}1}$, modify only the fourth line of this definition, as follows:

$$X_{2m\text{-}1}(\omega_n) = 0 \text{ for all } n < k$$

$$X_{2m\text{-}1}(B_{2m\text{-}1} \cap t_{n1}) = (1\text{-}\mathbf{p})^{-(n+1)}$$

$$X_{2m\text{-}1}(B_{2m\text{-}1} \cap t_{n2}) = 0$$

$$X_{2m\text{-}1}(B_1 \cap t_{n1}) = X_{2m\text{-}1}(B_1 \cap t_{n2}) = (1\text{-}\mathbf{p})^{-k}$$

for other states, $(j \neq 2^{m-1}, 1)$ $\quad X_{2m\text{-}1}(B_j \cap t_{n1}) = X_{2m\text{-}1}(B_j \cap t_{n2}) = 0$

and $\quad X_{2m\text{-}1}(B^c \cap t_{n1}) = X_{2m\text{-}1}(B^c \cap t_{n2}) = 0.$

It is obvious that the X_i are pairwise equivalent random variables. In order to verify that each is equivalent to Z, note that $\mathbf{Prob}(X_i = (1\text{-}\mathbf{p})^{-k}) = (1\text{-}\mathbf{p})\mathbf{p}\sum_{i=k\text{-}1}(1\text{-}\mathbf{p})^i = \mathbf{p}\sum_{i=k}(1\text{-}\mathbf{p})^i$ $= 1\text{-}\ \mathbf{Prob}\{\cup_{i \leq k}\ \omega_i) = (1\text{-}\mathbf{p})^k$ as is needed for equivalence with Z. Following the identical reasoning used in sub-section 3.1 we obtain the equation (*), as was to be shown.

References

de Finetti, B. (1974) *Theory of Probability*, Vols. I and II. New York: Wiley.

Dunford, N. and Schwartz, J.T. (1988) *Linear Operators Part 1: General Theory*. New York: Wiley.

Fishburn, P.C. (1979) *Utility Theory for Decision Making*. New York: Krieger.

Fishburn, P.C. (1982) *Foundations of Expected Utility*. Dordrecht: Reidel.

Royden, H.L. (1968) *Real Analysis*. New York: Macmillan .

Savage, L.J. (1954) *The Foundations of Statistics*. New York: Wiley.