11-2015

Non-conglomerability for countably additive measures that are not κ-additive

Teddy Seidenfeld
Carnegie Mellon University, teddy@stat.cmu.edu

Mark J. Schervish
Carnegie Mellon University, mark@stat.cmu.edu

Joseph B. Kadane
Carnegie Mellon University, kadane@stat.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/philosophy

Part of the Philosophy Commons
Non-conglomerability for countably additive measures that are not $\kappa$-additive*
T.Seidenfeld, M.J.Schervish, and J.B.Kadane

Abstract
Let $\kappa$ be an uncountable cardinal. Using the theory of conditional probability associated with de Finetti (1974) and Dubins (1975), subject to several structural assumptions for creating sufficiently many measurable sets, and assuming that $\kappa$ is not a weakly inaccessible cardinal, we show that each probability that is not $\kappa$-additive has conditional probabilities that fail to be conglomerable in a partition of cardinality no greater than $\kappa$. This generalizes our (1984) result, where we established that each finite but not countably additive probability has conditional probabilities that fail to be conglomerable in some countable partition.

Key Words: $\kappa$-additive probability, non-conglomerability, conditional probability, regular conditional probability distribution, weakly inaccessible cardinal.

1. Introduction. Consider a finitely, but not necessarily countably additive probability $P(\cdot)$ defined on a $\sigma$-field of sets $\mathcal{E}$, with sure-event $\varnothing$. That is, $<\Omega, \mathcal{E}, P>$ is a (finitely additive) measure space.
Let $B, C, D, E, F, G \in \mathcal{E}$, with $B \neq \varnothing$ and $F \cap G \neq \varnothing$.

Definition 1. A finitely additive conditional probability function $P(\cdot \mid B)$, satisfies the following three conditions:
(i) $P(C \cup D \mid B) = P(C \mid B) + P(D \mid B)$, whenever $C \cap D = \varnothing$;
(ii) $P(B \mid B) = 1$
Moreover, following de Finetti (1974) and Dubins (1975), in order to regulate conditional probability given a non-empty event of unconditional or conditional probability 0, we require the following.
(iii) $P(E \cap F \mid G) = P(E \mid F \cap G)P(F \mid G)$.
As is usual, we identify the unconditional probability function $P(\cdot)$ with $P(\cdot \mid \Omega)$ and refer to $P(\cdot)$ as a probability function.

This account of conditional probability is not the usual theory from contemporary Mathematical Probability. It differs from the received theory of Kolmogorovian regular conditional distributions in four ways:

1. The theory of regular conditional distributions requires that probabilities and conditional probabilities are countably additive. The de Finetti/Dubins theory requires only that probability and conditional probability is finitely additive. In this paper, we bypass most of this difference by exploring de Finetti/Dubins conditional probabilities associated with countably additive unconditional probabilities. Specifically, we do not require that conditional probabilities are countably additive.

2. When $P(B) = 0$ and $B$ is not empty, a regular conditional probability given $B$ is relative to a sub-$\sigma$-field $\mathcal{H} \subseteq \mathcal{E}$, where $B \in \mathcal{H}$. In the Finetti/Dubins theory of conditional probability, $P(\cdot \mid B)$, depends solely on the event $B$ and not on any

12/2/15
Non-conglomerability for $\kappa$-non-additive, $\sigma$-additive probabilities
sub-\(\sigma\)-field that embeds it. Example 2, which we present in Section 4 after Lemma 3, illustrates this difference.

3. Some countably additive probabilities do not admit regular conditional distributions relative to a particular sub-\(\sigma\)-field, even when both \(\sigma\)-fields are countably generated. (See Corollary 1 in our [2001].) In contrast, Dubins (1975) establishes the existence of full conditional probability functions: where, given a set \(\Omega\) of arbitrary cardinality, a conditional probability satisfying Definition 1 is defined with respect to each non-empty element of its powerset, i.e., where \(\mathcal{B}\) is the powerset of \(\Omega\). Hereafter, we require that each probability function includes its conditional probabilities (in accord with Definition 1) given each non-empty event \(B \in \mathcal{B}\). However, because we investigate conditional probabilities for a countably additive unconditional probability, in light of Ulam’s Theorem [1930], we do not require that \(\mathcal{B}\) is the powerset of the state space \(\Omega\).

4. Our focus in this paper is a fourth feature that distinguishes the de Finetti/Dubins theory of conditional probability and the Kolmogorovian theory of regular conditional probability. This aspect of the difference involves conglomerability of conditional probability functions.

Let \(I\) be an index set and let \(\pi = \{h_i; i \in I\}\) be a partition of the sure event where the conditional probabilities, \(P(E \mid h_i)\) are well defined for each \(E \in \mathcal{B}\) and \(i \in I\).

**Definition 2:** The conditional probabilities \(P(E \mid h_i)\) are conglomerable in \(\pi\) provided that, for each event \(E \in \mathcal{B}\) and real constants \(k_1\) and \(k_2\),

\[
\text{if } k_1 \leq P(E \mid h_i) \leq k_2 \text{ for each } i \in I, \text{ then } k_1 \leq P(E) \leq k_2.
\]

That is, conglomerability requires that the unconditional probability for event \(E\), \(P(E)\) lies within the (closed) interval of conditional probability values, \(\{P(E \mid h_i) | i \in I\}\), with respect to elements \(h\) of a partition \(\pi\).

In our (1984) we show that if \(P\) is merely finitely additive (i.e., if \(P\) is finitely but not countably additive) with conditional probabilities that satisfy Definition 1, and \(P\) is defined on a \(\sigma\)-field of sets, then \(P\) fails conglomerability in some countable partition. That is, for each merely finitely additive probability \(P\) there is an event \(E\), an \(\varepsilon > 0\), and a countable partition of measurable events \(\pi = \{h_n: n = 1, \ldots\}\), where

\[
P(E) > P(E \mid h_n) + \varepsilon \quad \text{for each } h_n \in \pi.
\]

Also, in our (1984) we establish sharp bounds for how large \(\varepsilon\) might be when \(P\) fails conglomerability, what we call the extent of non-conglomerability of \(P\).

**Definition 3:** The extent of non-conglomerability for a probability \(P\) is the supremum over different partitions and events of the values \(\varepsilon\) that satisfy (*)

\[
supremum_{\pi, E \in \mathcal{B}} | P(E) - \supremum_{h \in \pi} \{P(E \mid h)\} |
\]
A probability is *purely finitely additive* when, for each $\delta > 0$ there exists a countable partition $\{h_j: j = 1, \ldots\}$, where $\sum P(h_j) < \delta$. Then, as an illustration of the extent of non-conglomerability, a purely finitely additive $P$ satisfies:

$$\sup_{\pi \in \mathcal{E}} P(E) - \sup_{h \in \pi} \{P(E \mid h)\} = 1.$$ 

In other words, for each $\varepsilon > 0$ there exists an event $E$ and countable partition $\pi = \{h_n: n = 1, \ldots\}$, with $P(E) > 1 - \varepsilon$ and $\varepsilon > P(E \mid h_n)$ for each $h_n \in \pi$. Thus, the failure of conglomerability is as large as possible in this case, since all probabilities are confined to the $[0,1]$ interval.

The following example illustrates a failure of conglomerability for a merely finitely additive probability $P$ in a countable partition $\pi = \{h_n: n \in \{1, 2, \ldots\}\}$, where each element of the partition is not null, i.e., $P(h_n) > 0$ for each $n \in \{1, 2, \ldots\}$. Then, by both the theory of conditional probability according to Definition 1 and the theory of regular conditional distributions (ignoring the requirement that probability is countably additive), $P(E \mid h_n) = P(E \cap h_n)/P(h_n)$ is well defined. Thus, the failure of conglomerability in this example is due to the failure of countable additivity, rather than to a difference in how conditional probability is defined.

**Example 1** (Dubins, 1975): Let the sure event $\Omega = \{(i, n): i \in \{1, 2\} \text{ and } n \in \{1, 2, \ldots\}\}$ and $\mathcal{E}$ be the powerset of $\Omega$. Let $E = \{(1, n): n \in \{1, 2, \ldots\}\}$ and $h_n = \{(1,n), \{2,n\}$, and partition $\pi = \{h_n: n \in \{1, 2, \ldots\}\}$. Partially define the finitely additive probability $P$ by:

- (i) $P(i,n) = 1/2^{n+1}$ if $i = 1$, and $P(i,n) = 0$ if $i = 2$,
- (ii) $P(E) = 0.5$.

So $P$ is merely finitely additive over $E^c$ and $P(\cdot \mid E^c)$ is purely finitely additive. It follows easily that $P(h_n) = 1/2^{n+1} > 0$ for each $n \in \{1, 2, \ldots\}$. Thus, $P$ is not conglomerable in $\pi$ as: $P(E^c \mid h_n) = P(E^c \cap h_n)/P(h_n) = 0$, for each $n \in \{1, 2, \ldots\}$, whereas $P(E^c) = 0.5$. 

We note that Example 1 also illustrates the *extent of non-conglomerability* for each finitely additive probability that satisfies the two conditions (i) and (ii). That extent is $\frac{1}{2}$. Also, in our [1996], we discuss this example in connection with the value of information.

In the appendix to our (1986) we show that for a continuous, countably additive probability defined on the continuum, and assuming conditional probabilities that satisfy Definition 1 rather than being regular conditional distributions, then non-conglomerability results by considering a continuum-many set of different partitions of the continuum. These alternative partitions are generated by sets of equivalent (non-linearly transformed) measurable random variables. Conglomerability cannot be satisfied in all the partitions. Here we generalize that result to $\kappa$-non-additive probabilities that are countably additive.

Let $<\Omega, \mathcal{E}, P>$ be a measure space, with $P$ countably additive. That is, $\mathcal{E}$ is a $\sigma$-field of sets over $\Omega$. Set $B$ is *measurable* means that $B \in \mathcal{E}$. That $P$ is a countably additive probability is formulated with either of two equivalent, familiar definitions. That
these are equivalent definitions is immediate from the requirement that $\mathcal{E}$ is a $\sigma$–field of sets. (See, e.g., Billingsley, 1995, p. 25.)

**Definition 4a:** Let $\{A_i; i = 1, \ldots\}$ be a sequence of measurable, disjoint events, and let $A$ be their union. That is, $A_i \cap A_j = \emptyset$ if $i \neq j$, and $A = \bigcup_i A_i$. Then $P$ is **countably additive** (in the first sense) provided that $P(A) = \sum P(A_i)$.

**Definition 4b:** Let $\{B_i; i = 1, \ldots\}$ be an increasing sequence of measurable events, with $B$ their limit. That is, $B_i \subseteq B_j$ if $i \leq j$, and $B = \bigcup_i B_i$. Then $P$ is **countably additive** (in the second sense) provided that $P(B) = \lim P(B_i)$.

That is, $P$ is countably additive provided it is continuous over sequences that approximate events from below.

In this paper we examine non-conglomerability of a set of conditional probabilities $\{P(E \mid h)\}$ that satisfy (the de Finetti/Dubins) Definition 1, where these conditional probabilities are associated with a countably additive unconditional probability, $P$, that belongs to a measure space $<\Omega, \mathcal{E}, P>$. How large do we require the $\sigma$–field of sets $\mathcal{E}$ be in order to have available sufficiently many well defined conditional probabilities? By an important result of Ulam (1930), unless the cardinality of $\Omega$ is at least as great as some inaccessible cardinal, $\mathcal{E}$ cannot be the powerset of $\Omega$. However, without loss of generality, we may assume that the measure space is complete. That is, if $N \in \mathcal{E}$, $P(N) = 0$, and $E \subseteq N$, then $E \in \mathcal{E}$. See, e.g., Billingsley (1995, p. 44), Doob (1994, p. 37), or Halmos (1950, p. 55).

Our principal result here asserts that, subject to several structural assumptions to assure richness of $\mathcal{E}$ as explained in Section 3.1, the non-conglomerability of $P$ occurs in a partition by measurable events whose cardinality $\kappa$ is bounded above by the extent of non-additivity of the countably additive probability $P$.

There are two, parallel definitions for generalizing from countable additivity (also denoted $\sigma$-additivity) to $\kappa$-additivity. In the following, let $\alpha$, $\beta$, and $\gamma$ be ordinals and $\kappa$ a cardinal.

**Definition 5a:** Let $\{A_\alpha; \alpha < \gamma \leq \kappa\}$ be a $\gamma$-sequence of measurable, disjoint events, and let $A$ be their union, also presumed measurable. That is, $A_\alpha \cap A_\beta = \emptyset$ if $\alpha \neq \beta$ with $A = \bigcup_{\alpha < \gamma} A_\alpha$.

$P$ is **$\kappa$-additive** if $P(A) = \sum_{\alpha < \gamma} P(A_\alpha)$ for each such $\gamma$-sequence.

**Definition 5b** (Armstrong and Prikry, 1980): Let $\{B_\alpha; \alpha < \gamma \leq \kappa\}$ be an increasing $\gamma$-sequence of measurable events, where $B_\alpha \subseteq B_\beta$ whenever $\alpha \leq \beta < \gamma$ with $B = \bigcup_{\alpha < \gamma} B_\alpha$ also measurable.

$P$ is **$\kappa$-additive** if $P(B) = \sup_{\alpha < \gamma} P(B_\alpha)$ for each such $\gamma$-sequence.

That is, $P$ is $\kappa$-additive provided that probability is continuous from below over $\gamma$-long sequences that approach measurable events from below.
Next, with Lemma 1, we show that for a complete measure space, these two are equivalent concepts of \( \kappa \)-additivity. That is, a complete measure space is \( \kappa \)-additive\(_1\) if and only if it is \( \kappa \)-additive\(_2\). Then in the light of Lemma 1, if for each cardinal \( \kappa \), \( P \) is \( \kappa \)-additive, call \( P \) perfectly additive.

**Lemma 1:** Let \( \langle \Omega, \mathcal{S}, P \rangle \) be a complete (atomic) measure space with \( |\Omega| = \kappa \), where each point \( \omega \in \Omega \) is measurable. \( P \) is \( \kappa \)-additive\(_1\) if and only if it is \( \kappa \)-additive\(_2\).

**Proof:**

[\( \kappa \)-additive\(_1\) only if \( \kappa \)-additive\(_2\).] Let \( \{B_\alpha : \alpha < \gamma \le \kappa\} \) be an upward nested \( \gamma \)-length sequence of measurable events with measurable limit \( B = \bigcup_{\alpha<\gamma} B_\alpha \). Without loss of generality, let \( \gamma = \kappa \), as the sequence can be extended from \( \gamma \) to \( \kappa \) by repeating events arbitrarily many times.

Let \( \lambda = \text{cofinality of} \ \kappa \). Thus, let \( \{B_\beta : \beta < \lambda\} \) be a cofinal subsequence of the original sequence \( \{B_\alpha : \alpha < \kappa\} \). So, \( B = \bigcup_{\alpha<\kappa} B_\alpha = \bigcup_{\beta<\lambda} B_\beta \), where \( \sup_{\alpha<\kappa} P(B_\alpha) = \sup_{\beta<\lambda} P(B_\beta) \). We establish the lemma by showing that if \( P \) is \( \kappa \)-additive\(_1\), then \( P(B) = \sup_{\beta<\lambda} P(B_\beta) \).

Partition \( B \) into \( \lambda \)-many disjoint sets, \( \{C_\beta : 0 < \beta < \lambda\} \) using the increasing \( \lambda \)-sequence, \( \{B_\alpha : \beta < \lambda\} \), as follows.

Let \( C_0 = B_{\alpha_0} \).

- If \( \beta = \gamma+1 \) is a successor ordinal, \( C_{\gamma+1} = B_{\alpha_{\gamma+1}} - B_\gamma \).
- If \( \beta \) is a limit ordinal, \( C_\beta = B_\beta - \bigcup_{\gamma<\beta} B_\gamma \).

When cofinality(\( \kappa \)) = \( \lambda = \aleph_0 \), we are done by the familiar result equating the two senses of countable additivity. So, let \( \lambda \) be an uncountable regular cardinal – regular, since \( \lambda \) is a cofinality.

Since the sets \( B_{\alpha_\beta} \) are nested upwards, by finite additivity of \( P \), \( P \) strictly increases over this sequence at most countably many times. That is, consider the countable (and possibly empty), subsequence \( \{B_{\alpha_i} : i = 1, ...\} \), where \( P(B_{\alpha_{i+1}}) - P(B_{\alpha_i}) > 0 \). Because \( \lambda \) is regular and uncountable, this countable subsequence is not cofinal in \( \lambda \).

Let \( \delta \) be the least limit ordinal beyond this countable subsequence such that for each ordinal \( \eta \), \( \delta \le \eta < \lambda \), there is a real number \( c = P(B_{\alpha_{\eta}}) = \sup_{\beta<\lambda} P(B_{\alpha_{\beta}}) \). That is, \( P \) is constant on this \( \lambda \)-long tail of the \( B_{\alpha_\beta} \)-subsequence of the original \( B_\alpha \) sequence of events. So, for each pair of ordinals, \( (\mu, \nu) \), \( \delta \le \mu < \nu < \lambda \), \( P(B_{\alpha_{\nu}}) - P(B_{\alpha_{\mu}}) = 0 \).

Let \( D = B - B_{\alpha_\delta} \). So, \( D \) is measurable as each of \( B \) and \( B_{\alpha_\delta} \) is. Observe that \( D = \{\bigcup_{\eta} : \delta < \eta < \lambda\} \). We argue that \( P(D) = 0 \) by showing that for each \( \eta \), with \( \delta < \eta < \lambda \), \( C_\eta \) is...
measurable and $P(C_\eta) = 0$. Then, apply $\kappa$-additivity to this sequence.

For successor ordinals $\eta = \beta + 1$, we have that $C_\eta$ is measurable, since $\mathcal{E}$ is a $\sigma$-field. Then $P(C_\eta) = 0$, since $\eta > \delta$. That is, $P(B\alpha_{\beta + 1}) = P(B\alpha_\beta) = c$. For limit ordinals $\eta < \lambda$, note that $B\alpha_\delta \subseteq \cup_{\delta < \beta < \eta} B\alpha_\beta \subseteq B\alpha_\eta$ and $P(B\alpha_\eta) - P(B\alpha_\delta) = 0$. Thus, as $P$ is complete, since $\cup_{\delta < \beta < \eta} B\alpha_\beta$ differs from a measurable set by a set of measure 0, it too is measurable, from which we have $P(C_\eta) = 0$. Because $P$ is $\kappa$-additive, then $0 = \Sigma_{\delta < \eta < \lambda} P(C_\eta) = P(D)$. Then $P(B) = P(B\alpha_\beta) + P(D) = P(B\alpha_\delta) \leq \sup_{\beta < \kappa} P(B\alpha_\beta) = \sup_{\alpha < \kappa} P(B_\alpha)$, which establishes that $P$ is $\kappa$-additive.

$[\kappa$-additive only if $\kappa$-additive$]$. Let $\{A_\alpha; \alpha < \kappa\}$ be a $\kappa$-sequence of measurable, disjoint events, and let $A$ be their union, also presumed measurable. Since $P(A_\eta) > 0$ for at most countably many ordinals $\alpha$, and as each countable union of these countably many events is measurable (as $\mathcal{E}$ is a $\sigma$-field), without loss of generality assume that for each ordinal $\alpha < \kappa$, $P(A_\alpha) = 0$. That is, let $C = \cup_i A_{\alpha i}$ where $P(A_{\alpha i}) > 0$.

If $P(C) = 1$ we are done. So, assuming that $P(C) < 1$, we consider instead an $A_\alpha$-sequence of (measurable) null events defined with respect to the complete quotient-measure space $<\Omega\cdot C, \mathcal{E}/C, P(\cdot | C)>$.

Define the upward nested sequence $\{B_\alpha; \alpha < \kappa\}$ as follows.

$B_0 = A_0$

If $\alpha = \beta + 1$ is a successor ordinal, $B_\alpha = B_\beta \cup \{A_{\beta + 1}\}$

If $\alpha$ is a limit ordinal, $B_\alpha = \cup_{\beta < \alpha} A_\beta$.

Note that $A = \cup_{\alpha < \kappa} B_\alpha = \cup_{\alpha < \kappa} A_\alpha$. Thus, if each of the $\{B_\alpha; \alpha < \kappa\}$ is measurable, then since $P$ is hypothesized to be $\kappa$-additive, $0 = \sup_{\alpha < \kappa} P(B_\alpha) = P(A) = \Sigma_{\alpha < \kappa} P(A_\alpha)$ as required by $\kappa$-additivity. All that remains is to show that these $B_\alpha$ are measurable events. We argue by induction on $\kappa$. For $\kappa = \aleph_0$ we are done by the familiar result equating the two senses of $\sigma$-additivity. The argument for $\kappa = \aleph_1$ also is direct and worth stating: Each ordinal $\alpha < \aleph_1$ is countable. Hence, $|B_\alpha| \leq \aleph_0$, and as $\mathcal{E}$ is a $\sigma$-field, then each such $B_\alpha$ is measurable. For the argument by induction, similar reasoning applies. By hypothesis of induction, assume Lemma 1 holds for all cardinals $\lambda$ less than $\kappa$, and in particular then assume that each set $C$ with $|C| < \kappa$ is measurable whenever $P$ is $\lambda$-additive for each $\lambda < \kappa$. Then, for each $\alpha < \kappa$, $B_\alpha$ is measurable, with $P(B_\alpha) = 0$, since $|B_\alpha| < \kappa$.

Lemma 1
In the light of Lemma 1, in what follows we use $\kappa$-additivity to generalize countable additivity to larger cardinals.

Consider a complete (atomic) measure space $<\Omega, \mathcal{F}, P>$, where each point $\omega \in \Omega$, and where $P$ is countably additive but not $\kappa$-additive. Here we show the main Proposition of this paper:

- Subject to several structural assumptions on $\mathcal{F}$ (explained in Section 3.1) the probability $P$ fails to be conglomerable in some partition $\pi$ of measurable events, where the cardinality of $\pi$ at most $\kappa$.

Thus, rather than thinking that non-conglomerability is an anomalous feature of finite but not countably additive probabilities, and that non-conglomerability arises solely with finitely but not countably additive probabilities in countable partitions, here we argue for a different conclusion. Namely, we show that the (least) cardinal $\kappa$ of a partition where $P$ is non-conglomerable is the (least) cardinal for which $P$ is not $\kappa$-additive.

2. **Tiers of points.** The proof of the main Proposition is based on the structure of a linear order over equivalence classes (which we call tiers) of points in $\Omega$ defined by the following relation between pairs of points.

**Definition 6:** Consider the relation, $\sim$, of relative-non-nullity on pairs of points in $\Omega$. That is, for points, $\omega_\alpha$ and $\omega_\beta$, they bear the relation $\omega_\alpha \sim \omega_\beta$ provided that, either $\omega_\alpha = \omega_\beta$ or else $\omega_\alpha \neq \omega_\beta$ and $0 < P(\{\omega_\alpha\}|\{\omega_\alpha, \omega_\beta\}) < 1$.

**Lemma 2:** $\sim$ is an equivalence relation.

**Proof:** Only transitivity requires verification. Assume $\omega_1 \sim \omega_2 \sim \omega_3$. That is, assume $0 < P(\{\omega_1\} | \{\omega_1, \omega_2\}), P(\{\omega_2\} | \{\omega_2, \omega_3\}) < 1$. Then by condition (iii) of Definition 1 of coherent conditional probabilities:

$P(\{\omega_1\} | \{\omega_1, \omega_2, \omega_3\}) = P(\{\omega_1\} | \{\omega_1, \omega_2\}) P(\{\omega_1, \omega_2\} | \{\omega_1, \omega_2, \omega_3\})$. Similarly,

$P(\{\omega_3\} | \{\omega_1, \omega_2, \omega_3\}) = P(\{\omega_3\} | \{\omega_2, \omega_3\}) P(\{\omega_2, \omega_3\} | \{\omega_1, \omega_2, \omega_3\})$.

Now argue indirectly by cases.

- If $P(\{\omega_1\} | \{\omega_1, \omega_3\}) = 0$, then $P(\{\omega_1\} | \{\omega_1, \omega_2, \omega_3\}) = 0$ and $P(\{\omega_1, \omega_2\} | \{\omega_1, \omega_2, \omega_3\}) = 0$, since by assumption $P(\{\omega_1\} | \{\omega_1, \omega_2\}) > 0$. Then $P(\{\omega_2\} | \{\omega_1, \omega_2, \omega_3\}) = 0 = P(\{\omega_2\} | \{\omega_2, \omega_3\})$, which contradicts $\omega_2 \sim \omega_3$.

- If $P(\{\omega_1\} | \{\omega_1, \omega_3\}) = 1$, then $0 = P(\{\omega_3\} | \{\omega_1, \omega_3\}) = P(\{\omega_3\} | \{\omega_1, \omega_2, \omega_3\})$. Then $0 = P(\{\omega_2, \omega_3\} | \{\omega_1, \omega_2, \omega_3\})$, since $0 < P(\{\omega_3\} | \{\omega_2, \omega_3\})$.

So, $0 = P(\{\omega_2\} | \{\omega_1, \omega_2, \omega_3\}) = P(\{\omega_2\} | \{\omega_1, \omega_2\})$, which contradicts $\omega_1 \sim \omega_2$.

Hence $0 < P(\{\omega_1\} | \{\omega_1, \omega_3\}) < 1$, as required.

Lemma 2

The equivalence relation $\sim$ partitions $\Omega$ into disjoint tiers $\tau$ of relative non-null pairs of points. Evidently, if $P(\{\omega_2\} | \{\omega_1, \omega_2\}) = P(\{\omega_3\} | \{\omega_2, \omega_3\}) = 1$, then $P(\{\omega_3\} | \{\omega_1, \omega_3\}) = 1$. Thus, the tiers are linearly ordered by the relation $\uparrow$, defined as follows:

**Definition 7a:** $\tau_1 \uparrow \tau_2$ if for each pair $\{\omega_1, \omega_2\}, \omega_i \in \tau_i (i = 1, 2), P(\{\omega_2\} | \{\omega_1, \omega_2\}) = 1$. 

12/2/15

Non-conglomerability for $\kappa$–non-additive, $\sigma$–additive probabilities
Since the reverse ordering also is linear, we express this as:

**Definition 7b**: \( \tau_2 \downarrow \tau_1 \) if for each pair \( \{\omega_1, \omega_2\}, \omega_i \in \tau_i \ (i = 1, 2) \), \( P(\{\omega_2\} | \{\omega_1, \omega_2\}) = 1 \), i.e., if and only if \( \tau_1 \uparrow \tau_2 \).

There is a “top” tier of non-null points in this linear ordering, which we label \( \hat{\tau} \).

**Definition 8**: Let \( \hat{\tau} = \{\omega : P(\omega) > 0\} \).

Since \( |\hat{\tau}| \leq \aleph_0 \), as \( \mathcal{E} \) is a \( \sigma \)-field, \( \hat{\tau} \) is measurable. Moreover, for each \( \tau \neq \hat{\tau} \), \( \hat{\tau} \downarrow \tau \).

Evidently, this idea of tiers can be used to create an inductive hierarchy. Let the tiers of points be the type1 elements of this hierarchy. For tier2, in parallel with Definition 7, define an equivalence class among type1 tiers, \( \approx \), as follows:

**Definition 8**: \( \tau_\alpha \approx \tau_\beta \) provided that, either \( \tau_\alpha = \tau_\beta \) or else

\[ \tau_\alpha \neq \tau_\beta \text{ and } 0 < P(\{\tau_\alpha\} | \{\tau_\omega, \tau_\beta\}) < 1. \]

Then, just as above, this produces a type2 linear ordering over sets of type1 tiers. However, in this paper, our reasoning about non-conglomerability relies solely on the type1 linear order of tiers of points, without requiring measurability of the sets from the higher types of tiers.

### 3. The Main Proposition and its Proof.

#### 3.1 Structural assumptions for the Proposition.

The **Proposition** asserts that, subject to the assumptions explained in this subsection, when \( P \) is not-\( \kappa \)-additive, then non-conglomerability obtains in some partition whose cardinality is bounded above by the same cardinal, \( \kappa \). The **Proposition** is established with five lemmas: Lemmas 3-7. In Section 3.2, we explain how the proof is organized using these five lemmas.

We make assumptions about the measure space \( \langle \Omega, \mathcal{E}, P \rangle \). Regarding the cardinality \( \kappa \) of \( P \)'s non-additivity, we assume that \( \kappa \) is not a weakly inaccessible cardinal. This is to insure that the set of cardinals less than \( \kappa \) has cardinality less than \( \kappa \), which condition we use in the proof of Lemma 6.

Next, we state six structural assumptions that we impose on \( \mathcal{E} \) in order to secure sufficiently many measurable events for proving the central proposition.

**Definition 9**: When \( T \) is a set of tiers, denote by \( \bigcup T \) the subset of \( \Omega \) formed by the union of the elements in \( T \), the union of the tiers in \( T \).

**Structural Assumptions**:

- **SA1**: Each point, \( \omega \in \Omega \), is measurable. (Used with each of the five Lemmas.)
- **SA2**: Each tier, \( \tau \), is measurable. (Used with each of the five Lemmas.)
- **SA3**: Intervals of tiers form measurable sets. For each tier \( \tau' \), \( \bigcup \{\tau : \tau \downarrow \tau'\} \) and \( \bigcup \{\tau : \tau \uparrow \tau'\} \in \mathcal{E} \). In this sense, “Dedekind cuts” in the linear order of tiers create measurable sets. (Used in proving Lemmas 5, 6, and 7.)
SA4: Splitting non-null tiers. If $P(\tau) > 0$, there exist disjoint, measureable events $S_1 \cap S_2 = \emptyset, S_1 \cup S_2 = \tau$, where $|\tau| = |S_1| = |S_2|$. (Used with Lemma 3.)

SA5: Splitting a (non-null) linear order of tiers when the linear order is a well order.

Suppose that $T$ is a measurable set, with $|T| = \kappa$ and $\downarrow$ or $\uparrow$ is a well-order of the set of tiers in $T$, indexed by the ordinals less than or equal to $\kappa$.

Then the union of points in each of the following two “successor” sets of tiers is measurable. (Used with Lemmas 5, 6, and 7.)

(Recall that if $P(T) = 0$, since $P$ is complete, each subset of $T$ is measurable.)

$T_{odd}$ is the set of tiers with “odd” ordinal index, ending “2n–1” for a positive integer $n > 0$. Then $\cup T_{odd}$ is measurable.

$T_{even}$ is the set of tiers with “even” ordinal index, ending “2n” for a positive integer $n > 0$. Then $\cup T_{even}$ is measurable.

Moreover, when $P(T) > 0$, the two “successor” sets are not both null:

$P(\cup T_{odd} \cup \cup T_{even}) > 0$.

SA6: Splitting the linear order of tiers by cardinality. For each cardinal $\lambda \leq \kappa$,

define the set $h_\lambda = \{\tau; |\tau| = \lambda\}$. Then $\cup h_\lambda$ is measurable. (Used with Lemma 6.)

It is immediate from SA5 that when $\downarrow$ or $\uparrow$ is a well-order of the set of tiers in $T$ then the set of points in tiers of $T$ with limit ordinal index, $\cup T_{limit}$ also is measurable – since $\{T_{odd}, T_{even}, T_{limit}\}$ forms a partition of $T$.

3.2 The Proposition and its Proof.

**Proposition:** Let $<\Omega, \mathcal{F}, P>$ be an (atomic) complete, countably additive measure space with conditional probabilities satisfying Definition 1 and which satisfies the six Structural Assumptions of Section 3.1. Assume that $P$ is not perfectly additive and fails to be $\kappa$-additive for a cardinal $\kappa$ that is not weakly inaccessible. Then, there is a partition $\pi = \{h_i; i \in I\}$ of measurable events, where $|\pi| \leq \kappa$ and where $P$ fails to be conglomerable in $\pi$. That is, there exists a measurable event $E$, and an $\varepsilon > 0$ where:

$$P(E) > P(E \mid h_j) + \varepsilon$$

for each $h_j \in \pi$. ∗ Proposition

The proof of the Proposition proceeds in sequence through five lemmas: Lemmas 3-7. Lemmas 3 and 4 address, respectively, one of two non-exclusive, non-exhaustive, Sufficient Conditions for non-conglomerability of $P$. That is, it is consistent that neither (or both) of these two Sufficient Conditions obtains for the linear order of tiers.

**Sufficient Condition 1:** There is a tier $\tau$ below the top ($\tau \neq \tau$) that is not null, $P(\tau) > 0$.

Lemma 3 establishes that then $P$ is non-conglomerable.

**Sufficient Condition 2:** There exist two sets of tiers, $U$ and $V$, with $P(UV) > 0$ and $|U \cup V| = |U \lor V|$, but where $U$ is above $V$ in the linear ordering of tiers. That is, for each tier $\tau_1$ in $U$ and each tier $\tau_2$ in $V$, $\tau_1 \uparrow \tau_2$: Lemma 4 establishes then $P$ is non-conglomerable.
Lemmas 5-7 address, respectively, one of the following three exclusive and mutually exhaustive Cases for the linear order of tiers.

Case 1: The linear order ↑ is a well order on the set of tiers. Lemma 5 establishes that P is non-conglomerable in this case.

Case 2: The linear order ↓ is a well order on the set of tiers. Lemma 6 establishes that P is non-conglomerable in this case.

Case 3: There are two countable subsets \( L_1 = \{ \tau_1, \ldots, \tau_{n^*} \} \) and \( M_1 = \{ \tau'_1, \ldots, \tau'_{n^*} \} \) of the set of tiers, each well ordered as the natural number \( (N <) \), respectively, by ↓ and ↑. Lemma 7 establishes that P is non-conglomerable in this case.

The proofs of Lemmas 5, 6, and 7 rely on the two facts established by Lemmas 3 and 4 that, if either Sufficient Condition obtains within one of the three Cases, then P is non-conglomerable. The main Proposition is established by these five lemmas since the three Cases are jointly exhaustive.

Proof of the Main Proposition:

From the assumption that P is not perfectly additive \( (P(\tau) < 1) \), let \( \kappa \) be the least cardinal for which P is not \( \kappa \)-additive. \( \kappa \) is a regular cardinal since, otherwise, it would fail to be \( \lambda \)-additive on a subsequence of upward-converging sets, where \( \lambda < \kappa \) and \( \lambda = \text{cofinality}(\kappa) \). Without loss of generality, by considering the conditional probability \( P(\cdot | \tau^c) \) as the unconditional probability for the Proposition defined on the measureable space \( <\Omega-\tau^c, \mathcal{E}/\tau^c> \), we proceed assuming that \( P(\tau) = 0 \).

Lemma 3: Suppose there exists a non-null tier (of null points), \( P(\tau) > 0 \) – Situation 1 – then P is not conglomerable.

Proof: Since \( P(\omega) = 0 \) for each \( \omega \in \Omega \), and P is \( \lambda \)-additive for each cardinal \( \lambda < \kappa \), \( |\tau| = \kappa \). By the splitting condition, SA4, partition \( \tau \) into two disjoint measurable sets, \( T_0 \cap T_1 = \emptyset \) with \( T_0 \cup T_1 = \tau \); each with cardinality \( \kappa \), \( |T_0| = |T_1| = \kappa \). Label them so that \( P(T_0) = P(T_1) = d > 0 \).

We identify a partition of cardinality \( \kappa \), which we write as \( \pi = \{ h_\alpha : \alpha < \kappa \} \), where \( P(T_1 | h) < d/2 \) for each \( h \in \pi \). Each element \( h \in \pi \) is a finite set. Each element \( h_\alpha \) contains exactly one point from \( T_1 \), and some positive finite number of points from \( T_0 \), selected to insure that \( P(T_1 | h) < d/2 \).

By the Axiom of Choice, consider a \( \kappa \)-long well ordering of \( T_1 \), \( \{ \omega^1_\beta : \beta < \kappa \} \). We define \( \pi \) by induction. Consider the countable partition of \( T_0 \) into sets:

\[
\rho_{1,n} = \{ \omega \in T_0: (n-1)/n \leq P(\{ \omega^1_1 \} | \{ \omega^1_1, \omega \}) < n/(n+1) \}, \quad \text{for } n = 1, 2, \ldots .
\]

Observe that \( \cup_n \rho_{1,n} = T_0 \). Since \( |T_0| = \kappa \geq \aleph_1 \), by the pigeon-hole principle consider the least \( n^* \) such that \( \rho_{1,n^*} \) is infinite. Let \( U_1 = \{ \omega^1_{1,1}, \ldots, \omega^1_{1,m} \} \) be \( m \)-many points chosen from \( \rho_{1,n^*} \). Note that \( P(\{ \omega^1_1 \} | U_1 \cup \{ \omega^1_1 \}) \leq n^*/(m+n^*) \). Choose \( m \) sufficiently large so that \( n^*/(m+n^*) < d/2 \). Let \( h_1 = U_1 \cup \{ \omega^1_1 \} \). Since \( h_1 \) is a finite set, it
is measurable.

For ordinals $1 < \beta < \kappa$, define $h_\beta$ by induction, as follows. Denoting $T_{0,1} = T_0$, let $T_{0,\beta} = T_0 - (\cup_{0<\alpha<\beta} h_\alpha)$. Since, for each $0 < \alpha < \beta$, by hypothesis of induction $h_\alpha$ is a finite set, then $|\cup_{0<\alpha<\beta} h_\alpha| < \kappa$. So, $|T_{0,\beta}| = \kappa$. Since $T_{0,\beta}$ is a subset of $\tau$, just as above, consider the countable partition of $T_{0,\beta}$ into sets

$$\rho_{\beta,n} = \{\omega \in T_{0,\beta}: (n-1)/n \leq P(\{\omega_1^1\} |\{\omega_1^1, \omega\}) < n/(n+1)\},$$

for $n = 1, 2, \ldots$. Again, by the pigeon-hole principle, consider the least integer $n^*$ such that $\rho_{\beta,n^*}$ is infinite. Let $U_\beta = \{\omega_1^0, \ldots, \omega_1^m\}$ be $m$-many points chosen from $\rho_{\beta,n^*}$. Just as above, $P(\{\omega_1^1\}|U_\beta \cup \{\omega_1^1\}) \leq n^*/(m+n^*)$. Choose $m$ sufficiently large that $n^*/(m+n^*) < d/2$. Let $h_\beta = U_\beta \cup \{\omega_1^1\}$, which also is finite, hence measurable.

Observe that $T_1 \subset \cup_{0<\beta<\kappa} h_\beta$ and that for each $0 < \beta < \kappa$, $P(T_1 | h_\beta) < d/2$. In order to complete the partition $\pi$, consider a catch-all set $S$ with all the remaining points $\omega_1^\beta \in \Omega - \cup_{0<\beta<\kappa} h_\beta$. Note that each point $\omega \in S$ is not a member of $T_1$. So, for each $\omega \in S$, $P(T_1 | \{\omega\}) = 0$. So, for each point, $\omega \in S$, add $\{\omega\}$ as a separate partition element of $\pi$. Hence, $P$ is not conglomerable in $\pi$ as $P(T_1) = d > 0$, yet for each $h \in \pi$, $P(T_1 | h) < d/2$. Lemma 3

Observe that, with the same construction as in Lemma 3, given the real number $r > 0$, we can create a $\kappa$-size partition $\pi^*$ so that for each $h \in \pi^*$, $P(T_1 | h) < d/r$ rather than bounding the conditional probabilities $P(T_1 | h)$ above at $d/2$. Hence, the extent of non-conglomerability for $P$ is bounded below by the quantity $m = \text{Maximum}_{\pi \in \pi^*} P(\tau)$. In Section 5, with Example 3, we illustrate the first Sufficient Condition and the argument of Lemma 3 using an ordinary continuous random variable. We use Example 3 to explain a difference between the de Finetti/Dubins’ theory of conditional probability (Definition 1), and the familiar theory of regular conditional distributions.

Next, addressing Sufficient Condition 2, Lemma 4 establishes that then $P$ is non-conglomerable in some $\kappa$-sized partition of measurable events. We use Lemma 4 frequently in the arguments that follow for Lemmas 5, 6, and 7.

Lemma 4: Let each of $U$ and $V$ be two disjoint sets of tiers, with $\cup V$ a measurable set. Assume $|U| = \kappa$, and with $U$ entirely above $V$ in the linear ordering of $\tau$ tiers. That is, for each pair $\tau_\omega \subseteq U$ and $\tau_\nu \subseteq V$, $\tau_\omega \nsubseteq \tau_\nu$. Then if $P$ is conglomerable, $P(V) = 0$.

Proof: This is a straightforward cardinality argument. Assume $|V| = \kappa$, otherwise $P(V) = 0$ since $P$ is $\lambda$-additive for cardinals $\lambda < \kappa$. Because $\tau_\omega \nsubseteq \tau_\nu$, for each two points $\omega_1 \in \tau_\omega \subseteq U$ and $\omega_\nu \in \tau_\nu \subseteq V$, $P(\{\omega_1\} | \{\omega_\nu\}) = 0$. Consider a 1-1 function to pair elements of $U \cup V$ and elements of $U \cup U$. Let these pair-sets be the
elements of a \( \kappa \)-size partition, \( \pi = \{ h_\alpha : 0 < \alpha < \kappa \} \) augmented by the singleton point sets, \( \{ \{ \omega \} : \omega \in \Omega - [(\cup U)(\cup V)] \} \), if this set is not empty. Then, for each \( h \in \pi \), \( P(\cup V | h) = 0 \). If \( P \) is conglomerable, then \( P(\cup V) = 0 \). \( \text{Lemma 4} \)

Note that under the premises of \( \text{Lemma 4} \), the extent of non-conglomerability of \( P \) is bounded below by the quantity \( P(\cup V) \). Also note that in the argument above, it is possible that \( \cup U \) is not measurable with respect to \( \mathcal{E} \).

Consider the linear orders \( \uparrow \) and \( \downarrow \) over the set of tiers, as defined in Section 2. Either \( \uparrow \) or (exclusively) \( \downarrow \) is a well order of the set of tiers, or (exclusively) there are two countable subsets \( L_\downarrow = \{ \tau_1, ..., \tau_n, ... \} \) and \( M_\uparrow = \{ \tau'_1, ..., \tau'_n, ... \} \) of the set of tiers, each well ordered as the natural number (\( N < \)), respectively, by \( \downarrow \) and \( \uparrow \). That is, then elements of \( L_\downarrow \) satisfy \( \tau_m \downarrow \tau_n \) and elements of \( M_\uparrow \) satisfy \( \tau'_m \uparrow \tau'_n \) whenever \( n > m \). These three \( \text{Cases} \) are addressed in Lemmas 5, 6, and 7, respectively.

\( \text{Lemma 5:} \) Suppose each tier in the linear order \( \uparrow \) is null and that \( \uparrow \) is a well order – \( \text{Case 1} \). Then \( P \) is not conglomerable.

\( \text{Proof:} \) We index the well order \( \uparrow \) with an initial segment of the ordinals. Let \( \beta \) be the least ordinal in the well order such that \( P(\cup_{\alpha < \beta} \tau_\alpha) > 0 \) and let \( R \) be this set of tiers. \( R = \{ \tau_\alpha : \alpha < \beta \} \). By \( \text{SA3,} \cup R \) is measurable. Evidently, we may assume that \( \beta \) is a limit ordinal with \( |\beta| = \kappa \), since \( P(\tau_\alpha) = 0 \) for each tier, and \( P \) is \( \lambda \)-additive, \( \lambda < \kappa \).

Use \( \text{SA5} \) to partition \( \mathcal{R} \) into two disjoint sets of tiers, \( T_1 \) and \( T_2 \), each with cardinality (of numbers of tiers) \( \kappa \). For example, \( T_1 \) might be the set of tiers with successor ordinal index – the union of \( T_{\text{odd}} \) and \( T_{\text{even}} \). And \( T_2 \) might be the set of tiers with limit ordinal index. Then each of \( T_1 \) and \( T_2 \) is cofinal in the well order, \( \uparrow \), of \( R \). It is then an elementary fact that, there exist a pair of injective (increasing) functions \( f : \cup T_1 \Rightarrow \cup T_2 \) and \( g : \cup T_2 \Rightarrow \cup T_1 \) where \( P(\{ \{ \omega \} \mid \{ \omega, f(\omega) \}) = 0 \) and \( P(\{ \{ \omega \} \mid \{ \omega, g(\omega) \}) = 0 \), whenever \( \omega \) is in the domain, respectively, of the function \( f \) or \( g \), i.e., whenever \( \omega \in \cup T_1 \) or \( \omega \in \cup T_2 \), respectively. That is, each of \( f \) and \( g \) maps each element of its domain into a distinct element of its range belonging to a higher tier in the well order \( \uparrow \). In other words, \( f \) pairs each point in \( \cup T_1 \) with a point in \( \cup T_2 \) having a higher tier under \( \uparrow \). Likewise, \( g \) pairs each point in \( \cup T_2 \) with a point in \( \cup T_1 \) having a higher tier under \( \uparrow \).

Use the functions \( f \) and \( g \) to create two \( \kappa \)-size partitions, \( \pi_f \) and \( \pi_g \), as defined below, and similar in kind to the partition used in \( \text{Lemma 3} \). Without loss of generality, when considering \( f \) (respectively, \( g \)), index its domain – for \( f \) that is the set of points \( \omega \in \cup T_1 \) (respectively for \( g \), that is the set of points \( \omega \in \cup T_2 \) – using an initial segment of ordinals running through \( \kappa \). That is, when considering \( f \), write \( \cup T_1 = \{ \omega_1^1, \omega_2^1, ..., \omega_\alpha^1, ... \} \) with \( 0 < \alpha < \kappa \). Similarly for \( g \). Write \( \cup T_2 = \{ \omega_1^2, \omega_2^2, ..., \omega_\alpha^2, ... \} \).

For each ordinal \( 0 < \alpha < \kappa \), define the partition element \( h_\alpha \) of \( \pi_f \) to be the pair-set
As before, define the catch-all set: \( T_3 = \Omega - [\bigcup T_1 \cup \text{Range}(f)] \). And if this set is not empty, add its elements as singleton sets to create the partition \( \pi_f = \{ h_1, \ldots, h_\alpha, \ldots \} \cup T_3 \). Then, for each \( h \in \pi_f \), \( P(T_1 | h) = 0 \). In parallel fashion, with respect to function \( g \), define \( \pi_g \) so that for each \( h \in \pi_g \), \( P(T_2 | h) = 0 \).

Since \( P(R) > 0 \), at least one of \( T_1 \) and \( T_2 \) is not null, that is since \( \text{maximum}\{P(T_1), P(T_2)\} > 0 \), \( P \) is not conglomerable in at least one of these two partitions, \( \pi_f \) and \( \pi_g \). **Lemma 5**

Note that the reasoning for **Case 1** gives a lower bound, the \( \text{maximum}\{P(T_1), P(T_2)\} \), for the extent of non-conglomerability of \( P \).

The following example alerts the reader that **Cases 1 and 2**, where respectively \( \uparrow \) and \( \downarrow \) well order the set of tiers, are sufficiently dissimilar that for a countable state space \( \Omega \) only one is consistent with \( P \) being countably additive.

**Example 2.** Let \( \Omega = \{ \omega_1, \omega_2, \ldots, \omega_m, \ldots \} \) be countable, which is not covered by the Proposition. Then there is no countably additive probability \( P \) corresponding to **Case 3**. Specifically, let each point of \( \Omega \) constitute its own tier with \( P(\{ \omega_m \} | \{ \omega_m, \omega_n \}) = 0 \) whenever \( m < n \). Then \( P(\{ \omega_i \}) = 0 \), \( i = 1, 2, \ldots \), contradicting the \( \sigma \)-additivity of \( P \). However, if as in **Case 4**, \( P(\{ \omega_m \} | \{ \omega_m, \omega_n \}) = 1 \) whenever \( m < n \), then this well ordering of the tiers corresponds to a perfectly additive (principal ultrafilter) 0-1 unconditional probability, where \( P \) has range \( \{0, 1\} \), and where \( P(\{ \omega_1 \}) = 1 \). Conditional probability also is 0-1, where, for each nonempty subset \( \emptyset \neq S \subseteq \Omega \), \( P(E | S) = 1 \) if and only if \( E \) includes the minimal element of \( S \). **Example 2**

In the light of **Example 2**, the proof of non-conglomerability when \( \downarrow \) is a well order (**Case 2 - Lemma 6**) uses different reasoning than when \( \uparrow \) is a well order (**Case 1 - Lemma 5**), and shows that where \( P \) is conglomerable, it is concentrated on tiers with limit ordinal indices. This contradicts \( SA_5 \), which requires that the union of points in tiers with successor ordinal indices have positive probability.

**Lemma 6:** Suppose \( \downarrow \) is a well order of the set of tiers, each of which is \( P \)-null - **Case 2**. Then \( P \) is non-conglomerable.

**Proof:** We index the well order \( \downarrow \) of tiers with an initial segment of the ordinals. Let \( \beta \) be the least ordinal in this well order such that \( P(\bigcup_{\alpha < \beta} \tau_\alpha) > 0 \) and let \( R \) be this interval of tiers, \( R = \{ \tau_\alpha : 0 < \alpha < \beta \} \). Then, as before, we may assume that \( \beta \) is a limit ordinal with \( | \beta | = \kappa \), since \( P(\tau) = 0 \) for each tier in \( R \), and \( P \) is \( \lambda \)-additive for each cardinal \( \lambda < \kappa \). Without loss of generality, by **Lemma 4**, assume that \( P(\bigcup R) = 1 \).

Consider the partition (a “histogram”) of \( R \) according to the cardinality of each tier. That is, let \( \pi_c = \{ h_\lambda \} \) where \( \tau \in h_\lambda \) if and only if \( |\tau| = \lambda \), for each \( \lambda < \kappa \). In the light of **Lemma 4**, each tier has cardinality less than \( \kappa \). So \( \pi_c \) is a partition, and \( (SA_6) \) by measurable sets of tiers. Thus \( h_1 \) is the set of those tiers with exactly one point, \( \{ \omega \} \); \( h_n \) is the set of those tiers with exactly \( n \)-points, etc. Since \( \kappa \) is not a weakly inaccessible

12/2/15

Non-conglomerability for \( \kappa \)-non-additive, \( \sigma \)-additive probabilities

13
cardinal, $|\pi_c| < \kappa$. As $P$ is $\lambda$-additive for each $\lambda < \kappa$, and since the elements of $\pi_c$ are measurable then $\sum_{h \in \pi_c} P(\cup h) = 1 = P(R)$.

For an $h \in \pi_c$ if $|h| = \lambda < \kappa$, then call it small. By $SA_6$, $\cup h$ is measurable. As small, $h$ contains fewer than $\kappa$-many tiers. And as $P$ is $\lambda$-additive and each tier is $P$-null, it follows that $P(\cup h) = 0$. That is, each small element of $\pi_c$ constitutes a $P$-null set of points.

Call $h \in \pi_c$ large if $|h| = \kappa$. So, as each element of $\pi_c$ is measurable and $\sum_{h \in \pi_c} P(\cup h) = 1$, at least one large element of $\pi_c$ is not null. It follows, then, that $P$ is supported by countably many large elements of $\pi_c$, and $P(\cup \{h : h \text{ is large}\}) = 1$. By $SA_5$, each large $h$ can be partitioned into three disjoint, measurable sets, where the first two (those with successor ordinal indices) are not both $P$-null. We complete the reasoning for Lemma 6 by arguing indirectly for the contradiction that, if $P$ is conglomerable, then the set of tiers with successor ordinal indices is a $P$-null set.

So, let $L = \cup \{h_\lambda \mid h_\lambda \in \pi_c \text{ with } |h_\lambda| = \kappa\}$ $L$ is the set of points belonging to tiers in a large element of $\pi_c$. That is, each point $\omega$ in $L$ belongs to a tier $\tau$ that is an element of some large set, $h_\lambda \in \pi_c$, i.e., where there are $\kappa$-many tiers in $h_\lambda$. Since this is a subset of the well-ordering of $R$ under $\downarrow$, it too is well-ordered by $\downarrow$. For the reasons already given, we may index the elements of $h_\lambda$ by the ordinals less than $\kappa$. That is, write $h_\lambda = \{\tau^{h_\lambda}_\alpha : \alpha < \kappa\}$. We partition $h_\lambda$ into three sets, each a measurable set of points by $SA_5$.

(A) the set of successor tiers in $h_\lambda$ each with an even ordinal index ending "+2n" for integer, $n = 1, 2, \ldots$. This set has cardinality $\kappa$ and is cofinal in $h_\lambda$.

(B) the set of successor tiers in $h_\lambda$ each with an odd ordinal index ending "+2n-1" for integer, $n = 1, 2, \ldots$. This set has cardinality $\kappa$ and is cofinal in $h_\lambda$.

(C) the set of tiers in $h_\lambda$ each with a limit ordinal as its index. For convenience, since $0$ has no predecessor, we include the singleton set $\tau^{h_\lambda}_0$.

This set has cardinality $\kappa$ and is cofinal in $h_\lambda$.

We construct two partitions. The first partition shows that if $P$ is conglomerable, then $P(\cup B) = 0$. The second partition shows that if $P$ is conglomerable, then $P(\cup C) = 0$. Together, this contradicts the final clause of $SA_5$.

To create the first partition, pair each tier in the set $A$ 1-1 with its immediate predecessor tier in $h_\lambda$. Since each tier in $h_\lambda$ has a common cardinality, $\lambda$, then pair, 1-1, each element of each tier in $A$ with an element of its predecessor tier. Let $f$ be this 1-1 pairing of points in $UA$ with points in the $\cup$ (predecessors-to-$A$). Write these pairs as $\{\omega, f(\omega)\}$ where $\omega \in \tau^{h_\lambda}_\alpha \in A \subset h_\lambda$. Then, $P(\{\omega\} | \{\omega, f(\omega)\}) = 0$ for each such pair, since $f$ is regressive on the ordinals indexing tiers in $A$. Complete the partition by adding all the singleton sets $\{\omega\}$ for $\omega \in \cup R - (\cup A \cup \cup (\text{range } f))$ and denote an arbitrary element.
of this partition \( h_B \). Then, \( P(\mathbb{U}A \mid h_B) = 0 \), which gives us \( P(A) = 0 \) by conglomerability of \( P \).

Similarly, to create the partition targeted at showing \( P(\mathbb{U}B) = 0 \), use a 1-1 regressive function pair each element of the set of tiers \( B \) with its immediate predecessor tier in \( h_{\lambda} \) and continue the reasoning just as in the previous paragraph.

The upshot is that if \( P \) is conglomerable in each of these partitions, we have the a contradiction with \( S_5 \) that requires that at least one of sets \( \mathbb{U}A \) and \( \mathbb{U}B \) is not \( P \)-null.

**Lemma 6**

Note that the extent of non-conglomerability of \( P \) (in any one large partition element) is bounded below by the maximum of \( P(\mathbb{U}A) \) and \( P(\mathbb{U}B) \).

**Remark:** Lemma 6 is established by finding two, 1-1 regressive functions for the ordinals, respectively, indexing set \( A \) and indexing set \( B \). But set \( C \) is stationary; hence, by Fodor’s (1956) “Pressing Down” lemma, there is no such 1-1 regressive function on \( C \). We do not know whether, if \( P(\mathbb{U}C) > 0 \), \( P \) is non-conglomerable for a measurable event that is a subset of \( \mathbb{U}C \).

**Lemma 7:** Assume that there are two countable sets of tiers \( M_{\uparrow} = \{\tau'_1, \ldots, \tau'_{n}, \ldots\} \) and \( N_{\uparrow} = \{\tau''_1, \ldots, \tau''_{n}, \ldots\} \) well ordered respectively as the natural numbers, \((N, <)\).

That is, the elements of \( M_{\uparrow} \) satisfy: \( \tau_m \downarrow \tau_{n} \) and elements of \( N_{\uparrow} \) satisfy \( \tau''_{m} \uparrow \tau''_{n} \) whenever \( n > m - Case 3 \). Then \( P \) is not conglomerable.

**Proof:** Combine the two sequences \( M_{\uparrow} \) and \( N_{\uparrow} \) to form a single countable set \( L \), linearly ordered, either by \( \uparrow \) or by \( \downarrow \). Using the positive and negative rational numbers \( \mathbb{Z} \), we can represent this linear order \( L \) as one of five varieties, each variety corresponding to a subset of \( \mathbb{Z} \) under its natural order.

**L1:** Set \( M_{\downarrow} \) lies entirely below set \( N_{\uparrow} \) in \( L \). Then the order of tiers in \( L \) may be represented by the negative and positive integers. That is, \( M_{\downarrow} \) has tiers \( \tau_i \), for \( i = -1, -2, \ldots \), and \( N_{\uparrow} \) has tiers \( \tau_i \) for \( i = 1, 2, \ldots \).

**L2:** Set \( M_{\downarrow} \) lies entirely above set \( N_{\uparrow} \) in \( L \). Then the order in \( M_{\downarrow} \) may be represented by a set of rational numbers, \( \{q_i = 1+1/i : i = 1, 2, \ldots\} \) and the order in \( N_{\uparrow} \) may be represented by a set of rational numbers, \( \{q_i = -1+1/i : i = 1, 2, \ldots\} \).

**L3:** A tail of the sequence \( M_{\downarrow} \) lies between two elements of \( N_{\uparrow} \) but the tail of \( N_{\uparrow} \) is entirely above \( M_{\downarrow} \).

**L4:** A tail of the sequence \( N_{\uparrow} \) lies between two elements of \( M_{\downarrow} \) but the tail of \( M_{\downarrow} \) is entirely below \( N_{\uparrow} \).
L5: A tail of the sequence $M_\uparrow$ lies between two elements of $N_\uparrow$ and a tail of $N_\uparrow$ is lies between two elements of $M_\downarrow$.

In each case, the countably many tiers in the linear order $L$ create a countable partition of all the tiers, $R$: Partition the linear order $R$ by using the elements of $L$ to form cuts, in the fashion of Dedekind Cuts. These cuts produce measurable intervals in $R$, since each such interval is defined using no more than countably many elements of $L$.

By Lemma 4, if $P$ is conglomerable, and as it is countably additive, then one and only one of these intervals is not null. Denote that interval $I^*_{\alpha}$. That is, $P(R) = P(I^*_{\alpha})$. Thus $P$ is remote on these countably many intervals. Moreover, if we denote by $I^*_{\uparrow 0}$ the interval of tiers above $I^*_{\alpha}$, then also by Lemma 4, we have that $|\cup I^*_{\uparrow 0}| < \kappa$. Then as $P$ is complete and $\lambda$-additive for $\lambda < \kappa$, $\cup I^*_{\uparrow 0}$ is measurable with $P(\cup I^*_{\uparrow 0}) = 0$.

The linear order of tiers within the interval $I^*_{\alpha}$ is again one of the three types, corresponding to Cases 1, 2, or 3. If $I^*_{\alpha}$ produces a linear order that is a well order, corresponding to either Case 1 or 2, complete the argument by duplicating Lemma 5 or Lemma 6 (respectively) applied to the interval $I^*_{\alpha}$. If the linear order within $I^*_{\alpha}$ is also an instance of Case 3, then repeat the reasoning to produce a subinterval, $I^*_{\alpha} \subset I^*_{\uparrow 0}$, where $P(R) = P(I^*_{\uparrow 0})$.

Next we continue the argument, assuming that at each stage the interval $I^*_{\alpha}$ has an internal linear structure corresponding to Case 3. Define the intervals $I^*_{\alpha}$ inductively. At successor ordinals $\beta = \alpha + 1$, the process is as above. At limit ordinals $\beta < \kappa$, let $I^*_{\beta} = \cap I^*_{\alpha}$ for $\alpha < \beta$. To see that these are measurable sets, define the two sequences of increasing "tail" intervals

\[ I^*_{\uparrow 0} \subset I^*_{\uparrow 1} \subset ... \]

and

\[ I^*_{\downarrow 0} \subset I^*_{\downarrow 1} \subset ... \]

Since $P$ is complete and $\lambda$-additive for each $\lambda < \kappa$, for each $\alpha < \kappa$, the two sets $I^*_{\uparrow \alpha}$ and $I^*_{\downarrow \alpha}$ are measurable and null, $P(\cup I^*_{\uparrow \alpha}) = P(\cup I^*_{\downarrow \alpha}) = 0$. Since $I^*_{\alpha} = (I^*_{\uparrow \alpha} \cup I^*_{\downarrow \alpha})^c$, then $P(\cup I^*_{\alpha}) = 1$.

Continue in this fashion until the resulting interval $I^*$ is a P-null event. Then there is a sequence of nested subintervals $I^*_{\alpha} \supset I^*_1 \supset I^*_2 \supset ... \supset I^*_\alpha \supset ...$, where $P(I^*_{\alpha}) = P(R) = 1$ for each P-non-null interval $I^*_\alpha$. That will occur at a limit ordinal whose cardinality is $\kappa$. Without loss of generality, this sequence can be shortened to terminate at the ordinal $\kappa$. This forms a $\kappa$-long, "descendingly incomplete" sequence of probability 1 intervals, where $I^*_\kappa$ is P-null, hence measurable. (See Chang, 1967, also Kunen and Prikry 1971 for related material on descendingly incomplete ultrafilters.)
Next, consider the two tail intervals $I_{\alpha}^+\kappa$ and $I_{\alpha}^-\kappa$, where $I_{\alpha}^+\kappa$ is entirely above $I_{\alpha}^-\kappa$ in the linear order of tiers. There are two subcases to consider.

Subcase 1, where $|\cup I_{\alpha}^-\kappa| = \kappa$. By Lemma 4, if $P$ is conglomerable, then $P(\cup I_{\alpha}^+\kappa) = 0$. So, in this subcase, we have that $P(R) = P(\cup I_{\alpha}^+\kappa) = 1$. Use the $\kappa$-long well ordered sequence $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$ to create a corresponding $\kappa$-long well-ordered sequence of disjoint, measurable (null) sets of tiers, $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$, as follows.

Let $J_{\alpha}^+0 = I_{\alpha}^+0$. For a successor ordinal, $\beta = \alpha+1$, let $J_{\alpha}^+\beta = I_{\alpha}^+\alpha+1 - I_{\alpha}^+\alpha$. At limit ordinals $J_{\alpha}^+\beta = I_{\alpha}^+\beta - \cup_{\alpha<\beta} I_{\alpha}^+\alpha$. Then, for each $\alpha < \beta < \kappa$, the interval of tiers $J_{\alpha}^+\alpha$ is entirely above the interval of tiers $I_{\alpha}^+\beta$. Note that these intervals, $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$, form a partition $\pi^*$ of $I_{\alpha}^+\kappa$. Then adapt Lemma 6 to this well order of intervals to show that $P$ is not conglomerable.

Subcase 2, where $|\cup I_{\alpha}^+\kappa| < \kappa$. Then $P(\cup I_{\alpha}^-\kappa) = 0$, and so $P(\cup I_{\alpha}^+\kappa) = 1$. We adapt the reasoning of the previous subcase. Use the $\kappa$-long well ordered sequence $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$ to create a corresponding $\kappa$-long well-ordered sequence of disjoint, measurable (null) sets of tiers, $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$, as follows.

Let $J_{\alpha}^+0 = I_{\alpha}^+0$. For a successor ordinal, $\beta = \alpha+1$, let $J_{\alpha}^+\beta = I_{\alpha}^+\alpha+1 - I_{\alpha}^+\alpha$. At limit ordinals $J_{\alpha}^+\beta = I_{\alpha}^+\beta - \cup_{\alpha<\beta} I_{\alpha}^+\alpha$. Then, for each $\alpha < \beta < \kappa$, the interval of tiers $J_{\alpha}^+\alpha$ is entirely below the interval of tiers $I_{\alpha}^+\beta$. Note that these intervals, $\{I_{\alpha}^+\alpha; \alpha < \kappa\}$, form a partition $\pi^*$ of $I_{\alpha}^+\kappa$. Then adapt Lemma 5 to this well order of intervals to show that $P$ is not conglomerable. Lemma 7.

Note that this reasoning about non-conglomerability in Case 3 reduces it to Case 1 or Case 2 – assuming that neither Situation 1 nor 2 obtains. That provides a lower bound on the the extent of non-conglomerability relating to one of these previously analyzed Cases (or Situations).

The Proposition is immediate from the five Lemmas 3, 4, 5, 6, and 7. Proposition

5. An illustration of Sufficient Condition 1 – Lemma 3.

In this section we illustrate Sufficient Condition 1, and reasoning used in Lemma 3. We use this illustration to explain a difference between the de Finetti/Dubins theory of conditional probability used in this paper and the theory of regular conditional distributions from the received (Kolmogorovian) theory of Probability.

Example 3: Let $<\Omega, \Sigma, P>$ be the complete measure space of Lebesgue measurable...
subsets of the half--open unit interval of real numbers: \( \Omega = [0,1) \) and \( \mathcal{F} \) is its algebra of Lebesgue measurable subsets. Let \( P \) be the uniform, countably additive Lebesgue probability with constant density function \( f(\omega) = 1 \) for each real number \( 0 \leq \omega < 1 \), and \( f(\omega) = 0 \) otherwise. So \( P(\{\omega\}) = 0 \) for each \( \omega \in \Omega \). Evidently \( P \) is not \( \kappa = 2^{\aleph_0} \) additive, because \( \Omega \) is the union of \( 2^{\aleph_0} \)-many null sets.

As an illustration of **Sufficient Condition** 1 use the uniform density function \( f \) to identify conditional probability given finite sets as uniform over those finite sets, as well. That is, when \( F = \{\omega_1, \ldots, \omega_k\} \) is a finite subset of \( \Omega \) with \( k \)-many points, let \( P(\cdot \mid F) \) be the perfectly additive probability that is uniform on these \( k \)-many points. These conditional probabilities create a single tier \( \tau = \Omega \), as \( P(\{\omega_1\} \mid \{\omega_1, \omega_2\}) = 0.5 \) for each pair of points in \( \Omega \).

However, by the countable additivity of \( P \), it follows that each denumerable set of points is \( P \)-null. For example, with \( U = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\} \) (for \( n < \aleph_0 \)), then \( P(U) = 0 \). By Definition 1, then for each point \( \omega \in \Omega \), \( P(\{\omega\} \mid U) = 0 \) and the conditional probability \( P(\cdot \mid U) \) is a finite but not countably additive conditional probability function.

Next, consider the two events \( E = \{\omega: 0 \leq w < 0.9\} \) and its complement with respect to \( \Omega \), \( E^c = \{\omega: 0.9 \leq \omega < 1\} \), where \( P(E) = 0.9 \). This pair “splits” the sure event \( \Omega \). Let \( g \) be the 1-1 (continuous) map between \( E \) and \( E^c \) defined by \( g(\omega) = 0.9 + \omega/9 \), for \( \omega \in E \). Consider the \( \kappa \)-size partition of \( \Omega \) by pair-sets, \( \pi = \{\{\omega, g(\omega)\}: \omega \in E\} \). By assumption, \( P(\{\omega\} \mid \{\omega, g(\omega)\}) = 1/2 \) for each pair in \( \pi \). But then \( P \) is not conglomerable in \( \pi \).

The usual theory of regular conditional distributions treats the example differently. We continue the example from that point of view. Consider the measure space \( \langle \Omega, \mathcal{F}, P \rangle \) as above. Let the random variable \( X(\omega) = w \), so that \( X \sim U[0,1) \), \( X \) has the uniform distribution on \( \Omega \). In order to consider conditional probability given the pair of points \( \{\omega, g(\omega)\} \), let

\[
g(X) = \begin{cases} 
(\frac{X}{9}) + 0.9 & \text{if } 0 \leq X < 0.9 \\
9(X - 0.9) & \text{if } 0.9 \leq X < 1.
\end{cases}
\]

Define the random variable \( Y(\omega) = X(\omega) + g(X(\omega)) - 0.9 \).

Observe that \( Y \sim U[0,1.0) \). Also, note that \( Y \) is 2-to-1 between \( \Omega \) and \( [0.0, 1.0) \). That is \( Y = y \) is entails that either \( \omega = 0.9y \) or \( \omega = 0.1(y + 9) \).

Let the sub-\( \sigma \)-sigma field \( \mathcal{A} \) be generated by the random variable \( Y \). The regular conditional distribution relative to this sub-\( \sigma \)-sigma field, \( P(\mathcal{E} \mid \mathcal{A})(\omega) \), is a real-valued function defined on \( \Omega \) that is \( \mathcal{A} \)-measurable and satisfies the integral equation

\[
\int_A P(B \mid \mathcal{A})(\omega) \, dP(\omega) = P(A \cap B)
\]

whenever \( A \in \mathcal{A} \) and \( B \in \mathcal{E} \).

In our case, then \( P[B \mid \mathcal{A}](\omega) \) almost surely satisfies:

\[
P(X = 0.9Y \mid Y)(\omega) = 0.9
\]

and
P(X = 0.1(Y+9.0) | Y)(ω) = 0.1. Thus, relative to the random variable Y, this regular conditional distribution assigns conditional probabilities as if P({ω} | {ω, g(ω)}) = 0.9 for almost all pairs {ω, g(ω)} with 0 ≤ ω < 0.9. However, just as in the Borel "paradox" (Kolmogorov, 1933), for a particular pair {ω, g(ω)}, the evaluation of P({ω} | {ω, g(ω)}) is not determinate and is defined only relative to which sub-σ-sigma field $\mathcal{A}$ embeds it.

For an illustration of this last feature of the received theory of regular conditional distributions, consider a different pair of complementary events with respect to $\Omega$. Let $F = \{ω: 0 ≤ ω < 0.5\}$ and $F^c = \{ω: 0.5 ≤ ω < 1\}$. So, P(F) = 0.5. Let 

\[
\begin{align*}
    f(X) &= 1.0 - X & \text{if } 0 < X < 1. \\
    &= 0 & \text{if } X = 0.
\end{align*}
\]

Analogous to the construction above, let $Z(ω) = |X(ω) - f(X(ω))|$. So Z is uniformly distributed, $Z \sim U[0,1]$, and is 2-to-1 from $Ω$ onto $[0,1]$. Consider the sub-σ-sigma field $\mathcal{A}'$ generated by the random variable Z. Then the regular conditional distribution $P(B | \mathcal{A}')(ω)$, almost surely satisfies:

\[
P(X = 0.5 - Z/2 | Z \neq 0)(ω) = 0.5 \quad \text{and} \quad P(X = 0.5 + Z/2 | Z \neq 0)(ω) = 0.5
\]

and for convenience, $P(X = 0 | Z = 0) = P(X = 0.5 | Z = 0) = 0.5$.

However, $g(.09) = .91 = f(.09)$ and $g(.91) = .09 = f(.91)$. That is, $Y = 0.1$ if and only if $Z = 0.82$. So in the received theory, it is permissible to have $P(ω = .09 | Y = 0.1) = 0.9$ as evaluated with respect to the sub-σ-sigma field generated by Y, and also to have $P(ω = .09 | Y = 0.82) = 0.5$ as evaluated with respect to the sub-σ-sigma field generated by Z, even though the conditioning events are the same event. (Example 3)

6. Conclusion. Given a probability P that satisfies the five structural assumptions of the Proposition, we show that non-conglomerability of its coherent conditional probabilities is linked to the index of non-additivity of P. Specifically, as P is not $κ$-additive (and $κ$ is not an inaccessible cardinal) then there is a $κ$-size partition $π = \{h_α: α < κ\}$ where the coherent conditional probabilities $P(\cdot | h_α)$ are not conglomerable. Namely, there exists an event $E$ and a real number $ε > 0$ where, for each $h_α \in π$, $P(E > P(E | h_α) + ε$.

This permits us to conclude that the anomalous phenomenon of non-conglomerability is a result of adopting the de Finetti/Dubins theory of coherent conditional probability instead of the rival Kolmogorovian theory of regular conditional distributions. It is not a result of the associated debate over whether probability is allowed to be merely finitely additive rather than satisfying countable additivity. Restated, our conclusion is that even when P is $λ$-additive for each $λ < κ$, if P is not $κ$-additive and has coherent conditional probabilities, then P will experience non-conglomerability in a $κ$-sized partition. The received theory of regular conditional distributions sidesteps non-conglomerability by allowing conditional probability to depend upon a sub-sigma field, rather than being defined given an event.
*Acknowledgments:* In writing this paper we have benefitted from constructive comments and corrections, in particular about Lemma 6, from K.P. Hart. Also, we thank Jeremy Avigad, Jessi Cisewski, Paul Pedersen, Rafael Stern, Wilfried Sieg and an anonymous reader for their helpful advice.

References


