

Chapter Two: Probability Triples

March 23, 2016

Abstract

This chapter covers probability triples, the foundation of the theory, and how to construct them. In particular, a great deal of effort is spent setting the foundations for what to do in the case of uncountable sample spaces, as the basic theory does not generalize easily when taking the obvious choices for a sigma-algebra.

1 Basic Definition

Definition. A *probability triple* is compromised of three components:

1. A *sample space* Ω which is the set of all values we are considering.
2. A σ -*algebra* \mathcal{F} which is a sigma algebra over the sample space (recall this is a collection of sets which is closed under complement and countable union, as well as containing \emptyset).
3. A probability measure \mathbb{P} which is a mapping $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ that is also countably additive and such that $P(\emptyset) = 0$, $P(\Omega) = 1$.

Note that the fact that \mathcal{F} is closed under countable unions and complements implies it is closed under countable intersections as well. By deMorgan's Laws for set operations, we have that $(A^c \cup B^c)^c = A \cap B$, and so we can apply this to countably infinite operations via a reasonable limiting process.

Also note that there are several basic familiar results we can derive regarding this definition of a probability triple.

1. Complements: Consider $P(A \cup A^c) = 1 = P(A) + P(A^c)$. Thus the familiar 1 - rule for probability complements.
2. Setminus: Consider arbitrary A and B . Then

$$\begin{aligned} P(A) &= P((A \setminus (B \cap A)) \cup (B \cap A)) = P[A \setminus (B \cap A)] + P[B \cap A] \\ &= P[A \setminus B] + P[B \cap A] \Rightarrow P[A \setminus B] = P(A) - P(B \cap A) \end{aligned}$$

3. Inclusion/Exclusion: We want to derive a result for $P(A \cup B)$. Note that

$$P(A \cup B) = P[(A \setminus B) \cup (B \cap A) \cup (B \setminus A)] = P(A) - 2P(A \cap B) + P(A \cap B) + P(B)$$

4. Obviously countable subadditivity follows as well.

With these rules much of probability theory can be derived. Indeed, these are more or less the start point for Jaynes in his book, with a few exceptions.

2 Constructing Probability Triples

Now the question becomes how to apply these rules to actual scenarios. In the case of countable Ω , this is simple: just provide a probability measure which adds to 1 on the entire set, and the sigma algebra will be the set of all subsets of Ω . Take for example the Poisson(5) distribution, and let A be an arbitrary subset. This yields the following measure:

$$P(A) = \sum_{k \in A} e^{-5} 5^k / k!$$

Also we can easily recover the uniform distribution over any finite set from this method as well.

The challenge now comes from how to create a valid and meaningful probability triple over an uncountable set. We examine $[0, 1]$ for simplicity. First one might consider the set

$$\mathcal{J} = \text{all intervals contained in } [0, 1].$$

Unfortunately this cannot be considered a sigma algebra, as the union of two intervals is not an interval. However, it is clearly a semi-algebra: an essential tool for building the theory.

Definition. A *semi-algebra* is a collection of subsets of a sample space Ω such that it is closed under finite intersection, contains Ω and \emptyset , and the complement of every element is a disjoint union of elements in the set.

This is more or less a generalization of the example provided: in all likelihood, the historical origin of this notion comes from considering all intervals of a set. We prove this connection in the exercises.

However, it is not easy to generalize to a probability triple from this semi-algebra collection: a few obvious extensions fall markedly short, including considering the set of all finite unions of the semi-algebra and all countable unions (exercise 2.4.7). Thus...the next section.

3 The Extension Theorem

The extension theorem is the primary method we have for creating probability triples on uncountable sample spaces. It essentially begins with a semi-algebra and a function on that algebra (a proto-measure, so to speak) satisfying certain key conditions.

Theorem. Consider any sample space Ω , a semi-algebra \mathcal{J} on that sample space and a function $P : \mathcal{J} \rightarrow [0, 1]$ such that

1. Finite Additivity: for any finite collection of disjoint sets in \mathcal{J} we have that

$$P(\cup A_i) = \sum P(A_i)$$

where the sum and unions are taken to be over all i .

2. $P(\Omega) = 1$ and $P(\emptyset) = 0$.
3. Countable monotonicity property: if we have a collection of sets $\{A_i\}$ of cardinality \aleph_0 , and set which is a subset of a union of these sets, then the sum of $P(A_i)$ over all the i

Then there exists a σ -algebra $\mathcal{M} \supseteq \mathcal{J}$ and a probability measure P^* on \mathcal{M} which agrees with the function P for all $A \in \mathcal{J}$.

The proof of this theorem is complex, and involves many different results which we go onto prove here.

Definition. 2.3.4. The *outer measure* extension of the function P given in the extension theorem is denoted as P^* and defined as the infimum of the sum over all subcollections of sets in \mathcal{J} , where the union of these sub-collection is a superset of the input set. Basically we go from above and get smaller.

Now that we have this definition for the outermeasure, we can move forward. Essentially, it is the measure we will define for the extension, but in order for it to be appropriate we will need to restrict it to a certain σ -algebra. In the mean time, we can prove a necessary property which holds for any set in \mathcal{J} .

Lemma. 2.3.5. $P^*(A) = P(A)$ for all $A \in \mathcal{J}$.

Proof. Consider any arbitrary $A \in \mathcal{J}$. Then $P^*(A) \leq P(A)$ because we can just set the sum to be the single set A , so we know the infimum is at most as large as $P(A)$. The reverse holds by countable monotonicity. □

Next we prove another necessary result which holds regardless of the the input sets, and without the restriction: countable subadditivity.

Lemma. 2.3.6. P^* is countably subadditive:

$$P^*(\cup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} P^*(B_n) \quad \text{for any } B_1, B_2, \dots \subseteq \Omega$$

Proof. This proof essentially involves a simple bounding argument. Let $\{A_{n,i}\}_{i=0}^{\infty}$ be a collection of subsets of Ω where

$$P^*(B_n) + \frac{\epsilon}{2^n} > \sum P(A_i)$$

where $\epsilon > 0$. Such a collection must exist by the definition of the outer measure as an infimum. Summing over all the n we have the following relation:

$$\sum_{n=1}^{\infty} P^*(B_n) + \epsilon > \sum_{\forall n} \sum_{\forall i} P(A_{i,n}).$$

But since the union of the overall collection $\{A_{n,i}\}_{i=0}^{\infty}$ contains the union of the B_n what we really have is

$$\sum_{n=1}^{\infty} P^*(B_n) + \epsilon > \sum_{\forall n} \sum_{\forall i} P(A_{i,n}) \geq P^*(\cup_{n=1}^{\infty} B_n)$$

which holds for any arbitrary epsilon and therefore completes the proof. □

Definition. 2.3.7 The Lebesgue σ -algebra is

$$\mathcal{M} = \{A \subseteq \Omega \mid P^*(A \cap E) + P^*(A^c \cap E) = P(E)\}.$$

Lemma 2.3.9. If $A_1, A_2, \dots \in \mathcal{M}$ are disjoint, then $P^*(\cup_n A_n) = \sum_n P^*(A_n)$.

Proof. Letting $A_1 \cup A_2 = E$, by definition we have

$$P(A_1 \cap E) + P(A_2 \cap E) = P(A_1 \cup A_2) \Rightarrow P(A_1) + P(A_2) = P(A_1 \cup A_2)$$

where the implication follows since A_1 and A_2 are disjoint. So this holds for *any* finite collection, by induction.

Now we generalize to an arbitrary countable collection. Consider any $m \in \mathbb{N}$. Then by all our previous work/definitions:

$$\sum_{\forall n} P^*(A_n) \geq P^*(\cup_{\forall n} A_n) \geq P^*(\cup_{n \leq m} A_n) = \sum_{n \leq m} P^*(A_n).$$

This holds for any m , so taking the limit we have equality at infinity. □

The above result demonstrates why defining \mathcal{M} was important in the first place: because countable additivity over disjoint sets might not have held otherwise. Now that we have finished proving results for P^* on \mathcal{M} , we need only demonstrate that \mathcal{M} has the algebraic closure properties we desire (interestingly enough, the properties we demonstrate are true here are set-analogous to those in Jaynes for a Boolean algebra which covers all possible logic functions: but I digress).

Lemma 2.3.10 \mathcal{M} is an algebra (only proved for finite intersection).

Proof. First note closure under complements is trivial, and that \mathcal{M} obviously contains Ω . Consider $A, B \in \mathcal{M}$. By countable subadditivity of the outer-measure we have that

$$P^*((A \cap B) \cap E) + P^*((A \cap B)^c \cap E) \geq P^*(E).$$

We just need to prove the reverse direction and we are finished.

$$P^*((A \cap B) \cap E) + P^*((A \cap B)^c \cap E) = P^*(A \cap B \cap E) + P^*((A^c \cap B \cap E) \cup (A \cap B^c \cap E) \cup (A^c \cap B^c \cap E)).$$

Now we can apply subadditivity of the measure and the fact that $A, B \in \mathcal{M}$.

$$\begin{aligned} & P^*(A \cap B \cap E) + P^*((A^c \cap B \cap E) \cup (A \cap B^c \cap E) \cup (A^c \cap B^c \cap E)) \\ & \leq P^*(A \cap B \cap E) + P^*((A^c \cap B \cap E)) + P^*(A \cap B^c \cap E) + P^*(A^c \cap B^c \cap E) \\ & = P^*(B \cap E) + P^*(B^c \cap E) = P(E). \end{aligned}$$

□

Now the following lemma is really just mathematical machinery we need for our subsequent proof that \mathcal{M} is a σ -algebra. But it makes intuitive sense, because of the way the outer-measure is defined: we want the outer-measure to be closed under taking the intersection with some arbitrary set. That is to say, if the E were removed here it would follow clearly, but because the E is there it is because of the way it is defined.

Lemma 2.3.11 Let $A_1, A_2, \dots \in \mathcal{M}$ be disjoint. For each $m \in \mathbb{N}$, let $B_m = \cup_{n \leq m} A_n$. Then for all $m \in \mathbb{N}$, and for all $E \subseteq \Omega$,

$$P^*(E \cap B_m) = \sum_{n \leq m} P^*(E \cap A_n).$$

Proof. We prove this inductively. The base case, $m = 1$, holds trivially. Now we extrapolate and assume the proposition holds true for $m = k$.

$$P^*(E \cap B_{m+1}) = P^*(E \cap B_{m+1} \cap B_m) + P^*(E \cap B_{m+1} \cap B_m^c)$$

where this equality follows since $B_m \in \mathcal{M}$. But we are essentially done now because by assumption we have

$$P^*(E \cap B_{m+1} \cap B_m^c) = P^*(E \cap A_m)$$

and the other portion is just the original sum (by the inductive hypothesis). □

Now using this result we can prove that elements of \mathcal{M} are closed under countable union.

Lemma 2.3.13 Let $A_1, A_2, \dots \in \mathcal{M}$ be disjoint. Then $\cup_n A_n \in \mathcal{M}$.

Proof. First let $B_m = \cup_{n \leq m} A_n$. Then by preceding lemma we have

$$P^*(E) = P^*(B_m \cap E) + P^*(B_m^c \cap E) \geq \sum_{n \leq m} P^*(A_n \cap E) + P^*((\cup_n A_n)^c \cap E)$$

where the inequality follows since $B_m^c \supset (\cup_n A_n)^c$. This inequality follows for any m , and so we can extend it to

$$\left(\sum_{n \leq m} P^*(A_n \cap E) \right) + P^*((\cup_n A_n)^c \cap E) \geq \left(\sum_n P^*(A_n \cap E) \right) + P^*((\cup_n A_n)^c \cap E)$$

which finally yields the desired result:

$$P^*(E) \geq P^*(\cup_n A_n \cap E) + P^*((\cup_n A_n)^c \cap E).$$
□

Lemma 2.3.14 \mathcal{M} is a σ -algebra.

Proof. By the previous lemma, we can demonstrate \mathcal{M} is countably subadditive. The rest follows from the fact that \mathcal{M} is an algebra (lemma 2.3.10).

To demonstrate that Lemma 2.3.13 will hold even for non-disjoint sets, consider any arbitrary sequence of sets. Then let every n^{th} set in a new sequence of sets created from this given sequence be whatever new elements the original set in the sequence contributes. These sets are disjoint but the union is the same. □

Lemma 2.3.15 $\mathcal{J} \subseteq \mathcal{M}$.

Proof. Consider any $J \in \mathcal{M}$. We must demonstrate that for any set $E \in \mathcal{M}$, we have

$$P^*(J^c \cap E) + P^*(J \cap E) = P^*(E).$$

This in turn means that we must have

Extension Theorem *The original theorem stated at the beginning of this subsection*

4 Constructing the Uniform[0,1] distribution

We apply the extension theorem to an example set which will be used often throughout the text: the $[0, 1]$ interval.

An important concept that must be brought up is that of the Borel Sets.

Definition. For any given sample space Ω , the Borel σ -algebra of sets is the smallest sigma algebra containing the intervals of Ω , or $\sigma(\mathcal{J})$. We denote this σ -algebra as \mathcal{B} .

5 Extensions of the Extension Theorem

In general, we will generally be using the Extension Theorem to verify the existence of complex probability triples. However, while finite superadditivity is easily verifiable, countable monotonicity is not. Therefore, we provide some alternate formulations here which might be easier to deal with in certain scenarios.

Corollary 2.5.1 Let \mathcal{J} be a semi-algebra of subsets of Ω . Let $P : \mathcal{J} \rightarrow [0, 1]$ with $P(\emptyset) = 0$ and $P(\Omega) = 1$ satisfying the finite superadditivity property of the original extension theorem, as well as the property

$$P(A) \leq P(B) \text{ whenever } A, B \in \mathcal{J} \text{ with } A \subseteq B$$

and also the “countable subadditivity property on \mathcal{J} ” where

$$P\left(\bigcup\right)$$