Infinite product measures

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto
May 10, 2015

1 Introduction

The usual proof that the product of a collection of probability measures exists uses Fubini’s theorem. This is unsatisfying because one ought not need to use Fubini’s theorem to prove things having only to do with $\sigma$-algebras and measures. In this note I work through the proof given by Saeki of the existence of the product of a collection of probability measures.\(^1\) We speak only about the Lebesgue integral of characteristic functions.

2 Rings of sets and Hopf’s extension theorem

If $X$ is a set and $\mathcal{R}$ is a collection of subsets of $X$, we call $\mathcal{R}$ a ring of sets when (i) $\emptyset \in \mathcal{R}$ and (ii) if $A$ and $B$ belong to $\mathcal{R}$ then $A \cup B$ and $A \setminus B$ belong to $\mathcal{R}$. If $\mathcal{R}$ is a ring of sets and $A, B \in \mathcal{R}$, then $A \cap B = A \setminus (A \setminus B) \in \mathcal{R}$. Equivalently, one checks that a collection of subsets $\mathcal{R}$ of $X$ is a ring of sets if and only if (i) $\emptyset \in \mathcal{R}$ and (ii) if $A$ and $B$ belong to $\mathcal{R}$ then $A \triangle B$ and $A \cap B$ belong to $\mathcal{R}$, where $A \triangle B = (A \setminus B) \cup (B \setminus A)$ is the symmetric difference. One checks that indeed a ring of sets is a ring with addition $\triangle$ and multiplication $\cap$. If $\mathcal{I}$ is a nonempty collection of subsets of $X$, one proves that there is a unique ring of sets $\mathcal{R}(\mathcal{I})$ that (i) contains $\mathcal{I}$ and (ii) is contained in any ring of sets that contains $\mathcal{I}$. We call $\mathcal{R}(\mathcal{I})$ the ring of sets generated by $\mathcal{I}$.

If $\mathcal{A}$ is a ring of subsets of a set $X$, we call $\mathcal{A}$ an algebra of sets when $X \in \mathcal{A}$. Namely, an algebra of sets is a unital ring of sets. If $\mathcal{I}$ is a nonempty collection of subsets of $X$, one proves that there is a unique algebra of sets $\mathcal{A}(\mathcal{I})$ that (i) contains $\mathcal{I}$ and (ii) is contained in any algebra of sets that contains $\mathcal{I}$. We call $\mathcal{A}(\mathcal{I})$ the algebra of sets generated by $\mathcal{I}$.

For a nonempty collection $\mathcal{G}$ of subsets of a set $X$, we denote by $\sigma(\mathcal{G})$ the smallest $\sigma$-algebra of subsets of $X$ such that $\mathcal{G} \subseteq \sigma(\mathcal{G})$.

If $R$ is a ring of subsets of a set $X$ and $\tau : R \to [0, \infty]$ is a function such that (i) $\mu(\emptyset) = 0$ and (ii) when $\{A_n\}$ is a countable subset of $R$ whose members are pairwise disjoint and which satisfies $\bigcup_{n=1}^{\infty} A_n \in R$, then

$$\tau\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \tau(A_n),$$

we call $\tau$ a measure on $R$. The following is Hopf’s extension theorem.²

**Theorem 1** (Hopf’s extension theorem). Suppose that $X$ is a set, that $R$ is a ring of subsets of $X$, and that $\tau$ is a measure on $R$. If there is a countable subset $\{E_n\}$ of $R$ with $\tau(E_n) < \infty$ for each $n$ and such that $\bigcup_{n=1}^{\infty} E_n = X$, then there is a unique measure $\mu : \sigma(R) \to [0, \infty]$ whose restriction to $R$ is equal to $\tau$.

3 **Semirings of sets**

If $X$ is a set and $S$ is a collection of subsets of $X$, we call $S$ a semiring of sets when (i) $\emptyset \in S$, (ii) if $A$ and $B$ belong to $S$ then $A \cap B \in S$, and (iii) if $A$ and $B$ belong to $S$ then there are pairwise disjoint $C_1, \ldots, C_n \in S$ such that

$$A \setminus B = \bigcup_{i=1}^{n} C_i.$$

If $S$ is a semiring of subsets of a set $X$, we call $S$ a semialgebra of sets when $X \in S$. One proves that if $S$ is a semialgebra, then the collection $A$ of all finite unions of elements of $S$ is equal to the algebra generated by $S$, and that each element of $A$ is equal to a finite union of pairwise disjoint elements of $S$.³

4 **Cylinder sets**

Suppose that $\{(\Omega_i, \mathcal{F}_i, P_i) : i \in I\}$ is a nonempty collection of probability spaces and let

$$\Omega = \prod_{i \in I} \Omega_i.$$

If $A_i \in \mathcal{F}_i$ for each $i \in I$ and $\{i \in I : A_i \neq \Omega_i\}$ is finite, we call

$$A = \prod_{i \in I} A_i$$

a cylinder set. Let $C$ be the collection of all cylinder sets. One checks that $C$ is a semialgebra of sets.⁴

---

Lemma 2. Suppose that $P : \mathcal{C} \to [0, 1]$ is a function such that
\[ \sum_{n=1}^{\infty} P(A_n) = 1 \]
whenever $A_n$ are pairwise disjoint elements of $\mathcal{C}$ whose union is equal to $\Omega$. Then there is a unique probability measure on $\sigma(\mathcal{C})$ whose restriction to $\mathcal{C}$ is equal to $P$.

Proof. Let $\mathcal{A}$ be the collection of all finite unions of cylinder sets. Because $\mathcal{C}$ is a semialgebra of sets, $\mathcal{A}$ is the algebra of sets generated by $\mathcal{C}$, and any element of $\mathcal{A}$ is equal to a finite union of pairwise disjoint elements of $\mathcal{C}$. Let $A \in \mathcal{A}$. There are pairwise disjoint $B_1, \ldots, B_j \in \mathcal{C}$ whose union is equal to $A$. Suppose also that $\{C_1\}$ is a countable subset of $\mathcal{C}$ with pairwise disjoint members whose union is equal to $A$. Moreover, as $\Omega \setminus A \in \mathcal{A}$ there are pairwise disjoint $W_1, \ldots, W_p \in \mathcal{C}$ such that $\Omega \setminus A = \bigcup_{i=1}^{p} W_i$. On the one hand, $W_1, \ldots, W_p, B_1, \ldots, B_j$ are pairwise disjoint cylinder sets with union $\Omega$, so
\[ \sum_{i=1}^{j} P(B_i) + \sum_{i=1}^{p} P(W_i) = 1. \]
On the other hand, $W_1, \ldots, W_p, C_1, C_2, \ldots$ are pairwise disjoint cylinder sets with union $\Omega$, so
\[ \sum_{i=1}^{\infty} P(C_i) + \sum_{i=1}^{p} P(W_i) = 1. \]
Hence,
\[ \sum_{i=1}^{j} P(B_i) = \sum_{i=1}^{\infty} P(C_i); \]
this conclusion does not involve $W_1, \ldots, W_p$. Thus it makes sense to define $\tau(A)$ to be this common value, and this defines a function $\tau : \mathcal{A} \to [0, 1]$. For $C \in \mathcal{C}$, $\tau(C) = P(C)$, i.e. the restriction of $\tau$ to $P$ is equal to $\mathcal{C}$.

If $\{A_n\}$ is a countable subset of $\mathcal{A}$ whose members are pairwise disjoint and $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, for each $n$ let $C_{n,1}, \ldots, C_{n,j(n)} \in \mathcal{C}$ be pairwise disjoint cylinder sets with union $A_n$. Then
\[ \{C_{n,i} : n \geq 1, 1 \leq i \leq j(n)\} \]
is a countable subset of $\mathcal{C}$ whose elements are pairwise disjoint and with union $A$, so
\[ \tau(A) = \sum_{n=1}^{\infty} \sum_{i=1}^{j(n)} P(C_{n,i}). \]
But for each $n$,
\[ \tau(A_n) = \sum_{i=1}^{j(n)} P(C_{n,i}), \]
\[ \tau(A) = \sum_{n=1}^{\infty} \tau(A_n). \]

This shows that \( \tau : \mathcal{A} \rightarrow [0,1] \) is a measure. Then applying Hopf’s extension theorem, we get that there is a unique measure \( \mu : \sigma(\mathcal{A}) \rightarrow [0,1] \) whose restriction to \( \mathcal{A} \) is equal to \( \tau \). It is apparent that the \( \sigma \)-algebra generated by a semialgebra is equal to the \( \sigma \)-algebra generated by the algebra generated by the semialgebra, so \( \sigma(\mathcal{A}) = \sigma(\mathcal{C}) \). Because the restriction of \( \tau \) to \( \mathcal{C} \) is equal to \( P \), the restriction of \( \mu \) to \( \mathcal{C} \) is equal to \( P \). Now suppose that \( \nu : \sigma(\mathcal{A}) \rightarrow [0,1] \) is a measure whose restriction to \( \mathcal{C} \) is equal to \( P \). For \( A \in \mathcal{A} \), there are disjoint \( C_1, \ldots, C_n \in \mathcal{C} \) with \( A = \bigcup_{i=1}^{n} C_i \). Then

\[ \nu(A) = \sum_{i=1}^{n} \nu(C_i) = \sum_{i=1}^{n} P(C_i) = \sum_{i=1}^{n} \mu(C_i) = \mu(A), \]

showing that the restriction of \( \nu \) to \( \mathcal{A} \) is equal to the restriction of \( \mu \) to \( \mathcal{A} \), from which it follows that \( \nu = \mu \). \( \square \)

## 5 Product measures

Suppose that \( \{ (\Omega_i, \mathcal{F}_i, P_i) : i \in I \} \) is a nonempty collection of probability spaces. The **product \( \sigma \)-algebra** is \( \sigma(\mathcal{C}) \), the \( \sigma \)-algebra generated by the cylinder sets. We define \( P : \mathcal{C} \to [0,1] \) by

\[ P(A) = \prod_{i \in I_A} P_i(A_i) = \prod_{i \in I} P_i(A_i), \]

for \( A \in \mathcal{C} \) and with \( I_A = \{ i \in I : A_i \neq \Omega_i \} \), which is finite.

**Lemma 3.** Suppose that \( I \) is the set of positive integers. If \( \{ A_n \} \) is a countable subset of \( \mathcal{C} \) with pairwise disjoint elements whose union is equal to \( \Omega \), then

\[ \sum_{n=1}^{\infty} P(A_n) = 1. \]

**Proof.** For each \( k \geq 1 \), there is some \( i_k \) and \( A_{k,1} \in \mathcal{F}_1, \ldots, A_{k,i_k} \in \mathcal{F}_{i_k} \) such that

\[ A_k = \prod_{i=1}^{\infty} A_{k,i}, \]

with \( A_{k,i} = \Omega_i \) for \( i > i_k \). Let \( m \geq 1 \), let \( x = (x_i) \in A_m \), and let \( n \geq 1 \). If \( n = m \),

\[ \left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i > i_m} P_i(A_{n,i}) \right) = 1 = \delta_{m,n}. \]
If \( m \neq n \) and \( y_i \in \Omega_i \) for each \( i > i_m \) and we set \( y_i = x_i \) for \( 1 \leq i \leq i_m \), then because \( A_m \) and \( A_n \) are disjoint and \( y \in A_m \), we have \( y \notin A_n \) and therefore there is some \( i, 1 \leq i \leq i_n \), such that \( y_i \notin A_{n,i} \). Thus

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} \chi_{A_{n,i}}(y_i) \right) = \prod_{i=1}^{\infty} \chi_{A_{n,i}}(y_i) = 0. \tag{1}
\]

Either \( i_n \leq i_m \) or \( i_n > i_m \). In the case \( i_n \leq i_m \) we have \( A_{n,i} = \Omega_i \) for \( i > i_m \) and thus

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} \chi_{A_{n,i}}(y_i) \right) = \prod_{i=i_m+1}^{i_n} \chi_{A_{n,i}}(x_i),
\]

hence by (1),

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} P_1(A_{n,i}) \right) = \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) = 0 = \delta_{m,n}.
\]

In the case \( i_n > i_m \), we have \( A_{n,i} = \Omega_i \) for \( i > i_n \) and thus

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} \chi_{A_{n,i}}(y_i) \right) = \left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=i_m+1}^{i_n} \chi_{A_{n,i}}(y_i) \right),
\]

hence by (1) we have that for \( y_i \in \Omega_i, i > i_m \),

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=i_m+1}^{i_n} \chi_{A_{n,i}}(y_i) \right) = 0.
\]

Therefore, integrating over \( \Omega_i \) for \( i = i_m + 1, \ldots, i_n \),

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=i_m+1}^{i_n} P_1(A_{n,i}) \right) = 0,
\]

so

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} P_1(A_{n,i}) \right) = 0 = \delta_{m,n}.
\]

We have thus established that for any \( m \geq 1, x \in A_m \), and \( n \geq 1, \)

\[
\left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} P_1(A_{n,i}) \right) = \delta_{m,n}. \tag{2}
\]

Suppose by contradiction that

\[
\sum_{n=1}^{\infty} P(A_n) < 1,
\]
i.e.

\[ \sum_{n=1}^{\infty} \prod_{i=1}^{\infty} P_i(A_{n,i}) < 1. \]  \hfill (3)

If

\[ \sum_{n=1}^{\infty} \chi_{A_{n,1}}(x_1) \prod_{i=2}^{\infty} P_i(A_{n,i}) = 1 \]

for all \( x_1 \in \Omega_1 \), then integrating over \( \Omega_1 \) we contradict (3). Hence there is some \( x_1 \in \Omega_1 \) such that

\[ \sum_{n=1}^{\infty} \chi_{A_{n,1}}(x_1) \prod_{i=2}^{\infty} P_i(A_{n,i}) < 1. \]

Suppose by induction that for some \( j \geq 1 \), \( x_1 \in \Omega_1, \ldots, x_j \in \Omega_j \) and

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{j} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=j+1}^{\infty} P_i(A_{n,i}) \right) < 1. \]

If

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{j+1} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=j+2}^{\infty} P_i(A_{n,i}) \right) = 1 \]

for all \( x_{j+1} \in \Omega_{j+1} \), then integrating over \( \Omega_{j+1} \) we contradict (3). Hence there is some \( x_{j+1} \in \Omega_{j+1} \) such that

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{j+1} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=j+2}^{\infty} P_i(A_{n,i}) \right) < 1. \]

Therefore, by induction we obtain that for any \( j \geq 1 \), there are \( x_1 \in \Omega_1, \ldots, x_j \in \Omega_j \) such that

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{j} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i=j+1}^{\infty} P_i(A_{n,i}) \right) < 1. \]  \hfill (4)

Write \( x = (x_1, x_2, \ldots) \in \Omega \). Because \( \Omega = \bigcup_{m=1}^{\infty} A_m \), there is some \( m \) for which \( x \in A_m \). For \( j = i_m \), (4) states

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} P_i(A_{n,i}) \right) < 1. \]

But (2) tells us

\[ \sum_{n=1}^{\infty} \left( \prod_{i=1}^{i_m} \chi_{A_{n,i}}(x_i) \right) \left( \prod_{i>i_m} P_i(A_{n,i}) \right) = \sum_{n=1}^{\infty} \delta_{m,n} = 1, \]
a contradiction. Therefore,
\[ \sum_{n=1}^{\infty} P(A_n) = 1, \]
proving the claim. \(\square\)

**Lemma 4.** Suppose that \(I\) is an uncountable set. If \(\{A_n\}\) is a countable subset
of \(\mathcal{C}\) with pairwise disjoint elements whose union is equal to \(\Omega\), then
\[ \sum_{n=1}^{\infty} P(A_n) = 1. \]

**Proof.** For each \(n\), there are \(A_{n,i} \in \mathcal{F}_i\) with \(A_{n,i} = \Omega_i\), and \(I_n = \{i \in I : A_i \neq \Omega_i\}\) is finite. Then \(J = \bigcup_{n=1}^{\infty} I_n\) is countable. Let \(\Omega_J = \prod_{i \in J} \Omega_i\), let \(\mathcal{C}_J\) be the collection of cylinder sets corresponding to the probability spaces \(\{(\Omega_i, \mathcal{F}_i, P_i) : i \in J\}\), and define \(P_J : \mathcal{C}_J \to [0,1]\) by
\[ P_J(B) = \prod_{i \in J_B} P_i(B_i) = \prod_{i \in J} P_i(B_i), \]
for \(B \in \mathcal{C}_J\) and with \(J_B = \{i \in J : B_i \neq \Omega_i\}\), which is finite. \(P_J\) satisfies
\[ P_J(B) = P\left( B \times \prod_{i \in I \setminus J} \Omega_i \right), \quad B \in \mathcal{C}_J. \]

Let \(B_n = \prod_{i \in J} A_{n,i}\), i.e. \(A_n = B_n \times \prod_{i \in I \setminus J} A_{n,i}\). Then \(\{B_n\}\) is a countable
subset of \(\mathcal{C}_J\) with pairwise disjoint elements whose union is equal to \(\Omega_J\), and applying Lemma 3 we get that
\[ \sum_{n=1}^{\infty} P_J(B_n) = 1, \]
and therefore
\[ \sum_{n=1}^{\infty} P(A_n) = 1. \]
\(\square\)

Now by Lemma 2 and the above lemma, there is a unique probability measure \(\mu\) on \(\sigma(\mathcal{C})\) whose restriction to \(\mathcal{C}\) is equal to \(P\). That is, when \(\{(\Omega_i, \mathcal{F}_i, P_i) : i \in I\}\) are probability spaces and \(\mathcal{C}\) is the collection of cylinder sets corresponding
to these probability spaces, with \(\Omega = \prod_{i \in I} \Omega_i\) and \(P : \mathcal{C} \to [0,1]\) defined by
\[ P(A) = \prod_{i \in I} P(\{A_i\}) \]
for $A = \prod_{i \in I} A_i \in \mathcal{C}$, then there is a unique probability measure $\mu$ on the the product $\sigma$-algebra such that $\mu(A) = P(A)$ for each cylinder set $A$. We call $\mu$ the **product measure**, and write

$$\bigotimes_{i \in I} \mathcal{F}_i = \sigma(\mathcal{C})$$

and

$$\prod_{i \in I} P_i = \mu.$$