A Conditional Product Measure Theorem

J.M. Swart

November 28, 2000

J.M. Swart[1]
Theoretical Physics, University of Groningen, Nijenborgh 4,
9747 AG Groningen, the Netherlands

Key words: conditional independence, product measure

Abstract

Under a suitable regularity assumption, a generalisation of the product measure theorem to the case of conditional independence is given. A counterexample is constructed to show that some assumption (e.g. regularity) is needed.

1 A conditional product measure theorem

In the framework of product spaces, the concept of independence is strongly linked to the product measure theorem. A well–known generalisation of independence is the concept of conditional independence. Although many of its elementary properties have been studied, nobody so far seems to have been concerned with generalising the product measure theorem to the case of conditional independence.

Assume that Ω_0, Ω_1 and Ω_2 are measurable sets, with σ–fields $\mathcal{S}_0$, $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Define projections $\pi_{012,i}(\omega_0, \omega_1, \omega_2) = \omega_i$ from the product space $\Omega_0 \times \Omega_1 \times \Omega_2$ on Ω_i. The product–σ–field $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$ is generated by the projections $\pi_{012,i}$, with $i = 0, 1, 2$. Assume that μ is some probability measure on $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$. The conditional probability of an event $A \in \mathcal{S}_1 \otimes \mathcal{S}_2$ given

[1] Jan M. Swart, Rhijnvis Feithplein 1^{A}, 9715 GH Groningen, the Netherlands

ω_0 is defined as an equivalence class $P(A)$ of functions $P(A)(\omega_0)$, satisfying

$$\int_B P(A)\,\mu\circ\pi_{012,0}^{-1}(d\omega_0) = \mu(B\times A) \tag{1}$$

for each $B \in \mathcal{S}_0$. Here, $B\times A$ is used to denote the set $\{(\omega_0,\omega_1,\omega_2) \mid \omega_0 \in B,\ (\omega_1,\omega_2) \in A\}$, $\pi_{012,0}^{-1}$ denotes the inverse image and $\mu\circ\pi_{012,0}^{-1}$ is the measure induced by $\pi_{012,0}$ on Ω_0. Denote the set of all equivalence classes of a.e.$[\mu\circ\pi_{012,0}^{-1}]$ equal non–negative measurable functions of the variable ω_0 by the symbol $L^+(\Omega_0)$. Then P is a vector measure on $\mathcal{S}_1\otimes\mathcal{S}_2$ with values in $L^+(\Omega_0)$.

We say that with respect to μ the variables ω_1 and ω_2 are conditionally independent given ω_0, if for all $A_1\in\mathcal{S}_1$, $A_2\in\mathcal{S}_2$:

$$P(A_1\times A_2) = P(A_1\times\Omega_2)P(\Omega_1\times A_2) \tag{2}$$

For a comparable definition, see Loève (1978).

If Ω_0 consists of just one point, the product σ–field $\mathcal{S}_0\otimes\mathcal{S}_1\otimes\mathcal{S}_2$ becomes isomorphic to $\mathcal{S}_1\otimes\mathcal{S}_2$, and conditional independence reduces to independence. In this case the well–known product measure theorem tells us that the measure μ is determined uniquely by its projections on Ω_1 and Ω_2: Given probability measures μ_1 and μ_2 on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, there exists a unique probability measure μ on $\mathcal{S}_1\otimes\mathcal{S}_2$ such that ω_1 and ω_2 are independent, and $\mu\circ\pi_{12,1}^{-1} = \mu_1$, $\mu\circ\pi_{12,2}^{-1} = \mu_2$; this is just the product measure $\mu = \mu_1\times\mu_2$.

In general, in the case of conditional independence, assume that we have probability measures μ_{01} and μ_{02} on $\mathcal{S}_0\otimes\mathcal{S}_1$ and $\mathcal{S}_0\otimes\mathcal{S}_2$ respectively, and that we are looking for a μ such that

$$\begin{aligned}\mu\circ\pi_{012,01}^{-1} &= \mu_{01} &(3)\\ \mu\circ\pi_{012,02}^{-1} &= \mu_{02} &(4)\end{aligned}$$

here, $\pi_{012,01}(\omega_0,\omega_1,\omega_2) = (\omega_0,\omega_1)$; $\pi_{012,02}$ is defined similarly. If such a μ exists then clearly we must have

$$\mu_{01}\circ\pi_{01,0}^{-1} = \mu\circ\pi_{012,0}^{-1} = \mu_{02}\circ\pi_{02,0}^{-1} \tag{5}$$

Put $\mu_0 = \mu_{01}\circ\pi_{01,0}^{-1} = \mu_{02}\circ\pi_{02,0}^{-1}$. Let $P^1(A_1)$ be the conditional probability of an event $A_1\in\mathcal{S}_1$ given ω_0, with respect to μ_{01}. Similarly let $P^2(A_2)$ be defined with respect to μ_{02}. If P is defined with respect to μ as above, then we see that $P^1(A_1) = P(A_1\times\Omega_2)$, and similarly $P^2(A_2) = P(\Omega_1\times A_2)$. From conditional independence it follows that for each $A_0\in\mathcal{S}_0\ldots A_2\in\mathcal{S}_2$:

$$\mu(A_0\times A_1\times A_2) = \int_{A_0} P^1(A_1)P^2(A_2)\,\mu_0(d\omega_0) \tag{6}$$

The cubes $A_0 \times A_1 \times A_2$ form a π–system in the sense of Billingsley (1986). Since it generates the σ–field $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$, any two probability measures that coincide on this class are equal. Inspection of formula (6) thus shows that on the assumption of conditional independence the measure μ is determined uniquely by its projections $\mu \circ \pi_{012,01}^{-1}$ and $\mu \circ \pi_{012,02}^{-1}$.

Conversely, if we can find a measure μ such that (6), then it is easy to see that (2), (3) and (4) are fulfilled. Thus, the measures μ_{01} and μ_{02} define a unique measure μ if and only if they satisfy the agreement requirement (5), and if the function μ defined on cubes $A_0 \times A_1 \times A_2$ by formula (6) can be extended to a measure on $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$.

For the existence of such a measure μ, it is sufficient if there exists a product vector measure $P : \mathcal{S}_1 \otimes \mathcal{S}_2 \to L^+(\Omega_0)$ such that $P(A_1 \times A_2) = P^1(A_1)P^2(A_2)$. For let P be such a vector measure, then we can define μ by the formula

$$\mu(A) = \int_{\Omega_0} P(A_{\omega_0})\mu_0(d\omega_0) \tag{7}$$

Here, A_{ω_0} denotes the section $\{(\omega_1, \omega_2) \mid (\omega_0, \omega_1, \omega_2) \in A\}$. It is easily verified that μ is a measure and fulfills (6), and therefore is the desired measure.

Thus, the problem of constructing the measure μ from its projections μ_{01} and μ_{02} is reduced to the problem of finding a product vector measure $P^1 \times P^2$. In general such a product vector measure need not exist. I will prove a theorem on the assumption that either P^1 or P^2 has a regular version. I will construct a counterexample to show that some extra condition is indeed needed.

Theorem. *Let Ω_0, Ω_1 and Ω_2 be measurable sets, with σ–fields $\mathcal{S}_0$, $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Define projections $\pi_{012,01}$, $\pi_{01,0}$ and so on as above. Let μ_{01} be a probability measure on $\mathcal{S}_0 \otimes \mathcal{S}_1$, and let μ_{02} be a probability measure on $\mathcal{S}_0 \otimes \mathcal{S}_2$, such that*

$$\mu_{01} \circ \pi_{01,0}^{-1} = \mu_{02} \circ \pi_{02,0}^{-1} \tag{8}$$

Let $P^1(A_1)$ be the conditional probability of an event $A_1 \in \mathcal{S}_1$ given ω_0, with respect to μ_{01}. Similarly let $P^2(A_2)$ be defined with respect to μ_{02}. Assume that either P^1 or P^2 has a regular version. Then there exists a unique probability measure μ on $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$ such that

$$\mu \circ \pi_{012,01}^{-1} = \mu_{01} \tag{9}$$

$$\mu \circ \pi_{012,02}^{-1} = \mu_{02} \tag{10}$$

and such that with respect to μ, ω_1 and ω_2 are conditionally independent given ω_0.

Proof. It is sufficient to construct a product vector measure P such that $P(A_1 \times A_2) = P^1(A_1)P^2(A_2)$. The principle of the construction is based on the idea of repeated integrals. Using a conditional expectation I will perform an integration over ω_1, keeping ω_2 fixed. Then I will use a Fubini–like formula.

The situation is symmetric between 1 and 2. Assume that $\tilde{P}^1$ is a regular version of the conditional probability P^1. This means that $\tilde{P}^1(A_1)(\omega_0)$, for any fixed ω_0, is a probability measure on $\mathcal{S}_1$ which I will write as $\tilde{P}^1(\cdot)(\omega_0)$. Define the conditional expectation

$$E^1(f)(\omega_0) = \int_{\Omega_1} f\, \tilde{P}^1(d\omega_1)(\omega_0) \tag{11}$$

$E^1(f)$ is well–defined for all non–negative measurable functions $f(\omega_1)$ that are functions of the variable ω_1. Now extend the definition of E^1 to functions $f(\omega_1, \omega_2)$ that depend also on ω_2 by letting it operate with ω_2 fixed, i.e. define

$$E(f)(\omega_0, \omega_2) = E^1(f(\cdot, \omega_2))(\omega_0) \tag{12}$$

where $f(\cdot, \omega_2)$ is the function f considered as a function of ω_1, with ω_2 fixed. If f is measurable with respect to $\mathcal{S}_1 \otimes \mathcal{S}_2$, then $E(f)$, considered as a function of ω_0 and ω_2, is measurable with respect to $\mathcal{S}_0 \otimes \mathcal{S}_2$. Here, regularity is crucial. We will only need the case where f is some characteristic function.

We apply the π/λ–theorem (see Billingsley, 1986, theorem 3.2). Define the π–system $\mathcal{P} = \{A_1 \times A_2 \mid A_1 \in \mathcal{S}_1,\ A_2 \in \mathcal{S}_2\}$. $\mathcal{P}$ generates $\mathcal{S}_1 \otimes \mathcal{S}_2$. Let I_A denote the characteristic function of a set A. Define $\mathcal{G}$ as the class of all sets A in $\mathcal{S}_1 \otimes \mathcal{S}_2$ such that $E(I_A)$ is measurable. We verify that $\mathcal{G}$ is a λ–system; for example, if the A_i form an increasing row of measurable sets, $A_i \uparrow A$, then $E(I_{A_i}) \uparrow E(I_A)$ by the continuity of the measure $\tilde{P}^1(\cdot)(\omega_0)$, so $E(I_A)$ is measurable. We verify furthermore that $\mathcal{G}$ contains $\mathcal{P}$: Let $A_1 \times A_2 \in \mathcal{P}$ then $E(I_{A_1 \times A_2})(\omega_0, \omega_2) = E^1(I_{A_1})(\omega_0) \cdot I_{A_2}(\omega_2)$, which is measurable. By the π/λ–theorem we conclude that $\mathcal{G}$ contains $\mathcal{S}_1 \otimes \mathcal{S}_2$.

Define the conditional expectation $E^2(f)$ with respect to μ_{02} by

$$\int_B E^2(f) \mu_{02} \circ \pi_{02,0}^{-1}(d\omega_0) = \int_{\Omega_0 \times B} f\, \mu_{02}(d\omega_0 d\omega_2) \quad \forall B \in \mathcal{S}_2 \tag{13}$$

E^2 is well-defined on non–negative measurable functions of the variables ω_0 and ω_2, has values in $L^+(\Omega_0)$, and is countably linear. We have $E^2(f_0 I_{A_2}) = f_0 P^2(A_2)$ for each $A_2 \in \mathcal{S}_2$, and f_0 an (equivalence class of) non–negative measurable function(s) of the variable ω_0 alone. Now define the vector measure P as follows

$$P(A) = E^2(E(I_A)) \tag{14}$$

We check that P is indeed a vector measure, and that for all $A_1 \in \mathcal{S}_1$, $A_2 \in \mathcal{S}_2$: $P(A_1 \times A_2) = E^2(E(I_{A_1 \times A_2})) = E^2(P^1(A_1) I_{A_2}) = P^1(A_1) P^2(A_2)$. Thus, P is the desired product vector measure.

2 Two measures with no common extension

Counterexample. *There exist measurable sets $\Omega_0, \Omega_1, \Omega_2$, with σ–fields $\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2$, such that there exist probability measures μ_{01} and μ_{02} on $\mathcal{S}_0 \otimes \mathcal{S}_1$ and $\mathcal{S}_0 \otimes \mathcal{S}_2$ respectively, that satisfy:*

$$\mu_{01} \circ \pi_{01,0}^{-1} = \mu_{02} \circ \pi_{02,0}^{-1} \tag{15}$$

while there exists no measure μ on $\mathcal{S}_0 \otimes \mathcal{S}_1 \otimes \mathcal{S}_2$ with the property that

$$\mu \circ \pi_{012,01}^{-1} = \mu_{01} \tag{16}$$

$$\mu \circ \pi_{012,02}^{-1} = \mu_{02} \tag{17}$$

Note that no reference is made to conditional independence here.

Proof The proof is an adaptation of a counterexample by Dudley (1973) concerning spectral measures. Choose for Ω_0 the interval $[0,1]$. Let X be a subset of $[0,1]$ that has Lebesgue inner measure zero, and outer measure 1: $\lambda_*(X) = 0$, $\lambda^*(X) = 1$. (See Halmos, 1950.) Put $Y = [0,1] - X$. Choose $\Omega_1 = X$, and $\Omega_2 = Y$. Let $\mathcal{S}_0 = \mathcal{B}[0,1]$, the Borel sets on $[0,1]$. Define $\mathcal{S}_1 = \mathcal{B}[0,1] \cap X$ and $\mathcal{S}_2 = \mathcal{B}[0,1] \cap Y$. Define measures μ_1 and μ_2 on $\mathcal{S}_1$, $\mathcal{S}_2$ respectively by putting $\mu_1(A \cap X) = \lambda(A)$ and $\mu_2(A \cap Y) = \lambda(A)$. Define the map $t_1 : \Omega_1 \to \Omega_0 \times \Omega_1$ by $t_1(\omega_1) = (\omega_1, \omega_1)$. Similarly put $t_2(\omega_2) = (\omega_2, \omega_2)$. Now define the measures μ_{01} and μ_{02} as follows

$$\mu_{01} = \mu_1 \circ t_1^{-1} \tag{18}$$

$$\mu_{02} = \mu_2 \circ t_2^{-1} \tag{19}$$

Consider the measurable sets $N_{01} = \{(\omega_0, \omega_1) \mid \omega_0 \neq \omega_1\}$ and $N_{02} = \{(\omega_0, \omega_2) \mid \omega_0 \neq \omega_2\}$. We see that $\mu_{01}(N_{01}) = \mu_1 \circ t_1^{-1}(N_{01}) = \mu_1(\emptyset) = 0$. Similarly $\mu_{02}(N_{02}) = 0$. Imagine that a measure μ exists such that (15) holds. Then on the one hand we must have $\mu(\pi_{012,01}^{-1}(N_{01})) = 0$ and $\mu(\pi_{012,02}^{-1}(N_{02})) = 0$, so that

$$\mu(\pi_{012,01}^{-1}(N_{01}) \cup \pi_{012,02}^{-1}(N_{02})) = 0 \tag{20}$$

While on the other hand we have $\pi_{012,01}^{-1}(N_{01}) \cup \pi_{012,02}^{-1}(N_{02}) = \Omega_0 \times \Omega_1 \times \Omega_2$, so that we must have

$$\mu(\pi_{012,01}^{-1}(N_{01}) \cup \pi_{012,02}^{-1}(N_{02})) = 1 \tag{21}$$

We arrive at a contradiction. The idea is that the measure μ_{01} implies $\omega_0 = \omega_1$ almost surely, while μ_{02} implies $\omega_0 = \omega_2$ almost surely. Since $\omega_1 \neq \omega_2$ this leads to a paradox.

Acknowledgments I would like to thank prof. dr. A.J. Stam and prof. dr. E.G.F. Thomas who showed me the way to proof and refutation. Moreover Aernout van Enter and David Atkinson who carefully read the manuscript.

References

Loève, M. (1978), *Probability Theory I & II* (Springer Verlag, New York, 4th ed.)

Billingsley, P. (1986), *Probability and Measure* (Wiley, New York)

Dudley, R.M. (1973), A note on products of spectral measures, in: D.H. Tucker and H.B. Maynard, eds., *Vector and Operator Valued Measures and Applications* (Academic Press, New York) pp. 125–126.

Halmos, P.R. (1950), *Measure Theory* (Van Nostrand Reinhold Company, New York)