Transformations

Continuing Example 13.6 we discussed that the hexadecimal digits were generated by an underlying source – the $8 \times 400 = 3200$ waiting times between successive counts by the Geiger counter close to a radioactive material.

Each 8-tuple of waiting times might be represented on $(\mathbb{R}^8, \mathbb{B}_8)$. An outcome is a vector

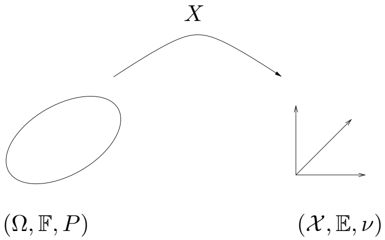
$$(x_1, \ldots, x_8).$$

What we actually considered was a transformed experiment where we consider a map

$$t : \mathbb{R}^8 \to \{0, 1, \ldots, 9, A, B, C, D, E, F\}$$

.

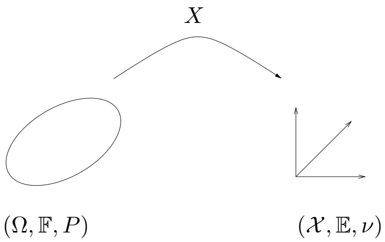
Abstract background representation



An experiment is represented by $(\mathcal{X}, \mathbb{E}, \nu)$. Imagine that there is a representation space $(\Omega, \mathbb{F}, P)$ and an underlying measurable map $X : \Omega \to \mathcal{X}$ with $X(P) = \nu$.

Then we call X a stochastic variable and we call the probability measure $\nu = X(P)$ the distribution (fordeling) of X.

One explanation

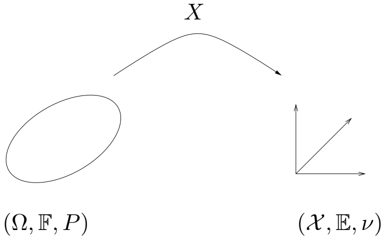


Ω is the space of all possible destinies for the entire world. The Goddess of Fortune chooses one from the probability measure P.

The map X represents the translation from the destiny for the entire world $\omega \in \Omega$ to the outcome $x = X(\omega)$ of our more humble experiment.

This is philosophical mumbo jumbo without substance.

Abstract background representation

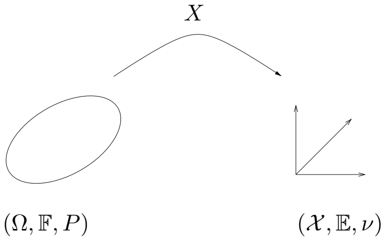


It's trivial that any given representation space $(\mathcal{X}, \mathbb{E}, \nu)$ can be cast into the abstract framework by $\Omega = \mathcal{X}$, $\mathbb{F} = \mathbb{E}$, $P = \nu$ and the identity map

$$X = \mathrm{Id} : x \mapsto x.$$

But one would get into trouble in the attempt to construct a single universal representation space $(\Omega, \mathbb{F}, P)$ that could be used for all purposes.

Abstract background representation

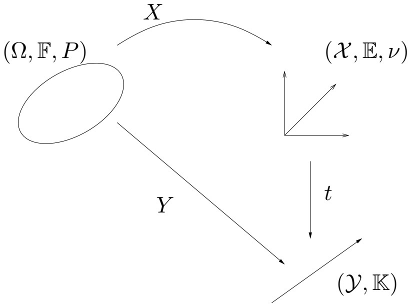


The better reason to introduce the abstract background Ω is the superior mathematical notation for handling transformations. We say

1) let X take values in the space $\mathcal{X}$
2) let the distribution of X be ν (the image measure is $X(P) = \nu$)

What we are really interested in is the distribution of $t(X) = t \circ X$ for some measurable $t : \mathcal{X} \to \mathcal{Y}$.

Abstract background representation



The transformed experiment given by $Y = t(X)$ has distribution

$$Y(P) = t \circ X(P) = t(X(P)) = t(\nu)$$

if the stochastic variable X has distribution ν.

Functions given as integrals

Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space, $\mathcal{Y}$ any set and $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ a function whose section function (snitfunktion)

$$x \mapsto f(x, y)$$

is either in $\mathcal{M}^+$ or in $\mathcal{L}$ for all $y \in \mathcal{Y}$ then we can define a function $\varphi : \mathcal{Y} \to \mathbb{R}^*$ by

$$\varphi(y) = \int f(x, y) \mathrm{d}\mu.$$

If $\mathcal{Y}$ is equipped with the σ-algebra $\mathbb{K}$ we ask if φ is $\mathbb{K}$-$\mathbb{B}$-measurable?

Theorem 4.22: If $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$ are two measurable spaces then the paving

$$\mathbb{E} \times \mathbb{K} = \{A \times B \mid A \in \mathbb{E}, \ B \in \mathbb{K}\}$$

is a generator for the product σ-algebra $\mathbb{E} \otimes \mathbb{K}$, which is stable under intersections.

Integrals of simple functions

Let $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$ be two measurable spaces.

The simplest simple, measurable functions on $\mathcal{X} \times \mathcal{Y}$ are indicators of product sets

$$f(x,y) = 1_{A \times B}(x,y) = 1_A(x) 1_B(y)$$

for $(x,y) \in \mathcal{X} \times \mathcal{Y}$, $A \in \mathbb{E}$ and $B \in \mathbb{K}$.

- f is $\mathbb{E} \otimes \mathbb{K}$-$\mathbb{B}$-measurable since $A \times B \in \mathbb{E} \otimes \mathbb{K}$.
- If μ is a measure on $(\mathcal{X}, \mathbb{E})$ the function

$$\varphi(y) = \int f(x,y) \mathrm{d}\mu(x) = \mu(A) 1_B(y)$$

is obviously $\mathbb{K}$-$\mathbb{B}$-measurable.

Extensions

Can we extend the result to get measurability of

$$\varphi(y) = \int 1_G(x, y) \mathrm{d}\mu(x)$$

for $G \in \mathbb{E} \times \mathbb{K}$? If we can, we find for a simple function $s \in \mathcal{S}^+(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ that the function

$$\varphi(y) = \int s(x, y) \mathrm{d}\mu(x) = \sum_{i=1}^{n} c_i \int 1_{G_i}(x, y) \mathrm{d}\mu(x)$$

is measurable. By approximating $f \in \mathcal{M}^+(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ by a sequence of simple functions $s_n \nearrow f$ we find by monotone convergence that

$$\varphi(y) = \int f(x, y) \mathrm{d}\mu(x) = \lim_{n \to \infty} \int s_n(x, y) \mathrm{d}\mu(x)$$

is in $\mathcal{M}^+(\mathcal{Y}, \mathbb{K})$.

Extensions

Finally by the decomposition

$$f(x,y) = f^+(x,y) - f^-(x,y)$$

for $f \in \mathcal{M}(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ with μ-integrable sections $x \mapsto f(x,y)$ for all $y \in \mathcal{Y}$ the function

$$\varphi(y) = \int f(x,y)\mathrm{d}\mu(x) = \int f^+(x,y)\mathrm{d}\mu(x) - \int f^-(x,y)\mathrm{d}\mu(x)$$

is in $\mathcal{M}(\mathcal{Y}, \mathbb{K})$.

Measurability of indicators

The real problem is in fact to establish measurability of

$$\varphi(y) = \int 1_G(x, y) \mathrm{d}\mu(x)$$

for $G \in \mathbb{E} \otimes \mathbb{K}$.

The argument is a Dynkin class argument that lifts the result from the obvious indicators $1_{A \times B}$ of products to all indicators 1_G for $G \in \mathbb{E} \otimes \mathbb{K}$.

Product space measurability

With $(\mathcal{X}_i, \mathbb{E}_i)$ for $i = 1, \ldots, k$ measurable spaces we recall that the product-σ-algebra $\mathbb{E}_1 \otimes \ldots \otimes \mathbb{E}_k$ on $\mathcal{X}_1 \times \ldots \mathcal{X}_k$ as the σ-algebra generated by the coordinate projections $\hat{X}_i : \mathcal{X}_1 \times \ldots \mathcal{X}_k \to \mathcal{X}_i$,

$$\hat{X}_i(x_1, \ldots, x_k) = x_i.$$

In symbols

$$\mathbb{E}_1 \otimes \ldots \otimes \mathbb{E}_k = \sigma(\hat{X}_1, \ldots, \hat{X}_k).$$

Observe that for $A_i \in \mathbb{E}_i$, $i = 1, \ldots, k$ then

$$X_i^{-1}(A_i) = \mathcal{X}_1 \times \ldots \mathcal{X}_{i-1} \times A_i \times \mathcal{X}_{i+1} \times \ldots \times \mathcal{X}_k$$

and

$$A_1 \times \ldots \times A_k = X_1^{-1}(A_1) \cap \ldots \cap X_k^{-1}(A_k).$$

Generator of the product sigma-algebra

Theorem 4.22: If $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$ are two measurable spaces then the paving

$$\mathbb{E} \times \mathbb{K} = \{A \times B \mid A \in \mathbb{E},\ B \in \mathbb{K}\}$$

is a generator for the product σ-algebra $\mathbb{E} \otimes \mathbb{K}$, which is stable under intersections.

In general,

$$\{A_1 \times \ldots \times A_k \mid A_i \in \mathbb{E}_i,\ i = 1, \ldots, k\}$$

is a generator for $\mathbb{E}_1 \otimes \ldots \otimes \mathbb{E}_k$, which is stable under intersections.

Product measures

Let $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$ be two measurable spaces.

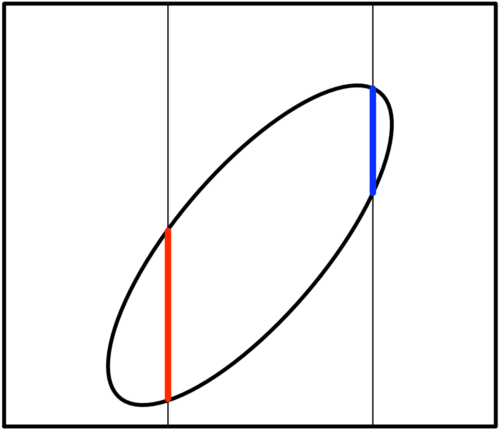
The next quest is to introduce the product measure of two measures μ and ν on $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$, respectively.

We want to introduce a measure $\mu \otimes \nu$ with

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$$

for all $A \in \mathbb{E}$ and $B \in \mathbb{K}$.

Measures of sections



Let ν be a measure on $(\mathcal{Y}, \mathbb{K})$. Then we can consider

$$x \mapsto \nu(G^x) = \int 1_G(x, y)\, d\nu(y)$$

Product measures

Let $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$ be two measurable spaces.

Theorem: If μ and ν are σ-finite measures on $(\mathcal{X}, \mathbb{E})$ and $(\mathcal{Y}, \mathbb{K})$, respectively, then there is a unique measure $\mu \otimes \nu$ on the product space $(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ satisfying that

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$$

for all $A \in \mathbb{E}$ and $B \in \mathbb{K}$.

Tonelli light

Observe that the existence proof was based on the construction of the measure λ by successive integration;

$$\lambda(G) = \int \nu(G^x)\mathrm{d}\mu(x) = \int \left(\int 1_G(x,y)\mathrm{d}\nu(y) \right) \mathrm{d}\mu(x).$$

An alternative candidate to the product measure is defined by integrating in the other order. Define

$$\tilde{\lambda}(G) = \int \mu(G^y)\mathrm{d}\nu(y) = \int \left(\int 1_G(x,y)\mathrm{d}\mu(x) \right) \mathrm{d}\nu(y).$$

Just as we show that $\lambda(A \times B) = \mu(A)\nu(B)$ we can show that $\tilde{\lambda}(A \times B) = \mu(A)\nu(B)$ so by uniqueness of the product measure we have shown the non-trivial result on interchange of the integration order for σ-finite measures:

$$\int \left(\int 1_G(x,y)\mathrm{d}\mu(x) \right) \mathrm{d}\nu(y) = \int \left(\int 1_G(x,y)\mathrm{d}\nu(y) \right) \mathrm{d}\mu(x)$$

Tonelli

Tonelli's Theorem: If $(\mathcal{X}, \mathbb{E}, \mu)$ and $(\mathcal{Y}, \mathbb{K}, \nu)$ are two σ-finite measure spaces and $f \in \mathcal{M}^+(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ then

$$\int f \mathrm{d}\mu \otimes \nu = \int \left(\int f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int \left(\int f(x,y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y).$$

One immediate consequence is that if $f \in \mathcal{M}^+(\mathcal{X}, \mathbb{E})$ and $g \in \mathcal{M}^+(\mathcal{Y}, \mathbb{K})$ then

$$\int fg \mathrm{d}\mu \otimes \nu = \left(\int f \mathrm{d}\mu \right) \left(\int g \mathrm{d}\nu \right),$$

or in words; the integral of a product w.r.t. a product measure is the product of the integrals.

Example 9.9

$$\int_0^1 \frac{1}{2}\frac{1}{\sqrt{1-y^2x}\sqrt{1-x}}\mathrm{d}x \overset{1}{=} \int_0^1 \frac{1}{1-z^2y^2}\mathrm{d}z \overset{2}{=} \sum_{n=0}^{\infty}\int_0^1 (zy)^{2n}\mathrm{d}z = \sum_{n=0}^{\infty}\frac{y^{2n}}{2n+1}$$

[1] a non-trivial but classical substitution

[2] interchanging the summation and integration by Cor. 6.20

Example 9.9

$$\int_0^1 \frac{1}{\sqrt{1-y^2x}}\mathrm{d}x \quad = \quad \frac{1}{\sqrt{x}}\arcsin(\sqrt{x})$$

since

$$\frac{\mathrm{d}}{\mathrm{d}w}\arcsin(w) = \frac{1}{\sqrt{1-w^2}}.$$

Tonelli

Tonelli's Theorem: If $(\mathcal{X}, \mathbb{E}, \mu)$ and $(\mathcal{Y}, \mathbb{K}, \nu)$ are two σ-finite measure spaces and $f \in \mathcal{M}^+(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ then

$$\int f \mathrm{d}\mu \otimes \nu = \int \left(\int f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int \left(\int f(x,y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y).$$

A product of Lebesgue measures is a Lebesgue measure:

$$m_k = \underbrace{m \otimes \ldots \otimes m}_{k \text{ times}}$$

and we can put parentheses as we like (associative law of $\otimes$). In particular $m_p \otimes m_q = m_{p+q}$, and we can use this to show that all hyperplanes in $\mathbb{R}^k$ are m_k-nullsets.

Fubini

Fubini's Theorem: If $(\mathcal{X}, \mathbb{E}, \mu)$ and $(\mathcal{Y}, \mathbb{K}, \nu)$ are two σ-finite measure spaces and if $f \in \mathcal{M}(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ is integrable w.r.t. $\mu \otimes \nu$ then

$$\int f \mathrm{d}\mu\otimes\nu = \int_A \left(\int f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int_B \left(\int f(x,y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y).$$

Here $A \in \mathbb{E}$ with $\mu(A^c) = 0$ and $B \in \mathbb{K}$ with $\nu(B^c) = 0$ are given by

$$A = \{x \in \mathcal{X} \mid \int |f(x,y)| \mathrm{d}\nu(y) < \infty\}$$

and

$$B = \{y \in \mathcal{Y} \mid \int |f(x,y)| \mathrm{d}\mu(x) < \infty\}.$$

Fubini in practice

Note that to use Fubini we need to first verify that f is integrable w.r.t. $\mu \otimes \nu$ and then compute the integral using either of the successive integration orders.

Verification of integrability follows from Tonelli's theorem.

- First verify that f is integrable by Tonelli

$$\int |f| \mathrm{d}\mu \otimes \nu = \int \left(\int |f(x,y)| \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) \leq \ldots < \infty$$

- then compute the value of the integral by Fubini

$$\int f \mathrm{d}\mu \otimes \nu = \int \left(\int f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x).$$

This is Fubinelli's theorem.

Sums and integrals

If $f_1, f_2, \ldots : \mathcal{X} \to \mathbb{R}$ are function we can define the function $f : \mathcal{X} \times \mathbb{N} \to \mathbb{R}$ by

$$f(x, n) = f_n(x).$$

Then f is $\mathbb{E} \otimes \mathbb{P}(\mathbb{N})$-$\mathbb{B}$-measurable if and only if all the functions f_n are $\mathbb{E}$-$\mathbb{B}$-measurable (Exercise 4.8).

Recall that integration w.r.t. the counting measure τ on $\mathbb{N}$ is the same as infinite sums. Thus

$$\sum_{n=1}^{\infty} f_n(x) = \int f(x, n) \mathrm{d}\tau(n)$$

whenever the sum and the integral make sense.

Interchanging sums and integrals

Theorem: If $f_1, f_2, \ldots : \mathcal{X} \to \mathbb{R}$ are measurable functions and μ is a measure on $(\mathcal{X}, \mathbb{E})$ then if

$$\sum_{n=1}^{\infty} \int |f_n| \mathrm{d}\mu < \infty$$

each f_n is integrable, the sum $\sum_{n=1}^{\infty} f_n(x)$ is finite for all $x \in A$ where

$$A = \{x \in \mathcal{X} \mid \sum_{n=1}^{\infty} |f_n(x)| < \infty\}$$

and $\mu(A^c) = 0$. Moreover, $\sum_{n=1}^{\infty} f_n(x)$ is μ-a.e. (on A) equal to an integrable function and

$$\sum_{n=1}^{\infty} \int f_n(x) \mathrm{d}\mu(x) = \int_A \sum_{n=1}^{\infty} f_n(x) \mathrm{d}\mu(x).$$

Interchanging sums and integrals

Proof: Fubinelli with the product measure $\mu \otimes \tau$. Catch, requires a priori μ to be σ-finite!

Workaround: The purpose of σ-finiteness is to assure uniqueness of the product measure as well as existence in terms of the technical measurability lemma 8.6. If we can check that by other means, Tonelli and Fubini applies! For all $A \in \mathbb{E} \otimes \mathbb{P}(\mathbb{N})$ we have the disjoint decomposition

$$A = \bigcup_{n=1}^{\infty} \mathcal{X} \times \{n\} \cap A = \bigcup_{n=1}^{\infty} A^n \times \{n\}, \quad A^n = \{x \in \mathcal{X} \mid (x, n) \in A\}$$

thus $\mu \otimes \tau(A) = \sum_{n=1}^{\infty} \mu(A^n)$, which defines this particular product measure and shows that it is uniquely specified by its values on product sets.

Alternative proof: Use dominated convergence as in the proof of Theorem 7.10.

Interchanging the order of integration

Using Fubini on

$$f(x,y) = \sin x \, e^{-xy} \qquad \text{for } x \in (0,K), y \in (0,\infty)$$

gives

$$\int_0^K \frac{\sin x}{x} \, dx = \int_0^\infty \frac{1}{1+y^2} \, dy + \dots$$

We find from this that

$$\int_0^K \frac{\sin x}{x} \, dx \to \frac{\pi}{2} \qquad \text{for} \quad K \to \infty$$

though

$$\int_0^\infty \left| \frac{\sin x}{x} \right| \, dx = \infty$$

Example: Marginalization

If λ is a measure on $(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ the image measures

$$\hat{X}(\lambda) \qquad \text{og} \qquad \hat{Y}(\lambda)$$

are the marginal measures of λ's.

Let $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, and consider $\hat{X} : \mathcal{X} \times \mathcal{X}_3 \to \mathcal{X}$, $\hat{X}_1 : \mathcal{X}_1 \times (\mathcal{X}_2 \times \mathcal{X}_3) \to \mathcal{X}_1$ and $\hat{X}_{01} : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_1$. Then

$$\hat{X}_1 = \hat{X}_{01} \circ \hat{X}$$

and by the theorem on successive transformation

$$\hat{X}_1(\lambda) = \hat{X}_{01}(\hat{X}(\lambda)).$$

Moral: For marginalization on multiple product spaces it does not matter if we do several succesive marginalizations or one combined marginalization.

Example: Marginalization

Example: Consider $m_2 = m \otimes m$ on $(\mathbb{R}^2, \mathbb{B}_2)$. Then

$$\hat{X}(m_2)(A) = m_2(A \times \mathbb{R}) = m(A)m(\mathbb{R}) = \begin{cases} \infty & \text{if } m(A) > 0 \\ 0 & \text{if } m(A) = 0 \end{cases}$$

Example: If λ is a probability measure on $(\mathcal{X} \times \mathcal{Y}, \mathbb{E} \otimes \mathbb{K})$ then the marginals are probability measures and λ is a product measure if and only if it is a product of its marginals;

$$\lambda = \hat{X}(\lambda) \otimes \hat{Y}(\lambda).$$