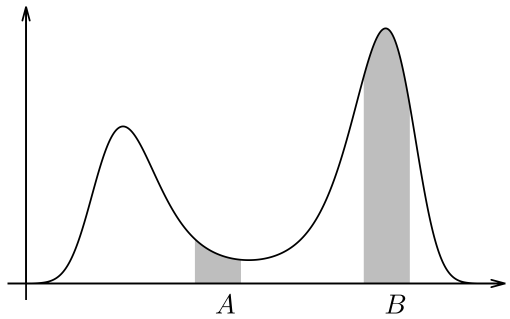
Measures with a density

Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space and $f \in \mathcal{M}^+$ a function. Define

$$\nu(A) = \int_A f \mathrm{d}\mu \qquad \text{for all } A \in \mathbb{E}$$



Measures with a density

Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space and $f \in \mathcal{M}^+$ a function. Define

$$\nu(A) = \int_A f \mathrm{d}\mu \qquad \text{for all } A \in \mathbb{E}$$

Lemma: ν is a measure on $(\mathcal{X}, \mathbb{E})$.

Proof: Check that $\nu(A) \in [0, \infty]$ and $\nu(\emptyset) = 0$ by the definition.

If $A_1, A_2, \ldots$ are pairwise disjoint $\mathbb{E}$-sets, then

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \int 1_{\bigcup_{n=1}^{\infty} A_n} f \mathrm{d}\mu = \int \sum_{n=1}^{\infty} 1_{A_n} f \mathrm{d}\mu \overset{\text{Cor. 6.20}}{=} \sum_{n=1}^{\infty} \int 1_{A_n} f \mathrm{d}\mu = \sum_{n=1}^{\infty} \nu(A_n)$$

We write $\nu = f \cdot \mu$ and say that ν has density (tæthed) f w.r.t. μ.

Restrictions

Example: Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space and $A \in \mathbb{E}$. Let $\nu = 1_A \cdot \mu$.

We find that

$$\nu(B) = \int_B 1_A \mathrm{d}\mu = \int 1_B 1_A \mathrm{d}\mu = \mu(B \cap A)$$

That is, ν is the restriction of μ to A.

Point: The restriction (Exercise 2.3) of a measure μ to a measurable set $A \in \mathbb{E}$ is the measure with density 1_A w.r.t. μ.

Integration w.r.t. measures with a density

Theorem: Let $\nu = f \cdot \mu$ on $(\mathcal{X}, \mathbb{E})$. For all $g \in \mathcal{M}^+(\mathcal{X}, \mathbb{E})$ it holds that

$$\int g \mathrm{d}\nu = \int g\, f \mathrm{d}\mu$$

Proof: The standard scheme for these kinds of proof: Show that the result holds for

1) indicator functions
2) simple functions
3) $\mathcal{M}^+$-functions

Point 3) is shown from 2) using monotone convergence.

Integration w.r.t. measures with a density

Theorem: Let $\nu = f \cdot \mu$ on $(\mathcal{X}, \mathbb{E})$. For a function $g \in \mathcal{M}(\mathcal{X}, \mathbb{E})$ it holds that g is ν-integrabel if and only if

$$\int |g\, f| \mathrm{d}\mu < \infty,$$

in which case

$$\int g \mathrm{d}\nu = \int g\, f \mathrm{d}\mu$$

Proof: First check the integrability condition using $\mathcal{M}^+$-integration of $|g|$ w.r.t. $f \cdot \mu$.

Second apply the $\mathcal{M}^+$-formula for g^+- and g^--integration.

Restrictions

Example: Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space and $A \in \mathbb{E}$. The restriction, $\mu|_A$, of μ to A has density 1_A w.r.t. μ.

For $f \in \mathcal{M}^+$ we have that

$$\int f \mathrm{d}\mu|_A = \int f\, 1_A \mathrm{d}\mu = \int_A f \mathrm{d}\mu$$

and if $f \in \mathcal{M}$ then f is integrable w.r.t. $\mu|_A$ if and only if

$$\int_A |f| \mathrm{d}\mu < \infty$$

in which case

$$\int f \mathrm{d}\mu|_A = \int_A f \mathrm{d}\mu.$$

Familiar example

Consider $\nu = e^x \cdot m$ on $(\mathbb{R}, \mathbb{B})$.

Then for $x \in \mathbb{R}$

$$\nu((-\infty, x]) = \int_{-\infty}^{x} e^y \mathrm{d}m(y) = e^x.$$

We have seen this measure before – via transformations – now we have another explicit characterization of the measure; for all $A \in \mathbb{B}$

$$\nu(A) = \int_A e^y \mathrm{d}m(y).$$

Successive densities

Theorem: Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a measure space, and let $f, g \in \mathcal{M}^+(\mathcal{X})$. Then

$$g \cdot (f \cdot \mu) = (f\, g) \cdot \mu$$

Proof: For $A \in \mathbb{E}$ it holds that

$$\Big(g \cdot (f \cdot \mu)\Big)(A) = \int 1_A\, g \mathrm{d}(f \cdot \mu) = \int (1_A\, g)\, f \mathrm{d}\mu = \int_A (f\, g) \mathrm{d}\mu$$

Uniqueness

Theorem: Let $(\mathcal{X}, \mathbb{E}, \mu)$ be a σ-finite measure spaces and let $f, g \in \mathcal{M}^+$. If $f \cdot \mu = g \cdot \mu$ then $f = g$ μ-almost everywhere.

Proof: Define

$$A = \{x \mid f(x) < g(x) < \infty\}, \quad B = \{x \mid g(x) < f(x) < \infty\},$$

$$C = \{x \mid f(x) < g(x) = \infty\}, \quad D = \{x \mid g(x) < f(x) = \infty\}.$$

and we show that they all have μ-measure 0.

□

Product measures

Let $(\mathcal{X}, \mathbb{E}, \mu)$ and $(\mathcal{Y}, \mathbb{K}, \nu)$ be two measure spaces and $f \in \mathcal{M}^+(\mathcal{X}, \mathbb{E})$ and $g \in \mathcal{M}^+(\mathcal{Y}, \mathbb{K})$.

Theorem: If $f \cdot \mu$ and $g \cdot \nu$ are σ-finite then

$$(f \cdot \mu) \otimes (g \cdot \nu) = (f \otimes g) \cdot (\mu \otimes \nu)$$

with $f \otimes g(x, y) = f(x)g(y)$.

Lemma: If $(\mathcal{X}, \mathbb{E}, \mu)$ is σ-finite and $f \in \mathcal{M}^+(\mathcal{X}, \mathbb{E})$ then $f \cdot \mu$ is σ-finite if and only if $f < \infty$ μ-a.e.

Product measures

Corollary: If $\lambda = h \cdot (\mu \otimes \nu)$ then $\lambda = (f \cdot \mu) \otimes (g \cdot \nu)$ if and only if $h = f \otimes g$ $\mu \otimes \nu$-a.e., that is there is a set $A \subset \mathcal{X} \times \mathcal{Y}$ such that for $(x, y) \in A$

$$h(x, y) = f(x)g(y)$$

and $A \subset B \in \mathbb{E}$ with $\mu \otimes \nu(B^c) = 0$, cf. also Exercise 11.6.

Important observation: If $C \in \mathbb{E} \otimes \mathbb{K}$ then the measure

$$(f \otimes g)1_C \cdot \mu \otimes \nu = (f \otimes g) \cdot \mu \otimes \nu|_C$$

is a product measure if $C = A \times B$ is a product set (up to a nullset).

Discrete integration

Let $\mathcal{X}$ be countable and τ the counting measure.

Let $p : \mathcal{X} \to [0, \infty]$ denote a map and consider the measure $\nu = p \cdot \tau$ on $(\mathcal{X}, \mathbb{P}(\mathcal{X}))$. Then

$$\nu(A) = \int_A p \, d\tau = \sum_{x \in A} p(x).$$

If $\mathcal{X}$ is infinite and $x_1, x_2, \ldots$ is a bijective counting of the elements in $\mathcal{X}$ we can write the sum as

$$\nu(A) = \lim_{n \to \infty} \sum_{i=1}^{n} 1_A(x_i)p(x_i).$$

Discrete integration

Let $\mathcal{X}$ be countable and τ the counting measure.

If $\nu = p \cdot \tau$ and $t : \mathcal{X} \mapsto \mathcal{Y}$ is a map we find for $y \in \mathcal{Y}$ that

$$t(\nu)(\{y\}) = \nu(t^{-1}(\{y\})) = \sum_{x \in t^{-1}(\{y\})} p(x).$$

Introducing

$$g(y) = \sum_{x \in t^{-1}(\{y\})} p(x)$$

we see that

$$t(\nu) = g \cdot \tau.$$

Example

Consider $t : \{0,1\}^n \to \{0,1,\ldots,n\}$ defined by

$$t(x_1,\ldots,x_n) = \sum_{i=1}^{n} x_i.$$

With τ the counting measure on $\{0,1\}^n$ we find for $m = 0,1,\ldots,n$ that

$$t(\tau)(\{m\}) = |t^{-1}(\{m\}| = \binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Thus $t(\tau)$ has density

$$g(m) = \binom{n}{m}$$

w.r.t. the counting measure on $\{0,1,\ldots,n\}$.

Density transformations

Let $(\mathcal{X},\mathbb{E},\mu)$ and $(\mathcal{Y},\mathbb{K},\xi)$ be two measure spaces. We think of μ and ξ as natural base measures.

Let $t : \mathcal{X} \to \mathcal{Y}$ be $\mathbb{E}$-$\mathbb{K}$-measurable.

The central problem: With $\nu = f \cdot \mu$ being a measure with density f on $(\mathcal{X},\mathbb{E})$.

1) Does $t(\nu)$ have a density w.r.t. ξ?

2) If so, how do we find the density?

Central question in probability theory as well as statistics.

Special transformations

Theorem: If $\nu = (g \circ t) \cdot \mu$ on $(\mathcal{X},\mathbb{E})$, and if $t : (\mathcal{X},\mathbb{E}) \to (\mathcal{Y},\mathbb{K})$ is measurable then

$$t(\nu) = g \cdot t(\mu).$$

Corollary: If $t : (\mathcal{X},\mathbb{E}) \to (\mathcal{Y},\mathbb{K})$ is bijective and bi-measurable and if $\nu = f \cdot \mu$ is a measure on $(\mathcal{X},\mathbb{E})$, then

$$t(\nu) = f \circ t^{-1} \cdot t(\mu).$$

Example

Recall that with $t : \{0,1\}^n \to \{0,1,\dots,n\}$ defined by $t(x_1,\dots,x_n) = \sum_{i=1}^n x_i$ then $t(\tau)$ has density $g(m) = \binom{n}{m}$ w.r.t. the counting measure on $\{0,1,\dots,n\}$. If $\nu = (p \circ t)\cdot\tau$ on $\{0,1\}^n$ then $t(\nu)$ has density

$$h(m) = \binom{n}{m} p(m)$$

w.r.t. τ on $\{0,1,\dots,n\}$.

Classical example: n independent flips of a (skewed) coin are modelled by the probability measure on $\{0,1\}^n$ with density

$$(x_1,\dots,x_n) \mapsto \gamma^{\sum_{i=1}^n x_i}(1-\gamma)^{n-\sum_{i=1}^n x_i} = \gamma^{t(x_1,\dots,x_n)}(1-\gamma)^{n-t(x_1,\dots,x_n)}$$

for $\gamma \in [0,1]$. The model for the sum of heads, say, is the probability measure with density

$$\binom{n}{m}\gamma^m(1-\gamma)^{n-m} \qquad \text{(the binomial distribution).}$$

Marginalization

Lemma: Let $(\mathcal{X}, \mathbb{E}, \mu)$ and $(\mathcal{Y}, \mathbb{K}, \lambda)$ be two σ-finite measure spaces.

Let $f \in \mathcal{M}^+(\mathcal{X} \times \mathcal{Y})$ and consider $\nu = f \cdot \mu \otimes \lambda$ on $\mathcal{X} \times \mathcal{Y}$.

Then $\hat{X}(\nu) = g \cdot \mu$ with

$$g(x) = \int f(x,y)\mathrm{d}\lambda(y) \qquad \text{for } x \in \mathcal{X}$$

Strategy for $\mathbb{R}^k$

Let $\nu = f \cdot m_k$ be a measure on $(\mathbb{R}^k, \mathbb{B}_k)$ with density f w.r.t. the Lebesgue measure m_k. Consider $t : \mathbb{R}^k \mapsto \mathbb{R}^k$. The strategy for finding $t(\nu)$ is as follows:

- Find disjoint $A_1,\dots,A_n \in \mathbb{B}_k$ with $\nu((\cup_i A_i)^c) = 0$ such that t is bijective and bi-measurable on each A_i.
- Observe that on A_i

$$t(\nu|_{A_i}) = (f \circ t^{-1}) t(m_k|_{A_i}).$$

- Find $t(m_k|_{A_i})$.

Moral: If the strategy works, the transformation problem boils down to transformations of the Lebesgue measure.

Example

Consider the map $t(x) = a + Bx$ for B and $k \times k$, invertible matrix. The (global) inverse map is

$$t^{-1}(y) = B^{-1}(y-a) \qquad \text{for } \; y \in \mathbb{R}^k$$

Due to the transformation results for the Lebesgue measure under affine transformations we have that $t(m_k) = |\det B^{-1}| m_k$, hence for $\nu = f \cdot m_k$ we have that $t(\nu) = g \cdot m_k$ where

$$g(y) = f\left(t^{-1}(y)\right)|\det B^{-1}| = f\left(B^{-1}(y-a)\right)|\det B^{-1}|.$$

For $k = 1$ this is

$$g(y) = f\left(\frac{y-a}{B}\right)\frac{1}{|B|} = f\left(\frac{y-a}{B}\right)|(t^{-1})'(y)|$$

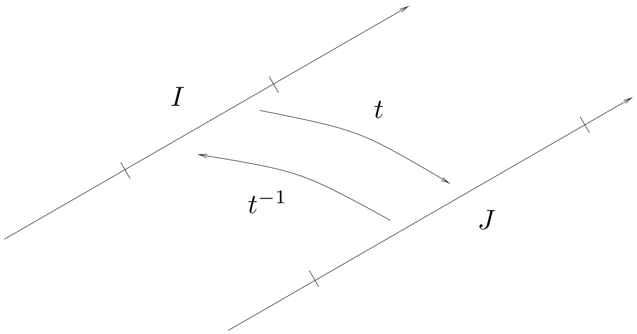
and we say $t(\nu)$ has location parameter a and scale parameter b w.r.t. μ.

Nice transformations

Consider a map $t : \mathbb{R} \to \mathbb{R}$ and open intervals I og J with

$1)$ t maps I bijectively on J,

$2)$ t^{-1} is a C^1-map on J.



Transformation of the Lebesgue measure

Theorem: For $t : \mathbb{R} \to \mathbb{R}$ a Borel measureable map that fulfills that

$1)$ t maps I bijectively on J,

$2)$ t^{-1} is a C^1-map on J.

it holds that $t(m|_I) = g \cdot m$ where g is

$$g(y) = \begin{cases} \left|(t^{-1})'(y)\right| & \text{for } y \in J\,, \\ 0 & \text{for } y \notin J\,. \end{cases}$$