Math 639: Lecture 2

Differentiation, product measures, independence

Bob Hough

January 26, 2017

Signed measures

Definition

A *signed measure* α on a measure space $(\Omega, \mathscr{F})$ is a set function which satisfies

1. α takes values in $(-\infty, \infty]$
2. $\alpha(\emptyset) = 0$
3. If $E = \bigsqcup_i E_i$ then $\alpha(E) = \sum_i \alpha(E_i)$ in the sense that
    - If $\alpha(E) < \infty$ then the sum converges absolutely
    - If $\alpha(E) = \infty$ then $\sum_i |\alpha(E_i)^-| < \infty$.

Signed measures

Example

Let μ be a measure, and f a function satisfying $\int |f^-| d\mu < \infty$. Then

$$\alpha(A) = \int_A f d\mu$$

is a signed measure.

Example

Let μ_1, μ_2 be measures with $\mu_2(\Omega) < \infty$. Then $\alpha(A) = \mu_1(A) - \mu_2(A)$ is a signed measure.

Positive sets

Definition

Given signed measure α and measurable set A, A is *positive* if every measurable $B \subset A$ has $\alpha(B) \geq 0$. A is *negative* if every measurable $B \subset A$ has $\alpha(B) \leq 0$.

Lemma

Every measurable subset of a positive set is positive. If the sets A_n are positive, then $A = \bigcup A_n$ is also positive.

Positive sets

Lemma

Let E be a measurable set with $\alpha(E) < 0$. Then there is a negative set $F \subset E$ with $\alpha(F) < 0$.

Positive sets

Proof.

Set $F_0 = E$, $i = 0$ and iterate the following process.

- Let $s_{i+1} = \sup\{\alpha(A) : A \subset F_i\}$. If $s_{i+1} = 0$, F_i is negative and we are done.
- Else, choose $E_{i+1} \subset F_i$ with $\alpha(E_{i+1}) > \frac{s_{i+1}}{2}$ and replace $F_{i+1} = F_i \setminus E_{i+1}$.

By additivity $s_1 < \infty$ and $s_{i+1} \leq \frac{s_i}{2}$. Hence if the process does not terminate, $s_i \downarrow 0$. In this case, set $F = \bigcap_i F_i$. Since

$$\alpha(E) = \alpha(F) + \sum_i \alpha(E_i)$$

converges absolutely, F cannot contain a set of positive measure or else one of the $\alpha(E_i)$ would need to be increased. Hence F is negative. □

Hahn decomposition

Theorem (Hahn decomposition)

Let α be a signed measure. Then there is a positive set A and a negative set B so that $\Omega = A \cup B$ and $A \cap B = 0$. Furthermore, if A', B' is another such decomposition, then $A \cap B'$ and $A' \cap B$ are null sets in the sense that all of their subsets have measure 0.

Hahn decomposition

Proof.

To prove the uniqueness statement, note that $A \cap B'$ is a positive and negative set, hence a null set, similarly $A' \cap B$. We prove the existence statement.

- Let $c = \inf\{\alpha(B) : B \text{ negative}\} \leq 0$. If $c = 0$ we are done.
- Otherwise, let B_i be negative sets with $\alpha(B_i) \downarrow c$, and set $B = \bigcup_i B_i$, which is negative. Since $\alpha(B) = \alpha(B - B_i) + \alpha(B_i) \leq \alpha(B_i)$ we have $\alpha(B) = c > -\infty$.
- We have $A = B^c$ is positive, since otherwise there exists $E \subset A$ which is negative, but then $B \cup E$ is negative and $\alpha(B \cup E) < c$, contradiction.

$\square$

Singular measures

Definition

Two measures μ_1 and μ_2 are *mutually singular* if there is a set A with $\mu_1(A) = 0$ and $\mu_2(A^c) = 0$. In this case we say μ_1 is *singular with respect to* μ_2 and write $\mu_1 \perp \mu_2$.

Example

The uniform measure on the Cantor set is singular with respect to Lebesgue measure.

Jordan decomposition

Theorem

Let α *be a signed measure. There are mutually singular measures* α_+ *and* α_- *so that* $\alpha = \alpha_+ - \alpha_-$. *Moreover, there is only one such pair.*

Jordan decomposition

Proof.

Let $\Omega = A \cup B$ be a Hahn decomposition. Define

$$\alpha_+(E) = \alpha(E \cap A), \qquad \alpha_-(E) = -\alpha(E \cap B).$$

This gives a decomposition as required. To prove the uniqueness, let ν_1 and ν_2 be singular measures, such that $\alpha = \nu_1 - \nu_2$. Let D be such that $\nu_1(D) = 0$ and $\nu_2(D^c) = 0$. By the uniqueness of the Hahn decomposition, A and D differ on a null set, so that

$$\alpha_+(E) = \alpha(E \cap A) = \alpha(E \cap D) = \nu_1(E),$$

which concludes the proof. $\square$

Lebesgue decomposition

Theorem (Lebesgue decomposition)

Let μ and ν be σ-finite measures. ν can be written as $\nu_r + \nu_s$, where ν_s is singular with respect to μ and

$$\nu_r(E) = \int_E g d\mu.$$

Lebesgue decomposition

Proof.

After making a countable decomposition, we may assume that both μ and ν are finite.

- Let $\mathscr{G}$ be the set of $g \geq 0$ such that for all E, $\int_E g d\mu \leq \nu(E)$.
- If $g, h \in \mathscr{G}$ then $\max(g, h) \in \mathscr{G}$. To check this, let $A = \{g > h\}$ and write

$$\int_E \max(g, h) d\mu = \int_{E \cap A} g d\mu + \int_{E \cap A^c} h d\mu \leq \nu(E).$$

□

Lebesgue decomposition

Proof.

- Let $\kappa = \sup\left\{\int g d\mu : g \in \mathscr{G}\right\}$. Choose $g_n \in \mathscr{G}$ such that $\int g_n d\mu \geq \kappa - \frac{1}{n}$, set $h_n = \max(g_1, ..., g_n)$, and let $h_n \uparrow h$. Then by monotone convergence, $h \in \mathscr{G}$ and $\int_\Omega h d\mu = \kappa$.
- Set $\nu_r(E) = \int_E h d\mu$ and $\nu_s(E) = \nu(E) - \nu_r(E)$.
- To check that ν_s is singular with respect to μ, let $\epsilon > 0$ and let $A_\epsilon \cup B_\epsilon$ be a Hahn decomposition for $\nu_s - \epsilon\mu$. Observe that

  $$\int_E (h + \epsilon \mathbf{1}_{A_\epsilon}) d\mu = \nu_r(E) + \epsilon\mu(A_\epsilon \cap E) \leq \nu(E).$$

  Hence $h + \epsilon \mathbf{1}_{A_\epsilon} \in \mathscr{G}$, but this implies that $\mu(A_\epsilon) = 0$.
- Let $A = \bigcup_n A_{\frac{1}{n}}$, with $\mu(A) = 0$. We have $\nu_s(A^c) = 0$, since otherwise, for some $\epsilon > 0$, $(\nu_s - \epsilon\mu)(A^c) > 0$, which contradicts that A^c is a negative set.

$\square$

Absolutely continuous measures

Definition

We say a measure ν is absolutely continuous with respect to μ, and write $\nu \ll \mu$ if $\mu(A) = 0$ implies $\nu(A) = 0$.

Radon-Nikodym theorem

Theorem (Radon-Nikodym theorem)

If μ and ν are σ-finite measures and ν is absolutely continuous with respect to μ, then there is a $g \geq 0$ so that $\nu(E) = \int_E g d\mu$. If there is another such function, h, then $h = g$ μ-a.e.. g is written $g = \frac{d\nu}{d\mu}$.

Radon-Nikodym theorem

Proof.

Let $\nu = \nu_r + \nu_s$ be a Lebesgue decomposition, and let A be such that $\nu_s(A^c) = 0$, $\mu(A) = 0$. By absolute continuity $\nu(A) = 0$ which implies $\nu_s \equiv 0$. Given two decompositions with functions g, h, one easily checks $\mu(g > h) = \mu(h > g) = 0$. $\square$

Product measures

Definition

Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be two measure spaces. The collection of *rectangles* of $\mathscr{A} \times \mathscr{B}$ is the empty set, together with

$$\mathscr{S} = \{A \times B : A \in \mathscr{A}, B \in \mathscr{B}\}.$$

The set of rectangles forms a semialgebra, since

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$
$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c).$$

Product measures

Theorem

Let $(X, \mathscr{A}, \mu_1)$ and $(Y, \mathscr{B}, \mu_2)$ be two σ-finite measure spaces. Set $\Omega = X \times Y$ and $\mathscr{F} = \sigma(\mathscr{S})$. There exists a unique measure μ on $\mathscr{F}$ such that for each rectangle $A \times B \in \mathscr{S}$,

$$\mu(A \times B) = \mu_1(A)\mu_2(B).$$

Product measures

Proof.

- By the Carathéodory extension theorem, it suffices to show that $\mu(A \times B) = \mu_1(A)\mu_2(B)$ extends to the algebra $\overline{\mathscr{S}}$ generated by $\mathscr{S}$.
- To do this, it suffices to check that if $A \times B = \bigsqcup_i A_i \times B_i$ is a finite or countable disjoint union, then

$$\mu(A \times B) = \sum_i \mu(A_i \times B_i).$$

- For $x \in A$ let $I(x) = \{i : x \in A_i\}$. We have $B = \bigsqcup_{i \in I(x)} B_i$, so

$$\mathbf{1}_A(x)\mu_2(B) = \sum_i \mathbf{1}_{A_i}(x)\mu_2(B_i).$$

Integrating with respect to μ_1 gives $\mu_1(A)\mu_2(B) = \sum_i \mu_1(A_i)\mu_2(B_i)$.

$\square$

Product measures

By the previous theorem and induction, it follows that if $\{(\Omega_i, \mathscr{F}_i, \mu_i) : i = 1, ..., n\}$ is a finite list of σ-finite measure spaces, then there is a unique measure μ on $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathscr{F} = \sigma(\{A_1 \times \cdots \times A_n : A_i \in \mathscr{F}_i\})$ such that

$$\mu(A_1 \times \cdots \times A_n) = \prod_{m=1}^{n} \mu_m(A_m).$$

The extension of this result to probability measures on infinite products is the subject of *Kolmogorov's extension theorem*.

Kolmogorov extension theorem

Let $\mathbb{N} = \{1, 2, 3, ...\}$ and let $\mathbb{R}^{\mathbb{N}} = \{(\omega_1, \omega_2, ...) : \omega_i \in \mathbb{R}\}$. Let $\mathscr{B}_{\mathbb{N}}$ be the σ-algebra generated by finite dimensional rectangles

$$\{\omega : \omega_i \in (a_i, b_i], i = 1, 2, ..., n\}$$

where $-\infty \leq a_i < b_i \leq \infty$.

Kolmogorov extension theorem

Theorem (Kolmogorov extension theorem)

Suppose we are given a sequence of probability measures $(\mathbb{R}^n, \mathscr{B}_{\mathbb{R}^n}, \mu_n)$*, which are consistent, in the sense that*

$$\mu_{n+1}((a_1, b_1] \times \cdots \times (a_n, b_n] \times \mathbb{R}) = \mu_n((a_1, b_1] \times \cdots \times (a_n, b_n]).$$

Then there is a probability measure Prob *on* $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}_{\mathbb{N}})$ *such that*

$$\mathrm{Prob}(\omega : \omega_i \in (a_i, b_i],\ 1 \leq i \leq n) = \mu_n((a_1, b_1] \times \cdots \times (a_n, b_n]).$$

Kolmogorov extension theorem

Example

Let $F_1, F_2, ...$ be distribution functions of measures $\mu_1, ..., \mu_n$, and let μ be the measure on $\mathbb{R}^n$ with

$$\mu((a_1, b_1] \times \cdots \times (a_n, b_n]) = \prod_{m=1}^{n} (F_m(b_m) - F_m(a_m)).$$

Thus μ is the product measure $\mu_1 \times \cdots \times \mu_n$. In particular, Kolmogorov's extension theorem gives a way of defining infinite products of probability measures.

Kolmogorov extension theorem

Proof of Kolmogorov's extension theorem.

Let $\mathscr{S}$ be the empty set, together with the collection of rectangles

$$\{\omega : \omega_i \in (a_i, b_i], 1 \leq i \leq n\}.$$

Define Prob on $\mathscr{S}$ according to the formula of the theorem. Since $\mathscr{S}$ is a semialgebra which generates $\mathscr{B}_{\mathbb{N}}$, it suffices to check that, if $A \in \mathscr{S}$ is the disjoint union of a sequence $\{A_i\}$ in $\mathscr{S}$ then

$$\mathrm{Prob}(A) = \sum_i \mathrm{Prob}(A_i).$$

in order to guarantee a unique extension of Prob to $\mathscr{B}_{\mathbb{N}}$. □

Kolmogorov extension theorem

Proof of Kolmogorov's extension theorem.

- It suffices to consider the case that $\{A_i\}$ is an infinite sequence, since any finite sequence of rectangles is determined in a finite number of coordinates.
- Set $B_n = A \setminus \bigcup_{i=1}^n A_i$. Thus B_n may be written as a finite disjoint union of rectangles, and so $\mathrm{Prob}(A) = \sum_{i=1}^n \mathrm{Prob}(A_i) + \mathrm{Prob}(B_n)$.
- Let $\mathscr{A}$ be the algebra formed from finite disjoint unions of rectangles of $\mathscr{S}$. The proof of the theorem is completed in the following lemma.

$\square$

Lemma

If $B_n \in \mathscr{A}$ and $B_n \downarrow \emptyset$, then $\mathrm{Prob}(B_n) \downarrow 0$.

Kolmogorov extension theorem

Proof.

The proof is a diagonalization argument.

- Suppose that $\mathrm{Prob}(B_n) \downarrow \delta > 0$. Possibly repeating sets, let

$$B_n = \bigcup_{k=1}^{K_n} \{\omega : \omega_i \in (a_i^k, b_i^k], 1 \leq i \leq n\}, \qquad -\infty \leq a_i^k < b_i^k \leq \infty.$$

- Choose $C_n \subset B_n$ of form

$$C_n = \bigcup_{k=1}^{K_n} \{\omega : \omega_i \in [\tilde{a}_i^k, \tilde{b}_i^k], 1 \leq i \leq n\}, \qquad -\infty < \tilde{a}_i^k < \tilde{b}_i^k < \infty$$

  such that $\mathrm{Prob}(B_n - C_n) \leq \frac{\delta}{2^{n+1}}$.

- Let $D_n = \bigcap_{m=1}^n C_n$ so $\mathrm{Prob}(B_n - D_n) \leq \sum_{m=1}^n \mathrm{Prob}(B_m - C_m) \leq \frac{\delta}{2}$.

□

Kolmogorov extension theorem

Proof.

- Thus $\mathrm{Prob}(D_n)$ converges to a limit $\geq \frac{\delta}{2}$.
- Let $D_n^* \subset \mathbb{R}^n$ be such that $D_n = D_n^* \times \mathbb{R}^{\mathbb{N}}$. Note that D_n^* is compact.
- Choose sequence $\omega_1, \omega_2, \ldots$ such that $\omega_i \in D_i$.
- By diagonalization, pick a subsequence $\omega_{n(i)}$ such that each coordinate of $\omega_{n(i)}$ converges (this is possible by compactness). Let the limit be θ. We have $(\theta_1, \theta_2, \ldots, \theta_n) \in D_n^*$ for each n, hence $\theta \in \bigcap_{n=1}^{\infty} D_n$, which provides the required contradiction.

$\square$

Fubini's theorem

Theorem

Let $(\Omega_1, \mathscr{F}_1, \mu_1)$ and $(\Omega_2, \mathscr{F}_2, \mu_2)$ be σ-finite measure spaces with product space $(\Omega, \mathscr{F}, \mu)$. Let f on Ω be measurable and satisfy either $f \geq 0$ or $\int |f| d\mu < \infty$. Then

$$\int_{\Omega_1} \int_{\Omega_2} f(x, y)\mu_2(dy)\mu_1(dx) = \int_{\Omega} f d\mu = \int_{\Omega_2} \int_{\Omega_1} f(x, y)\mu_1(dx)\mu_2(dy).$$

Fubini's theorem

Proof sketch.

- It suffices to prove the theorem when $f = \mathbf{1}_E$ is the indicator function of a measurable set, since then the usual method of approximation with simple functions concludes the argument.
- It suffices to check that the collection of E for which the theorem holds with $\mathbf{1}_E$ is a σ-algebra, since the theorem already holds for the semialgebra of rectangles.
- In fact, by the π-λ theorem, it suffices to show that this collection is a λ-system.
- Obviously Ω satisfies the condition. The set difference condition is met by linearity of the integral. The increasing set condition is met by monotone convergence.

$\square$

Differentiating under the integral

As an application of Fubini's theorem we prove several theorems on differentiating under the integral.

Theorem

Let $(S, \mathscr{S}, \mu)$ be a measure space. Let f be a complex-valued function defined on $\mathbb{R} \times S$. Let $\delta > 0$, and suppose that for $x \in (y - \delta, y + \delta)$ we have

1. $u(x) = \int_S f(x, s)\mu(ds)$ *with* $\int_S |f(x, s)|\mu(ds) < \infty$
2. *For fixed s, $\frac{\partial f}{\partial x}(x, s)$ exists and is a continuous function of x.*
3. $v(x) = \int_S \frac{\partial f}{\partial x}(x, s)\mu(ds)$ *is continuous at $x = y$.*
4. $\int_S \int_{-\delta}^{\delta} \left|\frac{\partial f}{\partial x}(y + \theta, s)\right| d\theta\mu(ds) < \infty$.

Then $u'(y) = v(y)$.

Differentiating under the integral

Proof.

For $|h| \leq \delta$, applying Fubini,

$$\begin{aligned} u(y+h) - u(y) &= \int_S f(y+h,s) - f(y,s)\mu(ds) \\ &= \int_S \int_0^h \frac{\partial f}{\partial x}(y+\theta,s) d\theta \mu(ds) \\ &= \int_0^h \int_S \frac{\partial f}{\partial x}(y+\theta,s)\mu(ds) d\theta. \end{aligned}$$

The last equation gives

$$\frac{u(y+h)-u(y)}{h} = \frac{1}{h}\int_0^h v(y+\theta)d\theta.$$

The claim follows from continuity, letting $h \to 0$. $\square$

Differentiating under the integral

The following variant of the above theorem is useful.

Theorem

Let $(S, \mathscr{S}, \mu)$ be a measure space. Let f be a complex valued function defined on $\mathbb{R} \times S$. Let $\delta > 0$, and suppose that for $x \in (y - \delta, y + \delta)$ we have

1. $u(x) = \int_S f(x,s)\mu(ds)$ *with* $\int_S |f(x,s)|\mu(ds) < \infty$.
2. *For fixed s, $\frac{\partial f}{\partial x}(x,s)$ exists and is continuous as a function of x.*
3. $\int_S \sup_{\theta \in [-\delta,\delta]} \left|\frac{\partial f}{\partial x}(y+\theta, s)\right| \mu(ds) < \infty$.

Then $u'(y) = v(y)$.

Differentiating under the integral

Proof.

To reduce to the previous theorem, it suffices to prove that

$$\int_S \frac{\partial f}{\partial x}(x, s)\mu(ds)$$

is continuous at $x = y$. This follows from the pointwise continuity for fixed s and dominated convergence. $\square$

Differentiating under the integral

Theorem

Let Z be a random variable. Suppose $\epsilon > 0$ and $\phi(\theta) = \mathrm{E}[e^{\theta Z}] < \infty$ for $\theta \in [-\epsilon, \epsilon]$. Then $\phi'(0) = \mathrm{E}[Z]$.

Proof.

Apply the previous theorem with μ the distribution of Z and $f(\theta, s) = e^{\theta s}$. □

Independence

Definition

Several σ-algebras $\mathscr{F}_1, \mathscr{F}_2, ..., \mathscr{F}_n$ are *independent* if, whenever $A_i \in \mathscr{F}_i$,

$$\mathrm{Prob}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathrm{Prob}(A_i).$$

Random variables $X_1, ..., X_n$ are *independent* if the σ-algebras $\sigma(X_1), ..., \sigma(X_n)$ are independent.
Sets $A_1, ..., A_n$ are *independent* if whenever $I \subset \{1, ..., n\}$ we have

$$\mathrm{Prob}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathrm{Prob}(A_i).$$

Pairwise independence

Definition

Several events $A_1, A_2, ..., A_n$ are *pairwise independent* if, for any $i \neq j$, $\text{Prob}(A_i \cap A_j) = \text{Prob}(A_i)\,\text{Prob}(A_j)$.

Pairwise independence does not imply independence, as the next example shows.

Pairwise independence

Example

Let X_1, X_2, X_3 be independent random variables with $\text{Prob}(X_i = 0) = \text{Prob}(X_i = 1) = \frac{1}{2}$. Let

$$A_1 = \{X_2 = X_3\}, \quad A_2 = \{X_1 = X_3\}, \quad A_3 = \{X_1 = X_2\}.$$

These events are pairwise independent, since if $i \neq j$, then

$$\text{Prob}(A_i \cap A_j) = \text{Prob}(X_1 = X_2 = X_3) = \frac{1}{4} = \text{Prob}(A_i)\,\text{Prob}(A_j).$$

They are not independent, since $\text{Prob}(A_1 \cap A_2 \cap A_3) = \frac{1}{4}$.

Independence of π-systems

Definition

Collections of sets $\mathscr{A}_1, \ldots, \mathscr{A}_n \subset \mathscr{F}$ are *independent* if whenever $A_i \in \mathscr{A}_i$ and $I \subset \{1, \ldots, n\}$ we have $\mathrm{Prob}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathrm{Prob}(A_i)$.

Recall that a π-system is a collection of sets closed under intersection.

Theorem

Suppose $\mathscr{A}_1, \ldots, \mathscr{A}_n$ are independent, and each $\mathscr{A}_i$ is a π-system. Then $\sigma(\mathscr{A}_1), \ldots, \sigma(\mathscr{A}_n)$ are independent.

Independence of π-systems

Proof.

- Let $A_2, ..., A_n$ be sets with $A_i \in \mathscr{A}_i$ and let F be the intersection of one or more of the A_i.
- Let $\mathscr{L} = \{A \in \mathscr{F} : \mathrm{Prob}(A \cap F) = \mathrm{Prob}(A)\,\mathrm{Prob}(F)\}$. Note that $\mathscr{A}_1 \subset \mathscr{L}$ by independence. We check that $\mathscr{L}$ is a λ-system.
  - $\Omega \in \mathscr{L}$
  - Let $A, B \in \mathscr{L}$ with $A \subset B$. Then $B - A \in \mathscr{L}$, since

$$\begin{aligned}\mathrm{Prob}((B-A)\cap F) &= \mathrm{Prob}(B \cap F) - \mathrm{Prob}(A \cap F)\\ &= (\mathrm{Prob}(B) - \mathrm{Prob}(A))\,\mathrm{Prob}(F)\\ &= \mathrm{Prob}(B-A)\,\mathrm{Prob}(F).\end{aligned}$$

  - If $\{B_k\} \subset \mathscr{L}$ and $B_k \uparrow B$ then $\mathrm{Prob}(B \cap F) = \lim \mathrm{Prob}(B_k \cap F) = \mathrm{Prob}(B)\,\mathrm{Prob}(F)$ so $B \in \mathscr{L}$.

$\square$

Independence of π-systems

Proof.

- By the π-λ theorem, $\sigma(\mathscr{A}_1) \subset \mathscr{L}$ for each F, so $\sigma(\mathscr{A}_1)$ is independent of $\mathscr{A}_2, ..., \mathscr{A}_n$.
- Replacing $\mathscr{A}_1$ with $\sigma(\mathscr{A}_1)$, and rearranging the order and iterating, we reach the conclusion that $\sigma(\mathscr{A}_1), ..., \sigma(\mathscr{A}_n)$ are independent.

$\square$

Independence of random variables

Theorem

Let $X_1, ..., X_n$ be random variables which satisfy, for all $x_1, ..., x_n \in (-\infty, \infty]$

$$\mathrm{Prob}(X_1 \leq x_1, ..., X_n \leq x_n) = \prod_{i=1}^{n} \mathrm{Prob}(X_i \leq x_i).$$

Then $X_1, ..., X_n$ are independent.

Proof.

The sets $\mathscr{A}_i = \{X_i \leq x_i\}$ form a π-system, and $\sigma(\mathscr{A}_i) = \sigma(X_i)$. Choosing $x_i = \infty$ omits X_i from left and right side above. Hence, the claim follows from the previous theorem. $\square$

Independence of composites

Theorem

Suppose $\mathscr{F}_{i,j}$, $1 \leq i \leq n, 1 \leq j \leq m(i)$ *are independent* σ*-algebras, and let* $\mathscr{G}_i = \sigma\left(\bigcup_j \mathscr{F}_{i,j}\right)$. *Then* $\mathscr{G}_1, ..., \mathscr{G}_n$ *are independent.*

Proof.

The collection of sets $\mathscr{A}_i = \bigcap_j A_{i,j}$ where $A_{i,j} \in \mathscr{F}_{i,j}$ form a π-system generating $\mathscr{G}_i$. The claim follows. □

Independence of composites

Theorem

Let $X_{i,j}$, $1 \leq i \leq n$, $1 \leq j \leq m(i)$ be independent, and $f_i : \mathbb{R}^{m(i)} \to \mathbb{R}$ be measurable. Then $f_i(X_{i,1}, ..., X_{i,m(i)})$ are independent.

Proof.

Let $\mathscr{F}_{i,j} = \sigma(X_{i,j})$ and $\mathscr{G}_i = \sigma\left(\bigcup_j \mathscr{F}_{i,j}\right)$. The result follows from the previous theorem, since $f_i(X_{i,1}, ..., X_{i,m(i)})$ is $\mathscr{G}_i$-measurable. □

Independent distributions

Theorem

Suppose $X_1, \ldots, X_n$ are independent random variables and X_i has distribution μ_i. Then $(X_1, \ldots, X_n)$ has distribution $\mu_1 \times \cdots \times \mu_n$.

Proof.

Calculate

$$\mathrm{Prob}((X_1, \ldots, X_n) \in A_1 \times \cdots \times A_n) = \prod_{i=1}^{n} \mathrm{Prob}(X_i \in A_i)$$

$$= \prod_{i=1}^{n} \mu_i(A_i) = \mu_1 \times \cdots \times \mu_n(A_1 \times \cdots \times A_n).$$

Since the distribution of $(X_1, \ldots, X_n)$ and $\mu_1 \times \cdots \times \mu_n$ agree on the π-system of sets $A_1 \times \cdots \times A_n$ which generates $\mathscr{B}_{\mathbb{R}^n}$, they agree. $\square$

Independence and expectation

Theorem

Suppose X and Y are independent and have distributions μ and ν. If $h : \mathbb{R}^2 \to \mathbb{R}$ is a measurable function with $h \geq 0$ or $\mathrm{E}[|h(X, Y)|] < \infty$, then

$$\mathrm{E}[h(X, Y)] = \iint h(x, y)\mu(dx)\nu(dy).$$

In particular, if $h(x, y) = f(x)g(y)$ where $f, g : \mathbb{R} \to \mathbb{R}$ are measurable functions with $f, g \geq 0$ or $\mathrm{E}[|f(X)|]$ and $\mathrm{E}[|g(Y)|] < \infty$, then

$$\mathrm{E}[f(X)g(Y)] = \mathrm{E}[f(X)]\,\mathrm{E}[g(Y)].$$

This follows from the previous theorem and Fubini's Theorem.

Independence and expectation

Theorem

If $X_1, \ldots, X_n$ are independent and satisfy either $X_i \geq 0$ for all i, or $\mathrm{E}[|X_i|] < \infty$ for all i, then

$$\mathrm{E}\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n \mathrm{E}[X_i].$$

This follows from the previous result and induction.

Correlation

Definition

Two random variables X and Y which satisfy $\mathrm{E}\left[X^2\right], \mathrm{E}\left[Y^2\right] < \infty$ are *uncorrelated* if $\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y]$.

Two random variables can be uncorrelated without being independent.

Sums of independent random variables

Theorem

Let X and Y be independent random variables with distributions μ and ν. Then $X + Y$ has distribution $\mu * \nu$ defined by

$$\mu * \nu((a, b]) = \iint_{x+y\in(a,b]} \mu(dx)\nu(dy).$$

Proof.

This follows from the fact that (X, Y) have distibution $\mu \times \nu$. $\square$

Sums of independent random variables

We record several consequences of the previous theorem.

1. If $F(x) = \mathrm{Prob}(X \leq x)$ then $X + Y$ has distribution function
$$\mathrm{Prob}(X + Y \leq z) = \int F(z - y)\nu(dy).$$
2. If X has density $f(x)$ then $X + Y$ has density
$$h(x) = \int f(x - y)\nu(dy).$$
3. In particular, if Y has density g then
$$h(x) = \int f(x - y)g(y)dy = f * g(x).$$

The Gamma distribution

The *Gamma distribution* with parameters $\alpha > 0$ and $\lambda > 0$ has density

$$f(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Theorem

If X and Y are independent, with X distributed gamma(α, λ) and Y distributed gamma(β, λ) then $X + Y$ is distributed gamma$(\alpha + \beta, \lambda)$.

The Gamma distribution

Proof.

For $x \geq 0$, the density of $X + Y$ at x is

$$\begin{aligned} f_{X+Y}(x) &= \int_0^x \frac{\lambda^\alpha (x-y)^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda(x-y)} \frac{\lambda^\beta y^{\beta-1}}{\Gamma(\beta)} e^{-\lambda y} dy \\ &= \frac{\lambda^{\alpha+\beta} e^{-\lambda x}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x (x-y)^{\alpha-1} y^{\beta-1} dy \end{aligned}$$

The latter integral is

$$x^{\alpha+\beta-1} \int_0^1 (1-u)^{\alpha-1} u^{\beta-1} du = x^{\alpha+\beta-1} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

$\square$

The normal distribution

The normal distribution with mean μ and variance a has density

$$\eta(\mu, a; x) = \frac{\exp\left(-\frac{(x-\mu)^2}{2a}\right)}{\sqrt{2\pi a}}.$$

Theorem

If $X = \eta(\mu, a)$ and $Y = \eta(\nu, b)$ are independent, then $X + Y = \eta(\mu + \nu, a + b)$.