# Probability Notes

Duncan MacIntyre

September 27, 2023

## Contents

# 1 Some measure theory

**Definition 1.1.** Let $\Omega$ be a set. A topology on $\Omega$ is a collection $\mathscr{A} \subset P(\Omega)$ that is closed under unions and finite intersections with $\Omega, \emptyset \in \mathscr{A}$. For $E \subset \Omega$, if $E \in \mathscr{A}$, we call $E$ open, and if $E^C \in \mathscr{A}$, we call $E$ closed.

For example, if $(X, d)$ is a metric space, $\mathscr{A} = \{E | E = \cup_i N_{r_i}(x_i)\}$ is a topology on $X$.

*Gap!!! The example should be proven.*

**Definition 1.2.** Let $\Omega$ be a set and $\mathscr{S} \subset P(\Omega)$. Then the topology generated by $\mathscr{S}$ is

$$\mathscr{A} = \{E \subset \Omega | E \text{ is a union of finite intersections of sets in } \mathscr{S}.\}$$

*Gap!!! The definition should be more explicit.*

*Gap!!! Examples and Furstenberg's theorem should be added here*

**Definition 1.3.** Let $\Omega$ be a set. An algebra is a collection $\mathscr{A} \subset P(\Omega)$ that is closed under finite unions and compliments with $\Omega \in \mathscr{A}$. If an algebra is also closed under countable unions, we call it a $\sigma$-algebra.

**Theorem 1.4.** *Algebras are closed under finite intersections and $\sigma$-algebras are closed under countable intersections.*

*Proof.* Let $\mathscr{A}$ be an algebra and let $A_1, \ldots, A_n \in \mathscr{A}$. Then $A_1^C, \ldots, A_n^C \in \mathscr{A}$, so $(\cap_{i=1}^n A_i)^C = \cup_{i=1}^n A_i^C \in \mathscr{A}$, so $\cap_{i=1}^n A_i \in \mathscr{A}$. The proof for $\sigma$-algebras is similar. □

**Definition 1.5.** Let $\mathscr{F}$ be a $\sigma$-algebra on a set $\Omega$. A measure on $\mathscr{F}$ is a function $\mu : \mathscr{F} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and for all disjoint $(A_i)_{i \in \mathbb{N}} \in \mathscr{F}$ we have $\mu(\cup_i A_i) = \sum_i \mu(A_i)$. We call $(\Omega, \mathscr{F}, \mu)$ a measure space.

*Gap!!! Add examples of measures and probability measures.*

**Theorem 1.6.** *An intersection of $\sigma$-algebras is also a $\sigma$-algebra.*

*Proof.* Let $(\mathscr{A}_i)$ be uncountably many $\sigma$-algebras. Let $A \in \cap \mathscr{A}_i$. Then $A^C \in \mathscr{A}_i$ for all $\mathscr{A}_i$, so $A^C \in \cap \mathscr{A}_i$. Similarly, let $(A_j)_{j \in \mathbb{N}} \in \cap \mathscr{A}_i$. Then $\cup_j A_j \in \mathscr{A}_i$ for all $\mathscr{A}_i$, so $\cup_j A_j \in \cap \mathscr{A}_i$. □

Theorem 1.6 motivates the following definition.

**Definition 1.7.** Let $\mathscr{A} \subset \Omega$. The $\sigma$-algebra generated by $\mathscr{A}$ is the intersection of all $\sigma$-algebras that contain $\mathscr{A}$. We denote it by $\sigma(\mathscr{A})$.

We can think of $\sigma(\mathscr{A})$ as the smallest $\sigma$-algebra containing $\mathscr{A}$.

**Notation 1.8.** A Borel $\sigma$-algebra is a $\sigma$-algebra generated by a topology. We will use $\mathscr{B}$ to denote the Borel $\sigma$-algebra of $\mathbb{R}$ that is generated by the usual open set topology.

*Gap!!! This section needs to be continued with product spaces, the upper/lower continuity of measures, the construction of the Lebesgue measure, the Caratheodory theorem, and the Dynkin uniqueness theorem.*

## 2 Probability spaces

**Definition 2.1.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a measure space. If $\mathbb{P}(\mathscr{F}) = 1$, we call $\mathbb{P}$ a probability measure, we call $(\Omega, \mathscr{F}, \mathbb{P})$ a probability space, and we call an element of $\mathscr{F}$ an event.

**Notation 2.2.** It is common to use a shorthand notation for events. We often write {condition} to mean $\{\omega \in \Omega \mid \text{ condition}\}$ and we often write $\mathbb{P}(\text{condition})$ to mean $\mathbb{P}\left(\{\omega \in \Omega \mid \text{condition}\}\right)$.

**Definition 2.3.** If $(A_i)_{i\in\mathbb{N}}$ are a sequence of events, "$A_i$ occurring infinitely often" is the event

$$\{A_i \text{ i.o.}\} = \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} A_n$$

and "$A_i$ occurring eventually" is the event

$$\{A_i \text{ eventually}\} = \bigcup_{m\in\mathbb{N}} \bigcap_{n\geq m} A_n.$$

*Remark.* $\{A_i \text{ i.o.}\}$ is the set of elements $\omega \in \Omega$ that belong in infinitely many of the events $(A_i)$ and $\{A_i \text{ eventually}\}$ is the set of elements $\omega \in \Omega$ that belong in all $(A_i)$ after a certain point. We note that $\{A_i \text{ eventually}\} \subset \{A_i \text{ i.o.}\}$.

*Gap!!! Add the Borel-Cantelli Lemma here.*

**Definition 2.4.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Events $(A_i)_{i\in I} \in \mathscr{F}$ are called independent if for any finite $J \subset I$ we have $\mathbb{P}\left(\cap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}\left(A_j\right)$.

**Theorem 2.5.** *If $(A_n)$ are independent and for all $n$ we define either $B_n = A_n$ or $B_n = A_n^C$ then all $(B_n)$ are independent.*

*Gap!!! Prove the above theorem.*

*Gap!!! Add the Goldberg conjecture discussion here, as an example.*

**Definition 2.6.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Collections $(\mathscr{A}_i) \subset \mathscr{F}$ are called independent if for any sequence $(A_i) \in \mathscr{F}$ with $A_i \in \mathscr{A}_i$ the $(A_i)$ are independent.

## 3 Random variables

**Definition 3.1.** Let $(\Omega_1, \mathscr{F}_1, \mu_1)$ and $(\Omega_2, \mathscr{F}_2, \mu_2)$ be measure spaces. A function $g : \Omega_1 \to \Omega_2$ is called measurable if for all $A \in \mathscr{F}_2$ we have $g^{-1}(A) \in \mathscr{F}_1$.

*Remark.* This is similar to the topological definition of continuous functions.

**Definition 3.2.** A random variable on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a measurable function $\mathsf{X} : \Omega \to \mathbb{R}$, where we use the usual Borel $\sigma$-algebra as the $\sigma$-algebra on $\mathbb{R}$.

**Notation 3.3.** When writing down an event involving a random variable $\mathsf{X}$, it is common to write $\mathsf{X}$ when we really mean $\mathsf{X}(\omega)$. For example, we might write $\mathbb{P}(\mathsf{X} = 1)$ to mean $\mathbb{P}\left(\{\omega \in \Omega : \mathsf{X}(\omega) = 1\}\right)$.

*Gap!!! Add some examples of random variables.*

*Remark.* We can form an equivalence relation on random variables by writing $\mathsf{X} \sim \mathsf{Y}$ if $\mathbb{P}(X = Y) = 1$, that is, if $\{\omega \in \Omega : \mathsf{X}(\omega) \neq \mathsf{Y}(\omega)\}$ has measure zero. Equivalent random variables have the same probabilistic properties. Some may prefer to think of "random variables" as these equivalence classes of functions. In this sense, it doesn't make sense to ask what a random variable's value is for a certain $\omega \in \Omega$; it only makes sense to look at the random variable's values on events with measure greater than zero.

**Definition 3.4.** The distribution function of a random variable $\mathsf{X}$ is the function $F_{\mathsf{X}} : \mathbb{R} \to \mathbb{R}$ given by $F(t) = \mathbb{P}(X \leq t)$.

We immediately see that distribution functions are monotonically increasing and right continuous (i.e., $\lim_{t \to s^+} F_{\mathsf{X}}(t) = F_{\mathsf{X}}(s)$ for all $s \in [0,1)$). Also, $F_{\mathsf{X}}$ has a jump discontinuity at $s$ if and only if $\mathbb{P}(\mathsf{X} = s) > 0$ (and if so, the jump is of size $\mathbb{P}(\mathsf{X} = s)$). There is an interesting theorem that provides something like a converse.

**Theorem 3.5.** *Let $F : \mathbb{R} \to [0,1]$ be monotonically increasing and right continuous (i.e., $\lim_{t \to s^+} F(t) = F(s)$ for all $s \in [0,1)$). Then there exists a probability space $(\Omega, \mathscr{f}, \mathbb{P})$ and a random variable $\mathsf{X} : \Omega \to \mathbb{R}$ such that $F = F_{\mathsf{X}}$.*

*Gap!!! Add the two theorem proofs.*

**Definition 3.6.** Given random variables $(\mathsf{X}_i)$ let $\mathscr{A}_i = \{{\mathsf{X}_i}^{-1}(A) : A \in \mathscr{B}\}$. We say that $(\mathsf{X}_i)$ are independent if $(\mathscr{A}_i)$ are independent. (Here, $\mathscr{B}$ is the usual Borel $\sigma$-algebra on $\mathbb{R}$.)

**Theorem 3.7.** *Let $(\mathsf{X}_i)$ be random variables and let $\mathscr{A}_i = \{\{\mathsf{X}_i \leq t\} : t \in \mathbb{R}\}$. Suppose $(\mathscr{A}_i)$ are independent. Then $(\mathsf{X}_i)$ are independent.*

*Gap!!! Prove the theorem.*

**Definition 3.8.** The joint distribution function of random variables $\mathsf{X}_1, \ldots, \mathsf{X}_n$ is the function $F : \mathbb{R}^n \to \mathbb{R}$ given by

$$F(\langle t_1, \ldots, t_n \rangle) = \mathbb{P}(\mathsf{X}_i \leq t_i \;\forall\, i \in \{0, \ldots, n\}).$$

*Gap!!! Add notes on stochastic domination, percolation, etc.*