4. Random Variables

In general, a random variable is a measurable function from $(\Omega, \mathcal{F})$ to some measurable space $(S, \mathcal{G})$, but we have agreed to reserve the unqualified term for the case $(S, \mathcal{G}) = (\mathbb{R}, \mathcal{B})$.

If $(S, \mathcal{G}) = (\mathbb{R}^d, \mathcal{B}^d)$, we will say that X is a *random vector*.

We now collect some results that will help us establish that various quantities of interest are indeed random variables.

Through a slight abuse of notation, we sometimes write $X \in \mathcal{F}$ to indicate that X is $(\mathcal{F}\text{-}\mathcal{B})$-measurable.

Theorem 4.1. *If $\mathcal{A}$ generates $\mathcal{G}$ (in the sense that $\mathcal{G}$ is the smallest σ-algebra containing $\mathcal{A}$) and $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$ for all $A \in \mathcal{A}$, then X is measurable.*

Proof. Because $X^{-1}\left(\bigcup_i E_i\right) = \bigcup_i X^{-1}(E_i)$ and $X^{-1}\left(E^C\right) = X^{-1}(E)^C$, we have that $\mathcal{E} = \{E \subseteq S : X^{-1}(E) \in \mathcal{F}\}$ is a σ-algebra. Thus, since $\mathcal{A} \subseteq \mathcal{E}$ and $\mathcal{A}$ generates $\mathcal{G}$ by assumption, $\mathcal{G} \subseteq \mathcal{E}$, so X is measurable. □

The fact that inverses commute with set operations also shows that for any function $X : \Omega \to S$, if $\mathcal{G}$ is a σ-algebra on S, then $\sigma(X) = \{X^{-1}(E) : E \in \mathcal{G}\}$ is a σ-algebra on Ω (called the *σ-algebra generated by* X). By construction, it is the smallest σ-algebra on Ω that makes X measurable with respect to $\mathcal{G}$.

Proposition 4.1. *If $\mathcal{A}$ generates $\mathcal{G}$, then $X^{-1}(\mathcal{A}) = \left\{X^{-1}(A) : A \in \mathcal{A}\right\}$ generates $\sigma(X)$.*

Proof. (Homework)

Since $\mathcal{A} \subseteq \mathcal{G}$, the definition of $\sigma(X)$ implies that $X^{-1}(\mathcal{A}) \subseteq \sigma(X)$ and thus $\sigma\left(X^{-1}(\mathcal{A})\right) \subseteq \sigma(X)$.

On the other hand, Theorem 4.1 shows that X is measurable as a map from $\left(\Omega, \sigma\left(X^{-1}(\mathcal{A})\right)\right)$ to $(S, \mathcal{G})$, so we must have that $\sigma(X) \subseteq \sigma\left(X^{-1}(\mathcal{A})\right)$. □

Example 4.1. If $(S, \mathcal{G}) = (\mathbb{R}, \mathcal{B})$, some useful generating sets are

$$\mathcal{A}_1 = \{(-\infty, x] : x \in \mathbb{R}\}, \quad \mathcal{A}_2 = \{(a, b) : a, b \in \mathbb{Q}\}.$$

Example 4.2. If $(S, \mathcal{G}) = (\mathbb{R}^d, \mathcal{B}^d)$, a convenient choice is

$$\mathcal{A} = \{(a_1, b_1] \times \cdots \times (a_d, b_d] : -\infty < a_i < b_i < \infty\}.$$

More generally, given an indexed collection of measurable spaces $\{(S_\alpha, \mathcal{G}_\alpha)\}_{\alpha \in A}$, the *product σ-algebra*, $\bigotimes_{\alpha \in A} \mathcal{G}_\alpha$, on $S = \prod_{\alpha \in A} S_\alpha$ is generated by $\{\pi_\alpha^{-1}(G_\alpha) : G_\alpha \in \mathcal{G}_\alpha, \alpha \in A\}$ where $\pi_\alpha : S \to S_\alpha$ is projection onto the α coordinate.

In other words, the product σ-algebra is the smallest σ-algebra for which the projections are measurable. This is because we want a function taking values in the product space to be measurable precisely when its components are measurable.

This is analogous to the definition of the product topology as the initial topology with respect to the coordinate projections.

Proposition 4.2. *If A is countable, then $\bigotimes_{\alpha\in A}\mathcal{G}_\alpha$ is generated by the rectangles $\{\prod_{\alpha\in A}G_\alpha : G_\alpha\in\mathcal{G}_\alpha\}$. If, in addition, $\mathcal{G}_\alpha$ is generated by $\mathcal{E}_\alpha\ni S_\alpha$ for every $\alpha\in A$, then $\bigotimes_{\alpha\in A}\mathcal{G}_\alpha$ is generated by $\left\{\prod_{\alpha\in A}E_\alpha : E_\alpha\in\mathcal{E}_\alpha\right\}$.*

Proof. If $G_\alpha\in\mathcal{G}_\alpha$, then $\pi_\alpha^{-1}(G_\alpha)=\prod_{\beta\in A}G_\beta$ where $G_\beta=S_\beta$ for all $\beta\neq\alpha$, hence

$$\sigma\left(\left\{\pi_\alpha^{-1}(G_\alpha):G_\alpha\in\mathcal{G}_\alpha,\alpha\in A\right\}\right)\subseteq\sigma\left(\left\{\prod_{\alpha\in A}G_\alpha:G_\alpha\in\mathcal{G}_\alpha\right\}\right).$$

On the other hand, $\prod_{\alpha\in A}G_\alpha=\bigcap_{\alpha\in A}\pi_\alpha^{-1}(G_\alpha)$, so

$$\sigma\left(\left\{\prod_{\alpha\in A}G_\alpha:G_\alpha\in\mathcal{G}_\alpha\right\}\right)\subseteq\sigma\left(\left\{\pi_\alpha^{-1}(G_\alpha):G_\alpha\in\mathcal{G}_\alpha,\alpha\in A\right\}\right).$$

The second statement will follow from the above argument once we show that $\bigotimes_{\alpha\in A}\mathcal{G}_\alpha$ is generated by $\mathcal{F}_1=\{\pi_\alpha^{-1}(E_\alpha):E_\alpha\in\mathcal{E}_\alpha,\alpha\in A\}$. To this end, observe that $\mathcal{F}_1\subseteq\{\pi_\alpha^{-1}(G_\alpha):G_\alpha\in\mathcal{G}_\alpha,\alpha\in A\}$ by definition, so $\sigma(\mathcal{F}_1)\subseteq\bigotimes_{\alpha\in A}\mathcal{G}_\alpha$.

On the other hand, arguing as in the proof of Theorem 4.1, we see that for each $\alpha\in A$, $\left\{E\subseteq S_\alpha:\pi_\alpha^{-1}(E)\in\sigma(\mathcal{F}_1)\right\}$ is a σ-algebra containing $\mathcal{E}_\alpha$ (and thus $\mathcal{G}_\alpha$), so $\pi_\alpha^{-1}(E)\in\sigma(\mathcal{F}_1)$ for all $E\in\mathcal{G}_\alpha$, hence $\sigma\left(\left\{\pi_\alpha^{-1}(G_\alpha):G_\alpha\in\mathcal{G}_\alpha,\alpha\in A\right\}\right)\subseteq\sigma(\mathcal{F}_1)$ as well. □

Because the product and metric topologies coincide for $\mathbb{R}^n$, Proposition 4.2 justifies Example 4.2.

(In general, one can show that if $S_1,...,S_n$ are separable metric spaces and $S=\prod_{i=1}^n S_i$ equipped with the product metric, then $\bigotimes_{i=1}^n\mathcal{B}_{S_i}=\mathcal{B}_S$.)

A simple but extremely useful observation is

Theorem 4.2. *If $X:(\Omega,\mathcal{F})\to(S,\mathcal{G})$ and $f:(S,\mathcal{G})\to(T,\mathcal{E})$ are measurable maps, then $f(X):(\Omega,\mathcal{F})\to(T,\mathcal{E})$ is measurable.*

Proof. For any $B\in\mathcal{E}$, $f^{-1}(B)\in\mathcal{G}$ since f is measurable, thus

$$\{\omega\in\Omega:f(X(\omega))\in B\}=\left\{\omega\in\Omega:X(\omega)\in f^{-1}(B)\right\}\in\mathcal{F}$$

since X is measurable. □

Theorem 4.2 is the familiar statement that compositions of measurable maps are measurable.

Thus if $f:\mathbb{R}\to\mathbb{R}$ is measurable (e.g. if f is any continuous function) and X is a random variable, then $f(X)$ is also a random variable.

An important application of Theorem 4.2 is given by

Theorem 4.3. *If $X_1,...,X_n$ are random variables and $f:(\mathbb{R}^n,\mathcal{B}^n)\to(\mathbb{R},\mathcal{B})$ is measurable, then $f(X_1,...,X_n)$ is a random variable.*

Proof. In light of Theorem 4.2, it suffices to show that $(X_1,...,X_n)$ is a random vector. To this end, observe that if $A_1,...,A_n$ are Borel sets, then

$$\{(X_1,...,X_n)\in A_1\times\cdots\times A_n\}=\bigcap_{i=1}^n\{X_i\in A_i\}\in\mathcal{F}.$$

Since $\mathcal{B}^n$ is generated by $\{A_1\times\cdots\times A_n:A_i\in\mathcal{B}\}$, the result follows from Theorem 4.1. □

Corollary 4.1. *If $X_1, ..., X_n$ are random variables, then so are $S_n = \sum_{i=1}^n X_i$ and $V_n = \prod_{i=1}^n X_i$.*

It is sometimes convenient to allow random variables to assume the values $\pm\infty$, and we observe that almost all of our results generalize easily to $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$ where $\mathbb{R} = \mathbb{R} \cup \{\pm\infty\}$ and $\overline{\mathcal{B}} = \{E \subseteq \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}\}$, which is generated, for example, by rays of the form $[-\infty, a)$ with $a \in \mathbb{R}$.

Theorem 4.4. *If $X_1, X_2, ...$ are random variables, then so are*

$$\inf_{n\in\mathbb{N}} X_n, \quad \sup_{n\in\mathbb{N}} X_n, \quad \liminf_{n\to\infty} X_n, \quad \limsup_{n\to\infty} X_n.$$

Proof. For any $a \in \mathbb{R}$, the infimum of a sequence is strictly less than a if and only if some term is strictly less than a, hence

$$\left\{\inf_{n\in\mathbb{N}} X_n < a\right\} = \bigcup_{n\in\mathbb{N}} \{X_n < a\} \in \mathcal{F},$$

hence $\inf_{n\in\mathbb{N}} X_n$ is measurable since $\{[-\infty, a) : a \in \mathbb{R}\}$ generates $\overline{\mathcal{B}}$.

To see that $\sup_{n\in\mathbb{N}} X_n$ is a random variable, note that $\sup_{n\in\mathbb{N}} X_n = -\inf_{n\in\mathbb{N}} -X_n$ and $f : x \mapsto -x$ is measurable.

Arguing as in the first case, $\inf_{m\geq n} X_m$ is measurable for all $m \in \mathbb{N}$, so it follows from the second case that

$$\liminf_{n\to\infty} X_n = \sup_{n\in\mathbb{N}} \left(\inf_{m\geq n} X_m\right)$$

is a random variable. The lim sup case is similar. □

It follows from Theorem 4.4 that

$$\left\{\lim_{n\to\infty} X_n \text{ exists}\right\} = \left\{\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n\right\} = \left\{\liminf_{n\to\infty} X_n - \limsup_{n\to\infty} X_n = 0\right\}$$

is measurable since it is the preimage of $\{0\} \in \mathcal{B}$ under the map $(\liminf_{n\to\infty} X_n) - (\limsup_{n\to\infty} X_n)$, which is the difference of measurable functions and thus measurable.

When $P(\{\lim_{n\to\infty} X_n \text{ exists}\}) = 1$, we say that the sequence converges almost surely to $X := \limsup_{n\to\infty} X_n$, and write $X_n \to X$ a.s.