November 7, 2021

RANDOM VARIABLES I

ABSTRACT. It is often more easier to work with functions defined on probability spaces, rather than working with probability spaces directly.

1. INTRODUCTION

Suppose we wager on a finite sequence n of coin-flip: I win 2 dollars when the coin comes up heads, and lose 1 dollar if it comes up tails. One way to model this is to come up with a probability space consisting of the possible values of my possibly negative winnings after n flips. However, we already know how to construct a canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for a sequence of n independent coin-flips by taking a product measure for $\Omega = \{0, 1\}^n$. If we define $X : \{0, 1\}^n \to \mathbb{R}$ via

$$X(\omega) = \sum_{i=1}^{n} [3\omega_i - 1],$$

then X is of my total winnings. Notice the value of X depends on ω so it depends on a which coin-flips we get; thus X is called a random variable. This notation allows us to address many questions using probability theory in an easy way. For example, the number

$$\mathbb{P}\{\omega \in \Omega : X(\omega) \geq 0\}$$

is the probability that I do not lose any money.

For countable probability spaces, a random variable is simply a function (usually real-valued) defined on the probability space. However, we have to be a bit more careful in the case of uncountable probability spaces. In what follows we develop some of the basic terminology and machinery that allows us to deal with random variables.

2. MEASURABLE FUNCTIONS

Let $(\Omega, \mathcal{F})$ and $(R, \mathcal{G})$ be a measurable spaces. We say that a function $f : \Omega \to R$ is **measurable** if for every $B \in \mathcal{G}$ we have

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

In many cases we are interested in real-valued measurable functions in which case by default we endow $R = \mathbb{R}$ with the the Borel sigma-algebra so that $\mathcal{G} = \mathcal{B}(\mathbb{R})$.

Remark 1. *An important elementary fact that is often useful in our discussion of measurable functions is that the inverse image preserves the set operation of union, intersection and complements; that is,*

$$f^{-1}\Big(\bigcup_{i \in I} A_i\Big) = \bigcup_{i \in I} f^{-1}(A_i)$$

$$f^{-1}\Big(\bigcap_{i \in I} A_i\Big) = \bigcap_{i \in I} f^{-1}(A_i)$$

$$f^{-1}(A^c) = \big(f^{-1}(A)\big)^c.$$

The simplest measurable functions are the **indicators**. For every $A \in \mathcal{F}$, we let $\mathbf{1}_A : \Omega \to \{0,1\}$ be given by $\mathbf{1}_A(\omega) = 1$ if $\omega \in A$, and $\mathbf{1}_A(\omega) = 0$ if $\omega \in A^c$.

Exercise 2.1. *Show that the indicators are measurable if we treat them as real-valued functions or if we treat them as functions to $\{0,1\}$ endowed with its power set as its sigma-algebra.*

Exercise 2.2. *Let $(\Omega, \mathcal{F})$ be a measurable space, and $f : \Omega \to \mathbb{R}$ be a measurable function. Show the set*

$$f^{-1}(\mathcal{B}) = \big\{ f^{-1}(B) : B \in \mathcal{B} \big\}$$

is a sigma-algebra for Ω.

Lemma 1. *Let $(\Omega, \mathcal{F})$ be a measurable space. Let $\mathcal{C}$ be a collection of subsets that generate $\mathcal{B}$. Show that $f : \Omega \to \mathbb{R}$ is measurable if and only if $f^{-1}(\mathcal{C}) \subseteq \mathcal{F}$.*

Proof. We prove the non-trivial direction. Consider

$$\mathcal{D} = \big\{ B \in \mathcal{B} : f^{-1}(B) \in \mathcal{F} \big\}.$$

By assumption, $\mathcal{C} \subseteq \mathcal{D}$; furthermore, it is easy to verify that $\mathcal{D}$ is a sigma-algebra. Since $\mathcal{C}$ generates $\mathcal{B}$, we are done. $\square$

Lemma 2. *Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f : \Omega \to \mathbb{R}$ is measurable if and only if for every $x \in \mathbb{R}$ we have*

$$f^{-1}(-\infty, x) = \{\omega \in \Omega : f(\omega) < x\} \in \mathcal{F}$$

and

$$f^{-1}(-\infty, x] = \{\omega \in \Omega : f(\omega) \leq x\} \in \mathcal{F}.$$

Exercise 2.3. *Show that a continuous function $f : \mathbb{R} \to \mathbb{R}$ is measurable. In this context, the Borel sigma-algebra is used for both the domain and co-domain.*

A direct proof of the following proposition is tricky.

Proposition 3. *Show that the sum of two real-valued measurable functions is again measurable.*

For $f : \Omega \to \mathbb{R}$ let

$$\{f < x\} = \{\omega \in \Omega : f(\omega) < x\}.$$

Proof. Let $(\Omega, \mathcal{F})$ be a measurable space. Let $f, g : \Omega \to \mathbb{R}$ be measurable functions. Let $x \in \mathbb{R}$. Set $h = f + g$. Observe that denseness of the rationals gives:

$$\{h < x\} = \bigcup_{r \in \mathbb{Q}} \{f < r\} \cap \{g < x - r\}$$

The measurability of f and g implies

$$\{f < r\} \cap \{g < x - r\} \in \mathcal{F}$$

for each $r \in \mathbb{Q}$. Since the union is countable, we have that $\{h < x\} \in \mathcal{F}$. By Lemma 2, this is enough. $\qquad\qquad\square$

Exercise 2.4. *Show that the multiplication of two real-valued measurable functions is again measurable.*

Exercise 2.5. *Show that the minimum of two real-valued measurable functions is again measurable.*

3. Random variables and their distributions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $X : \Omega \to \mathbb{R}$ be a measurable function; in this context, we say that X is a (real-valued) **random variable**. More generally, if $(R, \mathcal{G})$ is a measurable space, then $X : \Omega \to R$ is a **random variable** if X is a measurable function; usually we are interested in the case that R is a countable set and $R = \mathbb{R}^d$; in the case where R is countable, one can usually assume that $R = \mathbb{N}$. The **law** or the **distribution** of the random variable X is given by the measure

$$Q(B) := \mathbb{P}(X \in B)$$

for each $B \in \mathcal{B}$.

Exercise 3.1. *Check that Q is a probability measure for the measurable space $(\mathbb{R}, \mathcal{B})$.*

The **cumulative distribution function** for X is given by the function $F : \mathbb{R} \to [0, 1]$ with

$$F(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq x\} = \mathbb{P}(X \leq x).$$

Lemma 4. *If two random variables have the same cumulative distribution function, then they have the same law.*

The proof of Lemma 4 requires some tools from measure theory which we have not yet discussed.

Exercise 3.2. *Prove that a cumulative distribution function is non-decreasing and right continuous, with*

$$\lim_{x \to -\infty} F(x) = 0 \ \text{ and } \ \lim_{x \to \infty} F(x) = 1.$$

In general, if $F : \mathbb{R} \to [0, 1]$ has the properties in Exercise 3.2, we say that F is a **cumulative distribution function** (cdf). We will show later that given by a cdf F we can construct a random variable on a probability space with F as its cdf.

Theorem 5. *If F is a cdf, then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $X : \Omega \to \mathbb{R}$ such that X has F as its pdf.*

3.1. **Independence.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(R, \mathcal{G})$ be a measurable space. Let X and Y be random variables, We say that X and Y are **independent** if

$$\mathbb{P}(X \in A, Y \in B) \ := \ \mathbb{P}\{\omega \in \Omega : X(\omega) \in A \text{ and } Y(\omega) \in B\}$$
$$= \ \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

In the case of real-valued random variables, where $R = \mathbb{R}$, and $\mathcal{G} = \mathcal{B}$ one also has that X and Y are independent if and only if

$$\mathbb{P}(X \leq x, Y \leq y) \ = \ \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y)$$
$$= \ F_X(x)F_Y(y)$$

for all $x, y \in \mathbb{R}$.

If $X_1, X_2, \ldots,$ is a sequence of real-valued random variables, we say that the random variables are **independent** if for any finite distinct choice of indexes $i_1, \ldots, i_n$ and real numbers $x_1, \ldots, x_n$ we have

$$\mathbb{P}(X_{i_1} \leq x_1, \ldots, X_{i_n} \leq x_n) = \prod_{j=1}^{n} \mathbb{P}(X_{i_j} \leq x_j).$$

We say that the random variables are **identical** if they are have the same distribution. We will often work with independent and identical sequences of random variables, as called i.i.d. sequences. Let us stress here that if X and Y are identical random variables, this does *not* imply that $X(\omega) = Y(\omega)$; it only implies that they have an identical distribution.

Exercise 3.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let A and B be events. Show that A and B are independent if and only if $\mathbf{1}_A$ is independent of $\mathbf{1}_B$.*

In what follows we will discuss two important categories of real-valued random variables give some examples.

3.2. **Discrete random variables.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that a random variable X is **discrete** if there is a countable set C such that $\mathbb{P}(X \in C) = 1$. Sometimes the function given by

$$p(c) = \mathbb{P}(X = c)$$

is called a **probability mass function** or a **probability distribution function** (pmf, pdf). More generally, any $p : C \to [0, 1]$ such that

$$\sum_{c \in C} p(c) = 1$$

can be referred to as a pmf. Given a pmf p, the a random variable with pmf p is defined in the probability space $(C, \boldsymbol{\mathcal{P}}(C), \mathbb{P})$, where $\mathbb{P}(c) = p$—-just take $X : C \to C$ to be the identity map. Thus we have no theoretical issues with the existence of discrete random variables.

3.3. **Some important examples.** For what follows all random variables are assumed to be defined on a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$. We say that X is a **Bernoulli** random variable with parameter $p \in (0, 1)$, and we write $X \sim Bern(p)$, if $\mathbb{P}(X = 0) = 1 - p$ and $\mathbb{P}(X = 1) = p$. Let n be a positive integer. We say that S is a **binomial** random variable with parameter (n, p) and write $S \sim Bin(n, p)$, if $\mathbb{P}(0 \leq X \leq n) = 1$, and

$$\mathbb{P}(S = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

for all integers $0 \leq k \leq n$.

Exercise 3.4. *Let $p \in (0, 1)$. Let $(X_i)_{i=1}^n$ be i.i.d. random variables with $X_1 \sim Bern(p)$. Let $S = X_1 + \cdots + X_n$. Show that $S \sim Bin(n, p)$.*

Exercise 3.5. *Let $X_1, X_2, \ldots,$ be i.i.d. random variables with $X_1 \sim Bern(p)$. Let*

$$T = \inf \{n \geq 1 : X_n = 1\}.$$

Find the pmf for T.

Remark 2. *Notice that in Exercise 3.5, we assumed the existence of an infinite sequence of i.i.d. coin-flips. Technically, this is something that might not exist. Don't worry it does. We will prove this later.*

We say X is a **Poisson** random variable with mean $\lambda > 0$, and we write $X \sim Poi(\lambda)$ if

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

for all nonnegative integers k.

Exercise 3.6. *Let $n > 0$ and $\lambda > 0$. let $W \sim Poi(\lambda)$. Let $(X_i)_{i=1}^n$ be independent Bernoulli random variables with parameter $p = \lambda/n$. Let $S_n = X_1 + \cdots + X_n$. Show for all nonnegative integers k, we have $\lim_{n\to\infty} \mathbb{P}(S_n = k) = \mathbb{P}(W = k)$.*

Exercise 3.7 (Superposition). *Show that the sum of independent Poisson random variables is again a Poisson random variable.*

3.4. **Continuous random variables.** We say that a real-valued random variable X is a **continuous** random variable if there exists a nonnegative integrable function $f : \mathbb{R} \to [0, \infty)$ such that for all $x \in \mathbb{R}$ we have

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} f(v)dv;$$

and more generally,

$$\mathbb{P}(X \in A) = \int_{A} f(v)dv$$

for all $A \in \mathcal{B}$.

Thus the cdf F for X is (absolutely) continuous, and moreover; the fundamental theorem of calculus gives $F'(x) = f(x)$. Our density functions will usually be piecewise continuous bounded Riemann integrable functions.

The function f is called the **probability density function** (pdf) for X. In general, any integrable function $f : \mathbb{R} \to [0, \infty)$ with

$$\int_{-\infty}^{\infty} f(v)dv = 1$$

is a pdf. We will show later that given a pdf f we can construct a probability space in which we can define a random variable with f as its pdf.

Note that for a continuous random X with a pdf f we have

$$\mathbb{P}(a \le X \le b) = \int_{a}^{b} f(u)du$$

for all $a < b$ and for all $x \in \mathbb{R}$ we have

$$\mathbb{P}(X = x) = 0.$$

3.5. **Some important examples.** We say that a real-valued random variable U is **uniformly distributed** in the interval (a, b) if U has the density function $\frac{1}{b-a}\mathbf{1}_{(a,b)}$; thus if $(c, d) \subseteq (a, b)$, we have

$$\mathbb{P}(U \in (c, d)) = \mathbb{P}(c < U < d) = \frac{d - c}{b - a}.$$

Exercise 3.8. *Let U be uniformly distributed on $(0,1)$. Let $a < b$. Find a function $\phi : [0,1] \to \mathbb{R}$ so that $\phi(U)$ is uniformly distributed on (a,b).*

We say that a real-valued random variable X is an **exponential** random variable with parameter (mean) $\mu > 0$ if it has pdf given by

$$f(x) = \frac{1}{\mu} \exp(-x/\mu) \mathbf{1}[x > 0].$$

Exercise 3.9. *Check that the f defined above is indeed a pdf.*

Exercise 3.10. *Let X be an exponential random variable and $a > 0$. Check that aX is an exponential random variable.*

Exercise 3.11. *Show that the minimum of two independent exponential random variables is again an exponential random variable.*

Exercise 3.12. *Let U be uniformly distributed on $(0,1)$. Check that $-\log(U)$ is an exponential random variable.*

We say that X is a **normal** random variable with parameters (mean) $\mu \in \mathbb{R}$ and (variance) $\sigma^2 > 0$ if it has pdf given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-(x-\mu)^2/2\sigma^2].$$

Exercise 3.13. *Check that the f defined above is indeed a pdf.*

Exercise 3.14. *Let X be an normal random variable. Let $a, b \in \mathbb{R}$. Check that $aX + b$ is a again a normal variable.*