Math 525: Lecture 3

January 11, 2018

1 Random variables

Consider rolling two dice, corresponding to the sample space

$$\Omega = \{(m, n) \colon 1 \le m, n \le 6\}.$$

We can compute various numerical quantities based on the outcome of the rolls. For example, the sum of the two dice

$$X(m, n) = m + n$$

or their product

$$Y(m, n) = mn.$$

As we will see shortly, both X and Y are examples of random variables on a probability space. At least intuitively, a random variable is any function that maps the outcome of an experiment (e.g., (m, n)) to a numerical value (e.g., $m + n$ or mn).

Before we give a rigorous definition of a random variable, let's compute as a motivating example the probability that the two die sum to 5. Letting $X(m, n) = m + n$, this is simply

$$\mathbb{P}(\{\omega \in \Omega \colon X = 5\}) = \mathbb{P}(\{(1, 4), (2, 3), (3, 2), (4, 1)\}) = 4/36 = 1/9.$$

To give the rigorous definition of a random variable, we review the concept of an inverse map.

Definition 1.1. Let $f \colon A \to B$ be a function. Let $H \subset B$. Let

$$f^{-1}(H) = \{a \in A \colon f(a) \in H\}.$$

We call f^{-1} the *inverse map* of f. Note that the inverse map does not map points, but rather sets.

We are now ready to give the rigorous definition of a random variable. For the remainder, we will assume an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.2. A *random variable* is a function $X : \Omega \to \mathbb{R}$ which satisfies, for all $B \in \mathcal{B}(\mathbb{R})$,

$$X^{-1}(B) \equiv \{\omega \in \Omega \colon X(\omega) \in B\} \in \mathcal{F}.$$

The Borel σ-algebra is rather large, so the above definition is not easy to check. The following proposition makes our lives a bit simpler.

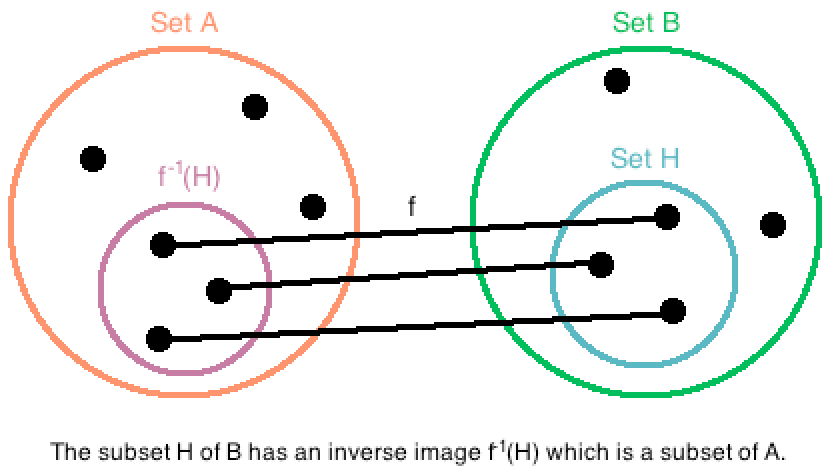

Figure 1: Inverse map

Proposition 1.3. *Suppose $\mathcal{G}$ generates the Borel σ-algebra (i.e., $\sigma(\mathcal{G}) = \mathcal{B}(\mathbb{R})$). Then, $X\colon \Omega \to \mathbb{R}$ is a random variable if and only if for each generating set $G \in \mathcal{G}$,*

$$X^{-1}(G) \in \mathcal{F}.$$

Proof. We prove only the nontrivial direction. Let

$$\mathcal{M} = \left\{B \subset \mathbb{R} \colon X^{-1}(B) \in \mathcal{F}\right\}.$$

By definition, we know that $\mathcal{G} \subset \mathcal{M}$. Therefore, $\sigma(\mathcal{G}) \subset \sigma(\mathcal{M})$. If $\mathcal{M}$ is a σ-algebra, it follows that

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{G}) \subset \sigma(\mathcal{M}) = \mathcal{M},$$

as desired. To check that $\mathcal{M}$ is a σ-algebra, verify the three properties:

1. $X^{-1}(\emptyset) = \emptyset \in \mathcal{F}$.
2. If $B \in \mathcal{M}$, then $X^{-1}(B^c) = (X^{-1}(B))^c \in \mathcal{F}$.
3. If $B_1, B_2, \ldots \in \mathcal{M}$, then $X^{-1}(\cup_{n\geq 1} B_n) = \cup_{n\geq 1} X^{-1}(B_n) \in \mathcal{F}$. □

The above proposition is particularly useful when we take $\mathcal{G}$ to be the set of intervals $\mathcal{G} = \{(-\infty, x] \colon x \in \mathbb{R}\}$. In this case...

Corollary 1.4. *$X\colon \Omega \to \mathbb{R}$ is a random variable if and only if for each $x \in \mathbb{R}$,*

$$\{\omega \in \Omega \colon X(\omega) \leq x\} \in \mathcal{F}.$$

Proof. If $G = (-\infty, x]$,

$$X^{-1}(G) = X^{-1}((-\infty, x]) = \{\omega \in \Omega \colon X(\omega) \leq x\}.$$ □

We will often use the above corollary to prove something is a random variable. To simplify notation, we often write

$$\{\omega \in \Omega \colon X(\omega) \leq x\} = \{X \leq x\}\,.$$

The above means that for a random variable X, $\mathbb{P}(\{X \leq x\})$ (i.e., the probability that X is at most x) is always well-defined. We define $\{X < x\}$, $\{X = x\}$, $\{X \geq x\}$, and $\{X > x\}$ similarly.

Proposition 1.5. *Let X be a random variable. Then, $\{X < x\}, \{X = x\}, \{X \geq x\}, \{X > x\} \in \mathcal{F}$.*

Proof. Let's just do the case of $\{X < x\}$. We can write

$$\{X < x\} = \bigcup_{\substack{q \in \mathbb{Q} \\ q > 0}} \{X \leq x - q\}\,,$$

in which case we see that $\{X < x\}$ is nothing other than a countable union of sets of the form $\{X \leq a\}$, which we know to be in $\mathcal{F}$. □

Proposition 1.6. *Let X and Y be random variables (on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$) and let a be a real number. Then, the following are also random variables:*

1. *aX.*
2. *$X + Y$.*
3. *XY.*
4. *Z defined by* $Z(\omega) = \begin{cases} Y(\omega)/X(\omega) & \text{if } X(\omega) \neq 0 \\ 0 & \text{if } X(\omega) = 0. \end{cases}$

Proof. Let $x \in \mathbb{R}$ be arbitrary.

1. Note that $\{aX \leq x\} = \{X \leq x/a\}$. Since X is a random variable, $\{X \leq x/a\} \in \mathcal{F}$.
2. It is sufficient to prove that $\{X + Y > x\} \in \mathcal{F}$ since $\{X + Y > x\} = \{X + Y \leq x\}^c$. Note that
$$\{X + Y > x\} = \bigcup_{q \in \mathbb{Q}} \{X > q\} \cap \{Y > x - q\}\,.$$
In this form, it is clear that $\{X + Y > x\} \in \mathcal{F}$.
3. For a real number a, define
$$a^+ = \max\{a, 0\} \qquad \text{and} \qquad a^- = \max\{-a, 0\}$$
as its positive and negative parts. Since for any real number a we have $a = a^+ - a^-$, it follows that
$$XY = \left(X^+ - X^-\right)\left(Y^+ - Y^-\right) = X^+Y^+ + X^+Y^- - X^-Y^+ + X^-Y^-.$$

Therefore, it is sufficient to consider the case in which X and Y are nonnegative. Moreover,

$$\{XY > x\} = \bigcup_{\substack{q\in\mathbb{Q} \\ q>0}} \{X > q\} \cap \{Y > x/q\}\,.$$

4. It is sufficient to consider the case of $Y = 1$ since $Y/X = Y(1/X)$. In this case,

$$\begin{aligned}\{Z \le z\} &= \Omega \cap \{Z \le z\} \\ &= (\{X \ne 0\} \cup \{X = 0\}) \cap \{Z \le z\} \\ &= (\{X \ne 0\} \cap \{Z \le z\}) \cup (\{X = 0\} \cap \{Z \le z\}) \\ &= (\{X \ne 0\} \cap \{1/X \le z\}) \cup (\{X = 0\} \cap \{0 \le z\})\,.\end{aligned}$$

First, note that $\{0 \le z\}$ is either equal to $\emptyset$ or Ω, and hence $\{X = 0\} \cap \{0 \le z\} \in \mathcal{F}$. We leave verifying that $\{X \ne 0\} \cap \{1/X \le z\} \in \mathcal{F}$ as an exercise. □

Definition 1.7. The *distribution function* of a random variable X is the function $F\colon \mathbb{R} \to [0,1]$ defined by

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

The distribution function is sometimes called the *cumulative distribution function* (CDF).

Example 1.8. Flip a coin twice. Let X be the number of heads H which show up. Since

$$\begin{aligned}\mathbb{P}(\{TT\}) &= 1/4 \\ \mathbb{P}(\{HT, TH\}) &= 2/4 \\ \mathbb{P}(\{HH\}) &= 1/4,\end{aligned}$$

we have that

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1/4 & \text{if } x < 1 \\ 3/4 & \text{if } x < 2 \\ 1 & \text{if } x \ge 2. \end{cases}$$

Remark 1.9. It now becomes clear why a random variable X requires $\{X \le x\} \in \mathcal{F}$! It is so that the dstribution function is well-defined at any point x.

In terms of the inverse function, note that

$$F(x) = \mathbb{P}(\, X^{-1}((-\infty, x])\,).$$

Proposition 1.10. *Let X be a random variable and F be its distribution function. Then,*

1. *F is nondecreasing.*
2. *F is right continuous. That is, $F(x) = \lim_{y\downarrow x} F(y)$ for each x.*
3. *$\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.*

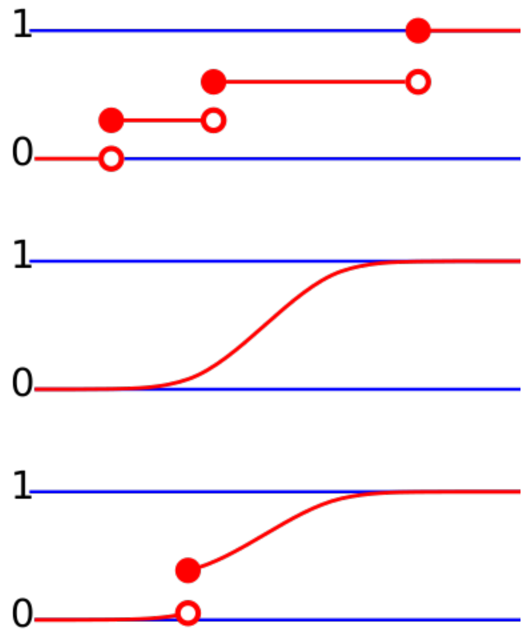

Figure 2: Examples of distribution functions

Proof. Let $A_x = \{X \le x\}$ so that $F(x) = \mathbb{P}(A_x)$.

1. If $x \le y$, then $A_x \subset A_y$ and hence $\mathbb{P}(A_x) \le \mathbb{P}(A_y)$.

2. Let $(y_n)_n$ be a sequence converging to x from above. Then, we have $A_{y_1} \supset A_{y_2} \supset \cdots$ and hence $\lim_{n\to\infty} \mathbb{P}(A_{y_n}) = \mathbb{P}(\cap_{n\ge1} A_{y_n}) = \mathbb{P}(A_x)$.

3. Note that $A_{-1} \supset A_{-2} \supset \cdots$ and hence $\lim_{n\to\infty} \mathbb{P}(A_{-n}) = \mathbb{P}(\cap_{n\ge1} A_{-n}) = \mathbb{P}(\emptyset) = 0$. Similarly, $A_1 \subset A_2 \subset \cdots$ and hence $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\cup_{n\ge1} A_n) = \mathbb{P}(\Omega) = 1$.

□

Since F is monotone, it follows that F has at most countably many discontinuities. We define

$$F(x-) = \lim_{y\uparrow x} F(x)$$

as the left-hand limit of F.

Proposition 1.11. *Let X be a random variable with distribution function F. Then,*

1. $\mathbb{P}(\{X < x\}) = F(x-)$.

2. $\mathbb{P}(\{X = x\}) = F(x) - F(x-)$.

3. *If* $a < b$, $\mathbb{P}(\{a < X \le b\}) = F(b) - F(a)$.

4. $\mathbb{P}(\{X > x\}) = 1 - F(x)$.

The above implies that if F is continuous, then $\mathbb{P}(\{X = x\}) = 0$ for all x! This means that there is zero probability of any particular realization of a random variable. Positive probability is assigned only to ranges (e.g., $\{a < X \leq b\}$).

Definition 1.12. Let $A \in \mathcal{F}$. The *indicator random variable on A* is the function

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

Note that for an indicator random variable I_A, we have $\mathbb{P}(\{I_A = 1\}) = \mathbb{P}(A)$ and $\mathbb{P}(\{I_A = 0\}) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Example 1.13. You play a game in which you roll a dice, and if the number you roll is greater than four, you get \$2. Otherwise, you lose \$1. How do we represent your winnings as a random variable?

Let $\Omega = \{1, \ldots, 6\}$. Since this is a finite sample space, we can safely take $\mathcal{F} = 2^\Omega$ and define $\mathbb{P}$ by $\mathbb{P}(\{\omega\}) = 1/6$ (each outcome is equally as likely). The random variable corresponding to your winnings is

$$X(\omega) = 2I_{\{\omega > 4\}}(\omega) - I_{\{\omega \leq 4\}}(\omega).$$

We'll often express this more succinctly as

$$X = 2I_{\{\omega > 4\}} - I_{\{\omega \leq 4\}}.$$

Remark 1.14. There are many other common notations for indicator functions. It helps to be aware of these:

$$I_A(\omega) \equiv \mathbf{1}_A(\omega) \equiv \chi_A(\omega) \equiv [A](\omega).$$

In closing this section, we introduce the notion of a Borel measurable function, which will give us our final piece of insight as to why we care about Borel σ-algebras.

Definition 1.15. We say $f \colon \mathbb{R} \to \mathbb{R}$ is a *Borel measurable function* (a.k.a. Borel function) if for each $B \in \mathcal{B}(\mathbb{R})$,

$$f^{-1}(B) \in \mathcal{B}(\mathbb{R}).$$

That is, the inverse image of a Borel set is also a Borel set.

Proposition 1.16. *If $f \colon \mathbb{R} \to \mathbb{R}$ is continuous, it is Borel measurable.*

Proof. Let

$$\mathcal{M} = \left\{B \subset \mathbb{R} \colon f^{-1}(B) \in \mathcal{B}(\mathbb{R})\right\}.$$

As usual, we can show that $\mathcal{M}$ is a σ-algebra (check). Let $\mathcal{G}$ be the set of open intervals in $\mathbb{R}$. Then, since f is continuous, $\mathcal{G} \subset \mathcal{M}$ and as such, $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{G}) \subset \sigma(\mathcal{M}) = \mathcal{M}$ as desired. $\square$

Exercise 1.17. Let X be a random variable and $f \colon \mathbb{R} \to \mathbb{R}$ be a Boreal measurable function. Show that Y defined by $Y(\omega) = f(X(\omega))$ is Borel measurable.

The above implies, most importantly, that taking the composition of a continuous function f and a random variable X yields a new random variable.