Math 525: Lecture 4

January 16, 2018

1 Uniform distribution

Definition 1.1. We say X is *uniformly distributed* on $[a, b]$ (written $X \sim U[a, b]$) if it is a random variable with distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \leq x < b \\ 1 & \text{if } x \geq b. \end{cases}$$

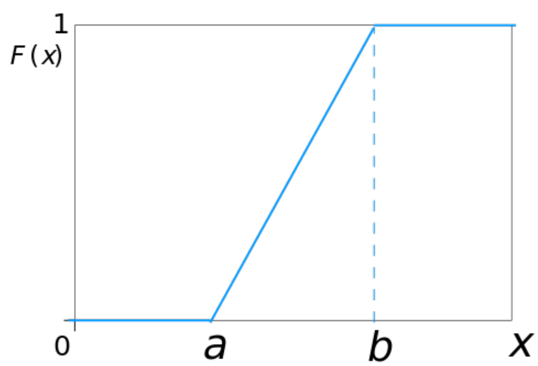

Figure 1: Uniform distribution

Intuitively, a uniform distribution tells us that any outcome in $[a, b]$ is "equally likely".

Remark 1.2. Actually, since F is continuous, no single outcome occurs with positive probability (recall that $\mathbb{P}(\{X = x\}) = F(x) - F(x-))$. What we really mean is that given two disjoint subsets A and B of $[a, b]$ which have the same "size", X is equally likely to be in either of them.

Example 1.3. The position of the pointer on a gameshow wheel can be modelled as a random variable uniformly distributed on $[0, 2\pi)$.

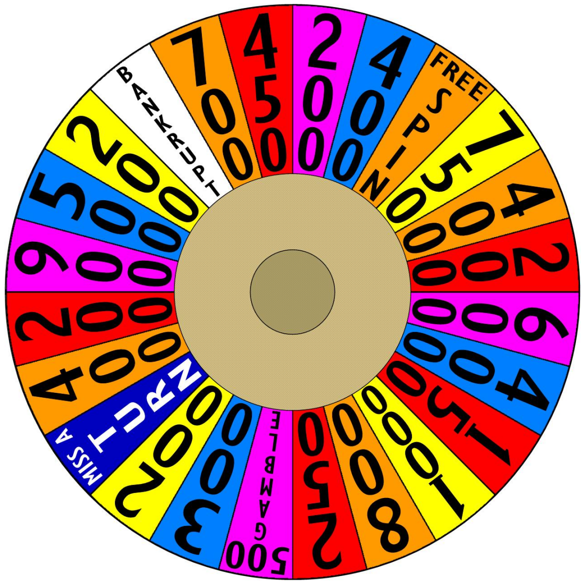

Figure 2: Game show wheel

How do we actually construct a random variable with a uniform distribution? There are a few ways to do this:

Example 1.4. Let $\Omega = \mathbb{R}$ and $A_x = \{\omega \in \Omega \colon \omega \leq x\}$. Define the probability measure $\mathbb{P} \colon \mathcal{B}(\mathbb{R}) \to [0, 1]$ by

$$\mathbb{P}(A_x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \leq x < b \\ 1 & \text{if } x \geq b. \end{cases}$$

$(\Omega, \mathcal{B}(\mathbb{R}), \mathbb{P})$ is a probability space. Moreover, the random variable X defined by $X(\omega) = \omega$ is uniformly distributed on $[a, b]$.

Remark 1.5. You may ask, at this point, why have we not defined $\mathbb{P}$ for sets not of the form A_x? Remember that $\mathcal{G} = \{A_x\}_{x \in \mathbb{R}}$ generates $\mathcal{B}(\mathbb{R})$. It turns out that to define a probability measure uniquely, it is sufficient to define it on generating sets (see, e.g., Corollary 1.8 of Walsh, John B. *Knowing the odds: an introduction to probability*. Vol. 139. American Mathematical Soc., 2012). The proof of this fact uses something called the *monotone class theorem*, which is outside of the scope of this course.

We could have also taken a slightly different approach in defining a uniformly distributed random variable:

Example 1.6. Let $\Omega = [a, b]$ and $\mathcal{B}([a, b]) = \sigma(\{[a, x] \colon x \in \mathbb{R}\})$ (compare this with the definition of $\mathcal{B}(\mathbb{R})$). Define A_x as before and the probability measure $\mathbb{P} \colon \mathcal{B}([a, b]) \to [0, 1]$ by

$$\mathbb{P}(A_x) = \frac{x-a}{b-a}.$$

$(\Omega, \mathcal{B}([a, b]), \mathbb{P})$ is a probability space. Moreover, the random variable Y defined by $Y(\omega) = \omega$ is uniformly distributed on $[a, b]$.

The probability spaces in the last two examples are, for all intents and purposes, identical... even though X and Y are technically not the same mathematical objects. The first probability space allows for points outside of $[a, b]$ to be outcomes, but they occur with zero probability. The second precludes them altogether.

Remark 1.7. For the remainder of the semester, unless we specify otherwise, we always assume that the relevant sample space is $\Omega = \mathbb{R}$ and the σ-algebra is $\mathcal{F} = \mathcal{B}(\mathbb{R})$.

2 Existence of random variables

In the last lecture, we showed that the distribution function F of any random variable is nondecreasing and right-continuous with

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

Today, we'll prove a "converse" of this fact.

Proposition 2.1. *Let $F \colon \mathbb{R} \to [0, 1]$ be nondecreasing and right-continuous with*

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

Then, there exists a random variable whose distribution function is F.

Note the subtlety here: in the previous lecture, we started out with a random variable and obtained a distribution function. The above proposition tells us we can go backwards: start with a distribution function and obtain a random variable.

Proof. We only consider the case in which F is a bijection. The general case is more challenging (see, e.g., Theorem 2.14 of Walsh, John B. *Knowing the odds: an introduction to probability.* Vol. 139. American Mathematical Soc., 2012).

Let $X \sim U[0, 1]$. In the case that F is a bijection, the inverse map F^{-1} maps singletons to singletons, and hence can be considered as a map from $\mathbb{R}$ to Ω. Therefore, we can define Y by $Y(\omega) = F^{-1}(X(\omega))$, or more succinctly, $Y = F^{-1} \circ X$. Since F is monotone, so too is F^{-1}. As a technical note, this implies that F^{-1} is Borel measurable and hence Y is indeed a random variable. Now, note that

$$\mathbb{P}(\{Y \leq y\}) = \mathbb{P}(\{F^{-1}(X) \leq y\}) = \mathbb{P}(\{X \leq F(y)\}) = \frac{F(y) - 0}{1 - 0} = F(y). \qquad \square$$

The proof above has a very important consequence for sampling from non-uniform distributions, as demonstrated below:

Example 2.2. You use a random number generator to generate n samples $U_1, \ldots, U_n \sim U[0, 1]$. You are given the distribution function F. Letting $X_i = F^{-1}(U_i)$, you obtain the samples $X_1, \ldots, X_n$, which all have the distribution function F.

This shows us that we can always turn the problem of sampling from a non-uniform distribution into one of sampling from a uniform distribution!

Example 2.3. Let U be a uniform random variable on $[0, 1]$. Let $X = U^2$. Let F denote the distribution function of X. Then, if $0 \leq x < 1$,

$$F(x) = \mathbb{P}(\{X \leq x\}) = \mathbb{P}(\{U^2 \leq x\}) = \mathbb{P}(\{U \leq \sqrt{x}\}) = \sqrt{x}.$$

3 Independence of random variables

In a previous lecture, we defined what it means for two events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to be independent (if $A, B \in \mathcal{F}$ and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, we say A and B are independent). We extend this definition now to random variables.

Definition 3.1. Two random variables X and Y are *independent* if for all $x, y \in \mathbb{R}$,

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

(i.e., the events $\{X \leq x\}$ and $\{X \leq y\}$ are independent).

Example 3.2. Let X be a random variable. Let $Y = X^2$. Suppose $0 < \mathbb{P}(\{X \leq a\}) < 1$ for some a. Then,

$$\mathbb{P}(\{X \leq a, Y \leq a^2\}) = \mathbb{P}(\{X \leq a, X^2 \leq a^2\}) = \mathbb{P}(\{X \leq a, X \leq a\}) = \mathbb{P}(\{X \leq a\}).$$

That is, X and Y are not independent.

The above definition concerns only sets of the form $\{X \leq x\}$ and $\{Y \leq y\}$. Can we extend it to other sets?

Proposition 3.3. *Let X and Y be independent random variables. Then,*

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(\{X \in A\})\mathbb{P}(\{Y \in B\})$$

whenever $A = (p, q]$ and $B = (r, s]$.

Proof. Note that

$$\begin{aligned}
\mathbb{P}\{X \in (-\infty, q], Y \in (r, s]\} &= \mathbb{P}\{X \leq q, r < Y \leq s\} \\
&= \mathbb{P}\{X \leq q, Y \leq s\} - \mathbb{P}\{X \leq q, Y \leq r\} \\
&= \mathbb{P}\{X \leq q\}\mathbb{P}\{Y \leq s\} - \mathbb{P}\{X \leq q\}\mathbb{P}\{Y \leq r\} \\
&= \mathbb{P}\{X \leq q\} \left(\mathbb{P}\{Y \leq s\} - \mathbb{P}\{Y \leq r\}\right) \\
&= \mathbb{P}\{X \leq q\}\mathbb{P}\{r < Y \leq s\}.
\end{aligned}$$

Now, use the same reasoning to get

$$\mathbb{P}\{p < X \leq q, r < Y \leq s\} = \mathbb{P}\{p < X \leq q\}\mathbb{P}\{r < Y \leq s\}. \qquad \square$$

Remark 3.4. The previous proposition can be extended more generally to the case of A and B in $\mathcal{B}(\mathbb{R})$ (see, e.g., Theorem 2.20 of Walsh, John B. *Knowing the odds: an introduction to probability*. Vol. 139. American Mathematical Soc., 2012). The proof uses, once again, the monotone class theorem.

4 Independence of multiple events

Let's generalize the concept of independence to families of random variables.

Definition 4.1. We say the family $\{A_\alpha\}_{\alpha\in\mathcal{A}} \in \mathcal{F}$ is independent if for each positive integer $k \le n$ and $\{A_{i_1}, \ldots, A_{i_k}\} \subset \{A_\alpha\}_{\alpha\in\mathcal{A}}$,

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_k}).$$

The notion of independence above is stronger than requiring each pair of events to be independent:

Example 4.2. Toss two fair coins at the same time. Let A be the event that the first coin is heads, B be the event that the second coin is heads, and C be the event that the first and second coins disagree (i.e., one is heads and the other is tails). Note that $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$.

Obviously, A and B are independent. To see that A and C are independent, note that

$$\mathbb{P}(A \cap C) = 1/4 = \mathbb{P}(A)\mathbb{P}(C).$$

Similarly, B and C are independent. This establishes that the three events are **pairwise** independent. However, note that $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) \neq 0$. Therefore, the events A, B, C are not independent, despite being pairwise independent.

Definition 4.3. We say the family of random variables $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ are independent if for each positive integer $k \le n$, $\{X_{i_1}, \ldots, X_{i_k}\} \subset \{X_\alpha\}_{\alpha\in\mathcal{A}}$, and $x_1, \ldots, x_k$,

$$\mathbb{P}(\{X_{i_1} \le x_1, \ldots, X_{i_k} \le x_n\}) = \mathbb{P}(\{X_{i_1} \le x_1\}) \cdots \mathbb{P}(\{X_{i_k} \le x_n\}).$$

Exercise 4.4. Let X and Y be i.i.d. integer-valued random variables (i.e., $\mathbb{P}(\{X \text{ is an integer}\}) = 1$ and similarly for Y). Let $p_n = \mathbb{P}(\{X = n\})$. Then,

$$\mathbb{P}(\{X = Y\}) = \sum_{n=-\infty}^{\infty} \mathbb{P}(\{X = n\})\mathbb{P}(\{Y = n\}) = \sum_{n=-\infty}^{\infty} p_n^2.$$

Similarly,

$$\mathbb{P}(\{X \le Y\}) = \sum_{n=-\infty}^{\infty} \mathbb{P}(\{X = n\}) \sum_{m=n}^{\infty} \mathbb{P}(\{Y = n\}) = \sum_{n=-\infty}^{\infty} p_n \sum_{m=n}^{\infty} p_m.$$

For example, suppose

$$p_n = \frac{6}{\pi^2}\frac{1}{n^2} \text{ if } n > 0 \qquad \text{and} \qquad p_n = 0 \text{ otherwise}$$

(you can check that $\sum_{n=1}^{\infty} p_n = 1$). Then,

$$\mathbb{P}(\{X = Y\}) = \sum_{n=1}^{\infty} p_n = \frac{36}{\pi^4} \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{36}{\pi^4}\frac{\pi^4}{90} = \frac{36}{90} = \frac{2}{5}.$$

and

$$\mathbb{P}(\{X \le Y\}) = \sum_{n=1}^{\infty} p_n \sum_{m=n}^{\infty} p_m = \frac{36}{\pi^4} \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{m=n}^{\infty} \frac{1}{m^2} = \frac{36}{\pi^4}\frac{7\pi^4}{360} = \frac{7}{10}.$$

5 Types of distributions

Definition 5.1. Let X be a random variable.

1. X has a *discrete distribution* if we can find a countable subset $\{x_n\}_n \subset \mathbb{R}$ for which

$$\sum_{n=1}^{\infty} \mathbb{P}\{X = x_n\} = 1.$$

2. X has a *continuous distribution* if its distribution function F is continuous.

3. X has an *absolutely continuous distribution* if its distribution function can be written

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

   for some integrable function f.

While many random variables fall into one of the above categories, there are still many which do not! For example...

Example 5.2. Consider flipping a coin. If the coin is heads, you receive one dollar. Otherwise, you receive Y dollars, where $Y \sim U[0, 1]$. This random variable is

$$X = I_{\{\text{tails}\}}Y + I_{\{\text{heads}\}}.$$

Its distribution function is

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \dfrac{x}{2} & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1. \end{cases}$$