Ryan Ta

Notes for Week 3 discussion on October 15

We addressed primarily only experiments with discrete outcomes. But please understand that there are experiments with continuous outcomes as well. Recall that C is the sample space that lists of all possible outcomes c of some given experiment. Let $\mathcal{D} := \{x \in \mathbb{R} : X(c) = x, c \in C\}$ be some range of real numbers, where $X : C \to \mathcal{D}$ is a function that we will define below.

Definition. *A function $X : C \to \mathcal{D}$ defined by $X(c) = x$ for some $c \in C$ and $x \in \mathcal{D}$ is called a* random variable.

Sometimes, $\mathcal{D}$ may be a finite space, where we can write $\mathcal{D} = \{d_1, \ldots, d_m\}$. In this case, we write $X(c) = d_i$ for any $i = 1, \ldots, m$, rather than $X(c) = x$ in general. Finite spaces appear when we have to deal with experiments—such as rolling dice—that produce discrete outcomes.

Definition. *A* cumulative distribution function *(c.d.f.) is a function $F_X : \mathbb{R} \to [0, 1]$ defined by*

$$\begin{aligned} F_X(x) &:= P_X((-\infty, x]) \\ &= P(\{c \in C : X(c) \le x\}) \\ &= P(X \le x), \end{aligned}$$

where the expression $P(X \le x)$ after the last equality is an informal but frequently-used version of $P(\{c \in C : X(c) \le x\})$.

Note that a c.d.f. always starts at $F_X(x) = 0$ from $x = -\infty$ and ends at $F_X(x) = 1$ towards $x = +\infty$, and that F_X is always increasing—in a continuous and/or discontinuous manner—from 0 to 1 as we traverse from left to right the values of $x \in \mathcal{D}$. For experiments with discrete outcomes such as rolling dice, F_X will always be increasing in a discontinuous manner with jumps. A good step towards finding a c.d.f. is to organize in a table all the values of x and $p_X(x)$, as will be done in the examples below. Our first two examples will focus on only discrete random variables.

Example (Rolling one six-sided die)**.** *The sample space for rolling one six-sided die is $C = \{1, 2, 3, 4, 5, 6\}$, with the 6 possible outcomes $c \in C$. In this example, let our range of numbers be exactly the outcomes of this experiment; in other words, here we have $\mathcal{D} = C = \{1, 2, 3, 4, 5, 6\}$. (Note: For other experiments, it is usually the case of $\mathcal{D} \neq C$.) This makes $\mathcal{D}$ a finite space, which means we can write $x = X(c) \in \{d_1, d_2, d_3, d_4, d_5, d_6\}$. Next, we can outline the range x and the probability $p_X(x)$ in the following table.*

x	1	2	3	4	5	6
$p_X(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Output values $p_X(x)$ of the probability mass function (p.m.f.)

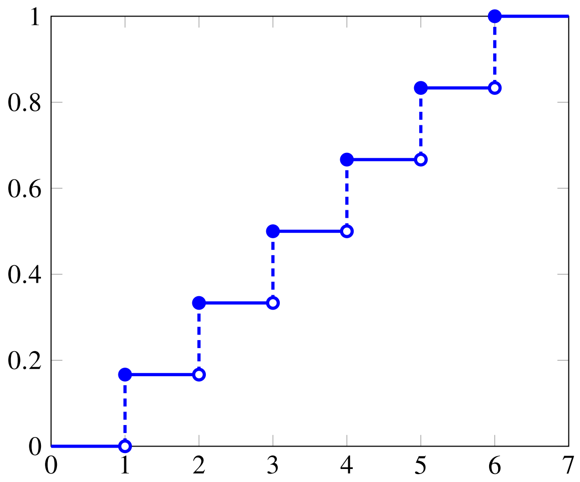
Consequently, our c.d.f. is given by

$$F_X(x) = \begin{cases} P(X \le x) = 0 & \text{if } x < 1, \\ P(X \le x) = \frac{1}{6} & \text{if } 1 \le x < 2, \\ P(X \le x) = \frac{2}{6} & \text{if } 2 \le x < 3, \\ P(X \le x) = \frac{3}{6} & \text{if } 3 \le x < 4, \\ P(X \le x) = \frac{4}{6} & \text{if } 4 \le x < 5, \\ P(X \le x) = \frac{5}{6} & \text{if } 5 \le x < 6, \\ P(X \le x) = 1 & \text{if } x \ge 6. \end{cases}$$

It is not required to write $P(X \le x)$ in our expression of the piecewise function F_X, but here we are only writing it in for purposes of clarity. Our c.d.f. for discrete random variables is obtained by adding up all the individual probabilities $p_X(x)$: for instance, for $P(X \le 5.8)$ (I chose 5.8 at random) we have

$$\begin{aligned} P(X \le 5.8) &= p_X(1) + p_X(2) + p_X(3) + p_X(4) + p_X(5) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= \frac{5}{6}. \end{aligned}$$

The graph of our c.d.f. is given below (next page). Note that the x-axis denotes $X(c) = x$ and the y-axis denotes $F_X(x)$.



Graph of the c.d.f. for the outcomes $i \in \{1,2,3,4,5,6\}$ when rolling one six-sided die

Example (Rolling two six-sided dice)**.** *The sample space for rolling two six-sided dice is* $C = \{(i,j) \in \mathbb{Z} \times \mathbb{Z} : 1 \le i,j \le 6\}$*, with the 36 possible outcomes* $c = (i,j) \in C$*. In this example, we would like to add up the numbers we rolled from the two six-sided dice. The lowest value of* $1+1 = 2$ *occurs when rolls a* 1 *and a* 1*, giving the outcome* $(1,1)$*, and the highest value of* $6+6 = 12$ *occurs when rolls a* 6 *and a* 6*, giving the outcome* $(6,6)$*. It is also possible to obtain numbers* $i+j$ *between* 2 *and* 12*. In other words, we have* $\mathcal{D} = \{2,3,4,5,6,7,8,9,10,11,12\}$*. This makes* $\mathcal{D}$ *a finite space, which means we can write* $x = X(c) \in \{d_1,d_2,d_3,d_4,d_5,d_6,d_7,d_8,d_9,d_{10},d_{11}\}$ *(there are* 11 *elements of* $\mathcal{D}$*). Next, we can display the frequencies of all the possible sums* $i+j$ *below.*

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

An array of all the sums $i+j$

If we count the number of appearances of each sum $i+j$ *(for example,* 7 *appears* 6 *times out of* 36 *outcomes, giving the probability* $p_X(7) = \frac{6}{36}$*), we can outline the range* x *and the probability* $p_X(x)$ *in the following table.*

x	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Output values $p_X(x)$ of the probability mass function (p.m.f.)

Consequently, our c.d.f. is given by

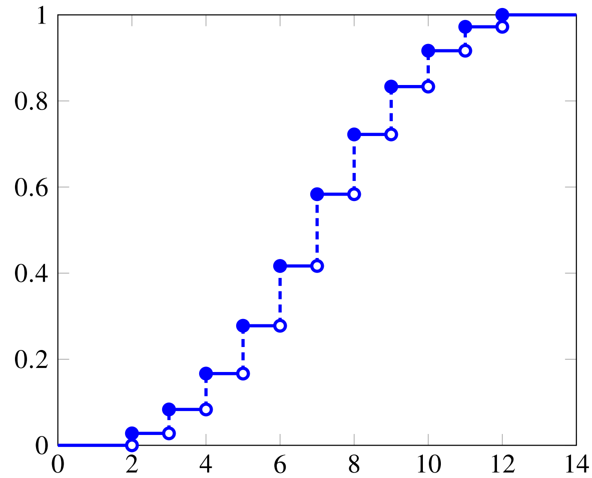
$$F_X(x) = \begin{cases} P(X \le x) = 0 & \text{if } x < 2, \\ P(X \le x) = \frac{1}{36} & \text{if } 2 \le x < 3, \\ P(X \le x) = \frac{3}{36} & \text{if } 3 \le x < 4, \\ P(X \le x) = \frac{6}{36} & \text{if } 4 \le x < 5, \\ P(X \le x) = \frac{10}{36} & \text{if } 5 \le x < 6, \\ P(X \le x) = \frac{15}{36} & \text{if } 6 \le x < 7, \\ P(X \le x) = \frac{21}{36} & \text{if } 7 \le x < 8, \\ P(X \le x) = \frac{26}{36} & \text{if } 8 \le x < 9, \\ P(X \le x) = \frac{30}{36} & \text{if } 9 \le x < 10, \\ P(X \le x) = \frac{33}{36} & \text{if } 10 \le x < 11, \\ P(X \le x) = \frac{35}{36} & \text{if } 11 \le x < 12, \\ P(X \le x) = 1 & \text{if } x \ge 12. \end{cases}$$

It is not required to write $P(X \le x)$ *in our expression of the piecewise function* F_X*, but here we are only writing it in for purposes of clarity. Our c.d.f. for discrete random variables is obtained by adding up all the individual probabilities* $p_X(x)$*: for instance, for*

$P(X \leq 9.4)$ (I chose 9.4 at random) we have

$$\begin{aligned} P(X \leq 9.4) &= p_X(2) + p_X(3) + p_X(4) + p_X(5) + p_X(6) + p_X(7) + p_X(8) + p_X(9) \\ &= 0 + \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} \\ &= \frac{30}{36}. \end{aligned}$$

The graph of our c.d.f. is given below. Note that the x-axis denotes $X(c) = x$ and the y-axis denotes $F_X(x)$.



Graph of the c.d.f. for the sum $i + j$ of the outcomes $\{(i, j) \in \mathbb{Z} \times \mathbb{Z} : 1 \leq i, j \leq 6\}$ when rolling two six-sided dice

Our last example concerns <u>continuous</u> random variables.

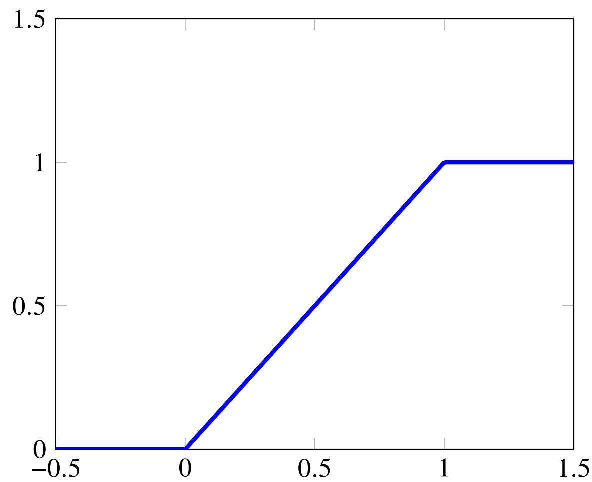
Example. *Let X be a real number chosen at random between 0 and 1. Then our c.d.f. is given by*

$$F_X(x) = \begin{cases} P(X \leq x) = 0 & \text{if } x < 0, \\ P(X \leq x) = x & \text{if } 0 \leq x < 1, \\ P(X \leq x) = 1 & \text{if } x \geq 1. \end{cases}$$

It is <u>not required</u> to write $P(X \leq x)$ in our expression of the piecewise function F_X, but here we are only writing it in for purposes of clarity. Our c.d.f. for continuous random variables is obtained by reasoniably assigning $P_X([a, b]) = b - a$ for any $0 \leq a < b \leq 1$ (see Example 1.5.2 on page 33-34 of the textbook); for instance, for $P(X \leq 0.75)$ (I chose 0.75 at random), then we have

$$\begin{aligned} P(X \leq 0.75) &= P(X < 0) + P(0 \leq x \leq 0.75) \\ &= 0 + (0.75 - 0) \\ &= 0.75. \end{aligned}$$

The graph of our c.d.f. is given below. Note that the x-axis denotes $X(c) = x$ and the y-axis denotes $F_X(x)$.



Graph of the c.d.f. of the real number chosen at random between 0 and 1