LECTURE 3
RANDOM VARIABLES, CUMULATIVE DISTRIBUTION FUNCTIONS (CDFs)

1 Random Variables

Random experiments typically require verbal descriptions, and arguments involving events are often cumbersome. It is much more convenient to work with numbers than sets, which means that we need numeric representations of random experiments and events. Another reason for developing such a representation is that data on the real world often comes in the form of numbers or variables. Hence, we need to develop a formal connection between variables and the probability model.

Definition 1. A random variable X is a function from the sample space Ω to the real line $\mathbb{R}$:

$$X : \Omega \to \mathbb{R}.$$

Thus, the random variable X assigns to *each* $\omega \in \Omega$ a real number $x = X(\omega)$. While $X(\omega)$ is a function, often to simplify the notation we will write X omitting the dependence on ω. Nevertheless, one should remember that a random variable is a mapping from the sample space to the real line. To distinguish between random variables (which are functions) and potential values they can take (realizations), we will use capital letters (X) to denote random variables and small letters (x) to denote realizations.

Table 1: States of the economy, probabilities, and stock prices of different companies

state of the economy ω	$P(\{\omega\})$	hi-tech $X(\omega)$	potatoes $Y(\omega)$	foie gras $Z(\omega)$
weak	0.2	90	20	20
somewhat weak	0.3	90	30	60
somewhat strong	0.3	100	30	95
strong	0.2	150	20	95

Examples of random variables defined on the same Ω are given in Table 1. Note that different random variables may carry different amount of information about the underlying random experiment, and there might be information loss. For example, someone who observes only X from 1 cannot tell whether state of the economy is “weak” or “somewhat weak” when $X = 90$.

Using the definition of random variables and a probability function P defined on events on Ω, we can can define the *distribution* of random variables.

Definition 2. The *distribution* of a random variable $X : \Omega \to \mathbb{R}$, is a *probability function* $P_X(\cdot)$ defined as follows. Let $B \subset \mathbb{R}$,

$$P_X(B) = P\left(\{\omega \in \Omega : X(\omega) \in B\}\right).$$

Since the distribution P_X of a random variable X is directly related to the probability function P, P_X is also a probability function, and therefore satisfies all the properties of probability functions derived earlier.

For example, consider $B = [60, 95]$, and the random variables in Table 1:

$$\{\omega \in \Omega : X(\omega) \in [60, 95]\} = \{\text{weak}, \text{somewhat weak}\}.$$

Hence,

$$P_X([60, 95]) = P\left(\{\text{weak}, \text{somewhat weak}\}\right) = 0.5.$$

Similarly, for the random variable Y and the same B,

$$\begin{aligned} P_Y([60, 95]) &= P\left(\{\omega \in \Omega : 60 \leq Y(\omega) \leq 95\}\right) \\ &= P\left(\emptyset\right) \\ &= 0. \end{aligned}$$

Lastly, for Z and the same B,

$$\begin{aligned} P_Z([60, 95]) &= P\left(\{\omega \in \Omega : 60 \leq Z(\omega) \leq 95\}\right) \\ &= P\left(\{\text{somewhat weak}, \text{somewhat strong,strong}\}\right) \\ &= 0.8. \end{aligned}$$

We can continue this exercise with different subsets B's of the real line, thus constructing for each random variable a mapping from a collection of subsets of $\mathbb{R}$ to [0,1]. While this collection cannot contain all possible subsets of the real line, it is extremely rich and includes all intervals and their unions, i.e. all the subsets of practical interest. The exact technical details are beyond the scope of this course.

In the above example, we have that in general $P_X \neq P_Y$, i.e. there are $B \subset \mathbb{R}$ such that $P_X(B) \neq P_Y(B)$. However, it is possible for two random variables to have the same distribution function.

Definition 3. Two random variables X and Y are equal in distribution (denoted $X =^d Y$) if for all $B \subset \mathbb{R}$,

$$P_X(B) = P_Y(B).$$

Note that $X =^d Y$ does not imply that $X = Y$. It is possible that X and Y have the same distribution, however, if we consider the event that they are equal

$$E = \{\omega \in \Omega : X(\omega) = Y(\omega)\},$$

it may occur only with probability zero, i.e. we can have $X =^d Y$ and $P(E) = 0$.

2 Cumulative distribution function (CDF)

2.1 Definition and examples

The distribution $P_X(B)$ can be computed for a variety of $B \subset \mathbb{R}$. If we restrict our attention to the collection of half-lines $\{(-\infty, u] : u \in \mathbb{R}\}$, the resulting distribution is called the *cumulative distribution function* or *CDF.*

Definition 4. The CDF of a random variable X is a map $F_X : \mathbb{R} \to [0, 1]$ defined as

$$F_X(u) = P_X((-\infty, u]) = P\left(\{\omega \in \Omega : X(\omega) \leq u\}\right)$$

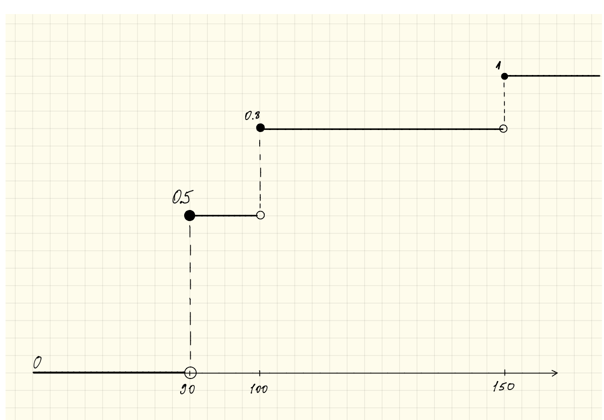
for *all* $u \in \mathbb{R}$.

Using a simplified notation,

$$F_X(u) = P(X \leq u) \text{ for all } u \in \mathbb{R}.$$

Note that $F_X(u)$ must be defined for all real numbers u.

Figure 1: The CDF of X with $P_X(90) = 0.5$, $P_X(100) = 0.3$, and $P_X(150) = 0.2$.



For example, the CDF of X from Table 1 is shown in Figure 1; it is given by

$$F_X(u) = \begin{cases} 0, & u < 90 \\ 0.5, & 90 \le u < 100 \\ 0.8, & 100 \le u < 150 \\ 1, & u \ge 150 \end{cases}$$

Note from the graph is that the CDF is a non-decreasing function, which reflects the fact that it shows cumulative probabilities. As a probability function, it is also bounded between zero and one. Note also that in this particular example, the CDF has discontinuity points - jumps. We will now discuss those properties in general.

2.2 Properties of CDFs

Since CDFs are defined through probability functions, the properties of CDFs follow directly from the properties of probabilities. In what follows below, let $F_X(u)$ denote the CDF of a generic random variable X that has a distribution function P_X.

Theorem 5. ***(Monotonicity)*** *$F_X(u)$ is a non-decreasing function: suppose that $a \le b$, then $F_X(a) \le F_X(b)$.*

Proof. By definition,

$$F_X(u) = P_X((-\infty, u]),$$

which is the probability of an event that $X \le u$. Since $a \le b$, we have the following relationship:

$$(-\infty, a] \subset (-\infty, b].$$

Hence, by Theorem 7(f) in Lecture 1,

$$P_X((-\infty, a]) \le P_X((-\infty, b]),$$

and the result follows. □

To derive other properties of CDFs, we need to discuss one more property of general probability functions: *continuity*. Recall that a function $g : \mathbb{R} \to \mathbb{R}$ is continuous if $\lim_{u \to x} g(u) = g(x)$. A similar result can be established for probability functions: for certain (monotone) sequences of sets $\{A_i : i = 1, 2, \dots\}$ converging (in the sense described below) to a set A, we will have that $\lim_{i \to \infty} P(A_i) = P(A)$, the result known as the continuity of probability functions. This is discussed in detail below.

To establish the continuity of probability functions, we will use the difference operation (between two sets):

Definition 6. Consider two sets G and H such that $G \subset H$. The difference between H and G is defined as

$$H - G = H \cap G^c.$$

Note that since

$$\begin{aligned} G \cap (H - G) &= G \cap (H \cap G^c) \\ &= 0, \end{aligned}$$

and because

$$G \cup (H - G) = H,$$

we have $P(G) + P(H - G) = P(H)$, or

$$P(H - G) = P(H) - P(G).$$

We will also need a notion of the limit of a sequence of sets. This can be naturally defined for *monotone* sequences of sets.

Definition 7. (Limits of monotone sequences sets) (a) Let $\{A_i : i = 1, 2, \ldots\}$ be a monotone *increasing* sequence of sets:

$$A_1 \subset A_2 \subset \ldots.$$

Its limit is defined as

$$\lim_{i\to\infty} A_i = A_1 \cup A_2 \cup \ldots = \cup_{i=1}^{\infty} A_i.$$

(b) Let $\{A_i : i = 1, 2, \ldots\}$ be a monotone *decreasing* sequence of sets:

$$A_1 \supset A_2 \supset \ldots.$$

Its limit is defined as

$$\lim_{i\to\infty} A_i = A_1 \cap A_2 \cap \ldots = \cap_{i=1}^{\infty} A_i.$$

The following theorem established the continuity of probabilities.

Theorem 8. (Continuity of probability) *(a) Let $\{A_i : i = 1, 2, \ldots\}$ be a monotone increasing sequence of events with $A = \lim_{i\to\infty} A_i = \cup_{i=1}^{\infty} A_i$. Then $\lim_{n\to\infty} P(A_n) = P(A)$.*

(b) Let $\{A_i : i = 1, 2, \ldots\}$ be a monotone decreasing sequence of events with $A = \lim_{i\to\infty} A_i = \cap_{i=1}^{\infty} A_i$. Then $\lim_{n\to\infty} P(A_n) = P(A)$.

Proof. To prove part (a), let's define

$$B_1 = A_1,$$
$$B_2 = A_2 - A_1,$$
$$B_3 = A_3 - A_2,$$

and etc. Note that the events B's are mutually exclusive, and

$$\begin{aligned}
A_1 &= B_1 \\
A_2 &= B_1 \cup B_2, \\
A_3 &= B_1 \cup B_2 \cup B_3, \\
&\dots \\
A_n &= B_1 \cup B_2 \cup B_3 \cup \ldots \cup B_n, \\
A &= B_1 \cup B_2 \cup B_3 \cup \ldots \cup B_n \cup \ldots.
\end{aligned}$$

Thus, by the third axiom of probability,

$$\begin{aligned}
P(A) &= P(B_1) + P(B_2) + \ldots \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\
&= \lim_{n\to\infty} P(\cup_{i=1}^{n} B_i) \\
&= \lim_{n\to\infty} P(A_n),
\end{aligned}$$

where the equality in the third line is also by the third axiom of probability.

To prove part (b), define

$$C_n = A_1 - A_n,$$
$$C = A_1 - A.$$

Since $\{A_i\}$ is monotone decreasing to A, the sequence $\{C_i\}$ is monotone increasing to C. By the result in part (a)

$$\begin{aligned}P(C) &= \lim_{n\to\infty} P(C_n) \\ &= \lim_{n\to\infty} P(A_1 - A_n) \\ &= \lim_{n\to\infty} (P(A_1) - P(A_n)) \\ &= P(A_1) - \lim_{n\to\infty} P(A_n).\end{aligned}$$

On the other hand,

$$P(C) = P(A_1 - A) = P(A_1) - P(A).$$

Hence,

$$P(A_1) - \lim_{n\to\infty} P(A_n) = P(A_1) - P(A),$$

or

$$\lim_{n\to\infty} P(A_n) = P(A).$$

□

We can now turn to remaining properties of CDFs.

Theorem 9. *(a)* $\lim_{u\to\infty} F_X(u) = 1$.

(b) $\lim_{u\to-\infty} F_X(u) = 0$.

(c) F_X *is a right-continuous function:* $\lim_{u\downarrow x} F_X(u) = F(x)$*, where* $u \downarrow x$ *denotes that* u *approaches* x *from above (from the right).*

(d) F_X *may not be left continuous:* $\lim_{u\uparrow x} F_X(u) \leq F(x)$*, where* $u \uparrow x$ *denotes that* u *approaches* x *from below (from the left).*

Proof. Since $F_X(u) = P_X(A)$ for the set $A = (-\infty, u]$, we can prove the results by using the continuity of P_X and appropriate choices of sequences of sets of the form $(-\infty, u]$.

For part (a), consider

$$A_n = (-\infty, n], \quad n = 1, 2, \ldots.$$

In this case we have a monotone *increasing* sequence of sets, and

$$A = \lim_{n\to\infty} A_n = \mathbb{R}.$$

Hence,

$$\begin{aligned}\lim_{n\to\infty} F_X(n) &= \lim_{n\to\infty} P_X(A_n) \\ &= P_X(\mathbb{R}) \\ &= 1.\end{aligned}$$

Here, the first equality holds by the definition of F_X, the second equality is by the continuity of P_X and because the sequence $A_n = (-\infty, n]$ expands to the entire real line. The last equality holds because

$$P_X(\mathbb{R}) = P(-\infty < X < \infty),$$

and the event $-\infty < X < \infty$ is always true because X is only allowed to take on real values (numbers), and ∞ is not a number.

The proof of part (b) is identical, however instead of an increasing sequence of events, we will use a decreasing sequence. Let

$$A_n = (-\infty, -n],$$

so that

$$A = \lim_{n\to\infty} A_n = \emptyset.$$

Then,

$$\begin{aligned}
\lim_{n\to\infty} F_X(n) &= \lim_{n\to\infty} P_X(A_n) \\
&= P_X(\emptyset) \\
&= 0.
\end{aligned}$$

For part (c), since $F_X(x) = P_X((-\infty, x])$ we need to construct a monotone sequence of sets approaching the set $(-\infty, x]$ "from the right". Let

$$A_n = (-\infty, x + 1/n].$$

Note that $\{A_n : n = 1, 2, \ldots\}$ is a monotone shrinking sequence of sets. Thus,

$$A = \lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n.$$

Since $x \in A_n$ for all $n = 1, 2, \ldots$, the limit of the sequence is

$$A = (-\infty, x].$$

Hence,

$$\begin{aligned}
\lim_{u\downarrow x} F_X(u) &= \lim_{n\to\infty} P_X((-\infty, x + 1/n]) \\
&= P_X(\lim_{n\to\infty} (-\infty, x + 1/n]) \\
&= P_X((-\infty, x]) \\
&= F_X(x).
\end{aligned}$$

For part (d), we need to construct a monotone sequence of sets approaching $(-\infty, x]$ "from the left". Let

$$A_n = (-\infty, x - 1/n].$$

Note that $\{A_n : n = 1, 2, \ldots\}$ is a monotone expanding sequence of sets. Thus,

$$A = \lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n.$$

Since $x \notin A_n$ for any $n = 1, 2, \ldots$, we have that

$$x \notin \cup_{n=1}^{\infty} A_n.$$

Hence, the limit of the sequence is

$$A = (-\infty, x).$$

Therefore,

$$\begin{aligned} \lim_{u\uparrow x} F_X(u) &= \lim_{n\to\infty} P_X((-\infty, x - 1/n]) \\ &= P_X(\lim_{n\to\infty}(-\infty, x - 1/n]) \\ &= P_X((-\infty, x)) \\ &= F_X(x) - P(X = x) \\ &\leq F_X(x). \end{aligned}$$

Note that

$$\lim_{u\uparrow x} F_X(u) = P_X((-\infty, x))$$

is different from the CDF of X at x unless the probability that $X = x$ is exactly zero. □

Remark. Any function that satisfies the properties described in Theorems 5 and 9 is a CDF function: Given a function $F(x)$ that satisfies the aforementioned properties, one can always construct a random variable X such that its CDF is equal to $F(x)$. We will demonstrate that later in the course.

Let's define the left limit of the CDF as

$$F_X^-(x) = \lim_{u\uparrow x} F_X(u) = P_X((-\infty, x)).$$

The CDF is continuous if for every $x \in \mathbb{R}$

$$F_X^-(x) = F_X(x).$$

Definition 10. A random variable is *continuous* or *continuously distributed* if its CDF is continuous everywhere.

Note that since

$$P(X = x) = F_X(x) - F_X^-(x), \tag{1}$$

the probability that a continuously distributed random variable takes on value x is exactly zero *for all* $x \in \mathbb{R}$.

Definition 11. A random variable is *discrete* if its CDF is a step function.

A discrete random variable takes a finite or countable number of values with strictly positive probabilities. In view of (1), those probabilities are captured by the jumps of the CDF.

The CDF provides a complete description for the behavior of a random variable: from the knowledge of the CDF, we can calculate all probabilities of practical importance concerning the behavior of a random variable. For example, for $a \le b$,

$$\begin{aligned} P(a < X \le b) &= P_X((-\infty, b]) - P_X((-\infty, a]) \\ &= F_X(b) - F_X(a), \\ P(a \le X \le b) &= P_X((-\infty, b]) - P_X((-\infty, a)) \\ &= F_X(b) - F_X^-(a). \end{aligned}$$

Example. Suppose X is distributed with a CDF

$$F_X(x) = \begin{cases} 0, & x \le 0 \\ x, & 0 < x \le 1 \\ 1, & x > 1 \end{cases}$$

This function satisfies the properties described in Theorems 5 and 9, so we have a CDF. (This distribution is known as *Uniform(0,1) distribution.*) Note also that the function is continuous everywhere:

$$F_X(x) - F_X^-(x) = 0$$

for all $x \in \mathbb{R}$. Hence, it does not matter whether intervals closed or open when computing probabilities. Now,

$$\begin{aligned} P(0 < X < 0.5) &= 0.5 - 0 = 0.5, \\ P(0.5 < X < 1) &= 1 - 0.5 = 0.5, \\ P(0.5 < X < 1.5) &= 1 - 0.5 = 0.5. \end{aligned}$$