Discrete random variables — Definitions and examples

MTH382 Probability Theory for Finance and Actuarial Science

Parin Chaipunya
KMUTT
└Mathematics @ Faculty of Science

Overview

The goal of this lecture is to intuitively introduce a random variable and formally introduce a discrete random variable. We shall go through the definition and a series of examples to get a better understanding of the concept.

Definition

What is the main idea of a random variable ?

- A random variable is **a representation of an uncertain event, often in a numerical manner.**
- Eventhough a random variable is not really a variable, it is usually treated as if it is a variable, but its value depends on a random event.
- We actually have used the idea of a random variable before, without realizing or calling it a random variable.

Definition of a random variable.

Always let $(\Omega, \mathcal{F})$ be a measurable space.

Definition 1.
A function $X : \Omega \to E$, with a countable set E, is called a **discrete random variable** (or briefly a **discrete r.v.**) on $(\Omega, \mathcal{F})$ if the preimage

$$\{X = x\} := X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$$

is an event in $\mathcal{F}$, for all $x \in E$. ▲

Remark.
We do not need a probability measure to discuss random variables, since this is just a tool to represent events in a simpler manner.

Some comments

- The countable set E could be any abstract set. In many applications (including **dice rolling** or **counting heads** in coin tosses), it is natural to set $E = \mathbb{Z}$ (or more commonly, $E = \mathbb{N}$) which has an advantage of being computable. In some other applications (like **head-or-tail** or **card faces**), there is no natural way to represent an event numerically. One would see that having $E = \mathbb{Z}$ (or $E = \mathbb{N}$) has a huge benefit that allows the study of **expectation**.
- The discrete (or later, continuous) nature of a random variable is usually understood directly from the context (the countability of E) and the term **random variable** (or **r.v.**) is usually used with its prefix omitted when there is no possible confusion. ▲

Dice rolling

Example 2.

Consider rolling a dice. The sample space is then $\Omega = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}$ and the σ-field of all events is the power set $\mathcal{F} = 2^{\Omega}$. We define $X : \Omega \to \{1, 2, 3, 4, 5, 6\}$, representing numerically an outcome, by

$$X(\epsdice{1}) = 1, \quad X(\epsdice{2}) = 2, \quad X(\epsdice{3}) = 3, \quad X(\epsdice{4}) = 4, \quad X(\epsdice{5}) = 5, \quad X(\epsdice{6}) = 6. \tag{1}$$

Then X is a r.v. ▲

Dice rolling II

In the following example, we still consider rolling a dice but we are only able to distinguish odd and even faces. One would see that the same r.v. function is no longer available as a r.v.

Example 3.
Consider $\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$ and $\mathcal{F} = \{\varnothing, \Omega, \{\text{⚀}, \text{⚂}, \text{⚄}\}, \{\text{⚁}, \text{⚃}, \text{⚅}\}\}$. Then the function $X : \Omega \to \{1, 2, 3, 4, 5, 6\}$ defined as (1) is not a r.v. For instance, one could notice that $X^{-1}(1) = \{\text{⚀}\} \notin \mathcal{F}$. ▲

The next is an example of a r.v. that does not take numerical values.

Example 4.
Consider tossing a coin twice, so that $\Omega = \{(\mathrm{H}, \mathrm{H}), (\mathrm{H}, \mathrm{T}), (\mathrm{T}, \mathrm{H}), (\mathrm{T}, \mathrm{T})\}$ and $\mathcal{F} = 2^{\Omega}$. Set $E = \{\mathrm{H}, \mathrm{T}\}$ and define $X_1, X_2 : \Omega \to E$ by

$$\begin{aligned} X_1((\mathrm{H}, \mathrm{H})) = X_1((\mathrm{H}, \mathrm{T})) = \mathrm{H}, \\ X_1((\mathrm{T}, \mathrm{H})) = X_1((\mathrm{T}, \mathrm{T})) = \mathrm{T}, \\ X_2((\mathrm{H}, \mathrm{H})) = X_2((\mathrm{T}, \mathrm{H})) = \mathrm{H}, \\ X_2((\mathrm{H}, \mathrm{T})) = X_2((\mathrm{T}, \mathrm{T})) = \mathrm{T}. \end{aligned}$$

We could see that X_1 and X_2 are r.v.s representing the outcomes of the first and second toss, respectively. These r.v.s are examples of those in which numerical representation is not natural. ▲

Under some circumstances, it also makes sense to assign numerical values to these non-numerical outcomes. We shall see this in the following example.

Example 5.
Consider tossing a coin twice and let Ω be the sample space with $\mathcal{F} = 2^{\Omega}$ being its σ-field of events. We shall now consider a r.v. X that represents the number of H occuring in the two tosses. This X is defined by

$$X((\mathrm{T},\mathrm{T})) = 0, \quad X((\mathrm{H},\mathrm{T})) = X((\mathrm{T},\mathrm{H})) = 1, \quad X((\mathrm{H},\mathrm{H})) = 2.$$

With this notion, we could consider probabilities like

$$P(X = 0) = P(\{X = 0\}) = P(\{\omega \in \Omega \mid X(\omega) = 0\}) = P(\{(\mathrm{T},\mathrm{T})\}) = \frac{1}{4}. \qquad \blacktriangle$$

Operations on random variables

Operations on random variables.

It is quite natural to ask which operations preserves the resulting function to be again a random variable.

We start with the ones that work for r.v.s taking numerical values.

Proposition 6.

(a) *If* $X, Y : \Omega \to \mathbb{R}$ *are discrete r.v.s, then* $Z := X + Y$ *a discrete r.v.*

(b) *If* $X : \Omega \to \mathbb{R}$ *is a discrete r.v. and* $\lambda \in \mathbb{R}$*, is* $Z := \lambda X$ *a discrete r.v.*

(c) *If* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *are discrete r.v.s and* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$*, then*

$$Z := \lambda_1 X_1 + \cdots + \lambda_n X_n$$

is a discrete r.v.

(d) *If* X *and* Y *are discrete r.v.s, is* $Z := XY$ *a discrete r.v.*

Examples of why we need to operate on r.v.s.

- If X_1 and X_2 represents the outcomes of the first and second dice rollings, then $X_1 + X_2$ states the sum of the two faces.
- A person invested in n assets. She invested λ_i unit in the i^{th} asset. If X_i represents the value of the i^{th} asset (assumed to be a discrete r.v. due to the uncertainty of the market), then

$$Z := \lambda_1 X_1 + \cdots + \lambda_n X_n$$

  represents her worth of the investment.

Operations on random variables.

Here, we present the ones that do not require the r.v.s to take values in $\mathbb{R}$.

Proposition 7.
Let E and F be two countable sets. If $X : \Omega \to E$ is a r.v. and $f : E \to F$, then $Y := f(X) = f \circ X$ is also a r.v.

Proposition 8.
Let E_1 and E_2 be two countable sets, and $X_1 : \Omega \to E_1$ and $X_2 : \Omega \to E_2$ are two r.v.s. Then $X := (X_1, X_2) : \Omega \to E_1 \times E_2$ is also a r.v.

Example of coupled r.v.s.

- Drawing a card of a standard deck has an outcome that could be represented by two discrete r.v.s. Let $X : \Omega \to \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$ be the r.v. representing the **suit**, and $Y : \Omega \to \{2, \ldots, 10, \mathrm{J}, \mathrm{Q}, \mathrm{K}, \mathrm{A}\}$ the **ranks**. Then to represent a card, we need both X and Y coupled into a new r.v. $Z = (X, Y)$.

Takeaways

Takeaways

- Random variables are used to represent events of interests.
- They are functions but treated more as variables.
- Several operations can be used to generate new random variables.


Bon journée!