Random Variables

This text is to help you understand what a random variable is conceptually and mathematically, and the notations.

Sample Space S is a set of items of interest. Each element ω in S is a sample (or outcome). $\omega \in S$.

A sample can be of any kind: an object, a person, a throw of dice, a time period, a number, or some combinations.

Sampling is to randomly select samples from the sample space, or to perform a series of "trials" through a random process to obtain outcomes.

e.g. If we throw two dices, the sample space is $(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,5),(6,6)$. $|S| = 36$.

ω can be any one of the 36 outcomes. ($(1,2)$ and $(2,1)$ are two different outcomes.)

Random Variable X is a real-valued function defined over the sample space S.

Given a sample in the sample space, the random variable will produce a real number for that sample.

i.e. Function X takes ω as argument and returns a real value.

$$\boxed{X: \ S \to \mathbb{R}, \quad \omega \to X(\omega), \quad \text{where } \omega \in S, \text{ a "simple event" (a sample which produces a possible outcome).}}$$

In the two-dice experiment, $X(\omega) = a + b$, where $\omega = (a, b)$.

Domain of Variation: All possible values of $X(\omega)$ (the image of X). $\boxed{S_X = \{x \in \mathbb{R} : x = X(\omega), \omega \in S\}.}$

In the two-dice experiment, $S_X = \{2, 3, \ldots, 12\}$. $|S_X| = 11$.

Events: An event is a subset of S of which the elements satisfy an assertion on X, and can be written as $(assertion)$.

For example, $(X = x) = \{\omega \in S : X(\omega) = x\}$, $(X \leq x) = \{\omega \in S : X(\omega) \leq x\}$.

In the two-dice experiment, the event of throwing a 5 is $\mathbf{E} = \{(1,4),(2,3),(3,2),(4,1)\}$. So $|(X = 5)| = 4$.

To throw 4 or less, $\mathbf{E} = \{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\}$. So $|(X \leq 4)| = 6$.

Probability: A ratio of the size of the event and the size of the sample space. $\boxed{P(\mathbf{E}) = \dfrac{|\mathbf{E}|}{|S|}.}$

$\boxed{P(\phi) = 0, \quad P(S) = 1.}$ $\because \phi \subseteq \mathbf{E} \subseteq S$, $\therefore \boxed{0 \leq P(\mathbf{E}) \leq 1.}$

In the two-dice experiment, $P\left((X=5)\right) = \dfrac{|(X=5)|}{|S|} = \dfrac{4}{36} = \dfrac{1}{9}$, $P\left((X \leq 4)\right) = \dfrac{|(X \leq 4)|}{|S|} = \dfrac{6}{36} = \dfrac{1}{6}$.

CDF: The Cumulative Distribution Function is defined as $\boxed{F_X(x) = P(X \leq x)}$, $x \in \mathbf{R}$. F_X is non-decreasing.

If there is only one random variable in the context, the subscribe X can be omitted.

$$\boxed{\text{For an interval } [a,b], \quad P(a < x \leq b) = F(b) - F(a), \quad 0 = \lim_{x \to -\infty} F(x) \leq F(a) \leq F(b) \leq \lim_{x \to +\infty} F(x) = 1.}$$