Experiments, Outcomes, Events and Random Variables: A Revisit

Berlin Chen
Department of Computer Science & Information Engineering
National Taiwan Normal University

Reference:

- D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability*

Experiments, Outcomes and Event

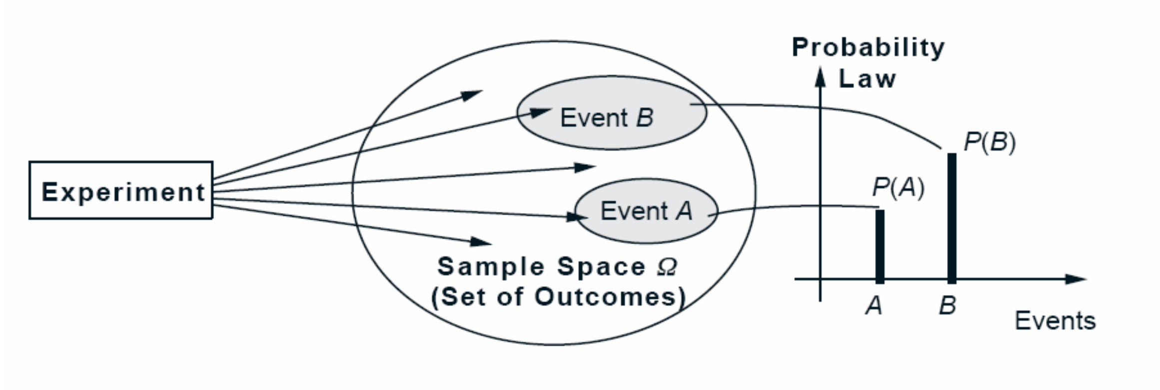
- An **experiment**
  - Produces exactly one out of several possible **outcomes**
  - The set of all possible outcomes is called the **sample space** of the experiment, denoted by
  - A subset of the sample space (a collection of possible outcomes) is called an **event**

- Examples of the **experiment**
  - A single toss of a coin (finite outcomes)
  - Three tosses of two dice (finite outcomes)
  - An infinite sequences of tosses of a coin (infinite outcomes)
  - Throwing a dart on a square (infinite outcomes), etc.

Probabilistic Models

- A probabilistic model is a mathematical description of an uncertainty situation or an experiment

- Elements of a probabilistic model
  - The **sample space**
    - The set of all possible outcomes of an experiment
  - The **probability law**
    - Assign to a set A of possible outcomes (also called an **event**) a nonnegative number $\mathbf{P}(A)$ (called the **probability** of A) that encodes our knowledge or belief about the collective "likelihood" of the elements of A



Three Probability Axioms

- **Nonnegativity**
  - $\mathbf{P}(A) \geq 0$, for every event A

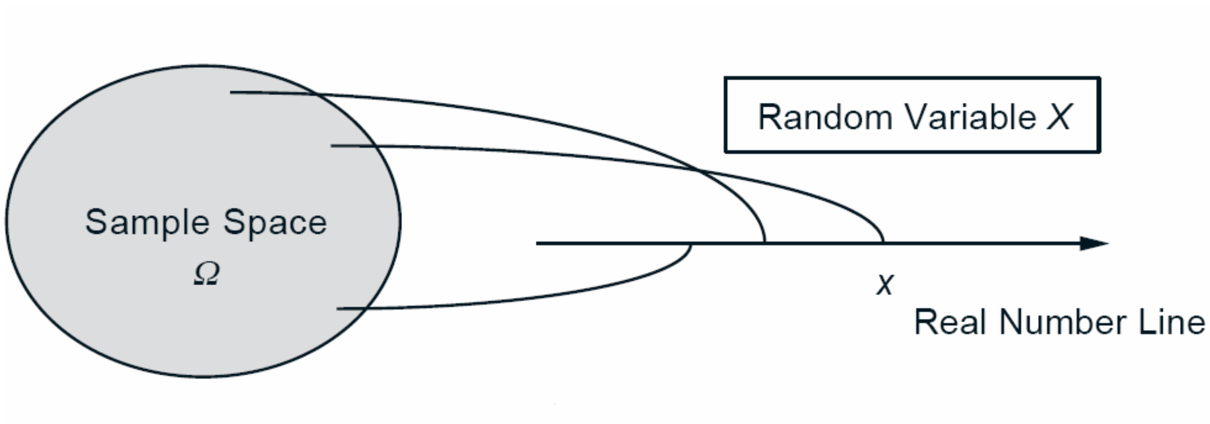
- **Additivity**
  - If A and B are two disjoint events, then the probability of their union satisfies

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$$

- **Normalization**
  - The probability of the entire sample space Ω is equal to 1, that is, $\mathbf{P}(\Omega) = 1$

Random Variables

- Given an experiment and the corresponding set of possible outcomes (the sample space), a random variable associates a particular number with each outcome
  - This number is referred to as the (numerical) value of the random variable
  - We can say a random variable is a real-valued function of the experimental outcome



Discrete/Continuous Random Variables (1/2)

- A random variable is called **discrete** if its **range** (the set of values that it can take) is finite or at most countably infinite

$$\text{finite} : \{1, 2, 3, 4\}, \text{countably infinite} : \{1, 2, \cdots\}$$

- A random variable is called **continuous (not discrete)** if its **range** (the set of values that it can take) is uncountably infinite
  - E.g., the experiment of choosing a point a from the interval $[-1, 1]$
    - A random variable that associates the numerical value a^2 to the outcome a is not discrete

Discrete/Continuous Random Variables (2/2)

- A discrete random variable X has an associated **probability mass function** (PMF), $p_X(x)$, which gives the probability of each numerical value that the random variable can take

- A continuous random variable X can be described in terms of a **nonnegative** function $f_X(x)$ $\left(f_X(x) \geq 0\right)$, called the **probability density function** (PDF) of X, which satisfies

$$\mathbf{P}(X \in B) = \int_B f_X(x)dx$$

for every subset *B* of the real line

Cumulative Distribution Functions (1/4)

- The cumulative distribution function (CDF) of a random variable X is denoted by $F_X(x)$ and provides the probability $\mathbf{P}(X \leq x)$

$$F_X(x) = \mathbf{P}(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(k), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{x} f_X(t)\,dt, & \text{if } X \text{ is continuous} \end{cases}$$

  - The CDF $F_X(x)$ accumulates probability up to x
  - The CDF $F_X(x)$ provides a unified way to describe all kinds of random variables mathematically

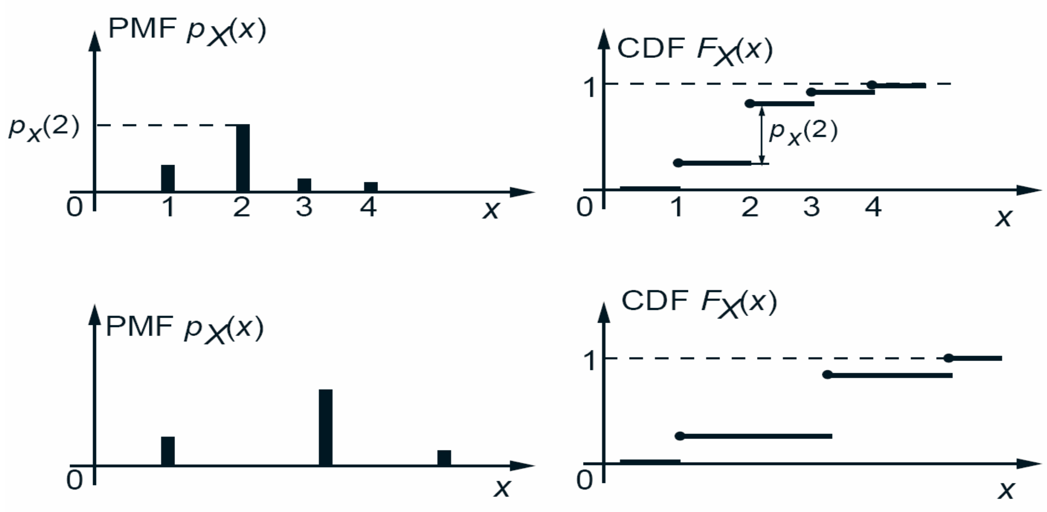
Cumulative Distribution Functions (2/4)

- The CDF $F_X(x)$ is monotonically non-decreasing

$$\text{if} \quad x_i \le x_j, \text{ then } F_X(x_i) \le F_X(x_j)$$

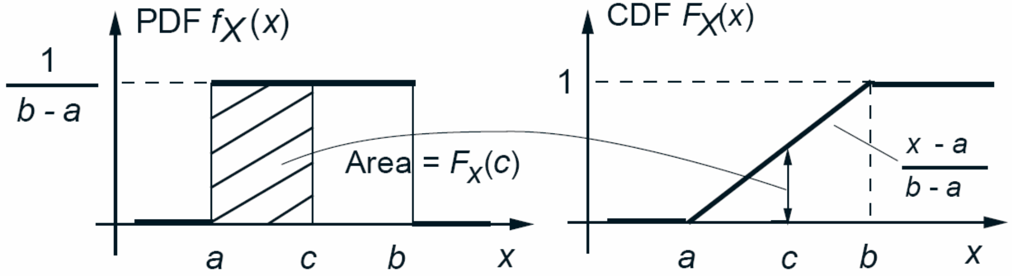
- The CDF $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$

- If X is discrete, then $F_X(x)$ is a piecewise constant function of x



Cumulative Distribution Functions (3/4)

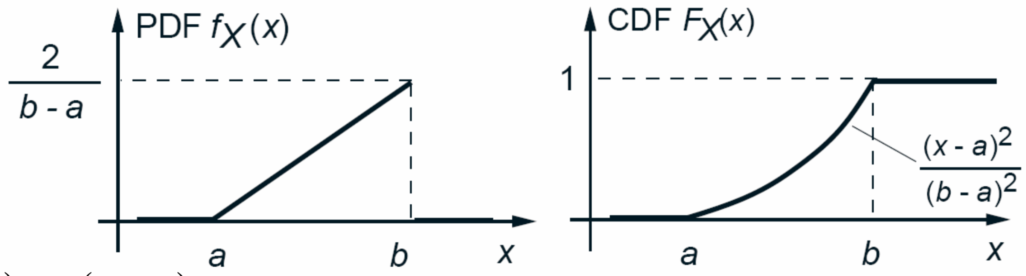
- If X is continuous, then $F_X(x)$ is a continuous function of x



$$f_X(x)=\frac{1}{b-a},\quad for\ a\le x\le b$$

$$F_x(X\le x)=\int_a^x f_X(t)dt=\int_a^x\frac{1}{b-a}dt$$

$$=\frac{x-a}{b-a}$$



$$f_X(x)=c(x-a),\quad for\ a\le x\le b$$

$$\Rightarrow \int_a^b c(x-a)dx=\frac{c}{2}(x-a)^2\Big|_a^b=1$$

$$\Rightarrow c=\frac{2}{(b-a)^2}$$

$$\Rightarrow f_X(b)=\frac{2(b-a)}{(b-a)^2}=\frac{2}{b-a}$$

$$F_x(X\le x)=\int_a^x f_X(t)dt=\int_a^x\frac{2(t-a)}{(b-a)^2}dt$$

$$=\frac{(x-a)^2}{(b-a)^2}$$

Cumulative Distribution Functions (4/4)

- If X is discrete and takes integer values, the PMF and the CDF can be obtained from each other by summing or differencing

$$F_X(k) = \mathbf{P}(X \le k) = \sum_{i=-\infty}^{k} p_X(i),$$

$$p_X(k) = \mathbf{P}(X \le k) - \mathbf{P}(X \le k-1) = F_X(k) - F_X(k-1)$$

- If X is continuous, the PDF and the CDF can be obtained from each other by integration or differentiation

$$F_X(x) = \mathbf{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\,dt,$$

$$p_X(x) = \frac{dF_X(x)}{dx}$$

  – The second equality is valid for those x for which the CDF has a derivative (e.g., the piecewise constant random variable)

Conditioning

- Let X and Y be two random variables associated with the same experiment
  - If X and Y are discrete, the conditional PMF of X is defined as (where $p_Y(y)$)

$$p_{X|Y}(x|y) = \mathbf{P}(X = x | Y = y) = \frac{\mathbf{P}(\{X = x\} \cap \{Y = y\})}{\mathbf{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

  - If X and Y are continuous, the conditional PDF of X is defined as (where $f_Y(y) > 0$)

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Independence

- Two **random variables** X and Y are **independent** if

$$p_{X,Y}(x, y) = p_X(x) p_Y(y), \quad \text{for all } x, y$$ (If X and Y are discrete)

$$f_{X,Y}(x, y) = f_X(x) f_Y(y), \quad \text{for all } x,y$$ (If X and Y are continuous)

- If two **random variables** X and Y are **independent**

$$p_{X|Y}(x|y) = p_X(x), \quad \text{for all } x, y$$ (If X and Y are discrete)

$$f_{X|Y}(x|y) = f_X(x), \quad \text{for all } x,y$$ (If X and Y are continuous)

Expectation and Moments

- The **expectation** of a random variable X is defined by

$$\mathbf{E}[X] = \sum_x x p_X(x) \qquad \text{(If } X \text{ is discrete)}$$

or

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx \qquad \text{(If } X \text{ is continuous)}$$

- The ***n*-th moment** of a random variable X is the expected value of a random variable X^n (or the random variable

$$\mathbf{E}[X^n] = \sum_x x^n p_X(x) \qquad \text{(If } X \text{ is discrete)}$$

or

$$\mathbf{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx \qquad \text{(If } X \text{ is continuous)}$$

  – The 1st moment of a random variable is just its mean

Variance

- The **variance** of a random variable X is the expected value of a random variable $(X - \mathbf{E}(X))^2$

$$\begin{aligned} \operatorname{var}(X) &= \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\ &= \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2 \end{aligned}$$

- The **standard derivation** is another measure of dispersion, which is defined as (a square root of variance)

$$\sigma_X = \sqrt{\operatorname{var}(X)}$$

  – Easier to interpret, because it has the same units as X

More Factors about Mean and Variance

- Let X be a random variable and let $Y = aX + b$

$$\mathbf{E}[Y] = a\mathbf{E}[X] + b$$

$$\operatorname{var}(Y) = a^2 \operatorname{var}(X)$$

- If X and Y are independent random variables

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$$

$$\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y)$$

$$\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)]\mathbf{E}[h(Y)]$$

g and h are functions of X and Y, respectively