Randomness

Outline

- Introduction to probability
- Conditional probability
- Random variables
- Some useful probability distributions

Random variables

Random variables

- A random variable is a numeric quantity whose value depends on the outcome of a random event.
    - We use a capital letter, like X, to denote a random variable
    - The values of a random variable are denoted with a lowercase letter, in this case x
    - For example, $P(X = x)$

Random variables

- Discrete random variables often take only integer values
  - Example: number of students present, gender of an unborn baby
- Continuous random variables take real (decimal) values
  - Example: tomorrow's PM 2.5 level, your final grade

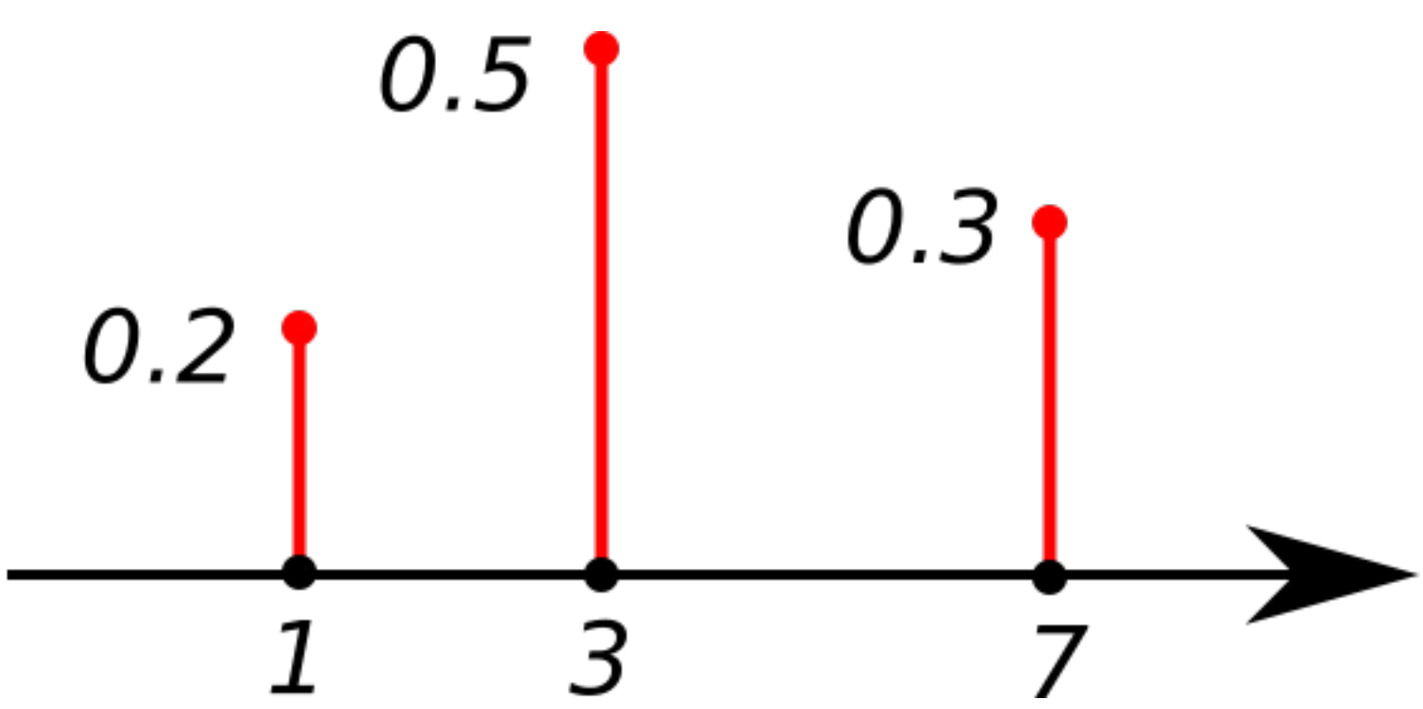
Probability Mass Functions

A probability mass function (pmf), $P(X = x) = f(x)$, of a discrete random variable X, is a function that gives the probability that a discrete random variable is exactly equal to some value. It should satisfy the following properties:

1. $P(X = x) = f(x) \geq 0$
2. $\sum f(x) = 1$

Probability Mass Functions

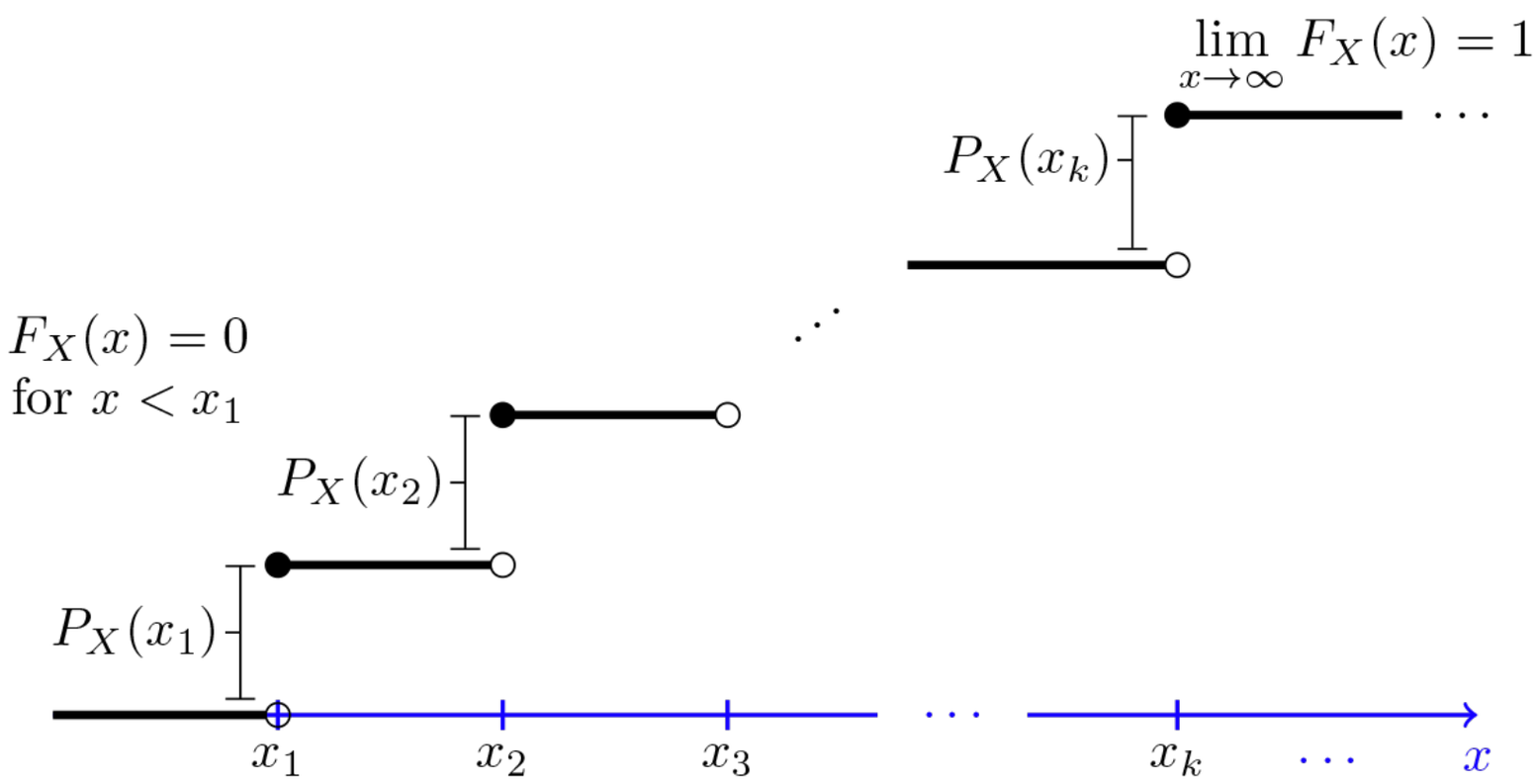
- Since the image of X is countable, the probability mass function f(x) is zero for all but a countable number of values of x.



Cumulative Distribution Function

- The cumulative distribution function (c.d.f.) for a discrete random variable is

$$F(x) = Prob(X \leq x) = \sum_{t \leq x} f(t)$$



Probability density function

A probability density function (pdf) of a continuous random variable is a function satisfying the following properties:

1. $f_X(x) \geq 0, -\infty < x < \infty$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$
3. $Prob(a < X \leq b) = \int_{a}^{b} f_X(x)dx$

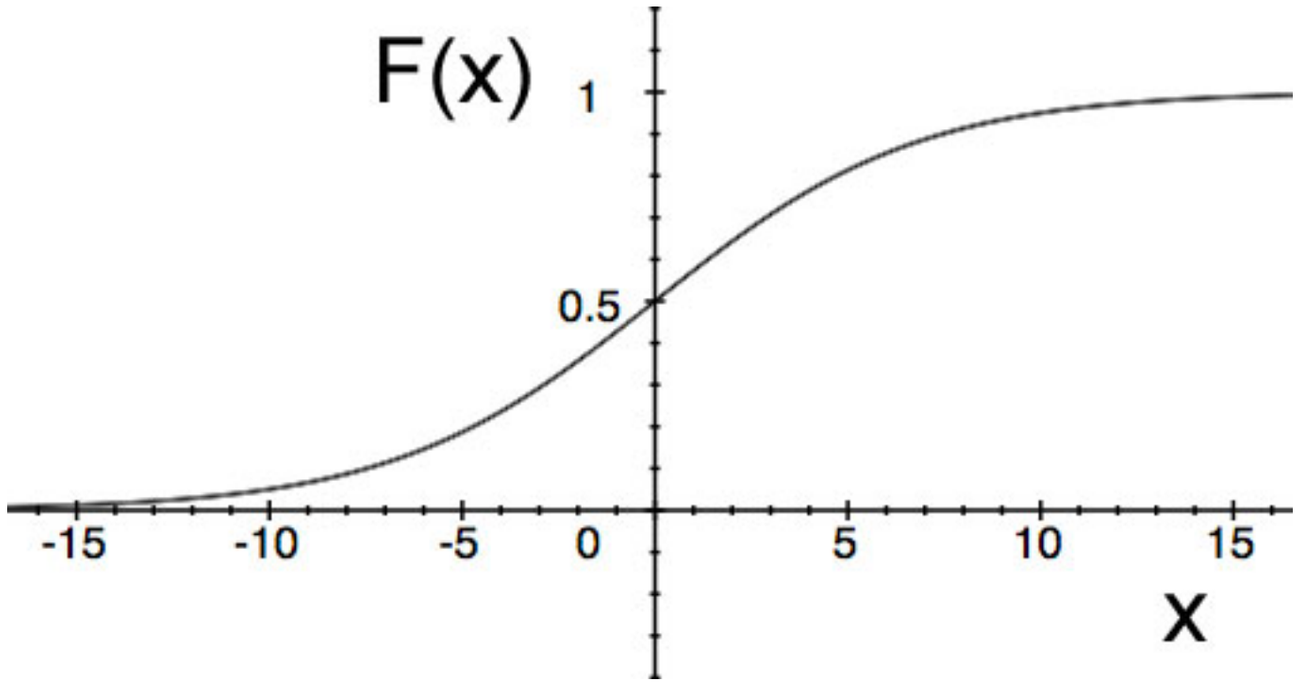
Probabilities from continuous distributions

- For a continuous random variable X, the probability $P(X = x)$ is defined as 0.
- $P(a < X \leq b) = F_X(a) - F_X(b) = \int_a^b f(x)dx$

C.d.f. for a continuous random variable

- The c.d.f. for a continuous random variable is

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

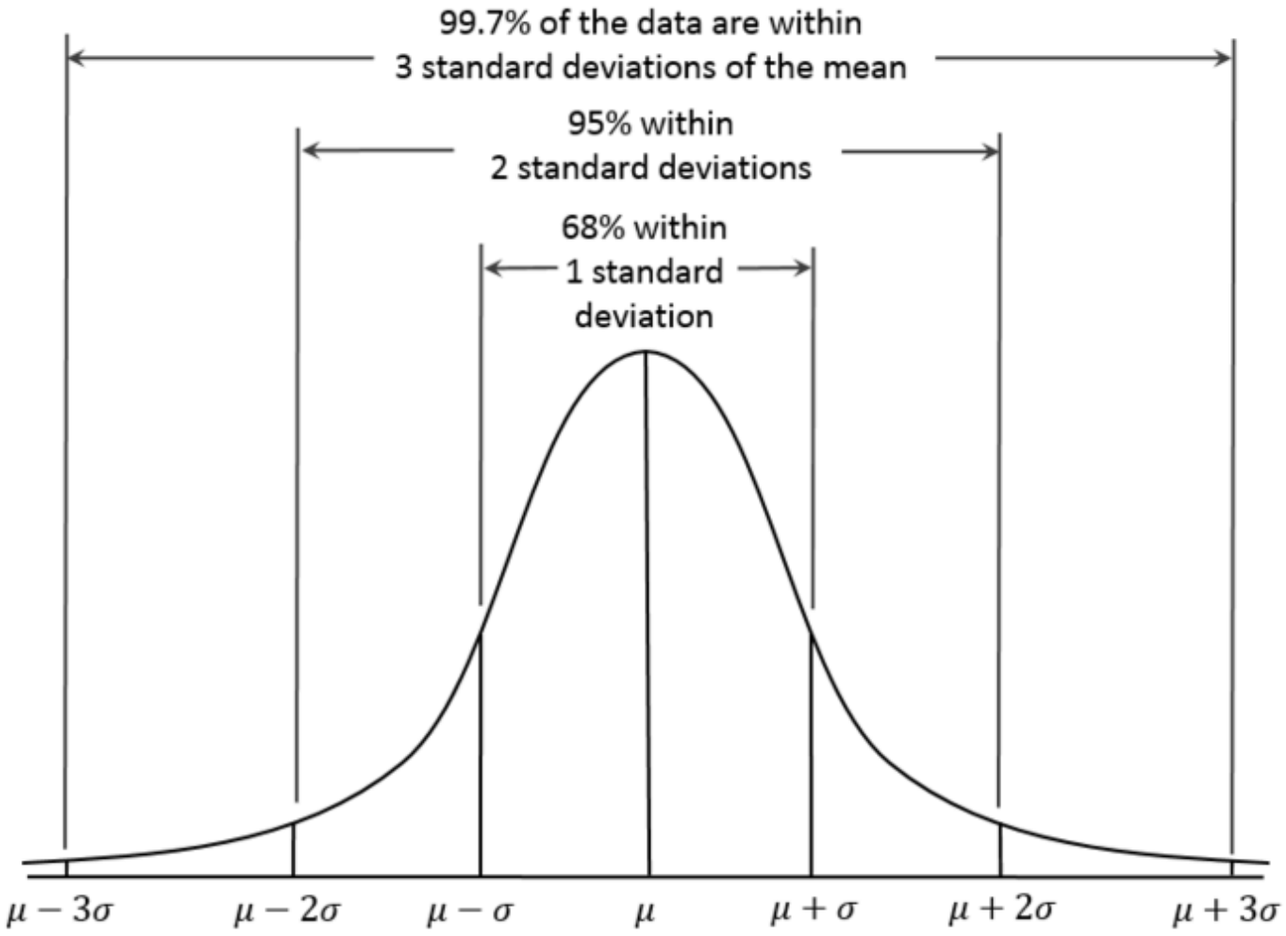


Useful probability distributions

Bernoulli distribution

Binomial distribution

Normal distribution



Central limit theorem

Distributions derived from normal random variables

Generating pseudo random numbers in Python

- random table
- numpy.random
- seed