# L4. Random Variable and Probability Functions

- Random Variable
- Discrete distribution (p.m.f.)
- Continuous distribution (p.d.f. and c.d.f.)
- expected value

## 1. Random Variable:

**Definition**,

Given a random experiment with an outcome space $S$, a function $X$ that assigns one and only one real number $X(s) = x$ to each element $s$ in $S$ is called a random variable. i.e.

$$s \xrightarrow{X} x = X(s)$$

$$s \in S \qquad\qquad x \in \mathbb{R}$$

How to understand this?

- r.v. $X$ represents all the $\boxed{\text{outcomes}}$ from an experiment

- If the outcomes are $\boxed{\text{numerical}}$, keep them how they are.

- If the outcomes are categorical, assign real numbers. (Encoding)

Type of Random Variables (Data):

$\boxed{\text{Discrete}}$
- Numerically Discrete: outcomes of rolling a die
- Binary: outcome is Yes or No ($\rightarrow$ 1 or 0)
- Categorical: outcome is mutually exclusive labels
- Ordinary: categorical data that admits a natural ordering

$\boxed{\text{Continuous}}$ - Outcomes that maps naturally on a real number

line. Height, weight, Age, Time
or any numerical measurements that
can not be concluded with limited
number of categories.

Remark: All statistical / machine learning methods are using
math equations. you need a NUMBER to represent
your data to plug in.

2. Probability Distribution.

When performing a random experiment, what's gonna happen?

- Possible Outcomes ( r.v. $X$ and its values)

- chance of each outcome ( probability)

A probability distribution fully describe the outcomes of an experiment
by giving

- $x \in \{ x : X(s) = x, \ s \in S \}$

- p.m.f or p.d.f. $f(x)$

## 3. Discrete Distribution

For a r.v. $X$ that has discrete values $\{x\}$,

the $\boxed{\text{probability mass function (p.m.f.)}}$

$$f(x) = \mathbb{P}(X = x)$$

that satisfies :

(a) $f(x) > 0, \quad x \in S_x$

(b) $\sum_{x \in S_x} f(x) = 1$

(c) $\mathbb{P}(X \in A) = \sum_{x \in S_A} f(x)$

$\boxed{\text{Ex 1.}}$ Roll a die twice and let $X$ equal the larger outcome when different or the common value when the same, give the distribution of $X$.

- Possible values (outcomes) of $X$:

$S = \{1, 2, 3, 4, 5, 6\}$

- p.m.f of each $x \in S$.

$f(x) = \mathbb{P}(X = x)$

$f(1) = \mathbb{P}(X = 1) = \frac{1}{36}$

$f(2) = \mathbb{P}(X = 2) = \frac{3}{36}$

$f(3) = \mathbb{P}(X = 3) = \frac{5}{36}$

$f(4) = \mathbb{P}(X = 4) = \frac{7}{36}$

$f(5) = \mathbb{P}(X = 5) = \frac{9}{36}$

$f(6) = \mathbb{P}(X = 6) = \frac{11}{36}$

The distribution of $X$:

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $f(x)$ | $\frac{1}{36}$ | $\frac{3}{36}$ | $\frac{5}{36}$ | $\frac{7}{36}$ | $\frac{9}{36}$ | $\frac{11}{36}$ |

or the p.m.f. of the distribution of $X$ is:

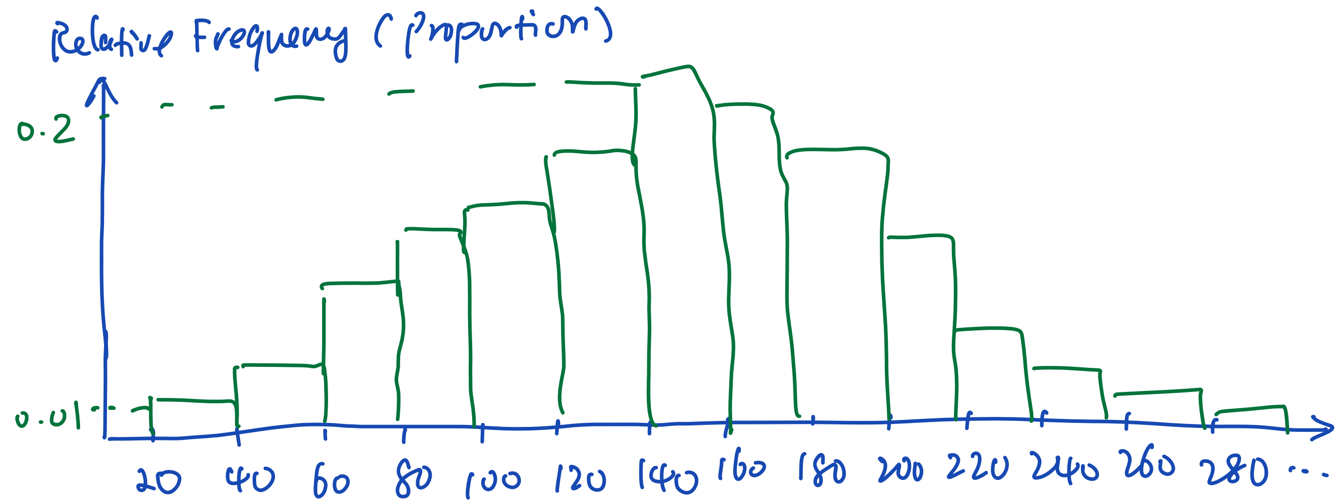
$$f(x) = \frac{2x-1}{36}, \quad x = 1, 2, 3, 4, 5, 6$$

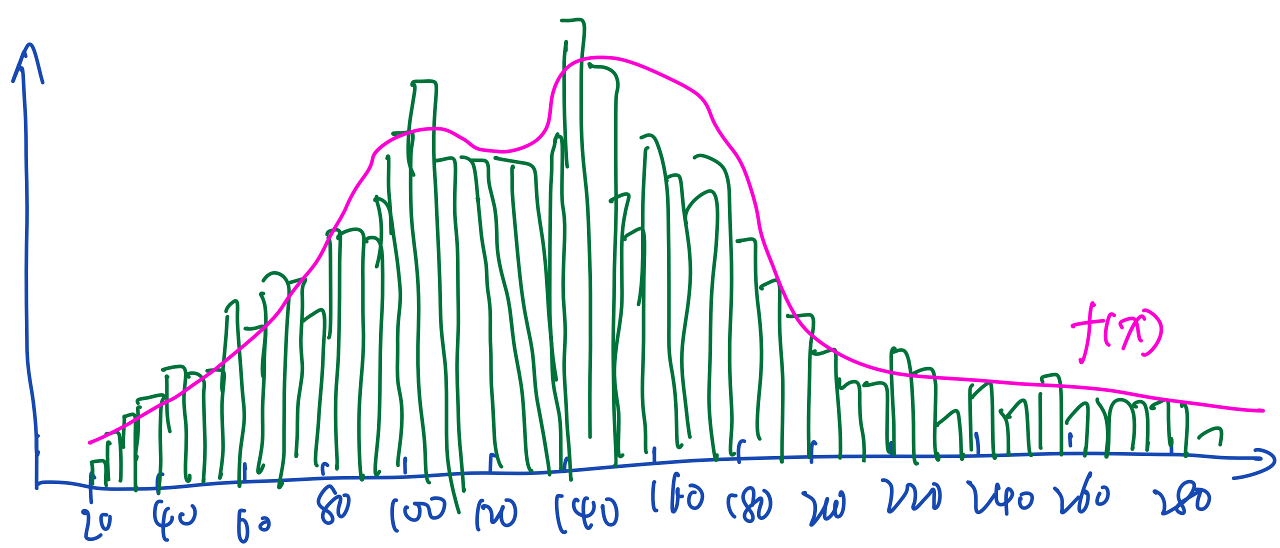
4. Continous Distribution

$\boxed{\text{Ex 2}}$ Consider the distribution of human weights

- we know the values are more or less continuous in a range

- Can we give $\mathbb{P}(X = a)$ for some $a$?



$$\mathbb{P}(a \le X \le b) = \int_a^b f(x)\, dx$$

For a r.v. $X$ that has continuous distribution,

the $\boxed{\text{probability density function}}$ (p.d.f.) is

an integrable function $f(x)$ satisfying:

(a) $f(x) > 0$

(b) $\int_S f(x) = 1$

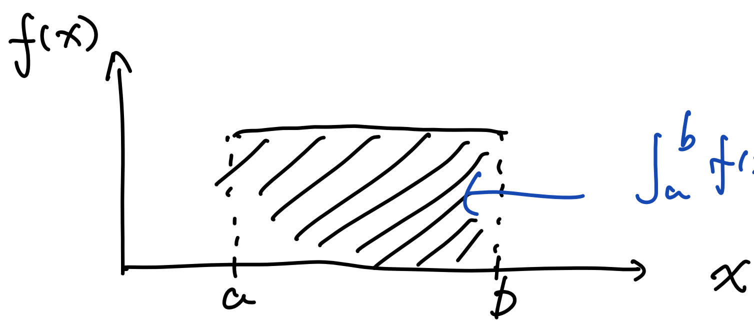
(c) $\forall\ (a,b) \subseteq S$, then

$$\mathbb{P}(a < X < b) = \int_a^b f(x)\, dx$$

In rare cases, you can figure out the p.d.f. directly:

$\boxed{\text{Ex 3}}$ Uniform distribution

Let r.v. X denote the outcome when a point is selected at random from an interval [a,b] and the prob. that each point get picked is the same:




$$\int_a^b f(x)\,dx = 1 = f(x)\cdot(b-a)$$

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

But most cases, the commonly used continuo distribution are pre-defined to try to match the shape of different data. (Instead of giving a experiment. figure out a new distribution). Those pre-definations rely on some characteristics that would help decide the shape / behavior of the data.

We call them $\boxed{\text{parameters}}$ of a distribution.

5. Expected Value

Informally, $\boxed{\text{expected value}}$ is the mean of the outcomes when repeating the experiment infinitely many times ( sample mean of a sample size $n \to \infty$)

Mathematically, this can be found using

- all the possible outcomes
- prob. of each outcome

$\boxed{\text{Def.}}$ The expected value of a r.v. $X$ is $E(X)$ or $\mu$

— when $X$ is discrete, and with p.m.f. $f(x)$

$$E(X) = \sum_{x \in S} x \cdot f(x)$$

— when $X$ is continuous and with p.d.f. $f(x)$

$$E(X) = \int_S x \cdot f(x)\, dx$$

$\boxed{\text{Ex 4}}$ Bernoulli Distribution :

$X$ is the outcome of a random experiment

that has binary outcomes. :

| $x$ | 0 | 1 |
|---|---|---|
| $f(x)$ | $1-p$ | $p$ |

parameter that decides the

bernoulli distributi

Expected value $E(X) = \Sigma x \cdot f(x)$

$= 0 \cdot (1-p) + 1 \cdot p$

$= p$

$\boxed{\text{Revisite Ex 3}}$ Uniform distribution

$E(X) = \int_S x \cdot f(x) \, dx$

$= \int_a^b x \cdot \frac{1}{b-a} \, dx$

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

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

$$= \frac{a+b}{2}$$

Properties of expected value (IMPORTANT!)

(i) For a function of the r.v. $X$,

i.e. $u(X)$,

$$E(u(X)) = \sum u(x) \cdot f(x)$$

$$\text{or} = \int_S u(x) f(x)\, dx$$

Ex: Bernoulli(p) or b(1, p)

$$E(X^2) = \sum x^2 \cdot f(x)$$

$$= (0)^2 \cdot (1-p) + (1)^2 \cdot (p)$$

$$= p$$

(2) If $c$ is a constant, $E(c) = c$

proof: $E(c) = \sum_{S} c \cdot f(x)$

$$= c \cdot \sum_{S} f(x)$$

$$= c \cdot 1$$

$$= c$$

(3) $E[c \cdot u(x)] = c \cdot E(u(x))$

proof?

(4) $E[c_1 u_1(x) + c_2 u_2(x)]$

$$= c_1 E(u_1(X)) + c_2 E(u_2(X))$$

6. c.d.f.

The cumulative distribution function (c.d.f.) is another way to present the information on probability in a distribution:

$$F(x) = \mathbb{P}(X \leq x)$$

discrete

$$= \begin{cases} \sum_{t \leq x} \mathbb{P}(X = t) = \sum_{t \leq x} f(t) \\ \int_{-\infty}^{x} f(t)\,dt \end{cases}$$

continuous

- For $x$ values for which $F'(x)$ exists, $F'(x) = f(x)$

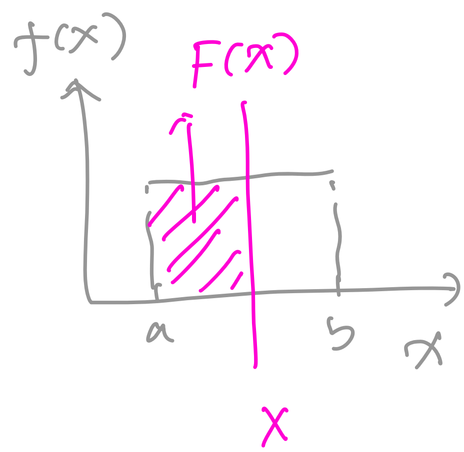
Revisit Ex. 3 Uniform distribution

$X \sim \text{Uniform}[a, b]$

so far we have:

p.d.f. $f(x) = \dfrac{1}{b-a}$ , $x \in [a, b]$

expected value $\mu = \dfrac{a+b}{2}$




$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$= \int_a^x \frac{1}{b-a}\,dt$$

$$= \frac{t}{b-a}\Big|_a^x$$

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

c.d.f. $F(x) = \dfrac{x-a}{b-a}$, $x \in [a, b]$

$\boxed{\text{Properties of c.d.f.}}$

(1) $F$ is a non-decreasing function, i.e.

if $s \le t$, $F(s) \le F(t)$

(2) $\lim_{t \to -\infty} F(t) = 0$

(3) $\lim_{t \to \infty} F(t) = 1$

$\boxed{\text{Application}}$ Simulate data from given distribution

- Idea: If $F(x) = \mathbb{P}(X \le x) = p$

then $x = F^{-1}(p)$

known c.d.f and its inverse

any given prob. $\in [0,1]$

To simulate data of size $n$ from given distribution (a.k.a. given c.d.f.)

(1) given $n$ random numbers from $[0,1]$:

$p_1, p_2, \ldots, p_n$ (with replacement)

(2) for each $p_i$, $x_i = F^{-1}(p_i)$

$\boxed{\text{Inverse Transformation Sampling}}$

For a r.v. $X$ with c.d.f. $F_X(\cdot)$, and

r.v. $U \sim$ Uniform $[0,1]$

$F_X^{-1}(U) \sim$ the distribution of $X$

$\boxed{\text{Ex}}$ Discrete Case — $F^{-1}$ is not well-defined

we want to generate (simulate) a sample of $(n)$ from a random variable that follows the distribution:

$$f(x) = \frac{4-x}{6}, \quad x = 1, 2, 3$$

From previous class, we know

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{5}{6} & 2 \le x < 3 \\ 1 & x \geq 3 \end{cases}$$

- generate $p \sim$ Uniform $[0, 1]$

- if $\begin{cases} 0 < p < \frac{1}{2}, & \text{fail} \\ \frac{1}{2} \le p < \frac{5}{6}, & \text{generate sample } x = 1 \\ \frac{5}{6} \le p < 1, & \text{generate } x = 2 \\ p = 1, & \text{generat } x = 3 \end{cases}$

$\boxed{\text{EX}}$ Continuous Case

$X \sim \exp(\lambda = 2)$, $f(x) = 2e^{-2x}$, $x \geq 0$

$F_X(x) = 1 - e^{-2x}$

- generate $p \sim \text{Uniform}(0, 1)$

- $X = F_X^{-1}(p) = -\dfrac{\ln(1-p)}{2}$