Recall: prob. space (LNp.3-3) $(\Omega, \mathcal{F}, P)$

Random Variables — 隨機變數

A Motivating Example

a special case of survey sampling

➢ Experiment: Sample k ($<n$) students without replacement from the population of all n students (labeled as 1, 2, …, n, respectively) in our class.

Ω: student 1 … student n $\to$ $\{\square, \square, \cdots, \square\}$ 1st 2nd kth

➢ Ω = {all combinations} = $\{\underbrace{\{i_1, \ldots, i_k\}}_{\omega}: 1 \le i_1 < \cdots < i_k \le n\}$

$\mathcal{F} = 2^{\Omega}$

➢ A probability measure P can be defined on Ω, e.g., when there is an equally likely chance of being chosen for each students,

$(\Omega, \mathcal{F}, P)$ is enough for discussing the prob. of ω's

small $p(\omega)$ — define (LNp.3-7)

$$P(\{i_1, \ldots, i_k\}) = 1/\binom{n}{k}.$$

$\#\Omega$ ($\because$ symmetric outcomes)

➢ For an outcome $\omega \in \Omega$, the experimenter may be more interested in some *quantitative attributes* of ω, rather than the ω itself, e.g.,

- The average weight of the k sampled students
- The maximum of their midterm scores
- The number of male students in the sample

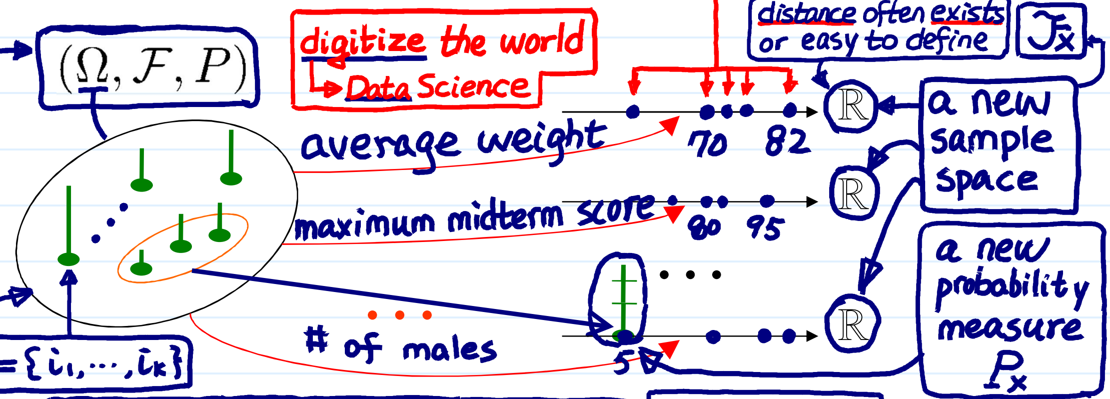
c.f. urn problem (LNp.4-5)

➢ **Q**: What mathematical structure would be useful to characterize the random quantitative attributes of ω's?

$\because$ the outcome ω is random

In some cases, this might be difficult to observe or characterize

may not have a (norm) distance definition btw 2 outcomes ω_1 & ω_2



can do numerical calculations on its outcomes

$(\mathbb{R}, \mathcal{F}_X, P_X)$

- Definition: A *random variable* X is a (measurable) function which maps the sample space Ω to the real numbers $\mathbb{R}$, i.e.,

ω_0 after the random outcome ω appears

$\mathbb{R} \ni X(\omega_0)$

$$X(\omega) = X : \Omega \to \mathbb{R}$$ — new sample space

➢ The P defined on Ω would be transformed into a *new* probability measure defined on $\mathbb{R}$ through the mapping X — P_X

- the outcome of X is random,
- but the map X is deterministic

$X(\omega): \Omega \to \mathbb{R}$

➢ Example (Coin Tossing): Toss a fair coin 3 times, and let

2 main aspects of X:
• the (random) value of X
• distribution of X

value: random
distribution: fixed

- X_1 = the total number of heads

  X_2 = the number of heads on the first toss

  X_3 = the number of heads minus the number of tails

operations on the outcomes of Ω

cf.

e.g., final outcome

- $\Omega = \{hhh, hht, hth, \boxed{thh}, htt, tht, tth, ttt\}$

$1/8 \quad 1/8 \quad 1/8 \quad 1/8 \quad 1/8 \quad 1/8 \quad 1/8 \quad 1/8 \Leftarrow P(\{\omega\})$

can do numerical calculations, e.g. $X_1 - X_3 = X_1(\omega) - X_3(\omega)$
"+", "−", "×", "÷", "exp", "log", … "<", ">", "=", …

X_1	:	3,	2,	2,	2,	1,	1,	1,	0.
X_2	:	1,	1,	1,	0,	1,	0,	0,	0.
X_3	:	3,	1,	1,	1,	−1,	−1,	−1,	−3.

$P_{X_1}(\{0\}) = 1/8$
$P_{X_1}(\{1\}) = 3/8$
$P_{X_1}(\{2\}) = 3/8$
$P_{X_1}(\{3\}) = 1/8$

➢ **Q**: Why particularly interested in functions that map to "$\mathbb{R}$"?

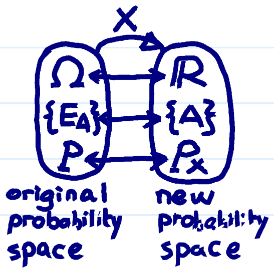
➢ **Q**: How to define the probability measure of X (i.e., P_X) from P?

P_X defined in the way is a probability measure (exercise)

Ans: For a (measurable) set (i.e., an event) $A \subset \mathbb{R}$,

$$P_X(X \in A) \equiv P(\{\omega : X(\omega) \in A\}).$$

A occurs $\qquad E_A \subset \Omega$



The P_X is often called the *distribution* of X.

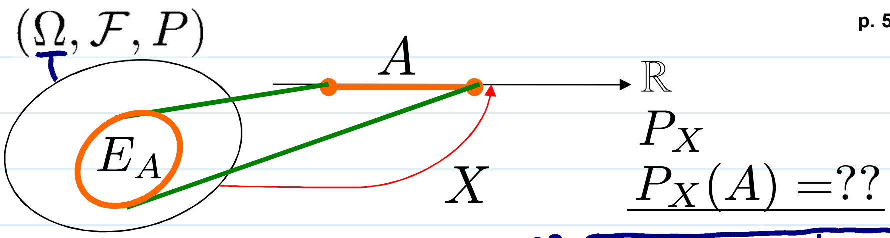
分配,分布

A is measurable

$E_A \in \mathcal{F}$

A occurs $\Leftrightarrow$ E_A occurs

$\boxed{P_X(A) = P(E_A)}$



continuous r.v. cf. 離散

Discrete Random Variables

cf. discrete sample space (LNp.3-6)

cf. continuous

- Definition: For a random variable (r.v.) X, let

new sample space

$$\mathbb{R} \supset \mathcal{X} = \{X(\omega) : \omega \in \Omega\},$$

sample space Ω: discrete or continuous
sample space $\mathcal{X}$: discrete.

be the range of X. Then, X is called *discrete* if $\mathcal{X}$ is a finite or countably infinite set, i.e.,

All possible values of X

$$\mathcal{X} = \{x_1, \ldots, x_n\} \text{ or } \mathcal{X} = \{x_1, x_2, \ldots\}.$$

X_1: $\mathcal{X} = \{0, 1, 2, 3\} \subset \mathbb{R}$
X_2: $\mathcal{X} = \{0, 1\} \subset \mathbb{R}$
X_3: $\mathcal{X} = \{-3, -1, 1, 3\} \subset \mathbb{R}$

➢ Example. The X_1, X_2, X_3 in the Coin Tossing example.

➢ Example. The number of coin tosses (X) until 1[st] head appears.

$\mathcal{X} = \{1, 2, 3, \ldots\} = \mathbb{Z}_+ \subset \mathbb{R}$

- The sample space of a r.v. is the **real line** $\mathbb{R}$. **Q**: For $\mathbb{R}$, are there some particular ways to define a probability measure (p.m.) on it? [cf., for general sample space Ω, a p.m. is defined on all (or any measurable) subsets of Ω]

P_X

prob. space: $(\Omega, \mathcal{F}, P)$ $\qquad 2^{\Omega}$