Lecture 10 Random Variables, PMF, Expectation & Variance

BIO210 Biostatistics

Xi Chen

Fall, 2024

School of Life Sciences
Southern University of Science and Technology

南方科技大学生命科学学院
SUSTech · SCHOOL OF
LIFE SCIENCES

What is a random variable (r.v.) ?

- An assignment of a value (a real number) to every possible outcome in the sample space.
- **Mathematically**: A real-valued **function** defined on a sample space Ω. In a particular experiment, a random variable (**r.v.**) would be some function that assigns a real number to each possible outcome.

Random Variable

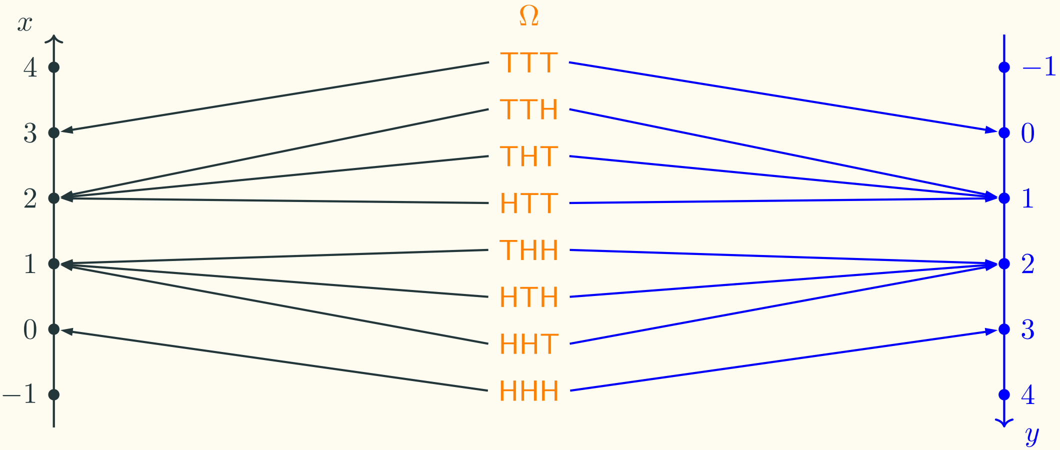
More about random variables

- Discrete or continuous.
- Can have several random variables defined on the same sample space.
- Notation
  - random variable $\boldsymbol{X}$: function $\Omega \mapsto \mathbb{R}$
  - numerical value: x : value $\in \mathbb{R}$

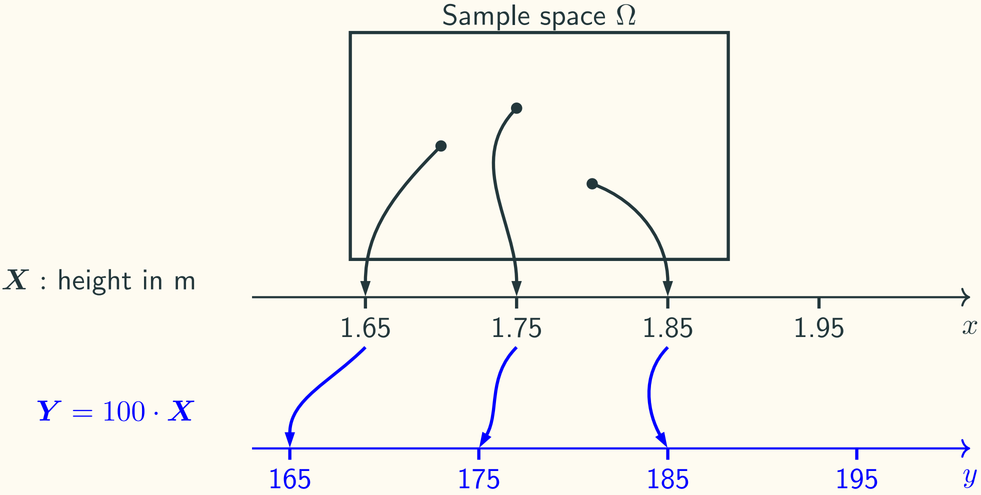
Different random variables on the same sample space

$\boldsymbol{X}$: number of tails

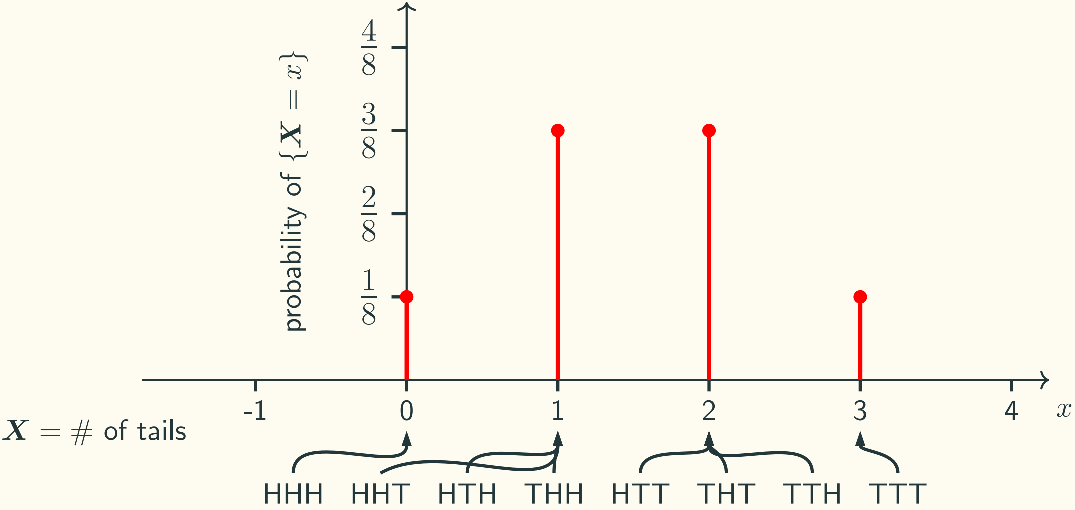
$\boldsymbol{Y}$: number of heads



Function of a random variable is an r.v.



Probability Mass Function (PMF)



Probability Mass Function (PMF)

The PMF of $\boldsymbol{X} =$ number of tails after three flips

x	$\mathbb{P}(\{\boldsymbol{X} = x\})$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$
otherwise	0

$$\mathbb{P}(\{\boldsymbol{X} = x\}) = \begin{cases} \dfrac{1}{8}, & x = 0, 3 \\ \dfrac{3}{8}, & x = 1, 2 \\ 0\,, & \text{otherwise} \end{cases}$$

PMF Notation

Probability Mass Function

- Notation

$$\mathbb{P}_{\boldsymbol{X}}(x) = \mathbb{P}(\{\boldsymbol{X} = x\}) = \mathbb{P}(\{\omega \in \Omega \mid \boldsymbol{X}(\omega) = x\})$$

- Properties

$$\mathbb{P}_{\boldsymbol{X}}(x) \geqslant 0$$
$$\sum_x \mathbb{P}_{\boldsymbol{X}}(x) = 1$$

ω	$\boldsymbol{X}(\omega) = x$	$\mathbb{P}_{\boldsymbol{X}}(x) = \mathbb{P}(\{\boldsymbol{X} = x\})$
HHH	0	$\frac{1}{8}$
THH, HTH, HHT	1	$\frac{3}{8}$
TTH, THT, THH	2	$\frac{3}{8}$
TTT	3	$\frac{1}{8}$

Geometric PMF

Experiment: keep flipping a coin ($\mathbb{P}(H) = p$) until a head comes up for the first time. Let the random variable $\boldsymbol{X}$ be the number of flips.

ω	$\boldsymbol{X}(\omega)$	$\mathbb{P}_{\boldsymbol{X}}(x)$
H	1	p
TH	2	$(1-p)p$
TTH	3	$(1-p)^2 p$
$\vdots$	$\vdots$	$\vdots$
$\underbrace{\text{TTT...TTT}}_{n-1}\text{H}$	n	$(1-p)^{n-1}p$

Geometric PMF. $\boldsymbol{X}$: geometric random variable.

How to compute a PMF $\mathbb{p}_{\boldsymbol{X}}(x)$

To compute a PMF $\mathbb{p}_{\boldsymbol{X}}(x)$:

1. Collect all possible outcomes for which $\boldsymbol{X} = x$;
2. add their probabilities;
3. repeat for all x.

Compute PMF

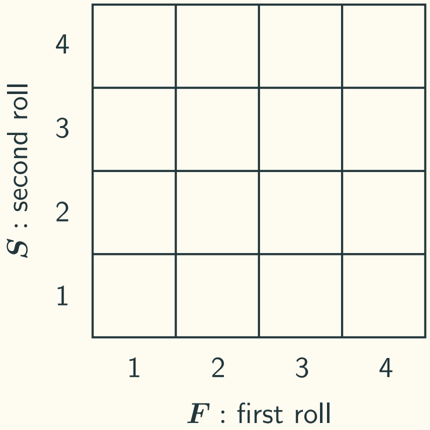
Experiment: two independent rolls of a fair tetrahedral die.

$\boldsymbol{F}$: outcome of the first roll

$\boldsymbol{S}$: outcome of the second roll

$\boldsymbol{X} = min(\boldsymbol{F}, \boldsymbol{S})$

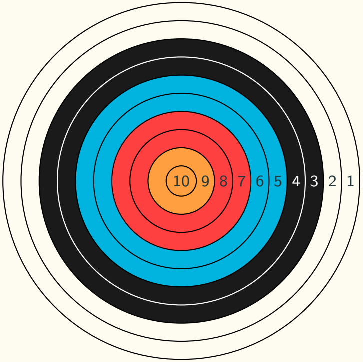
$\mathbb{P}_{\boldsymbol{X}}(x) = ?$



Expected value of a random variable (Expectation)

Experiment: archery

Let $\boldsymbol{X}$ be the score you get for each shot. What is the expected value of $\boldsymbol{X}$?



x	$\mathbb{P}_{\boldsymbol{X}}(x)$
1	0.19
2	0.17
3	0.15
4	0.13
5	0.11
6	0.09
7	0.07
8	0.05
9	0.03
10	0.01

Think: What is the average score you will get after a large number of trials?

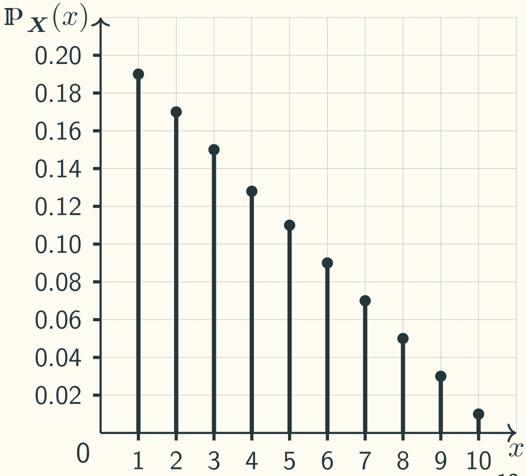
Expected value (Expectation)

Definition

$$\mathbb{E}\left[\boldsymbol{X}\right] = \sum_{x} x \mathbb{P}_{\boldsymbol{X}}(x)$$

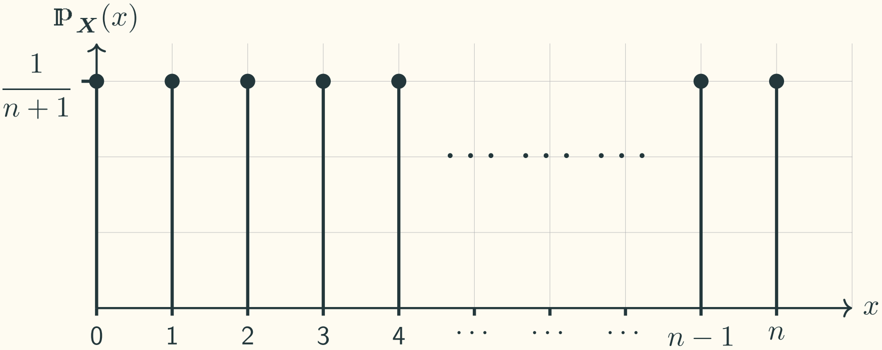
- Interpretation
  1. Centre of gravity of the PMF
  2. Average in large number of repetitions of the experiment

PMF of $\boldsymbol{X}$ from the archery experiment



Expectation of a Uniform Distribution

Example: a uniform discrete random variable $\boldsymbol{X}$ on $0, 1, 2, 3, ..., n$



What is $\mathbb{E}[\boldsymbol{X}]$?

Properties of expectations

Let X be a random variable, and let $\boldsymbol{Y} = g(\boldsymbol{X})$, what is $\mathbb{E}[\boldsymbol{Y}]$?

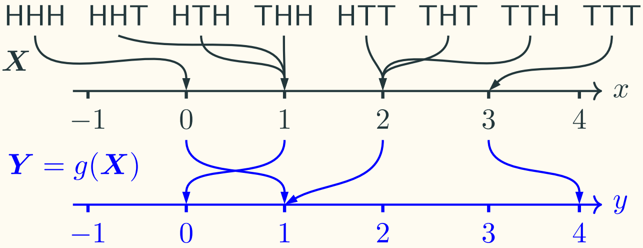
- The hard way:

$$\mathbb{E}[\boldsymbol{Y}] = \sum_y y \mathbb{P}_{\boldsymbol{Y}}(y)$$

- The easy way:

$$\mathbb{E}[\boldsymbol{Y}] = \sum_x g(x) \mathbb{P}_{\boldsymbol{X}}(x)$$

y	$\mathbb{P}_{\boldsymbol{Y}}(y)$
0	3/8
1	4/8
4	1/8



x	$g(x)$	$\mathbb{P}_{\boldsymbol{X}}(x)$
0	1	1/8
1	0	3/8
2	1	3/8
3	4	1/8

Expectation of a linear function of r.v.

- Caution: in general $\mathbb{E}\left[g(\boldsymbol{X})\right] \neq g(\mathbb{E}\left[\boldsymbol{X}\right])$
- Exception: if α, β are constants, then we have:

    - $\mathbb{E}\left[\alpha\right] = \ \alpha$
    - $\mathbb{E}\left[\alpha \boldsymbol{X}\right] = \ \alpha \mathbb{E}\left[\boldsymbol{X}\right]$
    - $\mathbb{E}\left[\alpha \boldsymbol{X} + \beta\right] = \ \alpha \mathbb{E}\left[\boldsymbol{X}\right] + \beta$

Variance and standard deviation of a random variable

Definition of Variance

$$\mathbb{V}\mathrm{ar}\left(\boldsymbol{X}\right) = \mathbb{E}\left[\left(\boldsymbol{X} - \mathbb{E}\left[\boldsymbol{X}\right]\right)^2\right]$$

Properties of Variance

- $\mathbb{V}\mathrm{ar}\left(\boldsymbol{X}\right) = \mathbb{E}\left[\boldsymbol{X}^2\right] - \left(\mathbb{E}\left[\boldsymbol{X}\right]\right)^2$
- If α, β are constants, then $\mathbb{V}\mathrm{ar}\left(\alpha\boldsymbol{X} + \beta\right) = \alpha^2\mathbb{V}\mathrm{ar}\left(\boldsymbol{X}\right)$

Definition of Standard Deviation

$$\sigma_{\boldsymbol{X}} = \sqrt{\mathbb{V}\mathrm{ar}\left(\boldsymbol{X}\right)}$$

Discrete Random Variables (Summary slide)