# Random Variables and Expectation

Josep Díaz Maria J. Serna Conrado Martínez
U. Politècnica de Catalunya

RA-MIRI 2022–2023

# Random variables

Flip 100 times a fair coin, each time if the outcome is $\mathrm{H}$ we give 1€, if it is $\mathrm{T}$ we get 1€. At the end, how much did we win or loose?. Notice $\Omega = \{\mathrm{T}, \mathrm{H}\}^{100}$

Given $\Omega$, a random variable is a function $X : \Omega \to \mathbb{R}$.
$X$ can be interpreted as a quantity, whose value depends on the outcome of the experiment.

**Example**

In the previous example, our total gain (or loss) is a random variable $X$,

$$X = \text{number of H's minus the number of T's.}$$

The number of heads $W$ and the number of tails $L$ are also random variables (and $X = W - L$).

# Events $\leftrightarrow$ random variables

Given a random variable $X$ on $\Omega$ and $a \in \mathbb{R}$ the event $X \geqslant a$ represents the set $\{\omega \in \Omega | X(\omega) \geqslant a\}$.

$$\mathbb{P}[X \geqslant a] = \sum_{\omega \in \Omega : X(\omega) \geqslant a} \mathbb{P}[\omega]$$

**Example**

In the previous example of 100 coin flips, for the event $W = 50$ we have $\mathbb{P}[W = 50] = \frac{\binom{100}{50}}{2^{100}}$ (*)

Given an event $A$ define the indicator r.v. $\mathbb{I}_A$:

$$\mathbb{I}_A = \begin{cases} 1 & \text{if } A \text{ true} \\ 0 & \text{otherwise} \end{cases}$$

**Example**

If $A$ = exactly 50 wins, $\mathbb{P}[A] = \mathbb{P}[\mathbb{I}_A = 1] = \mathbb{P}[W = 50]$, which is exactly (*)

# Expectation

The expectation $\mathbb{E}[X]$ of a r.v. $X : \Omega \to \mathbb{R}$ is defined as

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}[X = x] \,.$$

Expectation (mean, average) is just the weighted sum over all values of the r.v.

Notice: If $X$ is a r.v. then $\mathbb{E}[X] \in \mathbb{R}$.

**Example**

Let $X$ be an integer generated u.a.r. between 1 and 6. Then $\mathbb{E}[X] = \sum_{x=1}^{6} x \cdot \mathbb{P}[X = x] = \sum_{x=1}^{6} \frac{x}{6} = 3.5$, which is not a possible value for $X$.

# Third trick: Linearity of expectation

*Theorem*

1. *Given r.v.* $X, Y$, $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
2. *Given any constant* $c$, *and a rv* $X$, *then* $\mathbb{E}[cX] = c\,\mathbb{E}[X]$.
3. *More generally, given r.v.* $\{X_i\}_{i=1}^n$ *and* $n$ *real numbers* $\{a_i\}_{i=1}^n$, $\mathbb{E}[\sum_{i=1}^n a_i X_i] = \sum_{i=1}^n a_i\,\mathbb{E}[X_i]$.

*The proof is standard and relies on the fact that the sum of r.v. is a r.v.*

# Independent r.v.

Two random variables $X$ and $Y$ are said to be independent if

$$\forall x, y \in \mathbb{R}, \mathbb{P}[(X = x) \cap (Y = y)] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y] \,.$$

Two r.v. which are not independent are said to be dependent or correlated.

**Example**

Rolling two dice, let $X_1$ be a r.v. counting the pips in die 1, and let $X_2$ be a r.v. counting the pips in die 2. Then $X_1$ and $X_2$ are independent rv.

**Example**

Rolling two dice, let $X_1$ be a r.v. counting the pips in die 1, and let $X_3$ count the sum of pips in the two rollings, then $X_1$ and $X_3$ are correlated.

# Interesting Example

Given and array $A[1,\ldots,n]$ containing $n$ different keys, chosen u.a.r. from one permutation of the set of $n$ keys, let $a_i$, $1 \leqslant i \leqslant n$, be the key contained in $A[i]$. We say $a_i$ and $a_j$ are inverted if $i < j$ but $a_i > a_j$. Compute the expected number of inversions in $A$.

Let $X$ count the number of inversions in $A$.

For every pair $1 \leqslant i < j \leqslant n$ of positions in $A$ define an indicator r.v.:

$$X_{i,j} = \begin{cases} 1 & \text{if } a_i > a_j \\ 0 & \text{otherwise} \end{cases}$$

$X = \sum_{i<j} X_{i,j} \Rightarrow \mathbb{E}[X] = \sum_{i<j} \mathbb{E}\left[X_{i,j}\right] = \sum_{i<j} 1 \cdot \underbrace{\mathbb{P}\left[a_i > a_j\right]}_{=1/2}$

Notice $|\{(i,j)|1 \leqslant i < j \leqslant n\}| = (n-1) + (n-2) + \cdots + 2 + 1$

therefore, $\mathbb{E}[X] = \frac{1}{2}\sum_{i=1}^{n}(n-i) = \frac{1}{2}\sum_{i=1}^{n-1} i = \frac{n(n-1)}{4}$

# Deterministic algorithm to hire a student

We have $n$ students $\{1, \ldots, n\}$, we want to hire the best one to help us. For that we have to interview one by one, each time we find one that is more suitable that the previous ones, we preselect him. At the end we hire the last one pre-selected, but we indemnify with $S > 0$ €, each of the pre-selected not hired.
We want to minimize the number of students pre-selected.

**procedure** HIRING($n$)
  $best := 0$
  **for** $i := 1$ **to** $n$ **do**
    interview $i$-th candidate
    **if** $i$ is better than best **then**
      $best := i$ and pre-select $i$
    **end if**
  **end for**
**end procedure**

# Adversarial complexity


The adversary gives you a list of ordered students s.t. you are forced to pre-select each of them.

$$T(n) = \Theta(n).$$

# Average analysis of the hiring algorithm

The number of all possible order of the students is $n!$
We select u.a.r. an order with probability $= \frac{1}{n!}$.

*Lemma*

*The expected number of pre-selected is $\mathcal{O}(\log n)$.*

Proof

Let $X$ be a r.v. counting the number of pre-selected students.
For each $1 \leqslant i \leqslant n$ definean indicator r.v.

$$X_i = \begin{cases} 1 & \text{if } i \text{ is-pre-selected;} \\ 0 & \text{otherwise.} \end{cases}$$

Then, $X = \sum_{i=1}^{n} X_i \Rightarrow \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} 1 \cdot \underbrace{\frac{1}{i}}_{\text{why?}} = \ln n + \mathcal{O}(1)$. □

# Randomized algorithm for the hiring an student problem

To fool the input given by the adversary: Permute the input

```
procedure RAND-HIRE-STUDENT(n)
    Randomly permute the list [1, ..., n]
    best := 0
    for i := 1 to n do
        interview i-th candidate
        if i-th candidate is better than best then
            best := i and pre-select i-th candidate
        end if
    end for
end procedure
```

Let $X(n)$ a r.v. counting the number of pre-selections, on an input of $n$ students. Then $\mathbb{E}[X(n)] = \ln n + \mathcal{O}(1)$