## Introduction to Probability

Lecture 2: Random variables, probability mass function, expectation
Mateja Jamnik, Thomas Sauerwald
University of Cambridge, Department of Computer Science and Technology email: \{mateja.jamnik,thomas.sauerwald\}@cl.cam.ac.uk

## Outline

Random variable

## Probability mass function

## Cumulative distribution function

## Expectation

## What is a random variable?

## Random variable

A random variable $X$ is a function from the sample space to the real numbers.

- We can interpret $X$ as a quantity whose value depends on the outcome of an experiment (some probabilistic process).
- Roll two dice, $X$ : sum of dice
- Toss 3 coins, $X$ : number of heads
- Give a student a test, $X$ : score
- Stock market index
- Or can think of $X$ as a variable in a programming language that takes on values, has a type, and has a domain over which it is applicable.
- Many different types of RV: indicator, binary, choice, Bernoulli, etc.
- Random variable can be discrete or continuous:
- $X$ has finitely many possible values: discrete.
- $X$ has every integer as a possible value: discrete.
- $X$ amount of time it takes to finish a race: continuous (possible value: $\{t: 0 \leq t<\infty\}=[0, \infty)$ ).


## Examples of random variables

## Example

We toss 3 fair coins. Let a random variable $X$ be the total number of heads on the 3 coins. What are the probabilities of $X$ taking on the following values: $X=0, X=1, X=2, X=3, X \geq 4$ ?

1. $\mathbf{P}[X=0]=\frac{1}{8}$ where set of outcomes is $\{(T, T, T)\}$
2. $\mathbf{P}[X=1]=\frac{3}{8}$ where set of outcomes is $\{(H, T, T),(T, H, T),(T, T, H)\}$
3. $\mathbf{P}[X=2]=\frac{3}{8}$ where set of outcomes is $\{(H, H, T),(T, H, H),(H, T, H)\}$
4. $\mathbf{P}[X=3]=\frac{1}{8}$ where set of outcomes is $\{(H, H, H)\}$
5. $\mathbf{P}[X \geq 4]=0$ where set of outcomes is $\}$

## Random variables are NOT events

## random variables $\neq$ events

| Tossing 3 fair coins example |  |  |  |
| :--- | :---: | :--- | :--- |
| $X=x$ | $\mathbf{P}[X=x]$ | Set of outcomes | Possible event $E$ |
| $X=0$ | $\frac{1}{8}$ | $\{(T, T, T)\}$ | Toss 0 heads |
| $X=1$ | $\frac{3}{8}$ | $\{(H, T, T),(T, H, T),(T, T, H)\}$ | Toss exactly 1 head |
| $X=2$ | $\frac{3}{8}$ | $\{(H, H, T),(T, H, H),(H, T, H)\}$ | Event where $X=2$ |
|  |  |  | Toss exactly 2 heads |
| $X=3$ | $\frac{1}{8}$ | $\{(H, H, H)\}$ | Toss 0 tails |
| $X \geq 4$ | 0 | $\}$ | Toss 4 or more heads |

We can define events by condition of the value of a random variable (RV takes on values that satisfy a numerical test).

## Another example

## Example

Tossing a coin has the probability $p$ that it comes up heads. Toss a coin 5 times. Let $X$ : the number of heads in 5 tosses. What is the range of $X$ (i.e., what are the values that $X$ can take on with non-zero probability)? What is $\mathbf{P}[X=k]$ where $k$ is in the range of $X$ ?

- Notice that each coin toss is an independent trail.
- Recall $\mathbf{P}[2$ heads $]=\binom{5}{2} p^{2}(1-p)^{3}, \mathbf{P}[3$ heads $]=\binom{5}{3} p^{3}(1-p)^{2}$.
- Range of $X:\{0,1,2,3,4,5\}$
- $\mathbf{P}[X=k]=\binom{5}{k} p^{k}(1-p)^{5-k}$


## Outline

## Random variable

Probability mass function

## Cumulative distribution function

## Expectation

## Probability mass function definition (PMF)

## Discrete random variable

A random variable $X$ is discrete if its range has countably many values

$$
X=x \text { where } x \in\left\{x_{1}, x_{2}, x_{3}, \ldots\right\}
$$

## Probability mass function

The probability mass function (PMF) of a discrete random variable $X$ is a function $p(a)$ of $X$ that maps possible outcomes of a random variable to the corresponding probabilities:

$$
p(a)=\mathbf{P}[X=a]=p_{X}(a)
$$

Recall that probabilities must sum to $1: \sum_{i=1}^{\infty} p\left(a_{i}\right)=1$.

## Example for a single die

- Let $X$ be a RV representing a single die roll.
- Range of $X:\{1,2,3,4,5,6\}$, thus $X$ is a discrete RV.
- PMF of $X$ :

$$
p(x)=\mathbf{P}[X=x]= \begin{cases}\frac{1}{6} & x \in\{1,2,3,4,5,6\} \\ 0 & \text { otherwise }\end{cases}
$$



## Example for two dice

- Let $Y$ be a RV representing the sum of two independent dice rolls.
- Range of $Y:\{2,3, \ldots, 11,12\}$.
- PMF of $Y$ :

$$
p(y)=\mathbb{P}[Y=y]= \begin{cases}\frac{y-1}{36} & y \in \mathbb{Z}, 2 \leq y \leq 6 \\ \frac{13-y}{36} & y \in \mathbb{Z}, 7 \leq y \leq 12 \\ 0 & \text { otherwise }\end{cases}
$$

- Check $\sum_{y=2}^{12} p(y)=1$.



## Properties of PMF

Let possible values of $X=\left\{a_{1}, a_{2}, a_{3}, \ldots\right\}$.

1. By Axiom 1: $0 \leq p\left(a_{i}\right) \leq 1$.
2. $p(a)=0$ if $a$ is not a possible value.
3. By Axiom 3: $\sum_{i=1}^{\infty} p\left(a_{i}\right)=1$.

$$
\sum_{i=1}^{\infty} p\left(a_{i}\right)=\sum_{i=1}^{\infty} \mathbf{P}\left[X=a_{i}\right]=\mathbf{P}\left[\bigcup_{i=1}^{\infty}\left\{X=a_{i}\right\}\right]=\mathbf{P}[S]=1
$$

4. Notice that everything to do with discrete RVs is expressed in terms of (finite or infinite) sum.
5. For continuous RVs, these sums are replaced by integrals.

## Outline

## Random variable

## Probability mass function

Cumulative distribution function

## Expectation

## Cumulative distribution function definition (CDF)

Another useful way to analyse probabilities.

Cumulative distribution function
The cumulative distribution function (CDF) of a random variable $X$ is defined as

$$
F(a)=F_{X}(a)=\mathbf{P}[X \leq a] \text { where }-\infty<a<\infty
$$

For a discrete random variable $X$, the CDF is

$$
F(a)=\mathbf{P}[X \leq a]=\sum_{\text {all } x \leq a} p(x)
$$

Note that for a discrete RV the CDF is a step function, i.e., the value of $F$ is constant in the intervals ( $x_{i-1}, x_{i}$ ) and then takes a step of size $p\left(x_{i}\right)$ at $x_{i}$.

## Example

- Let the PMF for $X$ be given by $p(1)=\frac{1}{4}, p(2)=\frac{1}{2}, p(3)=\frac{1}{8}, p(4)=\frac{1}{8}$.
- Then CDF is:

$$
F(a)= \begin{cases}0 & a<1 \\ \frac{1}{4} & 1 \leq a<2 \\ \frac{3}{4} & 2 \leq a<3 \\ \frac{7}{8} & 3 \leq a<4 \\ 1 & 4 \leq a\end{cases}
$$

- Graphical depiction of function:



## Example for a single die



## Properties of CDF

1. $0 \leq F(x) \leq 1$ for all $x$
2. $\lim _{x \rightarrow-\infty} F(x)=0$
3. $\lim _{x \rightarrow \infty} F(x)=1$
4. $F(x)$ is a non-decreasing function of $x$ (if $x_{1}<x_{2}$ then $f\left(x_{1}\right) \leq f\left(x_{2}\right)$ )

## Outline

## Random variable

## Probability mass function

## Cumulative distribution function

Expectation

## Expectation - expected value

## Expectation

The expectation of a discrete random variable $X$ is defined as

$$
\mathbf{E}[X]=\sum_{x: P[x]>0} x \mathbf{P}[x]
$$

- Expectation is the average value of the random variable over many repetitions of the experiment it represents.
- It is the sum over all values of $X=x$ that have non-zero probability.
- AKA: mean, expected value, weighted average, centre of mass, first moment.


## Example of a die roll

What is the expected value of a 6 -sided die roll (i.e., what is the average value of a die roll)?

1. Define random variables:

$$
\begin{gathered}
X=\mathrm{RV} \text { for value of roll } \\
\mathbf{P}[X=x]= \begin{cases}\frac{1}{6} & x \in\{1, \ldots, 6\} \\
0 & \text { otherwise }\end{cases}
\end{gathered}
$$

2. Solve:

$$
E[X]=1\left(\frac{1}{6}\right)+2\left(\frac{1}{6}\right)+3\left(\frac{1}{6}\right)+4\left(\frac{1}{6}\right)+5\left(\frac{1}{6}\right)+6\left(\frac{1}{6}\right)=\frac{7}{2}
$$

## Example of school classes

## Example

A school has 3 classes with 5,10 and 150 students. What is the average class size?

Interpretation 1: Randomly choose a class with equal probability. Thus, $X=$ size of chosen class

$$
E[X]=5\left(\frac{1}{3}\right)+10\left(\frac{1}{3}\right)+150\left(\frac{1}{3}\right)=\frac{165}{3}=55
$$

Interpretation 2: Randomly choose a student with equal probability.
Thus, $Y=$ size of chosen class

$$
E[Y]=5\left(\frac{5}{165}\right)+10\left(\frac{10}{165}\right)+150\left(\frac{150}{165}\right)=\frac{22635}{165}=137
$$

This is a general phenomenon: it occurs because the more students are in a class, the more likely it is that a randomly chosen student would be in that class. As a result, bigger classes are given more weight than smaller classes.

## Example of Roulette Version 1

## Example

A roulette wheel has 36 places numbered from 1 to 36 . In addition, 18 of them are coloured red and the other 18 are coloured black. A ball is thrown to take one of 36 places. A gambler can bet:

- on the colour of the place that the ball takes. A correct, either red or black, place wins them a 1 to 1 ratio payout;
- on the number of the place that the ball takes. A correct number wins them a 35 to 1 ratio payout.
What is the expected value if a gambler bets on

1. the colour of the place in the roulette;
2. the number of the place in the roulette that the ball will fall.

Are the two different betting games fair?

## Example of Roulette Version 1 Cont.

## Example

What is the expected value if a gambler bets on

1. the colour of the place in the roulette;
2. the number of the place in the roulette that the ball will fall.

Are the two different betting games fair?

1. Let $E_{X}$ : bet on colour.

- If loose, then $X=-1$. Thus $\mathbf{P}\left[\right.$ loose $\left._{X}\right]=\frac{1}{2}$.
- If win, then $X=1$. Thus $\mathbf{P}\left[\operatorname{win}_{X}\right]=\frac{1}{2}$.
- Thus, $\mathbf{E}[X]=(-1)\left(\frac{1}{2}\right)+(1)\left(\frac{1}{2}\right)=0$, This game is "fair".

2. Let $E_{Y}$ : bet on number.

- If loose, then $Y=-1$. Thus $\mathbf{P}\left[\right.$ loose $\left._{Y}\right]=\frac{35}{36}$.
- If win, then $Y=35$. Thus $\mathbf{P}\left[\operatorname{win}_{Y}\right]=\frac{1}{36}$.
- Thus, $\mathbf{E}[Y]=(-1)\left(\frac{35}{36}\right)+(35)\left(\frac{1}{36}\right)=0$, This game is "fair" too.


## Example of Roulette Version 2

## Example

Change the game to add two green places, 0 and 00 . Now there are a total of 38 places. Payouts are the same as before. What are the expected values now?

1. Let $E_{X}$ : bet on red colour.

Thus, $\mathbf{E}[X]=(-1)\left(\frac{20}{38}\right)+(1)\left(\frac{18}{38}\right)=-\frac{1}{19}$.
2. Let $E_{Y}$ : bet on number 10 .

Thus, $\mathbf{E}[Y]=(-1)\left(\frac{37}{38}\right)+(35)\left(\frac{1}{38}\right)=-\frac{1}{19}$.

So, no, these games are not fair, as the gambler would loose $£ \frac{1}{19}=5.3$ pence per game.

