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

UNIVERSITY OF CAMBRIDGE

Outline

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: $X = 0, X = 1, X = 2, X = 3, X \geq 4$?

Answer

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?

Answer

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

Outline

Probability mass function definition (PMF)

Discrete random variable

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

$$X = x \text{ where } x \in \{x_1, x_2, x_3, \ldots\}$$

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:

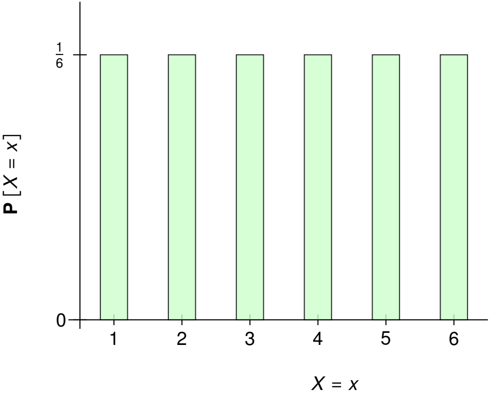
$$p(a) = \mathbf{P}[\, X = a \,] = p_X(a)$$

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

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:

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

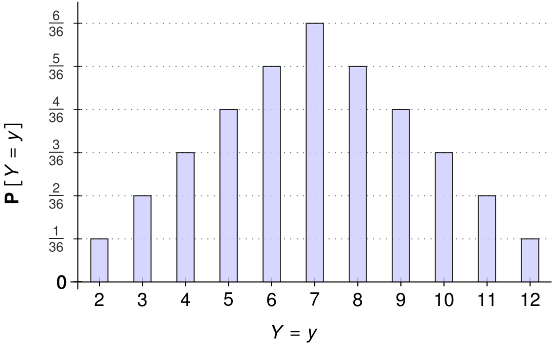


Example for two dice

- 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 $X = \{a_1, a_2, a_3, \ldots\}$.

1. By Axiom 1: $0 \le p(a_i) \le 1$.

2. $p(a) = 0$ if a is not a possible value.

3. By Axiom 3: $\sum_{i=1}^{\infty} p(a_i) = 1$.

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

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

$$F(a) = F_X(a) = \mathbf{P}[\,X \le a\,] \text{ where } -\infty < a < \infty$$

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

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

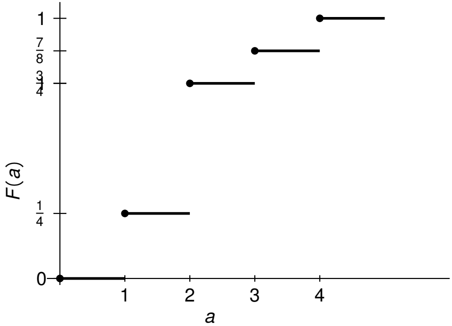
Note that for a discrete RV the CDF is a step function, i.e., the value of F is constant in the intervals (x_{i-1}, x_i) and then takes a step of size $p(x_i)$ at x_i.

Example

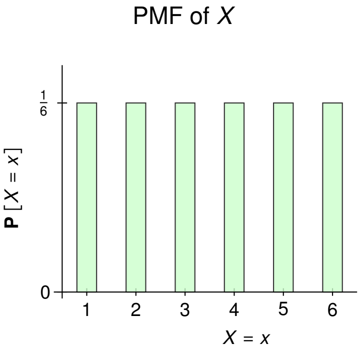
- Let the PMF for X be given by $p(1) = \frac{1}{4}, p(2) = \frac{1}{2}, p(3) = \frac{1}{8}, p(4) = \frac{1}{8}$.
- Then CDF is:

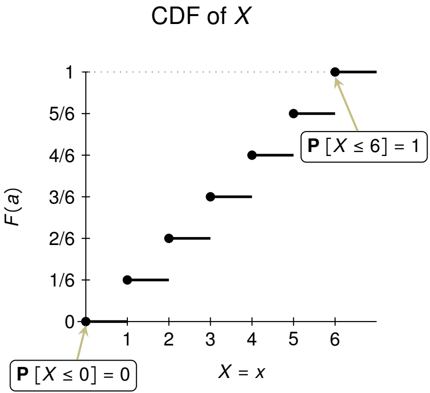
$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \le a < 2 \\ \frac{3}{4} & 2 \le a < 3 \\ \frac{7}{8} & 3 \le a < 4 \\ 1 & 4 \le a \end{cases}$$

- Graphical depiction of function:



Example for a single die





Properties of CDF

1. $0 \leq F(x) \leq 1$ for all x
2. $\lim_{x \to -\infty} F(x) = 0$
3. $\lim_{x \to \infty} F(x) = 1$
4. $F(x)$ is a non-decreasing function of x (if $x_1 < x_2$ then $F(x_1) \leq F(x_2)$)

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

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

- 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:

$$X = \text{RV for value of roll}$$

$$\mathbf{P}[X = x] = \begin{cases} \frac{1}{6} & x \in \{1, \ldots, 6\} \\ 0 & \text{otherwise} \end{cases}$$

2. Solve:

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

Example of school classes

Example

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

Answer

Interpretation 1: Randomly choose a <u>class</u> with equal probability. Thus, X = size of chosen class

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

Interpretation 2: Randomly choose a <u>student</u> with equal probability. Thus, Y = size of chosen class

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

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?

Answer

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?

Answer

1. Let E_X : bet on colour.
   - If win, then $X = 1$. Thus $\mathbf{P}[\, win_X \,] = \frac{1}{2}$.
   - If loose, then $X = -1$. Thus $\mathbf{P}[\, loose_X \,] = \frac{1}{2}$.
   - Thus, $\mathbf{E}[\, X \,] = (-1)(\frac{1}{2}) + (1)(\frac{1}{2}) = 0$, This game is "fair".
2. Let E_Y : bet on number.
   - If win, then $Y = 35$. Thus $\mathbf{P}[\, win_Y \,] = \frac{1}{36}$.
   - If loose, then $Y = -1$. Thus $\mathbf{P}[\, loose_Y \,] = \frac{35}{36}$.
   - Thus, $\mathbf{E}[\, Y \,] = (-1)(\frac{35}{36}) + (35)(\frac{1}{36}) = 0$, 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?

Answer

1. Let E_X : bet on red colour.
   Thus, $\mathbf{E}[\,X\,] = (-1)(\frac{20}{38}) + (1)(\frac{18}{38}) = -\frac{1}{19}$.

2. Let E_Y : bet on number 10.
   Thus, $\mathbf{E}[\,Y\,] = (-1)(\frac{37}{38}) + (35)(\frac{1}{38}) = -\frac{1}{19}$.

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