Elements of Probability

Fall semester 2019

Keivan Mallahi-Karai

25 September 2019

Jacobs University

When we use probability theorem in real life, our sample space can be viewed as the set of all possible scenarios.

Example: Modeling the stock market. Each point ω in the sample space can be viewed as a possible state of the world at some point in the future.

We are typically *not* interested in ω itself, but rather in quantities that depend on ω.

A typical example is the price of a stock S, which depends on the state of the world ω, and hence can be viewed as a function on the sample space Ω.

More generally, we are interested in assigning a numerical quantity to an outcome $\omega \in \Omega$ of the experiment that captures one particular aspect. This leads to the following definition.

Random variables: definition

Definition
Consider a probability space with the sample space Ω. A function

$$X : \Omega \to \mathbb{R}$$

is called a real valued *random variable*. Similarly, a function $X : \Omega \to \mathbb{R}^n$ is called a vector-valued random variable.

Example
Suppose that the flipping of a coin can result in heads with probability p and in tails with probability $1 - p$, This coin is tossed n times. For each outcome ω consider:

$$X_1(\omega) = \{\text{first head}\},$$
$$X_2(\omega) = \{\text{first tail}\},$$
$$X_3(\omega) = \{\text{total number of H}\},$$
$$X_4(\omega) = \{\text{total number of T}\}.X_5(\omega) \qquad = \{\text{total number of HH}\}.$$

Discrete random variables

Definition
A random variable X is called **discrete** if it takes a finite or countable number of values. The **probability mass function** of X is the function defined by

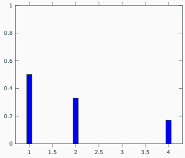
$$p(x) = \mathbb{P}[X = x].$$

Example

Suppose X takes values $1, 2, 4$ with probabilities:

$$\mathbb{P}[X = 1] = \frac{1}{2}, \quad \mathbb{P}[X = 2] = \frac{1}{3}, \quad \mathbb{P}[X = 4] = \frac{1}{6}.$$

x	1	2	4
$\mathbb{P}[X = x]$	1/2	1/3	1/6



Bernoulli random variables

The simplest discrete random variables are Bernoulli random variables.

Definition
A random variable X is called the *Bernoulli* random variable with parameter p if it only takes values 0 and 1, and

$$\mathbb{P}[X = 1] = p, \qquad \mathbb{P}[X = 0] = 1 - p.$$

A Bernoulli random variable X tells us whether something happened or not. The probability of happening $\mathbb{P}[X = 1]$ is called the parameter of X.

Example

A die is rolled. Let X be the random variable that tells us whether the outcome is larger than 4 or not. X has parameter $p = 2/6$.

⚀ ⚁ ⚂ ⚃ $\longrightarrow$ 0

⚄ ⚅ $\longrightarrow$ 1

Binomial distribution

Consider a coin that comes up heads with probability p. The coin is thrown n times. Suppose that the outcomes of different rounds are independent.

Suppose $n = 2$: Then

$$\begin{aligned}
\mathbb{P}[HH] &= \mathbb{P}[\text{ first} H]\,\mathbb{P}[\text{ second} H] = p^2. \\
\mathbb{P}[HT] &= \mathbb{P}[\text{ first} H]\,\mathbb{P}[\text{ second} T] = p(1-p). \\
\mathbb{P}[TH] &= \mathbb{P}[\text{ first} T]\,\mathbb{P}[\text{ second} H] = (1-p)p. \\
\mathbb{P}[TT] &= \mathbb{P}[\text{ first} T]\,\mathbb{P}[\text{ second} T] = (1-p)^2.
\end{aligned} \tag{1}$$

$$\begin{array}{ccccc}
\text{HH} & \longrightarrow & 2 & p^2 \\
\text{HT TH} & \longrightarrow & 1 & 2p(1-p) \\
\text{TT} & \longrightarrow & 0 & (1-p)^2
\end{array}$$

Binomial distribution

Suppose $n = 3$. Then the number of heads could be $0, 1, 2, 3$

$$\begin{array}{ccc} \text{HHH} & \longrightarrow & 3 & p^3 \\ \text{HHT HTH THH} & \longrightarrow & 2 & 3p^2(1-p) \\ \text{HTT THT TTH} & \longrightarrow & 1 & 3p(1-p)^2 \\ \text{TTT} & \longrightarrow & 0 & (1-p)^3 \end{array}$$

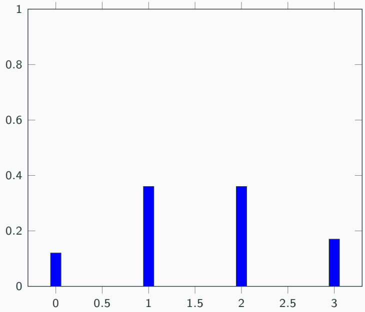
Definition
A random variable X has the Binomial distribution with parameters (n, p) if,

$$\mathbb{P}[X = k] = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$

Binomial distribution bar charts

Suppose $n = 3$ and $p = 1/2$. Then the values X can attain are $0, 1, 2, 3$. We have

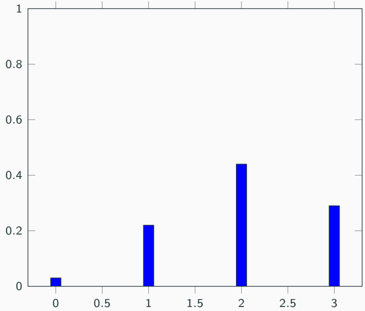
x	0	1	2	3
$\mathbb{P}[X = x]$	$1/8$	$3/8$	$3/8$	$1/8$



Binomial distribution bar charts

Suppose $n = 3$ and $p = 2/3$. Then the values X can attain are $0, 1, 2, 3$. We have

x	0	1	2	3
$\mathbb{P}[X = x]$	$1/27$	$6/27$	$12/27$	$8/27$



Review: Random variables and examples

Definition
In a probability space Ω, a function

$$X : \Omega \to \mathbb{R}$$

is called a real valued *random variable*.

Definition
A *Bernoulli* random variable with parameter p takes values 0 and 1, and

$$\mathbb{P}[X = 1] = p, \qquad \mathbb{P}[X = 0] = 1 - p.$$

Definition
A Binomial random variable X with parameters (n, p) takes values $0, 1, \ldots, n$ and

$$\mathbb{P}[X = k] = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$

Examples

Example

Werder Bremen football team wins each game with probability 20 percent and loses with probability 80 percent. What is the probability that they win exactly 4 games out of 34 games.

$$\mathbb{P}[E] = \binom{34}{4} \left(\frac{4}{10}\right)^4 \left(\frac{6}{10}\right)^{30} \approx 0.09.$$

Tie-breaking probability revisited

Consider $2n$ voters, each voting independently with probability p for candidate A and with probability $1-p$ for candidate B. What is the probability of a tie?

Denote the number of votes for candidate A by X.

X is a binomial random variable with parameters $(2n, p)$.

It follows that

$$\mathbb{P}[X = n] = \binom{2n}{n} p^n (1-p)^n.$$

$$\mathbb{P}[X = n] \approx \frac{(4p(1-p))^n}{\sqrt{\pi n}}.$$

	10	20	100	1000
0.5	0.17	0.12	0.05	0.01
0.45	0.16	0.10	0.02	7×10^{-7}
0.40	0.11	0.05	0.001	3×10^{-20}

A Bernoulli trial is repeated until a success occurs. Denote by X the number of needed trials. So X takes values $1, 2, 3, \ldots$.

$$\mathbb{P}[X = 1] = p.$$

$$\mathbb{P}[X = 2] = (1 - p)p.$$

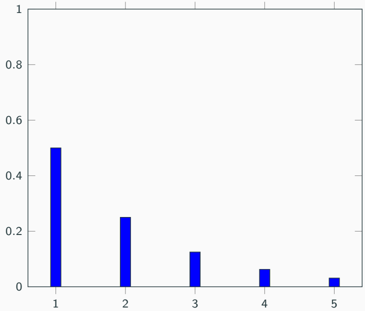
$$\mathbb{P}[X = 3] = (1 - p)^2 p.$$

Definition
We say that X is a *geometric random variable or that* X *has geometric distribution with parameter* p *if its probability mass function is given by*

$$p(k) = p(1 - p)^{k-1}, \qquad k = 1, 2, 3, \ldots$$

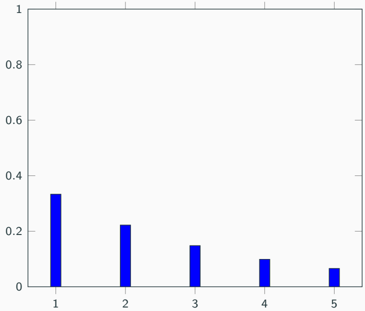
Geometric random variables

Let $p = 1/2$. Then the bar chart of a geometric random variable looks like this:



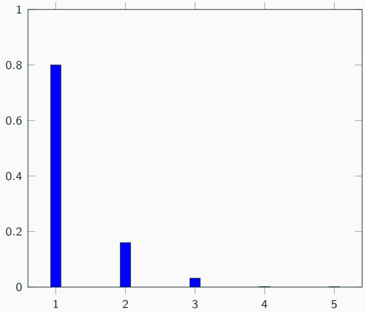
Geometric random variables

Let $p = 1/3$. Then the bar chart of a geometric random variable looks like this:



Geometric random variables

Let $p = 0.8$. Then the bar chart of a geometric random variable looks like this:



Another way of organizing information about a random variable is by using the distribution function.

Definition
Suppose X is a random variable. The *distribution function of* X *is the function defined by*

$$F_X(x) = \mathbb{P}\left[X \leq x\right].$$

Note the difference between the probability mass function and the distribution function of a discrete random variable X:

$$p_x(x) = \mathbb{P}\left[X = a\right].$$

$$F_X(x) = \mathbb{P}\left[X \leq x\right].$$

Example

A discrete random variable has the probability mass function given by

$$p_X(0) = \frac{1}{3}, \quad p_X(1) = \frac{1}{4}, \quad p_X(2) = \frac{5}{12}.$$

Find the distribution function of X.

It is clear that

$$F_X(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1/3 & \text{if } 0 \leq t < 1 \\ \frac{1}{3} + \frac{1}{4} & \text{if } 1 \leq t < 2 \\ 1 & \text{if } t \geq 2 \end{cases}$$

Example

Suppose X is a discrete random variable with the probability mass function given by

$$p_X(x) = \begin{cases} k \cdot x & \text{if } x = 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

1. Find the value of k.
2. Compute the probability of the event that X is even.
3. Compute $F_X(t)$.

x	1	2	3	4
$\mathbb{P}[X = x]$	k	2k	3k	4k

Clearly

$$k + 2k + 3k + 4k = 1 \Rightarrow k = 0.1.$$

x	1	2	3	4
$\mathbb{P}[X = x]$	1/10	2/10	3/10	4/10

Examples continued

x	1	2	3	4
$\mathbb{P}[X = x]$	$1/10$	$2/10$	$3/10$	$4/10$

$$\mathbb{P}[X \text{is even}] = \frac{2}{10} + \frac{4}{10} = \frac{6}{10}.$$

$$F_X(t) = \begin{cases} 0 & \text{if } t < 1 \\ 1/10 & \text{if } 1 \leq t < 2 \\ 3/10 & \text{if } 2 \leq t < 3 \\ 6/10 & \text{if } 3 \leq t < 4 \\ 1 & \text{if } t > 4 \end{cases}$$

Poisson random variables

Consider a binomial random variable with parameters p and n. Set

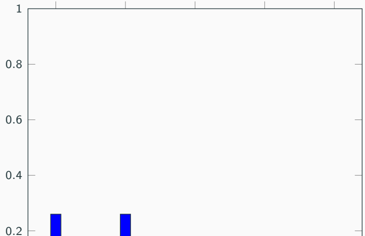
$$pn = \lambda$$

and let $n \to \infty$ and $p \to 0$. A simple computation shows that

$$\mathbb{P}[X = k] \to e^{-\lambda} \frac{\lambda^k}{k!}.$$

Definition
A random variable X is said to have *Poisson distribution* with parameter $\lambda > 0$ when

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$



Example

Suppose that the number of daily car accidents in a city has a Poisson distribution with parameter $\lambda = 2$. What is the probability that

1. There are no accidents on a certain day.
2. There are at most three accidents on a certain day.

Denote the number of accidents by X.

$$\mathbb{P}[X = 0] = e^{-2}\frac{2^0}{0!} = e^{-2} = 0.135.$$

$$\mathbb{P}[X \leq 3] = e^{-2} \cdot \frac{19}{3} = 0.85.$$

The notion of average

The average of a set of numbers $x_1, \ldots, x_n$ is

$$\frac{x_1 + \cdots + x_n}{n}$$

This is a useful notion of average only when numbers are equally significant, and hence must receive the same weight which is $1/n$.

What is the points are not equally important?

Expectation of a random variable

The expectation of a random variable gives us a weighted average of values of X.

Definition
For a discrete random variable X with values $x_1, x_2, \dots,$ obtained with probabilities $p_1, \dots, p_n$. Then the *expected value* of X is defined by

$$\mathbb{E}[X] = p_1 x_1 + \cdots + p_n x_n.$$

Example
A fair die is rolled. Suppose that X is the number shown. Compute $\mathbb{E}[X]$.

Clearly X takes value $1 \leq n \leq 6$, each with probability $1/6$. From here we have:

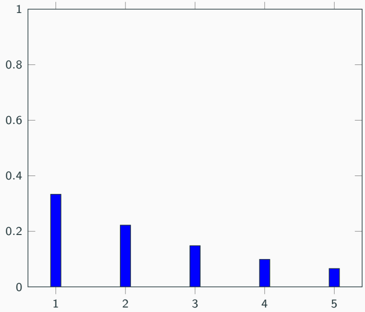
$$\mathbb{E}[X] = \frac{1}{6} \sum_{n=1}^{6} n = \frac{21}{6} = 3.5$$

Physical interpretation of the expected value

Expectation of a random variable $\equiv$ the center of mass of a finite set of points:

Put weights $m_1, \ldots, m_n$ at locations $x_1, \ldots, x_n$. Then the center of mass is at

$$\frac{m_1 x_1 + \cdots + m_n x_n}{m_1 + \cdots + m_n}.$$



Expected value of Bernoulli and binomial random variables

Example
Let X be a Bernoulli random variable with parameter p. Then

$$\mathbb{E}[X] = p \cdot 1 + (1-p) \cdot 0 = p.$$

Example
Let X be a Binomial random variable with the parameters (n, p). Then

$$\mathbb{E}[X] = \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j} = np.$$

Later we will see a better way of proving this.

Expected value of Poisson random variable

Suppose X is Poisson random variable with parameter λ. Then we have

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

$$\mathbb{E}[X] = 0 \cdot \mathbb{P}[X = 0] + 1 \cdot \mathbb{P}[X = 1] + 2 \cdot \mathbb{P}[X = 2] + 3 \cdot \mathbb{P}[X = 3] + \cdots$$

$$\mathbb{E}[X] + e^{-\lambda} \left(1 \frac{\lambda}{1!} + 2 \frac{\lambda^2}{2!} + 3 \frac{\lambda^3}{3!} + \cdots \right)$$

$$\mathbb{E}[X] = \lambda.$$

Expected value as the value of a game

Consider a lottery in which you can win values

$$x_1, \ldots, x_n$$

with probabilities $p_1, \ldots, p_n$.

If the lottery played many times then one expects the relative frequency of x_1 to be p_1, the relative frequency of x_2 to be p_2, etc. So, the average value of the lottery is

$$\mathbb{E}[X] = p_1 x_1 + \cdots + p_n x_n.$$

Chuck-a-Luck

- You pay 1 Euro to play the game.
- You roll three dice. If you roll at least one 6 then you receive 2 Euros.

Is this game fair? Denote by A the event that at least a 6 occurs and by X the value the game.

$$\mathbb{P}[A] = 1 - \frac{5^3}{6^3} = \frac{91}{216} \approx 0.42.$$

$$\mathbb{E}[X] = (-1) \cdot 0.58 + 1 \cdot 0.42 = -0.16.$$

Chuck-a-Luck: Second version

- You pay 1 Euro to play the game.
- You roll three dice. If you roll at least one 6 then you receive 2 Euros.
- If you roll three 6s, then you win an additional 25 Euros.

B: event that three 6s occur.

$$\mathbb{P}[B] = \frac{1}{216}.$$

x	-1	1	26
$\mathbb{P}[X = x]$	$\frac{125}{216}$	$\frac{90}{216}$	$\frac{1}{216}$

$$\mathbb{E}[X] = (-1)\frac{125}{216} + 1 \cdot \frac{90}{216} + 26 \cdot \frac{1}{216} \approx -0.04.$$

The idea of utility and its history

The idea of utility goes back to Daniell Bernoulli (1738).

Suppose a coin is thrown until a head occurs. If this happens at round n for the first time then the player gets 2^n dollars.

If head occurs in the first round, the player get 2 dollars.

If tails occurs in the first round, and heads in the second round then the player get 4 dollars.

How much are you willing to pay to play this game?

The value of the game is

$$2 \times \frac{1}{2} + 4 \times \frac{1}{4} + 8 \times \frac{1}{8} + \cdots = \infty$$

In reality people only pay around 3-4 dollars to play this game.

Bernoulli: it is not the absolute value of money but its utility

Solution to the paradox

If the utility is

$$u(x) = 2\log_2 x$$

then:

1. With probability $1/2$, one gets $u(2) = \log_2 2 = 2$
2. With probability $1/4$, one gets $u(4) = \log_2 4 = 4$
3. With probability $1/8$, one gets $u(8) = \log_2 8 = 6$.

So the expected utility is

$$\frac{2}{2} + \frac{4}{4} + \frac{6}{8} + \frac{8}{16} + \cdots = 4.$$

Attitude towards risk

Suppose you are offered the choice between the following alternatives:

1. 1 dollars.
2. first throwing a fair coin. If the outcomes is heads 2 dollars and if the outcome is tails zero.

1. 100000 dollars.
2. first throwing a fair coin. If the outcomes is heads 200000 dollars and if the outcome is tails zero.