Lecture III:
Famous Discrete Distributions

Carlo Cavicchia

carlo.cavicchia@uniroma1.it

Famous random variables

Some particular probability distributions occur often because they are useful description of certain chance phenomenon under study, sometimes is the experiment itself that allows us to determinate which probability distribution we have to take into account.

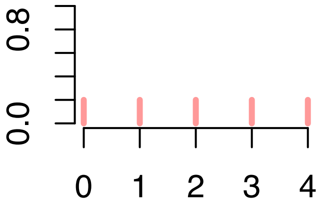
Uniform
idea

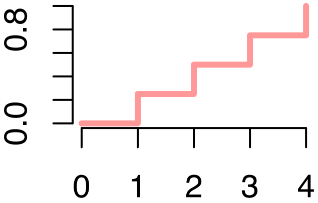
- The discrete uniform distribution is a symmetric probability distribution whereby a finite number of values are equally likely to be observed.
- Given n possible values for X, distributed uniformly, the probability to observe each value is equal to $1/n$.
- The discrete uniform distribution itself is inherently non-parametric. It is convenient, however, to represent its values generally by all integers in an interval $[a, b]$, so that a and b become the main parameters of the distribution.

Uniform

$X \sim \text{Uniform}(a, b)$

$$p_x = P(X \leq x) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1}$$





Formulas

expected value

$X \sim \text{Uniform}(a, b)$

$$p_x = P(X \leq x) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1}$$

$X \in [a, b] = [a, a + k, a + 2k, \dots, b]$ where $b = a + (n-1)k$

$$\mathbb{E}[X] = \sum_x x p_x = \sum_{l=0}^{n-1} x \frac{1}{n} = \frac{1}{n} \sum_{l=0}^{n-1} a + lk =$$

$$= \frac{1}{n}[na + k \sum_{l=0}^{n-1} l] = a + \frac{k(n-1)n}{2n} =$$

$$= a + \frac{k(n-1)}{2} = a + \frac{b-a}{2} =$$

$$= \frac{a+b}{2}$$

Formulas

expected value of X squared

$X \sim \text{Uniform}(a, b)$

$$p_x = P(X \leq x) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1}$$

$$\begin{aligned}
\mathbb{E}[X^2] = \sum_x x^2 p_x &= \frac{1}{b-a+1} \sum_{x=a}^{b} x^2 = \\
&= \frac{1}{b-a+1} \left(\frac{(b^2+b)(2b+1) - (a^2-a)(2a-1)}{6} \right) = \\
&= \frac{1}{b-a+1} \left(\frac{(2b^3+3b^2+b) - (2a^3-3a^2+a)}{6} \right)
\end{aligned}$$

Formulas

variance

$X \sim \mathsf{Uniform}(a, b)$

$$p_x = P(X = x) = \frac{\lfloor k \rfloor - a + 1}{b - a + 1}$$

$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - (\frac{a+b}{2})^2$

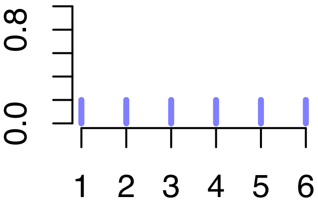
$$\begin{aligned}
\mathbb{V}[X] &= \frac{1}{b-a+1}(\frac{(2b^3 + 3b^2 + b) - (2a^3 - 3a^2 + a)}{6}) - (\frac{a+b}{2})^2 = \\
&= \cdots = \\
&= \frac{(b-a+1)^2 - 1}{12}
\end{aligned}$$

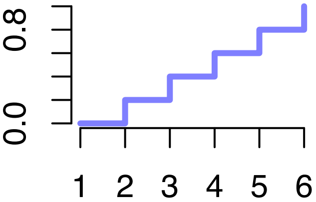
Example

Fair dice

- A simple example of the discrete uniform distribution is throwing a fair die. The possible values are $1, 2, 3, 4, 5, 6$, and each time the dice is thrown the probability of a given score is $1/6$. If two dice are thrown and their values added, the resulting distribution is no longer uniform since not all sums have equal probability.

$$P(X = x) = 1/6, \text{ with } x = 1, 2, 3, 4, 5, 6$$





Exercises

- A die is rolled.
  1. List the possible outcomes in the sample space.
  2. What is the probability of getting a number which is even?
  3. What is the probability of getting a number which is greater than 4?
  4. What is the probability of getting a number which is less than 3? What is its complement?
- Two dice are rolled.
  1. Construct the sample space. How many outcomes are there?
  2. Find the probability of rolling a sum of 7.
  3. Find the probability of getting a total of at least 10.
  4. Find the probability of getting a odd number as the sum.

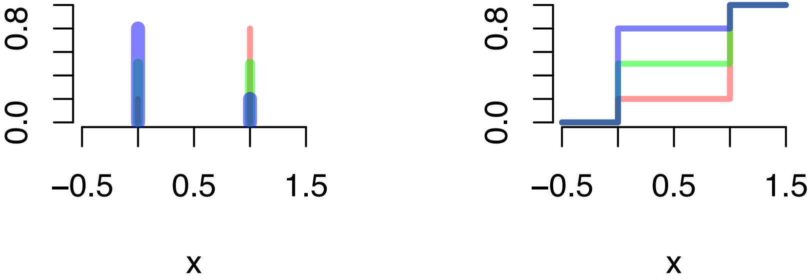
Bernoulli

idea

- Only to possible outcome (mutually exclusive and exhaustive): success and failure.
- Probability of success p is the only one parameter.
- Probability of failure is equal to $1 - p$.

Bernoulli

$$X = \begin{cases} 0 & \text{with probability } 1-p \\ 1 & \text{with probability } p \end{cases}$$



Formulas :: challenge

expected value and variance

$X \sim$ Bernoulli(p)

- who wants to try to compute $\mathbb{E}[X]$ and $\mathbb{V}[X]$?

Formulas
expected value

$X \sim \text{Bernoulli}(p)$

$$\begin{aligned}\mathbb{E}[X] &= \sum_x x p_x \\ &= 0 \times p_0 + 1 \times p_1 \\ &= 0 \times (1-p) + 1 \times p = p\end{aligned}$$

Formulas

expected value of X squared

$X \sim$ Bernoulli(p)

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_x x^2 p_x \\ &= 0 \times p_0 + 1 \times p_1 \\ &= 0 \times (1-p) + 1 \times p = p\end{aligned}$$

Formulas

variance

$X \sim \text{Bernoulli}(p)$

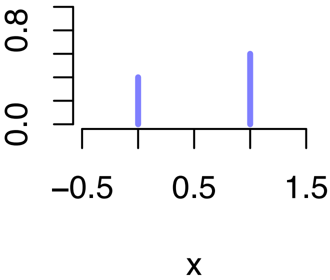
$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - p^2$

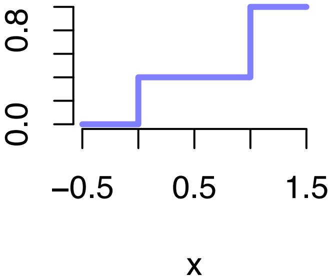
$$\begin{aligned}\mathbb{V}[X] &= p - p^2 \\ &= p(1-p)\end{aligned}$$

Example

- Let $X = 1$ if the price next month of Microsoft stock goes up and $X = 0$ if the price goes down (assuming it cannot stay the same). The probability of the event "the price next month of Microsoft stock goes up" is equal to $3/5$.

$$X = \begin{cases} 0 & \text{with probability } 2/5 \\ 1 & \text{with probability } 3/5 \end{cases}$$





Binomial

idea

Let $Y_1, Y_2, \dots, Y_n$ be n independent random variables identical distributed as a Bernoulli of parameter p: what does it happen to $X = Y_1 + Y_2 + \dots + Y_n$?

Binomial

idea

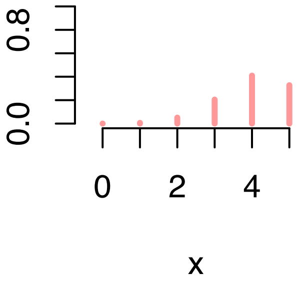
Conditions for Binomial Distribution

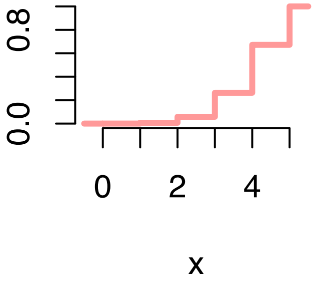
- Each of n trials has two possible outcomes. The outcome of interest is called a success and the other outcome is called a failure.
- Each trial has the same probability of a success. This is denoted by p, so the probability of a success is p and the probability of a failure is $1 - p$.
- The n trials are independent. That is, the result for one trial does not depend on the results of other trials.

The binomial random variable X is the number of successes in the n trials.

Binomial

$X \sim \text{Binomial}(n, p)$





Formulas

expected value

$X \sim$ Binomial(n, p)

$$p_x = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\begin{aligned}
\mathbb{E}[X] = \sum_x x p_x &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \binom{n}{x} p^x (n-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{z=0}^{n-1} \frac{n(n-1)!}{z!(n-z-1)!} p^{z+1} (1-p)^{n-z-1} \\
&= np \sum_{z=0}^{n-1} \frac{(n-1)!}{z!(n-z-1)!} p^z (1-p)^{n-1-z} = np
\end{aligned}$$

Formulas

expected value of X squared

$X \sim$ Binomial(n, p)

$$p_x = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\begin{aligned}
\mathbb{E}[X^2] = \sum_x x^2 p_x &= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{z=0}^{n-1} (z+1) \frac{n(n-1)!}{z!(n-z-1)!} p^{z+1} (1-p)^{n-z-1} \\
&= np[\sum_{z=0}^{n-1} z \binom{n-1}{z} p^z (1-p)^{n-1-z} + \sum_{z=0}^{n-1} \binom{n-1}{z} p^z (1-p)^{n-1-z}] \\
&= np(n-1)p + np
\end{aligned}$$

Formulas

variance

$X \sim \text{Binomial}(n, p)$

$$p_x = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - (np)^2$

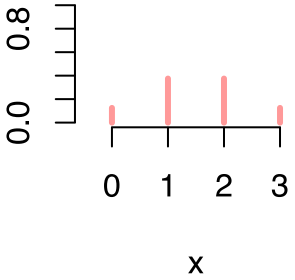
$$\begin{aligned} \mathbb{V}[X] &= np(n-1)p + np - (np)^2 \\ &= n^2p^2 - np^2 + np = \\ &= np(p-1) \end{aligned}$$

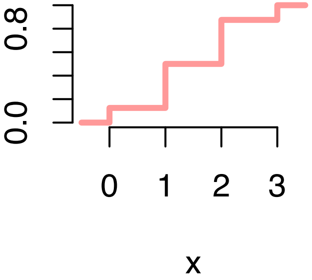
Example

- Mark flips a fair coin three times. X is the number of heads.

$$p_x = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

X may assume four different values: $0, 1, 2, 3$. $\mathbb{E}[X] = 1.5$ and $\mathbb{V}[X] = 0.75$.





Exercises

- A quiz in statistics course has four multiple-choice questions, each with five possible answers. A passing grade is three or more correct answers to the four questions. Allison has not studied for the quiz. She has no idea of the correct answer to any of the questions and decides to guess at random for each.
    1. Find the probability she lucks out and answers all four questions correctly.
    2. Find the probability that she passes the quiz.

- Each newborn baby has a probability of approximately 0.49 of being female and 0.51 of being male. For a family of four children, let X = number of children who are girls.
    1. Explain why the three conditions are satisfied for X to have the binomial distribution.
    2. Identify n and p for the binomial distribution.
    3. Compute the mean and the variance of X.
    4. Find the probability that the family has two girls and two boys.

Poisson

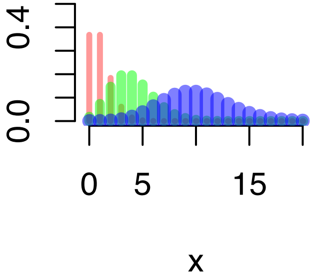
idea

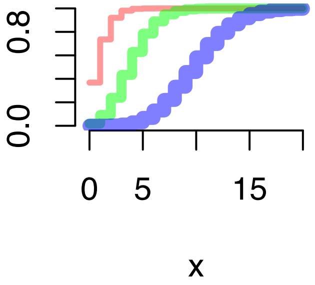
- Used to describe the number of events in a given interval of time or in a given space
    - $\#$ of clients calling a call-center
    - $\#$ of defects in a square meter of a manufactured good
    - $\#$ of patients arriving to the emergency hospital in the last hour

Poisson

$X \sim \text{Poisson}(\lambda)$

$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$





Formulas

expected value

$X \sim \text{Poisson}(\lambda)$

$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\begin{aligned}
\mathbb{E}[X] = \sum_x x p_x &= \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} = \\
&= \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x(x-1)!} = \\
&= \sum_{x=1}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-1)!} = \\
&= \sum_{z=0}^{\infty} \frac{\lambda^{z+1} e^{-\lambda}}{z!} = \\
&= \lambda \sum_{z=0}^{\infty} \frac{\lambda^z e^{-\lambda}}{z!} = \lambda
\end{aligned}$$

Formulas

expected value of X squared

$X \sim$ Poisson(λ)

$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\begin{aligned}
\mathbb{E}[X^2] = \sum_x x^2 p_x &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!} = \\
&= \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{(x-1)!} = \\
&= \sum_{z=0}^{\infty} (z+1) \frac{\lambda^{z+1} e^{-\lambda}}{z!} = \\
&= \lambda\left(\sum_{z=0}^{\infty} z \frac{\lambda^(z)e^{-\lambda}}{z!} + \sum_{z=0}^{\infty} \frac{\lambda^(z)e^{-\lambda}}{z!}\right) = \\
&= \lambda(\lambda + 1) = \lambda^2 + \lambda
\end{aligned}$$

Formulas

variance

$X \sim \text{Poisson}(\lambda)$

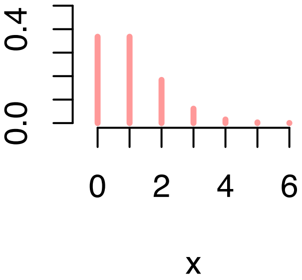
$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

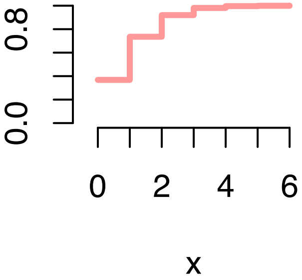
$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \lambda^2$

$$\begin{aligned} \mathbb{V}[X] &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

Example

- In average a call-center receives 1 per hour.





Exercises

- The average number of homes sold by the Acme Realty company is 2 homes per day. What is the probability that exactly 3 homes will be sold tomorrow?
- Suppose the average number of lions seen on a 1-day safari is 5. What is the probability that tourists will see fewer than 4 lions on the next 1-day safari?

Challenge

- In your pocket, there are 5 keys that all look the same. You need to find the right one to open the front door of your home. Compute the probability to find the right key on your second attempt assuming that you try them once at time, with no repetition.