Famous Discrete Random Variables

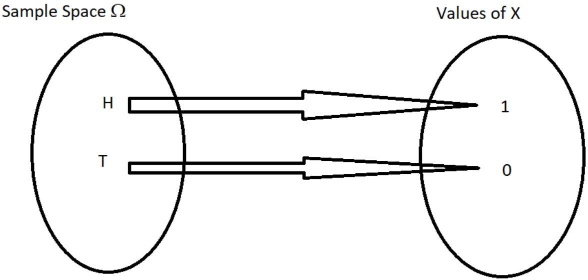
Lecture 3

Francesco Dotto
francesco.dotto@uniroma3.it

3/9/2021

A recap

A **random variable** is any function from the sample space to the real numbers.



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

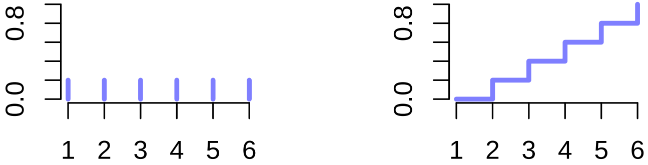
- The discrete uniform distribution is a symmetric probability distribution whereby a finite number of values are equally likely to be observed.
- Given n possible values for X, distributed uniformly, the probability to observe each value is equal to $1/n$.
- The discrete uniform distribution itself 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

An applied example

- 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.

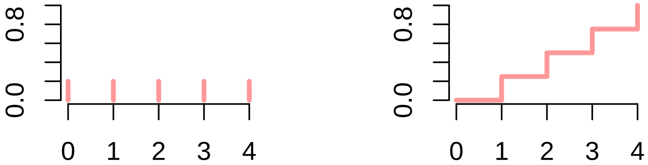
$$P(X = x) = 1/6, \text{ with } x = 1, 2, 3, 4, 5, 6$$



Uniform

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

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



Uniform Distribution

in a nutshell

Uniform Distribution

Let $X \sim Unif[a, b]$. Then, the following holds:

$$\mathbb{E}[X] = \frac{b+a}{2} \text{ and } \mathbb{V}[X] = \frac{(b-a+1)^2 - 1}{12}$$

Formulas

expected value

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

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

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

$$\begin{aligned}
\mathbb{E}[X] = \sum_x x p_x = \sum_{l=0}^{n-1} x \frac{1}{n} = \frac{1}{n} \sum_{l=0}^{n-1} a + lk = \\
= \frac{1}{n}[na + k \sum_{l=0}^{n-1} l] = a + \frac{k(n-1)n}{2n} = \\
= a + \frac{k(n-1)}{2} = a + \frac{b-a}{2} = \\
= \frac{a+b}{2}
\end{aligned}$$

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

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

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

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

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

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?

Bernoulli

idea

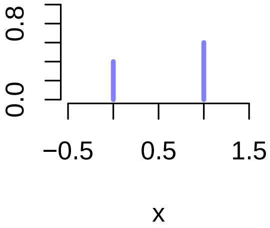
- Only to possible outcome (mutually exclusive and exhaustive): success and failure.
- Probability of success p is the only one parameter.
- Probability of failure is equal to $1 - p$.

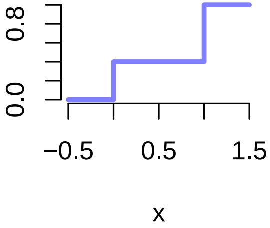
Bernoulli

An applied 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 $3/5$.

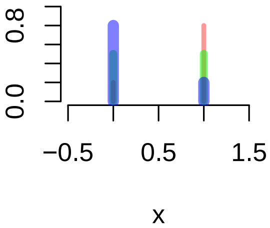
$$X = \begin{cases} 0 & \text{with probability } 2/5 \\ 1 & \text{with probability } 3/5 \end{cases}$$

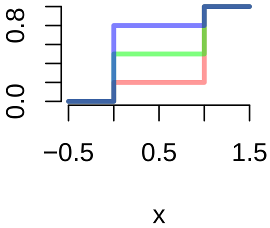




Bernoulli

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





Formulas :: challenge

expected value and variance

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

- 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

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

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

Formulas

variance

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

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

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

Exercise

In a population composed by 6503 women, 591 have passed the menopause. Determine the distribution of the random variable

$$X = \begin{cases} 0 & \text{the woman has passed the menopause} \\ 1 & \text{the woman has NOT passed the menopause} \end{cases}$$

From Bernoulli to Binomial

- A Bernoulli trial is a trial in which only two outcomes are possible:
  - 0: Failure
  - 1: Success
- If we perform a random experiment by repeating n independent Bernoulli trials, then the random variable X representing the number of successes in n trials is said to have a **Binomial** distribution

Binomial

Let us start from an applied example

Suppose 40% of a population supports Obama. A random sample of $n = 5$ voters is selected. Let X be the random variable representing the number of Obama supporters among $n = 5$ voters. Everytime we randomly select a voter we are doing a Bernoulli trial $Y_i \sim Be(0.4)$ with $i = 1, 2, \ldots, 5$.

- Compute $P(X = 0)$:
  $P(X = 0) = P(Y_1 = 0 \cap Y_2 = 0 \cap Y_3 = 0 \cap Y_4 = 0 \cap Y_5 = 0)$ and since $Y_i \sim Be(0.4)$ it yields $P(X = 0) = 0.6^6$
- What if I want to compute $P(X = 1)$? Any suggestions?

Binomial

Let us start from an applied example

- $P(X = 1)$:

$$
\begin{aligned}
P(X=1) = & \underbrace{(0.4)}_{\text{Y}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}} + \\
& \underbrace{(0.6)}_{\text{N}}\underbrace{(0.4)}_{\text{Y}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}} + \\
& \underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.4)}_{\text{Y}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}} + \\
& \underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.4)}_{\text{Y}}\underbrace{(0.6)}_{\text{N}} + \\
& \underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.6)}_{\text{N}}\underbrace{(0.4)}_{\text{Y}} = \\
& 5(0.6)^4(0.4)
\end{aligned}
$$

Binomial

Let us start from an applied example

The case of $P(X = 2)$ is slightly more complicate. We have to compute the number of possible configurations of 2 "Yes" and 3 "No". This is given by

$$\binom{5}{2} = \frac{5!}{(5-2)!2!}$$

where $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

Binomial distribution

Final remarks

Uniform Distribution

Given n Bernoulli trials with probability of succes p the random variable X representing the number of successes is called Binomial random variable ($Bin(n, p)$)and its probability mass function is given by

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

- If $X \sim (n, p)$ the following holds: $\mathbb{E}[X] = np$ and $\mathbb{V}[X] = np(1-p)$

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

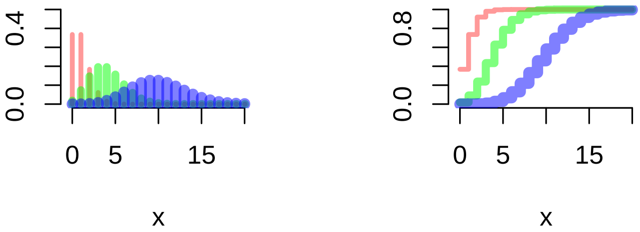
idea

- 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

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

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

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

Formulas

expected value of X squared

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

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

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

Formulas

variance

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

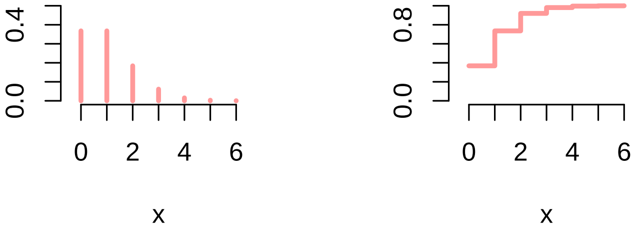
$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

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

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

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.