Discrete Random Variables

Pablo Arés Gastesi

Departamento de Matemáticas y Ciencia de Datos
Universidad San Pablo CEU
pablo.aresgastesi@ceu.es

Septembre 2024

Definition and Types

A **random variable** *is a function* X *on the sample space of a random event that takes numerical values.*

Example: we consider the experiment of tossing a coin twice. The possible basic outcomes are HH, HT, TH, TT. We define the random variable X as the number of heads. Thus

$$X(HH) = 2, \quad X(HT) = 1, \quad X(TH) = 1 \quad \text{and} \quad X(TT) = 0.$$

Random variables are of two types:

- Discrete: if X takes a finite (or infinite as the integers) number of values, as in the previous example: the number of heads obtained when tossing a coin N times.
- Continuous: if X takes all values in an interval, for example, the life time of a LED bulb.

Probability Distribution Function

The **probability distribution function** $P(x)$ *of a discrete random variable* X *gives us the probability that* X *takes the value* x*:* $P(x) = P(X = x)$.

In the above example, X takes the values $0, 1, 2$.

$$P(0) = 1/4 = 0.25, \quad P(1) = 2/4 = 0.5 \quad \text{and} \quad P(2) = 1/4 = 0.25$$

The **cumulative probability distribution** $F(x)$ *is given by the probability that* X *does not exceed the value* x*, that is*

$$F(x) = P(X \leq x)$$

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \leq x < 1 \\ 0.75 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

Expected Value

The **expected value** *$E[X]$ of a discrete random variable, also known as the* **mean**, *and denote by μ, is given by*

$$E[X] = \mu = \sum xP(x)$$

That is, μ is equal to the sum of all possible values multiplied by their probabilities.

For our previous example we have

$$E[X] = 0\,(0.25) + 1\,(0.5) + 2\,(0.25) = 1$$

Variance

The **variance** *of a discrete random variable X is the expected value of $(X-\mu)^2$:*

$$\sigma^2 = E[(X-\mu)^2] = \sum (x-\mu)^2\, P(x)$$

The **standard deviation** *σ is the positive square root of the variance.*

Above example:

$$\sigma^2 = (0-1)^2\,(0.25) + (1-1)^2\,(0.5) + (2-1)^2\,(0.25) = 0.25 + 0 + 0.25 = 0.5$$

and

$$\sigma = \sqrt{0.5} = 0.707$$

Variance

Important

The variance of X can also be computed as

$$\sigma^2 = E[X^2] - \mu^2$$

Example:

$$\begin{aligned} E[X^2] &= 0^2\,(0.25) + 1^2\,(0.5) + 2^2\,(0.25) = 0 + 0.5 + 1 = 1.5 \\ \sigma^2 &= E[X^2] - \mu^2 = 1.5 - 1^2 = 1.5 - 1 = 0.5 \end{aligned}$$

Linear Combinations of a Variable

Important

If X is a discrete random variable with mean μ_X and variance σ_X^2, and $Y = a + bX$, with a and b two numbers, then

$$\mu_Y = a + b\mu_X, \qquad \sigma_Y = |b|\, \sigma_X.$$

Discrete Uniform Distribution

This is the distribution of an experiment with a finite number of outcomes, all equally likely to take place, like the outcomes of throwing a (fair) dice. If there are n outcomes, $x_1, x_2, \dots, x_n$, then the probability distribution function is

$$P(X = x_1) = P(X = x_2) = \cdots = P(X = x_n) = \frac{1}{n}.$$

If X takes the values $a, a+1, \dots, b$ then we have

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

For a dice, $a = 1, b = 6, E[X] = 3.5, V[X] = 2.91\bar{6}$

Bernoulli Distribution

A Bernoulli experiment has two possible outcomes, one called success and another called failure. The probability of success is p $(0 < p < 1)$, and the probability of failure is $1 - p$. A **Bernoulli random variable** *X takes the value 1 on a successful outcome, and 0 in a failure. We write $X \sim B(p)$.*

Important

If X is a Bernoulli random variable, then we have

$$\mu_X = E[X] = p, \qquad \sigma_X^2 = p(1-p)$$

Binomial Distribution

A **Binomial distribution***, with parameters n and p, consists on counting the number of successful outcomes in the repetition of a Bernoulli experiment with probability p (independent repetitions). We write $X \sim B(n, p)$.*

Observe that we count just the number of successful outcomes, no the order. If $0 \leq k \leq n$, how many different cases of denoted by C_k^n or $\binom{n}{k}$, and it is given by

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

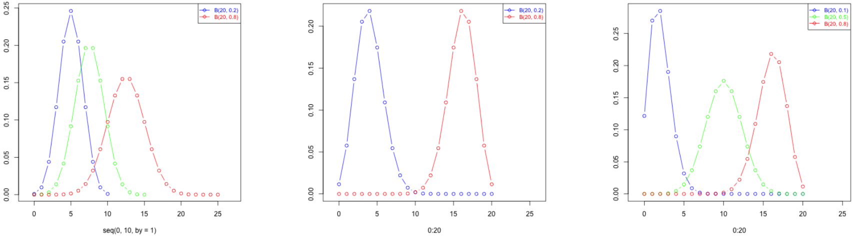
Binomial Distribution

Important (Binomial Distribution)

If $X \sim B(n, p)$ then we have:

- *Probability distribution function:* $P(x) = \binom{n}{x} p^x (1-p)^{n-x}$
- *Mean:* $E[X] = np$
- *Variance:* $V[X] = np(1-p)$

Binomial Distribution



Multinomial Distribution

Suppose we have $X_1, X_2, \ldots, X_k$ binomial random variables with $X_1 \sim B(n, p_1), X_2 \sim B(n, p_2), \ldots, X_k \sim B(n, p_k)$, and $p_1 + p_@ + \cdots + p_k = n$. Then we say that the joint variables of random vector $X = (X_1, \ldots, X_k)$ follows a **multinomial** *distribution, $X \sim M(n, p_1, \ldots, p_k)$.*

We have:

Important (Multinomial Distribution)

Suppose $X \sim M(n, p_1, \ldots, p_k)$.

- *Probability distribution function:* $P(X_1 = n_1, \ldots, X_k = n_k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}$, *with* $n_1 + \cdots n_k = n$.
- *Mean of each variable:* $E[X_i] = np_i$
- *Variance of each variable¿* $V[X_i] = np_i(1 - p_i)$

Poisson Distribution

The **Poisson distribution** *expresses the probability of a given number of events occurring in a fixed interval (of time or space) if these events occur with a known constant mean rate and independently (an occurence in one inteval does not influence the probability of an occurrence in another interval). The rate of occurrence in an interval is denoted by λ, and we write $X \sim P(\lambda)$.*

The interval mentioned in the above description does not need to have length 1. For example we can study the number of patients admitted into a hospital is 4 patients in 2 hours. By the scaling property we have that the number of patients admitted in one hour will be 2, and in one day we will have 48.

Poisson Distribution

Important

If $X \sim P(\lambda)$ *we have*

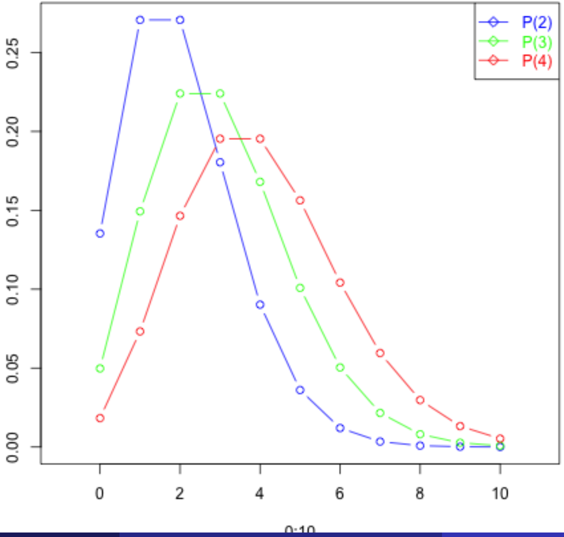
- *Probability distribution function:* $P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ *with* x *non-negative integer.*
- *Mean:* $E[X] = \lambda$
- *Variance:* $V[X] = \lambda$

We have that the sum of independent Poisson variables is also a Poisson variable.

If $X_1 \sim P(\lambda_1), X_2 \sim P(\lambda_2), \ldots, X_k \sim P(\lambda_k)$ *are independent random variables, then*

$$X_1 + X_1 + \cdots + X_k \sim P(\lambda_1 + \lambda_2 + \cdots + \lambda_k)$$

Poisson Distribution



The formula for the distribution function of the binomial is a little complicated to compte, while that for the Poisson is simpler. Thankfully we have that the binomial distribution $B(n, p)$ can be approximated by a Poisson distribution $P(\lambda)$, with $\lambda = np$, when $n \geq 50$ and $p \leq 0.1$.
Example: take $X \sim B(100, 0.01)$; then $\lambda = 100 \cdot 0.01 = 1$. Set $Y \sim P(1)$.
We consider, for example, $x = 2$. We have:

$$P_X(2) = \binom{100}{2} (0.01)^2 (0.99)^{98} \approx 0.1849$$

$$P_Y(2) = \frac{1^2}{2!} e^{-1} \approx 0.1839$$