

Seminar 4

Random variables

A **random variable** is a function from sample space to the real numbers $X : \mathcal{S} \rightarrow \mathbb{R}$.

It means that for every outcome $\omega \in \mathcal{S}$ there is a real number $X(\omega)$.

Not all functions are allowed, but this is slightly beyond the scope of our course.

Distribution of a random variable

Consider random variable $X : \mathcal{S} \rightarrow \mathbb{R}$. We introduce **the distribution** (or distribution law) \mathcal{L} of random variable X . Distribution acts on numbers in \mathbb{R} in the same way as probability function P acts on outcomes. We will write $X \sim \mathcal{L}$.

Two types of distributions

A probability distribution can be **discrete** or **continuous**.

Discrete random variables can only take countably many values (think integers), continuous random variables can take uncountably many values (think reals).

There is also a third type of distributions, which you never encounter in practice; it's possible for a distribution to be a mix of several types, which you also do not normally encounter.

Example 1

Consider event A and a random variable $X = \mathbb{I}nd_A$, an indicator:

$$\mathbb{I}nd_A(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{else} \end{cases}$$

$$\mathbb{P}(X = 1) = \mathbb{P}(A) = p$$

$$\mathbb{P}(X = 0) = 1 - \mathbb{P}(A) = 1 - p$$

We say that X follows **Bernoulli distribution** with parameter p and write $X \sim Be(p)$.

We will call $\mathbb{P}_X(\omega)$ a **probability mass function** (PMF).

Bernoulli trial scheme

Previously we have worked with independent events that were happening in one probability space. But sometimes we want to have multiple trials, where for every trial the probability space is known, but we are interested in the probability space covering all the trials at once. We can achieve it via direct product of probability spaces.

If all probability spaces are the same and equal to:

- $S = \{0, 1\}$
- $\mathbb{P}(1) = p$ and $\mathbb{P}(0) = 1 - p$

Then we call such experiment a **Bernoulli trial scheme**, and the probability space of it is:

- $S = \{(i_1, \dots, i_n), i_j \in \{0, 1\}\}$
- $\mathbb{P}(i_1, \dots, i_n) = p^{\text{num } j \text{ such that } i_j=1} (1 - p)^{\text{num } j \text{ such that } i_j=0}$

Example 2

Consider $X_1, \dots, X_n \sim Be(p)$ independent random variables. Then $Y = \sum_{k=1}^n X_k$ follows **Binomial distribution** with parameters n and p , $Y \sim Bi(n, p)$. $\mathbb{P}(Y = k) = ?$

Solution 2

If $Y \sim \text{Bi}(n, p)$, then

$$\mathbb{P}(Y = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Example 3

We say X follows discrete uniform distribution $\text{DU}([1, n])$ and we write $X \sim \text{DU}([1, n])$ if

$$P(X = k) = \frac{1}{n}$$

Example 4

Consider an urn with w white balls and b black balls. We draw n balls out of the urn at random without replacement. Let X be the number of white balls in the sample. What is the distribution of X ? What is its PMF?

Solution 4

If $X \sim \text{HGeom}(w, b, n)$, then

$$\mathbb{P}(X = k) = \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}}$$

Example 5

What is the difference between hypergeometric and binomial distributions?

Reminder:

- Binomial story: Consider an urn with w white balls and b black balls. We draw n balls from the urn with replacement. Let X be the number of white balls in the sample.
- Hypergeometric story: Consider an urn with w white balls and b black balls. We draw n balls out of the urn at random without replacement. Let X be the number of white balls in the sample.

Bernoulli trials in Binomial story are independent. The Bernoulli trials in the Hypergeometric story are dependent, since the sampling is done without replacement.

Example 5

Consider X and Y independent \mathbb{Z} -valued random variables. $\mathbb{P}(X + Y = k) = ?$

Solution 5

$$\mathbb{P}(X + Y = k) = \sum_m \mathbb{P}(X = m) \mathbb{P}(Y = k - m)$$

Example 6

Let $X \sim \text{Bi}(n, p)$ and $Y \sim \text{Bi}(m, p)$ be independent. What is the distribution of $Z = X + Y$?

Solution 6

$$\begin{aligned} \mathbb{P}(X + Y = k) &= \sum_j \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-k+j} = \\ &= p^k (1-p)^{n+m-k} \sum_j \binom{n}{j} \binom{m}{k-j} = \\ &= \binom{n+m}{k} p^k (1-p)^{n+m-k} \\ Z &\sim \text{Bi}(n+m, p) \end{aligned}$$

Cumulative distribution function

Note that distribution of a discrete distribution of a random variable X is uniquely defined by its PMF $\mathbb{P}(X = x_i)$. In general, we define the distribution using cumulative distribution function (CDF):

$$F_X(x) = \mathbb{P}(X < x)$$

It has the following properties:

- F_X is non-decreasing
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$
- $\lim_{x \rightarrow +\infty} F_X(x) = 1$
- F_X is left continuous

Interesting enough, the converse is also true. Any function that conforms to the properties above defines some probability distribution on \mathbb{R} and this relation is unique.

Probability density function

- If X has a discrete distribution, then F_X has a countable number of jumps $p_i = \mathbb{P}(X = x_i)$ and at $x = x_i$ it is continuous
- If X has absolutely continuous distribution, then F_X is differentiable a.e. and can be recovered from its derivative:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

where $f_X(t)$ is the probability density function and $f_X(t) = F'_X(x)$ a.e.