WuS - Lecture Notes Week 5

Ruben Schenk, ruben.schenk@inf.ethz.ch

March 29, 2022

0.1 Examples of Continuous Random Variables

0.1.1 Uniform Distributions

Definition (Uniform distribution in $[a, b]$, $a < b$**):** A coninuous random variable X is said to be **uniform in** $[a, b]$ if its density is equal to

$$f_{a,b}(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

In this case, we write $X \sim \mathcal{U}([a, b])$.

Intuition: X represents a uniformly chosen point in the interval $[a, b]$.

Properties:

- The probability to fall in an interval $[c, c+l] \subset [a, b]$ depends only on its length l:

$$\mathbb{P}[X \in [c, c+l]] = \frac{l}{b-a}.$$

- The distribution function of X is equal to:

$$F_X(x) = \begin{cases} 0 & \text{if } x < a, \\ \frac{x-a}{b-a} & \text{if } x \in [a, b], \\ 1 & \text{if } x > b. \end{cases}$$

0.1.2 Exponential Distribution

The exponential distribution is the continuous analogue of the geometric distribution.

Definition (Exponential distribution with $\lambda > 0$**:** A continuous random variable T is said to be **exponential with parameter** $\lambda > 0$ if its density is equal to:

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & x > 0. \end{cases}$$

In that case, we write $T \sim \text{Exp}(\lambda)$.

Intuition: T represents the time of a "clock ring". For example, the time at which the first customer arrives in a shop is well modeled by an exponential random variable.

Properties:

- The waiting probability is exponentially small:

$$\forall t \geq 0 : \mathbb{P}[T > t] = e^{-\lambda t}.$$

- It has the absence of memory property:

$$\forall t, s \geq 0 : \mathbb{P}[T > t + s \,|\, T > t] = \mathbb{P}[T > s].$$

0.1.3 Normal Distribution

Definition: A continuous random variable X is said to be **normal with parameters** m **and** $\sigma^2 > 0$ if its density is equal to:

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

In that case, we write $X \sim \mathcal{N}(m, \sigma^2)$.

Properties:

- If $X_1, ..., X_n$ are independent random variables with parameters $(m_1, \sigma_1^2), ..., (m_n, \sigma_n^2)$ respectively, then
$$Z = m_0 + \lambda_1 X_i + \cdots \lambda_n X_n$$
is a normal random variable with parameters $m = m_0 + \lambda_1 m_1 + \cdots + \lambda_n m_n$ and $\sigma^2 = \lambda_1^2 \sigma_1^2 + \cdots \lambda_n^2 \sigma_n^2$.
- In particular, if $X \sin \mathcal{N}(0, 1)$ (in this case we say that X is a **standard normal random variable**), then
$$Z = m + \sigma \cdot X$$
is a normal random variable with parameters m and σ^2.

1 Expectation

1.1 Expectation for General Random Variables

Definition: Let $x : \Omega \to \mathbb{R}_+$ be a random variable with nonnegative values. The **expectation** of X is defined as
$$\mathbb{E}[X] = \sum_0^\infty (1 - F_X(x))\, dx.$$

Proposition: Let X be a nonnegative random variable. Then we have
$$\mathbb{E}[X] \geq 0$$
with equality if and only if $X = 0$ almost surely.

Definition: Let X be a random variable. If $E[|X|] < \infty$, then the expectation of X is defined by
$$\mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-].$$

1.2 Expectation of a Discrete Random Variable

Proposition: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable with values in W (finite or countable) almost surely. We have
$$\mathbb{E}[X] = \sum_{x \in W} x \cdot \mathbb{P}[X = x],$$
provided the sum is well defined.

Example 1 (Bernoulli): Let X be a Bernoulli random variable with parameter p. We have
$$\mathbb{E}[X] = p.$$

Example 2 (Poisson): Let X be a Poisson random variable with parameter $\lambda > 0$, then
$$\mathbb{E}[X] = \lambda.$$

Definition: Let $A \in \mathcal{F}$ be an event. Consider the **indicator function** $\mathbb{1}_A$ of A, defined by
$$\forall \omega \in \Omega : \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{if } \omega \notin A, \\ 1 & \text{if } \omega \in A. \end{cases}$$

Then $\mathbb{1}_A$ is a random variable. Ineed, we have:
$$\{\mathbb{1}_A \leq a\} = \begin{cases} \emptyset & \text{if } a > 0, \\ A^C & \text{if } 0 \leq a < 1, \\ \Omega & \text{if } a \geq 1, \end{cases}$$

and $\emptyset$, A^C, Ω are three elements of $\mathcal{F}$. Furthermore, writing $X = \mathbb{1}_A$, we have

$$\mathbb{P}[X=0] = 1 - \mathbb{P}[A] \quad \text{and} \quad \mathbb{P}[X=1] = \mathbb{P}[A].$$

Therefore, $\mathbb{1}_A$ is a Bernoulli random variable with parameter $\mathbb{P}[A]$. Hence,

$$\mathbb{E}[\mathbb{1}_A] = \mathbb{P}[A].$$

Proposition: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable with values in W (finite or countable) almost surely. For every $\phi : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbb{E}[\phi(X)] = \sum_{w \in W} \phi(x) \cdot \mathbb{P}[X = x],$$

provided the sum is well defined.

1.3 Expectation of a Continuous Random Variable

Proposition: Let X be a continuous random variable with density f. Then, we have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx,$$

provided the integral is well defined.

Example 1 (Uniform): We have

$$\mathbb{E}[X] = \frac{1}{b-a} \int_a^b x\, dx = \frac{1}{b-a} \cdot (\frac{1}{2}b^2 - \frac{1}{2}a^2).$$

Therefore,

$$\mathbb{E}[X] = \frac{a+b}{2}.$$

Example 2 (Exponential): By integration by parts, we have

$$\mathbb{E}[X] = \int_0^{\infty} x\lambda e^{-\lambda x}\, dx = [-xe^{-\lambda x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\, dx.$$

Therefore,

$$\mathbb{E}[X] = \frac{1}{\lambda}.$$

Proposition: Let X be a continuous random variable with density f. Let $\phi : \mathbb{R} \to \mathbb{R}$ be such that $\phi(X)$ is a random variable. Then we have

$$\mathbb{E}[\phi(X)] = \int_{\infty}^{\infty} \phi(x) f(x)\, dx,$$

provided the integral is well defined.

1.4 Calculus

Theorem (Linearity of the expectation): Let $X, Y : \Omega \to \mathbb{R}$ be random variables, let $\lambda \in \mathbb{R}$. Provided the expectations are well defined, we have:

1. $\mathbb{E}[\lambda \cdot X] = \lambda \cdot \mathbb{E}[X]$
2. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Application 1 (Binomial): Let S be a binomial random variable with parameters n and p. By definition we have

$$\mathbb{E}[S] = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}.$$

By linearity we have $\mathbb{E}[S_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]$, where $X_1, ..., X_n$ are n i.i.d. Bernoulli random variables. Using that $\mathbb{E}[X_i] = p$ for every p, we deduce directly

$$\mathbb{E}[S] = \mathbb{E}[S_n] = np.$$

Application 2 (Normal): By Proposition we have (with $Y \sim \mathcal{N}(0, 1)$)

$$\mathbb{E}[X] = \mathbb{E}[m + \sigma \cdot Y] = m + \sigma \cdot \mathbb{E}[Y],$$

hence it suffices to compute the expectation of Y. Writing $f_{0,1}$ for the density of Y, we have

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} x \cdot f_{0,1}(x)\, dx = 0$$

because $x \cdot f_{0,1}(x)$ is an odd function. Finally, we obtain

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

Theorem: Let X, Y be two random variables. If X and Y are *independent,* then

$$\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

1.5 Characterizations via Expectations

1.5.1 Density

Proposition: Let X be a random variable. Let $f : \mathbb{R} \to \mathbb{R}_+$ such that $\int_{-\infty}^{\infty} f(x)\, dx = 1$. Then the following are equivalent:

1. X is continuous with density f.
2. For every function $\phi : \mathbb{R} \to \mathbb{R}$ measurable bounded,

$$\mathbb{E}[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f(x)\, dx.$$

1.5.2 Independence

Theorem: Let X, Y be 2 discrete random variables. Then the following two are equivalent:

1. X, Y are independent.
2. For every $\phi : \mathbb{R} \to \mathbb{R}$, $\psi : \mathbb{R} \to \mathbb{R}$ (measurable) bounded

$$\mathbb{E}[\phi(X)\psi(Y)] = \mathbb{E}[\phi(X)]\mathbb{E}[\psi(Y)].$$