WuS - Lecture Notes Week 4

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

June 21, 2022

1 Discrete and Continuous Random Variables

1.1 Discontinuity & Continuity Points of F

We have seen that the distribution function $F = F_X$ of a random variable X is always *rigt continuous.* What about left continuous?

Example: For a Bernoulli random variable $X \sim \text{Ber}(p)$ with $p < 1$, we have $F_X(-h) = 0$ for every $h > 0$, but $F_X(0) = 1 - p \neq 0$. Therefore, F_X is not left continuous at 0, i.e.

$$\lim_{h \downarrow 0} F_X(-h) = 0 \neq F_X(0).$$

The following proposition gives an interpretation of the limit

$$F(a-) := \lim_{h \downarrow 0} F(a-h)$$

at a given point a for a general distribution function.

Proposition (probability of a given value): Let $X : \Omega \to \mathbb{R}$ be a random variable with distribution function F. Then for every a in $\mathbb{R}$ we have

$$\mathbb{P}[X = a] = F(a) - F(a-).$$

We give the following interpretation of the above introduce proposition. Fix some $a \in \mathbb{R}$. Then:

- If F is not continuous at a point $a \in \mathbb{R}$, then the "jump size" $F(a) - F(a-)$ is equal to the probability that $X = a$.
- If F is continuous at a point $a \in \mathbb{R}$, then $\mathbb{P}[X = a] = 0$.

1.2 Almost Sure Events

Definition: Let $A \in \mathcal{F}$ be an event. We say that A occurs **almost surely (a.s.)** if

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

Remark: This notion can be extended to any set $A \subset \Omega$: We say that A occurs almost surely if there exists an event $A' \in \mathcal{F}$ such that $A' \subset A$ and $\mathbb{P}[A'] = 1$.

1.3 Discrete Random Variables

Definition (Discrete Random Variables): A random variable $X : \Omega \to \mathbb{R}$ is said to be **discrete** if there exists some set $W \subset \mathbb{R}$ finite or countable such that

$$X \in W \quad a.s.$$

Remark: If the sample space Ω is finite or countable, then every random variable $X : \Omega \to \mathbb{R}$ is discrete.

Definition: Let X be a discrete random variable taking some values in some finite or countable set $W \subset \mathbb{R}$. The **distribution of** X is the sequence of numbers $(p(x))_{x \in W}$ defined by

$$\forall x \in W : p(x) := \mathbb{P}[X = x].$$

Proposition: The distribution $(p(x))_{w \in W}$ of a discrete random variable satisfies

$$\sum_{x \in W} p(x) = 1.$$

Example: Consider the random variable defined by

$$\forall \omega \in \Omega : X(\omega) := \begin{cases} -1, & \text{if } \omega = 1, 2, 3, \\ 0, & \text{if } \omega = 4, \\ 2, & \text{if } \omega = 5, 6. \end{cases}$$

Then X takes values in $W = \{-1, 0, 2\}$ almost surely and its distribution is given by

$$p(-1) = \frac{1}{2}, \quad p(0) = \frac{1}{6}, \quad p(2) = \frac{1}{3}.$$

Remark: Conversely, if we are given a sequence of numbers $(p(x))_{x \in W}$ with values in $[0, 1]$ and such that $\sum_{x \in W} p(x) = 1$, then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable X with associated distribution $(p(x))$. This observation is important in practice, it allows us to write: "*Let X be a discrete random variable with distribution $(p(x))_{x \in W}$.*"

1.3.1 From p to F_X

Proposition: Let X be a discrete random variable with values in a finite or countable set W almost surely, and distribution p. Then the distribution function of X is given by

$$\forall x \in \mathbb{R} : F_X(x) = \sum_{y \leq x,\, y \in W} p(y).$$

1.3.2 From F_X to p

Given a discrete random variable X. A random variable with a piecewise cosntant function F is discrete and W and p are given by:

- $W = \{\text{positions of the jumps of } F_X\}$
- $p(x) =$ "height of the jump" at $x \in W$

1.4 Examples of Discrete Random Variables

1.4.1 Bernoulli Distribution

Definition (Bernoulli): Let $0 \leq p \leq 1$. A random variable X is said to be a **Bernoulli random variable with parameter** p if it takes values in $W = \{0, 1\}$ and

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

In that case, we write $X \sim \text{Ber}(p)$.

1.4.2 Binomial Distribution

Definition (Binomial): Let $0 \leq p \leq 1$, let $n \in \mathbb{N}$. A random variable X is said to be a **binomial random variable with parameters** n **and** p if it takes values in $W = \{0, ..., n$ and

$$\forall k \in \{0, ..., n\} : \mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

In that case we write $X \sim \text{Bin}(n, p)$. This appears in applications when we consider the number of successes in a repetition of Bernoulli experiments.

> **Proposition (Sum of independent Bernoulli and binomial):** Let $0 \leq p \leq 1$, let $n \in \mathbb{N}$. Let $X_1, ..., X_n$ be independent Bernoulli random variables with parameter p. Then
>
> $$S_n := X_1 + \cdots + X_n$$
>
> is a binomial random variable with parameter n and p.

Remark: In particular, the distribution Bin(1, p) is the same as the distribution Ber(p). One can also check that if $X \sim \text{Bin}(m, p)$ and $Y \sim \text{Bin}(n, p)$ and X, Y are independent, then $X + Y = \text{Bin}(m+n, p)$.

1.4.3 Geometric Distribution

Definition (Geometric): Let $0 \leq p \leq 1$. A random variable X is said to be a **geometric random variable with parameter** p if it takes values in $W = \mathbb{N} \setminus \{0\}$ and

$$\forall k \in \mathbb{N} \setminus \{0\} : \mathbb{P}[X = k] = (1-p)^{k-1} \cdot p.$$

In that case, we write $X \sim \text{Geom}(p)$.

The geometric random variable appears naturally as the first success in an infinite sequence of Bernoulli experiments with parameter p. This is formalized by the following proposition.

> **Proposition:** Let $X_1, X_2, ...$ be a sequence of infinitely many independent Bernoulli r.v.'s with parameter p. Then
>
> $$T := \min\{n \geq 1 : X_n = 1\}$$
>
> is a geometric random variable with parameter p.

> **Proposition:** Let $T \sim \text{Geom}(p)$ for some $0 < p < 1$. Then
>
> $$\forall n \geq 0, \forall k \geq 1 : \mathbb{P}[T \geq n + k \,|\, T > n] = \mathbb{P}[T \geq k].$$

1.4.4 Poisson Distribution

Definition: Let $\lambda > 0$ be a positive real number. A random variable X is said to be a **Poisson random variable with parameter** λ if it takes values in $W = \mathbb{N}$ and

$$\forall k \in \mathbb{N} : \mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}.$$

In this case, we write $X \sim \text{Poisson}(\lambda)$.

The Poisson distribution appears naturally as an approximation of a binomial distribution when the parameter n is large and the parameter p is small, as stated formally in the following proposition.

> **Proposition (Poisson approximation of the binomial):** Let $\lambda > 0$. For every $n \geq 1$, consider a random variable $X_n \sim \text{Bin}(n, \frac{\lambda}{n})$. Then
>
> $$\forall k \in \mathbb{N} : \lim_{n \to \infty} \mathbb{P}[X_n = k] = \mathbb{P}[N = k],$$
>
> where N is a Poisson random variable with parameter λ.

1.5 Continuous Random Variables

Definition (Continuous Random Variables): A random variable $X : \Omega \to \mathbb{R}$ is said to be **continuous** if its distribution function F_X can be written as

$$F_X(a) = \int_{-\infty}^{a} f(x)\,dx \quad \text{for all } a \in \mathbb{R}$$

for some nonnegative function $f : \mathbb{R} \to \mathbb{R}_+$, called the **density** of X.

Intuition: $f(x)\,dx$ represents the probability that X takes a value in the infinitesimal interval $[x,\, x+dx]$.

Proposition: The density f of a random variable satisfies

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1.$$

1.5.1 From f to F_X

Let X be a continuous random variable with density f. By definition, the distribution function F_X can be calculated as the integral

$$F_X(x) = \int_{-\infty}^{x} f(y)\,dy.$$

1.5.2 From F_X to f

Since one goes from f to F_X by integrating, it is natural to expect that the reverse operation is to take the derivative. This is in general the case, provided F_X is regular enough. The following theorem will be useful in applications to calculate densities.

Theorem: Let X be a random variable. Assume that the distribution function F_X is continuous and piecewise $\mathcal{C}^1$, i.e. that there exists $x_0 = -\infty < x_1 < \cdots < x_{n-1} < x_n = +\infty$ such that F_X is $\mathcal{C}^1$ on every interval $(x_i,\, x_{i+1})$. Then X is a continuous random variable and a density f can be constructed by defining

$$\forall x \in (x_i,\, x_{i+1}) : f(x) = F_X'(x)$$

and setting arbitrary values at $x_1, \ldots, x_{n-1}$.