10 Random variables

The fundamental idea of modern probability theory is to model one or more random experiments as a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. The sets $A \in \mathcal{A}$ are called events. In most cases, the events of Ω are not observed directly. Rather, the observations are aspects of the single experiments that are coded as measurable maps from Ω to a space of possible observations. In short, every random observation (the technical term is random variable) is a measurable map. The probabilities of the possible random observations will be described in terms of the distribution of the corresponding random variable, which is the image measure of $\mathbf{P}$ under X. Hence we have to develop a calculus to determine the distributions of, for example, sums of random variables.

Definition 10.1 (Random variables)**.** Let $(\Omega', \mathcal{A}')$ be a measurable space and let $X : \Omega \to \Omega'$ be measurable.

(i) X is called a **random variable** with values in $(\Omega', \mathcal{A}')$. If $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then X is called a real random variable.

(ii) For $A' \in \mathcal{A}'$, we denote $[X \in A'] := X^{-1}(A')$ and $\mathbf{P}[X \in A'] := \mathbf{P}[X^{-1}(A')]$. In particular, we let $[X \geq 0] := X^{-1}([0, +\infty[)$ and define $[X \leq b]$ similarly and so on. □

Definition 10.2 (Distributions)**.** Let $(\Omega', \mathcal{A}')$ be a measurable space and let $X, X_i : \Omega \to \Omega'$ be random variables, $i \in I$.

(i) The probability measure $\mathbf{P}_X := \mathbf{P} \circ X^{-1}$ is called the **distribution** of X. We write $X \sim \mu$ if $\mu = \mathbf{P}_X$ and say that X has distribution μ.

(ii) The family $(X_i)_{i \in I}$ is called **identically distributed** if $\mathbf{P}_{X_i} = \mathbf{P}_{X_j}$ for all $i, j \in I$.

(iii) If X is a real random variable, then the map $F_X : x \mapsto \mathbf{P}[X \leq x]$ is called the **distribution function** of X (or, more accurately, of $\mathbf{P}_X$). □

Property 10.3. For any distribution function F, there exists a real random variable X with $F_X = F$.

Proof.

Exercise 10.1. □

Example 10.4. We present some prominent distributions of real random variables X. These are some of the most important distributions in probability theory, and we will come back to these examples in many places.

(i) Let $p \in [0, 1]$ and $\mathbf{P}[X = 1] = 1 - \mathbf{P}[X = 0] = p$. Then $\mathbf{P}_X =: \mathrm{Ber}_p$ is called the **Bernoulli distribution** with parameter p; formally

$$\mathrm{Ber}_p = (1 - p)\delta_0 + p\delta_1.$$

(ii) Let $p \in [0,1]$ and $\mathbf{P}_X =: \mathrm{Ber}_p$. The distribution $\mathbf{P}_Y =: \mathrm{Rad}_p$ of $Y = 2X - 1$ is called the **Rademacher distribution** with parameter p; formally

$$\mathrm{Rad}_p = (1-p)\delta_{-1} + p\delta_1.$$

(iii) Let $p \in [0,1]$, $n \in \mathbb{N}$ and

$$\mathbf{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k},$$

for any $k = 0, \dots, n$. Then $\mathbf{P}_X =: \mathrm{B}_{n,p}$ is called the **binomial distribution** with parameters n and p; formally

$$\mathrm{B}_{n,p} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k.$$

(iv) Let $p \in]0,1]$ and

$$\mathbf{P}[X = n] = p(1-p)^n,$$

for any $n \in \mathbb{N}_0$. Then $\mathbf{P}_X =: \gamma_p$ is called the **geometric distribution** with parameter p; formally

$$\gamma_p = \sum_{k=0}^{+\infty} p(1-p)^n \delta_n.$$

(v) Let $\lambda \geq 0$ and

$$\mathbf{P}[X = n] = e^{-\lambda} \frac{\lambda^n}{n!},$$

for any $n \in \mathbb{N}_0$. Then $\mathbf{P}_X =: \mathrm{Poi}_\lambda$ is called the **Poisson distribution** with parameter λ.

(vi) Let $\theta > 0$ and

$$\mathbf{P}[X \leq x] = \begin{cases} \int_0^x \theta e^{-\theta t} dt & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Then $\mathbf{P}_X =: \exp_\theta$ is called the **exponential distribution** with parameter θ.

(vii) Let $\mu \in \mathbb{R}$, $\sigma^2 > 0$ and

$$\mathbf{P}[X \leq x] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

for any $x \in \mathbb{R}$. Then $\mathbf{P}_X =: \mathcal{N}_{\mu,\sigma^2}$ is called the **normal distribution** with parameters μ and σ^2. In particular, $\mathcal{N}_{0,1}$ is called the standard normal distribution.