1. Introduction and some structural properties of empirical measures.

Many standard procedures in statistics are based on a random sample $x_1,\dots,x_n$ of i.i.d. observations, i.e., it is assumed that observations (or measurements) occur as realizations (or values) $x_i = \xi_i(\omega)$ in some sample space X of a sequence of independent and identically distributed (i.i.d.) random elements $\xi_1,\dots,\xi_n$ defined on some basic probability space (p-space for short) $(\Omega,\mathcal{F},\mathbb{P})$; here ξ is called a RANDOM ELEMENT in X whenever there exists a $(\Omega,\mathcal{F},\mathbb{P})$ such that $\xi: \Omega \to X$ is $\mathcal{F},\mathcal{B}$-measurable for an appropriate σ-algebra $\mathcal{B}$ in X, in which case the law $\mu \equiv \mathcal{L}\{\xi\}$ of ξ is a well defined p-measure on $\mathcal{B}$ ($\mu(B) = \mathbb{P}(\{\omega\in\Omega: \xi(\omega)\in B\}) \equiv \mathbb{P}(\xi\in B)$ for short, $B\in\mathcal{B}$).

In classical situations, the sample space X is usually the k-dimensional Euclidean space $\mathbb{R}^k$, $k\geq 1$, with the Borel σ-algebra $\mathcal{B}_k$. In the present notes, if not stated otherwise, the sample space X is always an arbitrary measurable space $(X,\mathcal{B})$.

Given then i.i.d. random elements ξ_i in $X = (X,\mathcal{B})$ with (common) law μ on $\mathcal{B}$ we can associate with each (sample size) n the so-called EMPIRICAL MEASURE

$$(1) \qquad \mu_n := \frac{1}{n}(\varepsilon_{\xi_1} + \dots + \varepsilon_{\xi_n}) \text{ on } \mathcal{B},$$

where

$$\varepsilon_x(B) := \begin{cases} 1 \text{ if } x\in B \\ 0 \text{ if } x\notin B \end{cases}, \quad B\in\mathcal{B}.$$

In other words, given the first n observations $x_i = \xi_i(\omega)$, $i=1,\dots,n$, $\mu_n(B) \equiv \mu_n(B,\omega)$ is the average number of the first n x_i's falling into B. (The notation $\mu_n(\cdot,\omega)$ should call attention to the fact that μ_n is a random p-measure on $\mathcal{B}$.)

μ_n may be viewed as the statistical picture of μ and we are thus interested in the connection between μ_n and μ, especially when n tends to infinity.