3. Distributions

In this subsection, we introduce the notion of distributions.

3.1. **Distributions on $\mathbb{R}$.**

Definition 3.1. Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ be a probability space. The distribution of $\mathbb{P}$ is a function F on $\mathbb{R}$ defined by $F(x) = \mathbb{P}((-\infty, x])$.

It follows immediately from the above definition that F is non-decreasing, right-continuous and satisfies

$$F(-\infty) := \lim_{x \to -\infty} F(x) = 0, \quad F(\infty) := \lim_{x \to \infty} F(x) = 1. \tag{3.1}$$

Definition 3.2. A non-decreasing, right-continuous function satisfying (3.1) is called a distribution function.

Theorem 3.1. *For any distribution function F, there is a unique probability $\mathbb{P}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that F is the distribution of $\mathbb{P}$.*

Proof. The proof is based on Carathéodory's extension theorem. We display the version in probability as follows but skip its proof.

Theorem 3.2 (Carathéodory's extension theorem)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be a finite probability space. If $\mathbb{P}_0$ is σ-additive, then there is a probability space $(\Omega, \sigma(\mathcal{F}), \mathbb{P})$ such that $\mathbb{P}(A) = \mathbb{P}_0(A)$ for $A \in \mathcal{F}$.*

Consider the following class.

$$\mathcal{C}_0 = \{(a, b] | -\infty \le a \le b \le \infty\}, \quad \mathcal{C} = \{\text{finite disjoint unions of sets in } \mathcal{C}_0\},$$

where $(a, a] := \emptyset$ for $a \in \mathbb{R}$ and $(a, \infty] := (a, \infty)$ for $a \in \mathbb{R} \cup \{-\infty\}$. For $A = \bigcup_{i=1}^n (a_i, b_i] \in \mathcal{C}$, define $\mathbb{P}_0(A) = \sum_{i=1}^n (F(b_i) - F(a_i))$, where $F(-\infty) := 0$ and $F(\infty) := 1$. One can prove that $\mathcal{C}$ is a field, $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$ and $\mathbb{P}_0$ is a finite probability on $(\mathbb{R}, \mathcal{C})$. By Theorem 3.2, it remains to show that $\mathbb{P}_0$ is σ-additive or, equivalently, $\mathbb{P}_0$ is continuous at $\emptyset$. Let A_n be a sequence in $\mathcal{C}$ satisfying $A_n \supset A_{n+1}$ and $\bigcap_n A_n = \emptyset$. Given $\epsilon > 0$, one may choose $N > 0$ such that $\mathbb{P}_0((-N, N]) \ge 1 - \epsilon/2$. By the right continuity of F, one may choose $B_n \in \mathcal{C}$ such that $\overline{B}_n \subset A_n \cap (-N, N]$ and $\mathbb{P}_0(B_n) \ge \mathbb{P}_0(A_n \cap (-N, N]) - \epsilon/2^{n+1}$ for $n \ge 1$. Set $C_n = \bigcap_{i=1}^n \overline{B}_i$. Since $\bigcap_n A_n = \emptyset$, $\bigcap_n C_n = \bigcap_n \overline{B}_n = \emptyset$. Note that $\overline{B}_n$ is compact. By the nested set property, $\bigcap_{n=1}^{n_0} C_n = \bigcap_{n=1}^{n_0} \overline{B}_n = \emptyset$ for some n_0 and, hence, $\bigcap_{n=1}^{n_0} B_n = \emptyset$. This implies

$$\begin{aligned}
\mathbb{P}_0(A_{n_0}) &\le \frac{\epsilon}{2} + \mathbb{P}_0(A_{n_0} \cap (-N, N]) = \frac{\epsilon}{2} + \mathbb{P}_0\left((A_{n_0} \cap (-N, N]) \setminus \bigcap_{n=1}^{n_0} B_n \right) \\
&\le \frac{\epsilon}{2} + \sum_{n=1}^{n_0} \mathbb{P}_0((A_n \cap (-N, N]) \setminus B_n) \le \frac{\epsilon}{2} + \sum_{n=1}^{n_0} \frac{\epsilon}{2^{n+1}} < \epsilon.
\end{aligned}$$

This proves that $\mathbb{P}_0$ is continuous at $\emptyset$. □

Remark 3.1. The probability corresponding to a distribution function is called a Lebesgue-Stieltjes probability.

3.2. **Distributions on $\mathbb{R}^n$.**

Definition 3.3. Fix $n \ge 1$ and let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P})$ be a probability space. The distribution of $\mathbb{P}$ is defined by $F(x_1, ..., x_n) = \mathbb{P}(\prod_{i=1}^n (-\infty, x_i])$.

For $1 \le i \le n$ and $-\infty \le a_i < b_i \le \infty$, set

$$\Delta^i_{a_i,b_i} F(x_1, ..., x_n) = F(x_1, ..., x_{i-1}, b_i, x_{i+1}, ..., x_n) - F(x_1, ..., x_{i-1}, a_i, x_{i+1}, ..., x_n).$$

Clearly, one has

$$\Delta^i_{a_i,b_i} \Delta^j_{a_j,b_j} F(x_1, ..., x_n) = \Delta^j_{a_j,b_j} \Delta^i_{a_i,b_i} F(x_1, ..., x_n).$$

Exercise 3.1. Let $\mathbb{P}$ be a probability on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with distribution F.

(1) Show that $\mathbb{P}(\prod_{i=1}^n (a_i, b_i]) = \Delta^1_{a_1,b_1} \cdots \Delta^n_{a_n,b_n} F(x_1, ..., x_n)$. (*Hint:* Prove that if $a_i = -\infty$ for $k < i \le n$, then $\mathbb{P}(\prod_{i=1}^n (a_i, b_i]) = \Delta^1_{a_1,b_1} \cdots \Delta^k_{a_k,b_k} F(x_1, ..., x_k, b_{k+1}, ..., b_n)$.
(2) We write $y_m = (y_{m,1}, ..., y_{m,n}) \downarrow x = (x_1, ..., x_n)$ if $y_{m,i} \downarrow x_i$ for $1 \le i \le n$. Prove that if $y_m \downarrow x$, then $F(y_m) \downarrow F(x)$.
(3) Prove that $F(x_1, ..., x_n) \to 1$ if $x_1, ..., x_n \to \infty$ and $F(x_1, ..., x_n) \to 0$ if $x_i \to -\infty$ for some $1 \le i \le n$.

Definition 3.4. A non-negative function F defined on $\mathbb{R}^n$ is called a distribution function if

(1) For $a_i < b_i$ and $1 \le i \le n$, $\Delta^1_{a_1,b_1} \cdots \Delta^n_{a_n,b_n} F(x_1, ..., x_n) \ge 0$.
(2) $F(x_1, ..., x_n) \to 1$ if $x_1, ..., x_n \to \infty$.
(3) $F(x_1, ..., x_n) \to 0$ if $x_i \to -\infty$ for some $1 \le i \le n$.
(4) $F(y_m) \downarrow F(x)$ if $y_m \downarrow x$.

Remark 3.2. Note that Definition 3.4(4) implies that if $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ satisfy $x_i \le y_i$ for all $1 \le i \le n$, then $F(x) \le F(y)$. As a consequence, for $i \ne j$,

$$\lim_{x_j \to \infty} \lim_{x_i \to \infty} F(x_1, ..., x_n) = \lim_{x_i \to \infty} \lim_{x_j \to \infty} F(x_1, ..., x_n).$$

Theorem 3.3. *For any distribution function F on $\mathbb{R}^n$, there is a unique probability $\mathbb{P}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that F is the distribution of $\mathbb{P}$.*

Proof. Let $\mathcal{C}_0$ be the following class

$$\mathcal{C}_0 = \left\{ \prod_{i=1}^n (a_i, b_i] \middle| a_i, b_i \in \mathbb{R} \cup \{\pm\infty\},\ 1 \le i \le n, n \ge 1 \right\},$$

where $(a, a] := \emptyset$ for $a \in \mathbb{R}$ and $(a, \infty] := (a, \infty)$ for $a \in \mathbb{R} \cup \{-\infty\}$, and $\mathcal{C}$ be the class consisting of finite unions of mutually disjoint sets in $\mathcal{C}_0$.

Exercise 3.2. Prove that $\mathcal{C}$ is a field.

Let $x = (x_1, ..., x_n) \in (\mathbb{R} \cup \{\pm\infty\})^n$. For $x_{i_1} = \cdots = x_{i_k} = \infty$, we define

$$F(x) := \lim_{y_{i_1}, ..., y_{i_k} \to \infty} F(y),$$

where $y = (y_1, ..., y_n)$ and $y_j = x_j$ for $j \notin \{i_1, ..., i_k\}$. If $x_i = -\infty$ for some $1 \le i \le n$, we set $F(x) = 0$. For $A = \prod_{i=1}^n (a_i, b_i] \in \mathcal{C}_0$, define

$$\mathbb{P}(A) := \Delta^1_{a_1,b_1} \cdots \Delta^n_{a_n,b_n} F(x_1, ..., x_n)$$

and, for $B = \bigcup_{i=1}^n A_i \in \mathcal{C}$ with $A_i \in \mathcal{C}_0$ and $A_i \cap A_j = \emptyset$ if $i \ne j$, define

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(A_i). \tag{3.2}$$

Exercise 3.3. Prove that $\mathbb{P}$ is well-defined.

It is easy to see from (3.2) that $\mathbb{P}$ is a finite probability on $(\mathbb{R}^n, \mathcal{C})$. Note that $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R}^n)$. By Theorem 3.2, it remains to show that $\mathbb{P}$ is continuous at $\emptyset$. This can be proved in a similar way as the one-dimensional case and is left for the reader. □

Exercise 3.4. Fix $n \in \mathbb{N}$ and let ρ be a permutation of $\{1, 2, ..., n\}$. For $x = (x_1, ..., x_n) \in \mathbb{R}^n$ and $A \subset \mathbb{R}^n$, define $\rho(x) = (x_{\rho(1)}, ..., x_{\rho(n)})$ and $\rho(A) = \{\rho(x)|x \in A\}$. Let F be a distribution function on $\mathbb{R}^n$ and set $G(x) = F(\rho(x))$.

(1) Prove that G is a distribution function on $\mathbb{R}^n$.
(2) Let $\mathbb{P}_F, \mathbb{P}_G$ be the probabilities on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with distributions F, G. Show that $\mathbb{P}_G(B) = \mathbb{P}_F(\rho(B))$.

Hint: The $\pi - \lambda$ lemma.

Exercise 3.5. Let F be a distribution function on $\mathbb{R}^n$ and set

$$H(x_1, ..., x_{n-1}) = \lim_{x_n \to \infty} F(x_1, ..., x_n).$$

Show that H is a distribution function on $\mathbb{R}^{n-1}$. Let $\mathbb{P}_F, \mathbb{P}_H$ be probabilities on $\mathcal{B}(\mathbb{R}^n), \mathcal{B}(\mathbb{R}^{n-1})$ with distributions F, H. Show that $\mathbb{P}_H(A) = \mathbb{P}_F(A \times \mathbb{R})$ for $A \in \mathcal{B}(\mathbb{R}^{n-1})$.

3.3. **Distributions on $\mathbb{R}^\infty$.** Recall the notation $\mathcal{C}(B) = \{x = (x_1, x_2, ...) \in \mathbb{R}^\infty | (x_1, ..., x_n) \in B\}$ for $B \in \mathcal{B}(\mathbb{R}^n)$. Note that if $B_1 \in \mathcal{B}(\mathbb{R}^m)$, $B_2 \in \mathcal{B}(\mathbb{R}^n)$ with $m \le n$ and $\mathcal{C}(B_1) = \mathcal{C}(B_2)$, then $B_2 = B_1 \times \mathbb{R}^{n-m}$.

Theorem 3.4 (Kolmogorov's extension theorem on $\mathbb{R}^\infty$)**.** *For $n \in \mathbb{N}$, let $\mathbb{P}_n$ be a probability on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Suppose that $\mathbb{P}_n$ satisfies the following consistency property*

$$\mathbb{P}_{n+1}(B \times \mathbb{R}) = \mathbb{P}_n(B), \quad \forall B \in \mathcal{B}(\mathbb{R}^n),\ n \in \mathbb{N}.$$

Then, there is a unique probability $\mathbb{P}$ on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ such that

$$\mathbb{P}(\mathcal{C}(B)) = \mathbb{P}_n(B), \quad \forall B \in \mathcal{B}(\mathbb{R}^n),\ n \in \mathbb{N}.$$

Proof. Let $\mathcal{F} = \{\mathcal{C}(B)|B \in \mathcal{B}(\mathbb{R}^n), n \ge 1\}$ and define $\mathbb{P}(\mathcal{C}(B)) = \mathbb{P}_n(B)$. Obviously, $\mathcal{F}$ is a field. Note that if $\mathcal{C}(B_1) = \mathcal{C}(B_2)$ with $B_1 \in \mathcal{B}(\mathbb{R}^m)$, $B_2 \in \mathcal{B}(\mathbb{R}^n)$ and $m \le n$, then $B_2 = B_1 \times \mathbb{R}^{n-m}$. This implies

$$\mathbb{P}(\mathcal{C}(B_1)) = \mathbb{P}_m(B_1) = \mathbb{P}_n(B_1 \times \mathbb{R}^{n-m}) = \mathbb{P}_n(B_2) = \mathbb{P}(\mathcal{C}(B_2)),$$

which proves that $\mathbb{P}$ is well-defined. Further, for $m < n$ and $A \in \mathcal{B}(\mathbb{R}^m), B \in \mathcal{B}(\mathbb{R}^n)$, if $\mathcal{C}(A) \cap \mathcal{C}(B) = \emptyset$, then $(A \times \mathbb{R}^{n-m}) \cap B = \emptyset$ and this yields

$$\begin{aligned}
\mathbb{P}(\mathcal{C}(A) \cup \mathcal{C}(B)) &= \mathbb{P}(\mathcal{C}((A \times \mathbb{R}^{n-m}) \cup B)) = \mathbb{P}_n((A \times \mathbb{R}^{n-m}) \cup B) \\
&= \mathbb{P}_n(A \times \mathbb{R}^{n-m}) + \mathbb{P}_n(B) = \mathbb{P}_m(A) + \mathbb{P}_n(B) \\
&= \mathbb{P}(\mathcal{C}(A)) + \mathbb{P}(\mathcal{C}(B)).
\end{aligned}$$

Hence, $\mathbb{P}$ is a finite probability on $(\mathbb{R}^\infty, \mathcal{F})$.

Note that $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R}^\infty)$. By Carathéodory's extension theorem, it remains to show the continuity of $\mathbb{P}$ at $\emptyset$. Let $\mathcal{C}(B_n)$ be a decreasing sequence in $\mathcal{F}$ satisfying $\bigcap_{n=1}^\infty \mathcal{C}(B_n) = \emptyset$. Without loss of generality, we may assume that $B_n \in \mathcal{B}(\mathbb{R}^n)$. Assume the inverse that $\mathbb{P}(\mathcal{C}(B_n)) \ge \epsilon > 0$.

Exercise 3.6. Let $\mathbb{P}$ be a probability on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Prove that, for any $S \in \mathcal{B}(\mathbb{R}^n)$ and $\epsilon > 0$, there are an open set $A \supset S$ and a closed set $B \subset S$ such that

$$\mathbb{P}(A \setminus S) < \epsilon, \quad \mathbb{P}(S \setminus B) < \epsilon.$$

Exercise 3.7. Let $\mathcal{F}$ be a field over Ω and $\mathbb{P}$ be a probability on $(\Omega, \sigma(\mathcal{F}))$. Show that, for any $A \in \sigma(\mathcal{F})$ and $\epsilon > 0$, there is $B \in \mathcal{F}$ such that $\mathbb{P}(A \Delta B) < \epsilon$, where $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

By Exercise 3.6, we may choose, for $n \geq 1$, a compact set $K_n \subset B_n$ such that $\mathbb{P}_n(B_n \setminus K_n) \leq \epsilon 2^{-n-1}$ or, equivalently, $\mathbb{P}(\mathcal{C}(B_n) \setminus \mathcal{C}(K_n)) \leq \epsilon 2^{-n-1}$. Set $A_n = \bigcap_{i=1}^n (K_i \times \mathbb{R}^{n-i})$. Clearly, $B_n \times \mathbb{R}^i \supset B_{n+i}$ for all $i \geq 1$. As a result, one has

$$\begin{aligned}\mathbb{P}(\mathcal{C}(B_n) \setminus \mathcal{C}(A_n)) &= \mathbb{P}_n(B_n \setminus A_n) \leq \sum_{i=1}^n \mathbb{P}_n(B_n \setminus (K_i \times \mathbb{R}^{n-i})) \\ &\leq \sum_{i=1}^n \mathbb{P}_n((B_i \times \mathbb{R}^{n-i}) \setminus (K_i \times \mathbb{R}^{n-i})) = \sum_{i=1}^n \mathbb{P}_i(B_i \setminus K_i) < \epsilon/2.\end{aligned}$$

Hence, $\mathbb{P}(\mathcal{C}(A_n)) > \epsilon/2$ and this implies $\mathcal{C}(A_n)$ is non-empty.

Next, for $n \in \mathbb{N}$, let $x_n = (x_{n,1}, x_{n,2}, ...) \in \mathcal{C}(A_n)$. Note that $\mathcal{C}(A_n) \supset \mathcal{C}(A_{n+1})$. This implies $\{(x_{i,1}, ..., x_{i,n}) | i \geq n\} \subset A_n$ for $n \geq 1$. For $n = 1$, since A_1 is compact, we may choose a subsequence $(k_{1,m})_{m=1}^\infty$ of $\mathbb{N}$ and $y_1 \in A_1$ such that $x_{k_{1,m},1} \to y_1$. Inductively, one may select a subsequence $(k_{i+1,m})_{m=1}^\infty$ of $(k_{i,m})_{m=1}^\infty$ and $y_{i+1} \in \mathbb{R}$ such that $x_{k_{i+1,m},i+1} \to y_{i+1}$. Set $n_i = k_{i,i}$. Obviously, one has $(x_{n_i,1}, ..., x_{n_i,m}) \to (y_1, ..., y_m)$ as $i \to \infty$ for any $m \geq 1$. By the compactness of A_m, $(y_1, ..., y_m) \in A_m$. Consequently, we obtain $\mathcal{C}(\{(y_1, ..., y_m)\}) \subset \mathcal{C}(A_m) \subset \mathcal{C}(B_m)$ for all $m \geq 1$ and this yields $(y_1, y_2, ...) \in \bigcap_{n=1}^\infty \mathcal{C}(B_n)$, which contradicts $\bigcap_{n=1}^\infty \mathcal{C}(B_n) = \emptyset$. $\square$

Definition 3.5. Let $\mathbb{P}$ be a probability on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$. For $n \geq 1$, the n-dimensional distribution of $\mathbb{P}$ is defined by

$$F_n(x_1, ..., x_n) = \mathbb{P}\left(\prod_{i=1}^n (-\infty, x_i] \times \mathbb{R}^\infty\right).$$

Remark 3.3. Note that the family of finitely dimensional distributions of $\mathbb{P}$ satisfying the following consistency property.

$$\lim_{y \to \infty} F_{n+1}(x_1, ..., x_n, y) = F_n(x_1, ..., x_n). \tag{3.3}$$

The following theorem is a simple corollary of Theorems 3.3, 3.4 and the π-λ lemma.

Theorem 3.5. *For $n \geq 1$, let F_n be a distribution function on $\mathbb{R}^n$. If the family $\{F_n | n = 1, 2, ...\}$ satisfies the consistency property in* (3.3)*, then there is a probability $\mathbb{P}$ on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ such that F_n is the n-dimensional distribution of $\mathbb{P}$ for all $n \in \mathbb{N}$.*

Example 3.1. Recall the model of independent tosses of a fair coin. Let F be a function on $\mathbb{R}$ defined by $F = \frac{1}{2}\mathbf{1}_{[0,\infty)} + \frac{1}{2}\mathbf{1}_{[1,\infty)}$. For $n \geq 1$, set $F_n(x_1, ..., x_n) = F(x_1) \times \cdots \times F(x_n)$. Then, F_n is the distribution of n independent tosses of fair coins, where $0, 1$ denote respectively for the tail and head. By Theorem 3.5, there is a probability on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ such that F_n is the corresponding n-dimensional distribution.

Exercise 3.8. Let ρ be a permutation of positive integers and $F_n : \mathbb{R}^n \to \mathbb{R}$ be a family of distribution functions satisfying (3.3). For $x = (x_1, x_2, ...) \in \mathbb{R}^\infty$ and $A \subset \mathbb{R}^\infty$, define $\rho(x) = (x_{\rho(1)}, x_{\rho(2)}, ...)$ and $\rho(A) = \{\rho(x) | x \in A\}$. For $n \geq 1$, let $m_n = \max\{\rho^{-1}(i) | 1 \leq i \leq n\}$ and set

$$G_n(x_1, ..., x_n) = \lim_{\substack{x_{\rho(i)} \to \infty \\ i \notin \rho^{-1}(\{1,...,n\})}} F_{m_n}(x_{\rho(1)}, ..., x_{\rho(m_n)}).$$

(1) Prove that G_n is a distribution function on $\mathbb{R}^n$ and satisfies (3.3).
(2) Let $\mathbb{P}_F, \mathbb{P}_G$ be the probabilities in Theorem 3.5 associated with F_n, G_n. Prove that $\mathbb{P}_G(A) = \mathbb{P}_F(\rho(A))$ for all $A \in \mathcal{B}(\mathbb{R}^\infty)$.

3.4. **Distributions on $\mathbb{R}^T$.** In this section, we introduce the extension theorem of Kolmogorov on $\mathbb{R}^T$. Let T be a set and $\mathcal{S}(T)$ be the set of all finite or infinite sequences in T with distinct terms. For $t \in T$, let $\Omega_t = \mathbb{R}$. Within this subsection, we set $x = (x_t)_{t\in T}$, $\Omega = \prod_{t\in T} \Omega_t$ and $\mathcal{B} = \mathcal{B}(\Omega)$. For $\tau = (t_n) \in \mathcal{S}(T)$, we write $x_\tau = (x_{t_n})$ and $\Omega_\tau = \prod_n \Omega_{t_n} := \{x_\tau | x_{t_n} \in \Omega_{t_n}, \forall n\}$ and $\mathcal{B}_\tau = \mathcal{B}(\Omega_\tau)$. For any (finite or infinite) sequence $\delta = (s_n) \in \mathcal{S}(\{t_n : n \geq 1\})$ and $B \subset \Omega_\delta$, define

$$\mathcal{D}_\tau(\delta, B) = \{x_\tau \in \Omega_\tau | x_\delta \in B\}, \quad \mathcal{D}(\delta, B) = \{x \in \Omega | x_\delta \in B\}.$$

We write $(s_n) \prec (t_n)$ if (s_n) is a subsequence of (t_n).

For $\tau = (t_n) \in \mathcal{S}(T)$, $B \subset \Omega_\tau$ and any permutation ρ of the subindex of τ, write

$$\rho(\tau) = (t_{\rho(n)}), \quad \rho(B) = \{x_{\rho(\tau)} \in \Omega_{\rho(\tau)} | x_\tau \in B\}.$$

Definition 3.6. Let T be a set. For $\tau = (t_1, ..., t_n) \in \mathcal{S}(T)$, let $\mathbb{P}_\tau$ be a probability on $(\Omega_\tau, \mathcal{B}_\tau)$. The family $\{\mathbb{P}_\tau | \tau \in \mathcal{S}(T),\ \tau \text{ is finite}\}$ is said to have the **consistency property** if

(1) For $\theta = (s_1, ..., s_m) \prec \tau = (t_1, ..., t_n) \in \mathcal{S}(T)$ and $B \in \mathcal{B}_\theta$,
$$\mathbb{P}_\tau(\mathcal{D}_\tau(\theta, B)) = \mathbb{P}_\theta(B).$$
(2) For $\tau = (t_1, ..., t_n) \in \mathcal{S}(T)$, $B \in \mathcal{B}_\tau$ and any permutation ρ of $\{1, ..., n\}$,
$$\mathbb{P}_{\rho(\tau)}(\rho(B)) = \mathbb{P}_\tau(B).$$

Theorem 3.6 (Kolmogorov's extension theorem on $\mathbb{R}^T$)**.** *Let T be a set and assume that $\{\mathbb{P}_\tau | \tau \in \mathcal{S}(T),\ \tau$ is finite$\}$ satisfies the consistency property in Definition 3.6. Then, there is a unique probability $\mathbb{P}$ on $(\Omega, \mathcal{B})$ such that*

$$\mathbb{P}(\mathcal{D}(\tau, B)) = \mathbb{P}_\tau(B).$$

for any finite sequence $\tau \in \mathcal{S}(T)$ and $B \in \mathcal{B}_\tau$.

Proof. By Theorem 3.4, for $\tau = (t_n)_{n=1}^\infty \in \mathcal{S}(T)$, there is a unique probability $\mathbb{P}_\tau$ on $(\Omega_\tau, \mathcal{B}_\tau)$ satisfying

$$\mathbb{P}_\tau(\mathcal{D}_\tau(t_1, ..., t_n, B)) = \mathbb{P}_{(t_1,...,t_n)}(B), \quad \forall B \in \mathcal{B}_{(t_1,...,t_n)},\ n \in \mathbb{N}.$$

First, we prove that, for $\tau = (t_n)_{n=1}^\infty \in \mathcal{S}(T)$, $\theta \prec \tau$ and any permutation ρ of $\mathbb{N}$,

(i) $\mathbb{P}_\tau(\mathcal{D}_\tau(\theta, B)) = \mathbb{P}_\theta(B)$ for all $B \in \mathcal{B}_\theta$.
(ii) $\mathbb{P}_{\rho(\tau)}(\rho(B)) = \mathbb{P}_\tau(B)$ for all $B \in \mathcal{B}_\tau$.

For (i), it is obviously true when θ is finite. We consider the case that θ is infinite in the following. Let $\theta = (s_n)_{n=1}^\infty$ and set

$$\mathcal{A} = \{A \in \mathcal{B}_\theta | \mathbb{P}_\tau(\mathcal{D}_\tau(\theta, A)) = \mathbb{P}_\theta(A)\}.$$

Clearly, $\mathcal{A}$ is a λ-system. By the π-λ lemma, it remains to show that $B \times \Omega_{(s_{n+1},s_{n+2},...)} \in \mathcal{A}$ for all $B \in \mathcal{B}_{(s_1,...,s_n)}$ and $n \geq 1$. Fix $n \geq 1$ and set $\theta_n = (s_1, ..., s_n)$. Since $\theta_n \prec \theta \prec \tau$, one has

$$\mathcal{D}_\tau(\theta_n, B) = \mathcal{D}_\tau(\theta, \mathcal{D}_\theta(\theta_n, B)), \quad \forall B \in \mathcal{B}_{\theta_n}.$$

As $\theta \prec \tau$, there are positive integers $k_1 < \cdots < k_n$ such that $s_i = t_{k_i}$ for $1 \leq i \leq n$. Set $M = k_n$ and $\tau_M = (t_1, ..., t_M)$. Since $\theta_n \prec \tau_M \prec \tau$, one has

$$\mathcal{D}_\tau(\tau_M, \mathcal{D}_{\tau_M}(\theta_n, B)) = \mathcal{D}_\tau(\theta_n, B).$$

As a consequence, this implies

$$\begin{aligned}\mathbb{P}_\tau(\mathcal{D}_\tau(\theta, \mathcal{D}_\theta(\theta_n, B))) &= \mathbb{P}_\tau(\mathcal{D}_\tau(\tau_M, \mathcal{D}_{\tau_M}(\theta_n, B))) = \mathbb{P}_{\tau_M}(\mathcal{D}_{\tau_M}(\theta_n, B)) \\ &= \mathbb{P}_{\theta_n}(B) = \mathbb{P}_\theta(\mathcal{D}_\theta(\theta_n, B)).\end{aligned}$$

This proves (i).

For (ii), let $\tau = (t_n)_{n=1}^{\infty}$ be a sequence, ρ be a permutation of $\mathbb{N}$. Consider the following class.

$$\mathcal{A}' := \{A \in \mathcal{B}_\tau | \mathbb{P}_{\rho(\tau)}(\rho(A)) = \mathbb{P}_\tau(A)\}.$$

By the π-λ lemma, as $\mathcal{A}'$ is a λ-system, it suffices to show that $C \times \Omega_{(t_{n+1}, t_{n+2}, ...)} \in \mathcal{A}'$ for all $C \in \mathcal{B}_{(t_1,...,t_n)}$ and $n \geq 1$. Fix $n \geq 1$. Let $\tau_n = (t_1, ..., t_n)$ and $\tau_n' = (t_{n+1}, t_{n+2}, ...)$ and write $\{b_1 < \cdots < b_n\} = \{\rho^{-1}(1), ..., \rho^{-1}(n)\}$. Set $N = b_n$ and let ρ' be a permutation of $\{1, ..., n\}$ defined by

$$\rho'(i) = \rho(b_i) \quad \forall 1 \leq i \leq n.$$

Note that $\rho'(\tau_n) = (t_{\rho'(1)}, ..., t_{\rho'(n)}) \prec (t_{\rho(1)}, ..., t_{\rho(N)}) \prec \rho(\tau)$. This implies, for $C \in \mathcal{B}_{\tau_n}$,

$$\rho\left(C \times \Omega_{\tau_n'}\right) = \mathcal{D}_{\rho(\tau)}(\rho'(\tau_n), \rho'(C)) = \mathcal{D}_{\rho(\tau)}\left(t_{\rho(1)}, ..., t_{\rho(N)}, C'\right),$$

where

$$C' = \mathcal{D}_{(t_{\rho(1)},...,t_{\rho(N)})}(\rho'(\tau_n), \rho'(C)).$$

By the consistency property, we obtain

$$\mathbb{P}_{\rho(\tau)}\left(\rho(C \times \Omega_{\tau_n'})\right) = \mathbb{P}_{(t_{\rho(1)},...,t_{\rho(N)})}(C') = \mathbb{P}_{\rho'(\tau_n)}(\rho'(C)) = \mathbb{P}_{\tau_n}(C) = \mathbb{P}_\tau(C \times \Omega_{\tau_n'}),$$

as desired. This proves (ii).

For $A \in \mathcal{B}$, one may select $\tau \in \mathcal{S}(T)$ and $B \in \mathcal{B}_\tau$ such that $A = \mathcal{D}(\tau, B)$. Define

$$\mathbb{P}(A) = \mathbb{P}_\tau(B).$$

Exercise 3.9. Prove that $\mathbb{P}$ is well-defined.

To prove that $\mathbb{P}$ is a probability on $(\Omega, \mathcal{B})$, it remains to show that $\mathbb{P}$ is σ-additive. Let $A_n \in \mathcal{B}$ be mutually disjoint sets. For $n \in \mathbb{N}$, one may choose $\eta_n \in \mathcal{S}(T)$ and $B_n \in \mathcal{B}_{\eta_n}$ such that $A_n = \mathcal{D}(\eta_n, B_n)$. Let $\eta \in \mathcal{S}(T)$ be a sequence consisting of terms in η_n for all n. For $n \in \mathbb{N}$, there is a permutation ρ_n of $\mathbb{N}$ such that $\rho_n(\eta_n) \prec \eta$. Note that

$$A_n = \mathcal{D}(\rho_n(\eta_n), \rho_n(B_n)) = \mathcal{D}(\eta, C_n), \quad C_n = \mathcal{D}_\eta(\rho_n(\eta_n), \rho_n(B_n)).$$

Clearly, C_n are mutually disjoint and, by the σ-additivity of $\mathbb{P}_\eta$, we have

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\mathcal{D}\left(\eta, \bigcup_{n=1}^{\infty} C_n\right)\right) = \sum_{n=1}^{\infty} \mathbb{P}_\eta(C_n) = \sum_{n=1}^{\infty} \mathbb{P}_{\rho_n(\eta_n)}(\rho_n(B_n)) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

This finishes the proof. $\square$

Definition 3.7. Let $\mathbb{P}$ be a probability on $(\Omega, \mathcal{B})$. For $\tau = (t_1, ..., t_n) \in \mathcal{S}(T)$, the function $F_\tau(x_\tau) := \mathbb{P}(\mathcal{D}(\tau, \prod_{i=1}^n (-\infty, x_{t_i}]))$ is called a **finite dimensional distribution** of $\mathbb{P}$.

Theorem 3.7. *Let T be a set. For $n \in \mathbb{N}$ and finite $\tau \in \mathcal{S}(T)$, let F_τ be a distribution function on Ω_τ. Suppose that, for $n \geq 1$, $\tau = (t_1, ..., t_n) \in \mathcal{S}(T)$, $x_\tau \in \Omega_\tau$ and any permutation ρ of $\{1, ..., n\}$,*

$$F_{\rho(\tau)}(x_{\rho(\tau)}) = F_\tau(x_\tau), \tag{3.4}$$

and

$$\lim_{x_{t_n} \to \infty} F_\tau(x_\tau) = F_{\check{\tau}}(x_{\check{\tau}}), \tag{3.5}$$

where $\check{\tau} = (t_1, ..., t_{n-1})$. Then, there is a unique probability $\mathbb{P}$ on $(\Omega, \mathcal{B})$ such that

$$F_\tau(x_\tau) = \mathbb{P}(\mathcal{D}(\tau, (-\infty, x_{t_1}] \times \cdots \times (-\infty, x_{t_n}]))$$

for all $\tau = (t_1, ..., t_n) \in T$ and $n \geq 1$.