

1 Kolmogorov’s extension theorem

We state and prove the Kolmogorov’s extension theorem when the index set is $T = \{1, 2, 3, \ldots \} = \mathbb{N}$.

**Theorem 1** (Theorem 1.4.22, Dembo’s Notes). Suppose we are give probability measures $\mu_n$ on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ that are consistent, that is,

$$\mu_{n+1}(B_1 \times \cdots \times B_n \times \mathbb{R}) = \mu_n(B_1 \times \cdots \times B_n) \quad \forall B_i \in \mathcal{B}, \ i = 1, \ldots, n < \infty. \quad (1)$$

Then, there exists a unique probability measure $\mathbb{P}$ on $(\mathbb{R}^N, \mathcal{B}_c)$ such that

$$\mathbb{P}(\{\omega : \omega_i \in B_i, i = 1, \ldots, n\}) = \mu_n(B_1 \times \cdots \times B_n) \quad \forall B_i \in \mathcal{B}, \ i = 1, \ldots, n < \infty.$$  

**Remark** Kolmogorov’s extension theorem builds the foundation on which stochastic processes are defined: namely, for any index set $T$, to define the distribution of a stochastic process $X_T$, it suffices to give a consistent collection of joint distributions of $(X_{t_1}, \ldots, X_{t_n})$ on finitely many coordinates. The measure of $X_T$ on $(\mathbb{R}^T, \mathcal{B}_c)$, then, by the extension theorem, is guaranteed to exist and is unique.

The theorem is trivial when $T = \{1, \ldots, n\}$ is finite: just take $\mathbb{P} = \mu_n$. $T = \mathbb{N}$ is the first non-trivial case of the theorem. This case can give us, for example, the probability measure of countably many i.i.d. R.V.-s $(X_1, X_2, \ldots)$.

**Proof of Theorem 1** The proof mainly follows that of [1, Chapter 36]. Let $\mathbb{R}_0^N$ be the collection of cylindrical sets of the form

$$A = \{x \in \mathbb{R}^N : (x_1, \ldots, x_n) \in H\}, \quad (2)$$

where $n \in \mathbb{N}$ and $H \in \mathcal{B}_{\mathbb{R}^n}$. That is, we consider sets that require the first $n$ coordinates lie in some Borel set $H \subset \mathbb{R}^n$. By definition of the cylindrical $\sigma$-algebra, we have $\mathcal{B}_c = \sigma(\mathbb{R}_0^N)$. On this collection, define the set function

$$\mathbb{P}(A) = \mu_n(H).$$

We are going to use Caratheodory’s extension theorem to extend $\mathbb{P}$ to $\mathcal{B}_c$, which we divide into the following steps.

**$\mathbb{P}$ is well-defined** To show this, we need to verify that if a cylindrical set $A$ has two representations of the form (2) then they give coinciding values of $\mathbb{P}(A)$. Consider

$$A = \{x : (x_1, \ldots, x_{n_1}) \in H_1\} = \{x : (x_1, \ldots, x_{n_2}) \in H_2\}$$
for some $n_1 \geq n_2$, then it is easy to see that $H_1 = H_2 \times \mathbb{R}^{n_1-n_2}$. (Check this!) It remains to show that
\[ \mu_{n_1}(H_1) = \mu_{n_1}(H_2) = \mu_{n_2}(H_2). \] (3)
Repeating the consistency condition \([1]\) gives that $\mu_{n_1}(B_1 \times \cdots \times B_{n_2} \times \mathbb{R}^{n_1-n_2}) = \mu_{n_2}(B_1 \times \cdots \times B_{n_2})$, and a standard extension argument shows that $\mu_{n_1}(\cdot \times \mathbb{R}^{n_1-n_2}) = \mu_{n_2}(\cdot)$, verifying \([3]\).

\[ \mathbb{R}^N_0 \text{ is an algebra; } P \text{ finitely additive on } \mathbb{R}^N_0 \] Clearly $\emptyset \in \mathbb{R}^N_0$. For any cylindrical set $A$, we have $A^c = \{ x \in \mathbb{R}^N : (x_1, \ldots, x_n) \in H^c \}$, so $A^c \in \mathbb{R}^N_0$. Let $A, B$ be two cylindrical sets:
\[ A = \{ x : (x_1, \ldots, x_n) \in H_1 \}, \quad B = \{ x : (x_1, \ldots, x_n) \in H_2 \}. \]
Without loss of generality, let $n_1 \geq n_2$. We then have
\[ A \cup B = \{ x : (x_1, \ldots, x_n) \in H_1 \cup (H_2 \times \mathbb{R}^{n_1-n_2}) \} \in \mathbb{R}^N_0. \] (4)
This shows that $\mathbb{R}^N_0$ is an algebra. If $A$ and $B$ are disjoint, then $H_2 \times \mathbb{R}^{n_1-n_2} \cap H_1 = \emptyset$, giving that
\[ P(A \cup B) = \mu_{n_1}(H_1 \cup (H_2 \times \mathbb{R}^{n_1-n_2})) = \mu_{n_1}(H_1) + \mu_1(H_2 \times \mathbb{R}^{n_1-n_2}) = P(A) + P(B), \]
so $P$ is finitely additive.

\[ P \text{ is a probability measure on } \mathbb{R}^N_0 \] Clearly $P \geq 0$ and $P(\emptyset) = 0$. Let $A$ be a cylindrical set, then
\[ P(A^c) = \mu_n(H^c) = 1 - \mu_n(H) = 1 - P(A). \]
It remains to show countable additivity. As it is finitely additive, it suffices to show that $A_k \in \mathbb{R}^N_0$ with $A_k \downarrow \emptyset$ implies $P(A_k) \to 0$. (See the Remark in Dembo notes, page 14). As we can always make the defining index non-decreasing, we can let
\[ A_k = \{ x : (x_1, \ldots, x_n) \in H_k \} \]
where $n_k \in \mathbb{N}$ is increasing and $H_k \subset \mathbb{R}^{n_k}$.

Suppose $P(A_k) \not\to 0$, then $P(A_k) \geq \varepsilon$ holds for all $k$, for some $\varepsilon > 0$. This means $\mu_{n_k}(H_k) \geq \varepsilon$. Applying \([1]\) Theorem 12.3, there exists compact sets $K_k \subset H_k$ such that $\mu_{n_k}(H_k \setminus K_k) \leq \varepsilon/2^{k+1}$. Define
\[ B_k = \{ x : (x_1, \ldots, x_{n_k}) \in K_k \}, \]
so $P(B_k) \leq \varepsilon/2^{k+1}$. Define $C_k = \bigcap_{j=1}^k B_j$, then we have $C_k \subset B_k \subset A_k$ and $P(A_k \setminus C_k) \leq \varepsilon/2$, so $P(C_k) \geq \varepsilon/2$, and thus $C_k$ is non-empty.

Now, for all $k$, choose a point $x^{(k)} \in C_k$. As $C_k$ is the intersection of $\{B_j\}_{j \leq k}$, we have $(x_1^{(k)}, \ldots, x_{n_j}^{(k)}) \in K_j$ for all $j \leq k$. In other words, the first $n_j$ indices of $\{x^{(k)}\}_{k \geq j}$ lie in the compact set $K_j$. Hence, there exists a subsequence $k_i$ such that $(x_1^{(k_i)}, \ldots, x_{n_j}^{(k_i)})$ converges. By the diagonal method, we can find a subsequence $k_i$ such that $(x_1^{(k_i)}, \ldots, x_{n_j}^{(k_i)})$ converges for all $j$. Let $x$ be the point in $\mathbb{R}^N$ such that $x_1, \ldots, x_{n_j}$ is the limit of the above sequence (as the limits are consistent, $x$ exists). The closedness of $K_j$ implies that $(x_1, \ldots, x_n) \in K_j$, so $x \in A_j$. Thus we have found a point $x \in \bigcap_{j=1}^\infty A_j$, contradictory to that $A_j \emptyset$. Hence our assumption is wrong so we must have $P(A_j) \to 0$. \(\square\)

References