1 The Portmanteau Theorem

Theorem 1.1 (The Portmanteau Theorem, Theorem 2.1, [1])

Suppose:

- (S, ρ) *is a metric space,* $\mathcal{B}(S)$ *its the Borel* σ*-algebra,* $(S, \mathcal{B}(S))$ *is the corresponding measurable space.*
- $P, P_1, P_2, \ldots \in \mathcal{M}_1(S, \mathcal{B}(S))$ *are probability measures on* $(S, \mathcal{B}(S))$.

Then, the following are equivalent:

(i) P_n *converges weakly to* P*, i.e. for each bounded continuous* $\mathbb{R}$*-valued function* $f : S \longrightarrow \mathbb{R}$*, we have*

$$\lim_{n\to\infty} \int_S f(s)\,\mathrm{d}P_n(s) \;=\; \int_S f(s)\,\mathrm{d}P(s).$$

(ii) *For each closed set* $F \subset S$*, we have*

$$\limsup_{n\to\infty} P_n(F) \;\leq\; P(F).$$

(iii) *For each open set* $G \subset S$*, we have*

$$\liminf_{n\to\infty} P_n(G) \;\geq\; P(G).$$

(iv) *For each* P*-continuity set* $A \in \mathcal{B}(S)$*, i.e.* $P(\partial A) = 0$*, we have*

$$\lim_{n\to\infty} P_n(A) \;=\; P(A).$$

Proof

(i) $\Longrightarrow$ (ii)

For each $\varepsilon > 0$, by Lemma A.2, choose a bounded continuous functions $f_\varepsilon : S \longrightarrow [0,1]$ such that

$$I_F \;\leq\; f_\varepsilon \;\leq\; I_{F^\varepsilon}.$$

This implies, for each $\varepsilon > 0$, we have

$$P_n(F) \;=\; \int_S I_F(x)\,\mathrm{d}P_n(x) \;\leq\; \int_S f_\varepsilon(x)\,\mathrm{d}P_n(x).$$

By (i), we thus have

$$\limsup_{n\to\infty} P_n(F) \;\leq\; \lim_{n\to\infty} \int_S f_\varepsilon(x)\,\mathrm{d}P_n(x) \;=\; \int_S f_\varepsilon(x)\,\mathrm{d}P(x) \;\leq\; \int_S I_{F^\varepsilon}(x)\,\mathrm{d}P(x) \;=\; P(F^\varepsilon)\,.$$

By Lemma A.2, we have $F^\varepsilon \downarrow F$ as $\varepsilon \downarrow 0$. Hence, $P(F^\varepsilon) \downarrow P(F)$ as $\varepsilon \downarrow 0$ (by Theorem 2.3, [2]). We may now conclude:

$$\limsup_{n\to\infty} P_n(F) \;\leq\; \lim_{\varepsilon\to 0^+} P(F^\varepsilon) \;=\; P(F)\,.$$

(ii) $\Longrightarrow$ (iii)

Assume (ii) holds. Let $G \subset S$ be a open subset. Then, $F := S \setminus G$ is closed. By (ii), we have:

$$\begin{aligned} 1 - \liminf_{n\to\infty} P_n(G) &= \limsup_{n\to\infty} \{ 1 - P_n(G) \} = \limsup_{n\to\infty} P_n(S \setminus G) = \limsup_{n\to\infty} P_n(F) \\ &\leq P(F) = P(S \setminus G) = 1 - P(G) , \end{aligned}$$

which yields

$$\liminf_{n\to\infty} P_n(G) \geq P(G) . \tag{1.1}$$

<u>(iii) $\Longrightarrow$ (ii)</u>

Assume (iii) holds. Let $F \subset S$ be an closed subset. Then, $G := S \setminus F$ is open. By (iii), we have:

$$\begin{aligned} 1 - \limsup_{n\to\infty} P_n(F) &= \liminf_{n\to\infty} \{ 1 - P_n(F) \} = \liminf_{n\to\infty} P_n(S \setminus F) = \liminf_{n\to\infty} P_n(G) \\ &\geq P(G) = P(S \setminus F) = 1 - P(F) , \end{aligned}$$

which yields

$$\limsup_{n\to\infty} P_n(F) \leq P(F) . \tag{1.2}$$

<u>(ii) and (iii) $\Longrightarrow$ (iv)</u>

Let $A \in \mathcal{B}(S)$. Then, by (ii) and (iii), we have:

$$P(A^\circ) \leq \liminf_{n\to\infty} P_n(A^\circ) \leq \liminf_{n\to\infty} P_n(A) \leq \limsup_{n\to\infty} P_n(A) \leq \limsup_{n\to\infty} P_n\big(\overline{A}\big) \leq P\big(\overline{A}\big) .$$

Hence, if $\partial A := \overline{A} \setminus A^\circ$ is a P-continuity set, i.e. $P(\partial A) = 0$, hence $P(A^\circ) = P(\, A\,) = P\big(\overline{A}\big)$, then (iv) follows.

<u>(iv) $\Longrightarrow$ (i)</u>

Let $f : S \longrightarrow \mathbb{R}$ be a bounded continuous $\mathbb{R}$-valued function on S. We need to show $\int_S f(s)\,\mathrm{d}P_n(s) \longrightarrow \int_S f(s)\,\mathrm{d}P(s)$. By linearity, we may assume $0 \leq f \leq 1$.

Claim:

$f^{-1}(\,(t,\infty)\,) = \{\, s \in S \mid f(s) > t \,\}$ is a P-continuity set, except for at most countably many $t \in [0,1]$.

Proof of Claim: First, note that the continuity of f implies that

$$\partial \{\, s \in S \mid f(s) > t \,\} \subset \{\, s \in S \mid f(s) = t \,\} , \quad \text{for each } t \in [0,1].$$

Indeed,

$$\begin{aligned} & s_0 \in \partial \{\, s \in S \mid f(s) > t \,\} \\ \Longleftrightarrow \quad & \text{every neighbourhood of } s_0 \text{ non-trivially intersects both } \{\, s \in S \mid f(s) > t \,\} \text{ and } \{\, s \in S \mid f(s) \leq t \,\} \\ \Longrightarrow \quad & \exists\, s_1, s_2, \ldots \in \{\, s \in S \mid f(s) > t \,\} ,\ s_1', s_2', \ldots \in \{\, s \in S \mid f(s) \leq t \,\} \text{ with } s_n \longrightarrow s_0,\ s_n' \longrightarrow s_0 \\ \Longrightarrow \quad & f(s_0) = f\Big(\lim_{n\to\infty} s_n\Big) = \lim_{n\to\infty} f(s_n) \geq t \ \text{ and } \ f(s_0) = f\Big(\lim_{n\to\infty} s_n'\Big) = \lim_{n\to\infty} f(s_n') \leq t \ \text{ (by continuity of } f) \\ \Longrightarrow \quad & f(s_0) = t, \ \text{ i.e. } \ s_0 \in \{\, s \in S \mid f(s) = t \,\} . \end{aligned}$$

Next, note that, since f is continuous, $f^{-1}(\{\,t\,\})$ is $\mathcal{B}(S)$-measurable for each $t \in [0,1]$. Thus,

$$S \;=\; \bigsqcup_{t\in[0,1]} \{\, s \in S \mid f(s) = t \,\} \;=\; \bigsqcup_{t\in[0,1]} f^{-1}(\{\,t\,\})$$

is a partition of S into uncountably many pairwise disjoint $\mathcal{B}(S)$-measurable subsets. By Lemma A.4,

$$P\big(\, f^{-1}(\{\,t\,\})\,\big) \;=\; 0, \;\; \text{for all but countably many } t \in [0,1],$$

which in turn implies

$$P(\,\partial\,\{\, s \in S \mid f(s) > t \,\}\,) \;=\; 0, \;\; \text{for all but countably many } t \in [0,1].$$

This completes the proof of the Claim.

The above Claim and (iv) together imply:

$$P_n(\,f > t\,) \;\longrightarrow\; P(\,f > t\,)\,, \;\; \text{for almost every } t \in [0,1].$$

By Lemma A.3 and the Lebesgue Dominated Convergence Theorem, we have

$$\begin{aligned}
\int_S f(s)\,\mathrm{d}P_n(s) \;&=\; \int_0^\infty P_n(\,f > t\,)\;\mathrm{d}t \\
&=\; \int_0^1 P_n(\,f > t\,)\;\mathrm{d}t \;\longrightarrow\; \int_0^1 P(\,f > t\,)\;\mathrm{d}t \\
&=\; \int_0^\infty P(\,f > t\,)\;\mathrm{d}t \;=\; \int_S f(s)\,\mathrm{d}P(s),
\end{aligned}$$

which proves that (iv) $\Longrightarrow$ (i). □

A Technical Lemmas

Lemma A.1 *Suppose (S,ρ) is a metric space, and $A \subset S$ is an arbitrary non-empty subset. Define*

$$\rho(\,\cdot\,,A) \;:\; S \;\longrightarrow\; \mathbb{R} \;:\; x \;\longmapsto\; \inf_{y\in A}\{\,\rho(x,y)\,\}$$

Then,

(i) *$\rho(\,\cdot\,,A)$ is a continuous $\mathbb{R}$-valued function on S.*

(ii) *For each $x \in S$, $\rho(x,A) = 0$ if and only if $x \in \overline{A}$.*

PROOF

(i) Suppose $x_n \longrightarrow x$. We need to prove $\rho(x_n,A) \longrightarrow \rho(x,A)$, which follows immediately from the following two Claims:

Claim 1: $\rho(x,A) \;\leq\; \liminf_{n\to\infty} \rho(x_n,A)$.

Claim 2: $\limsup_{n\to\infty} \rho(x_n,A) \;\leq\; \rho(x,A)$.

Proof of Claim 1: For each $y \in S$, we have:

$$\rho(x,y) \;\leq\; \rho(x,x_n) \;+\; \rho(x_n,y).$$

Hence,

$$\rho(x,A) \;=\; \inf_{y\in A}\rho(x,y) \;\leq\; \rho(x,x_n) \;+\; \inf_{y\in A}\rho(x_n,y) \;=\; \rho(x,x_n) \;+\; \rho(x_n,A).$$

Since $\rho(x,x_n) \longrightarrow 0$, the preceding inequality implies

$$\rho(x,A) \;\leq\; \liminf_{n\to\infty}\rho(x_n,A).$$

This proves Claim 1.

Proof of Claim 2: For each $y \in S$, we have:

$$\rho(x_n,y) \;\leq\; \rho(x_n,x) \;+\; \rho(x,y).$$

Hence,

$$\rho(x_n,A) \;=\; \inf_{y\in A}\rho(x_n,y) \;\leq\; \rho(x_n,x) \;+\; \inf_{y\in A}\rho(x,y) \;=\; \rho(x_n,x) \;+\; \rho(x,A).$$

Since $\rho(x,x_n) \longrightarrow 0$, the preceding inequality implies

$$\limsup_{n\to\infty}\rho(x_n,A) \;\leq\; \rho(x,A).$$

This proves Claim 2.

(ii)

$$\begin{aligned}\rho(x,A)=0 \quad&\Longleftrightarrow\quad \inf_{y\in A}\rho(x,y)=0\\ &\Longleftrightarrow\quad \text{For each } \varepsilon>0\text{, there exists } y\in A \text{ such that } \rho(x,y)<\varepsilon\\ &\Longleftrightarrow\quad y\in\overline{A}\end{aligned}$$

□

Lemma A.2 *Suppose (S,ρ) is a metric space, and $A \subset S$ is an arbitrary non-empty subset. For each $\varepsilon > 0$, define*

$$A^{\varepsilon} \;:=\; \{\; s\in S \;\mid\; \rho(s,A)<\varepsilon \;\}\,.$$

Then the following are true:

(i) *A^{ε} is an open subset of S. In particular, A^{ε} is a $\mathcal{B}(S)$-measurable subset of S.*

(ii) *$A^{\varepsilon} \downarrow \overline{A}$, as $\varepsilon \downarrow 0$.*

(iii) *There exists a bounded continuous $\mathbb{R}$-valued function $f : S \longrightarrow \mathbb{R}$ such that*

$$I_{\bar{A}}(x) \;\leq\; f(x) \;\leq\; I_{A^{\varepsilon}}(x)\,,\quad \text{for each } x\in S.$$

Proof

(i) Let $x \in A^{\varepsilon}$. Let $\delta := \varepsilon - \rho(x,A) > 0$. Let $U := \{\, y \in S \mid \rho(x,y) < \delta/2 \,\}$. Then, for each $y \in U$ and $a \in A$, we have

$$\rho(y,a) \;\leq\; \rho(y,x)+\rho(x,a) \quad\Longrightarrow\quad \rho(y,A) \;\leq\; \rho(y,x)+\rho(x,A) \;\leq\; \frac{\delta}{2}+\varepsilon-\delta \;=\; \varepsilon-\frac{\delta}{2},$$

which implies $\rho(y,A) \;\leq\; \varepsilon - \dfrac{\delta}{2} \;<\; \varepsilon$. Hence $U \subset A^{\varepsilon}$. Since U is an open subset of S, we may now conclude that A^{ε} is indeed an open subset of S.

(ii) First, note that $A^{\varepsilon_1} \subset A^{\varepsilon_2}$ whenever $\varepsilon_1 \leq \varepsilon_2$. Indeed, suppose $\varepsilon_1 \leq \varepsilon_2$. Then,

$$x \in A^{\varepsilon_1} \implies \rho(x,A) < \varepsilon_1 \implies \rho(x,A) < \varepsilon_2 \implies x \in A^{\varepsilon_2},$$

which proves $A^{\varepsilon_1} \subset A^{\varepsilon_2}$ whenever $\varepsilon_1 \leq \varepsilon_2$. Next,

$$\begin{aligned} x \in \bigcap_{\varepsilon>0} A^\varepsilon &\iff x \in A^\varepsilon, \text{ for each } \varepsilon > 0 \\ &\iff \rho(x,A) < \varepsilon, \text{ for each } \varepsilon > 0 \\ &\iff \rho(x,A) = 0 \\ &\iff x \in \overline{A} \text{ (by Lemma A.1)} \end{aligned}$$

Hence, we see that

$$\bigcap_{\varepsilon>0} A^\varepsilon \;=\; \overline{A}.$$

This proves completes the proof of (ii).

(iii) Define $f : S \longrightarrow \mathbb{R}$ as follows:

$$f(x) \;:=\; \max\left\{\, 0\,,\, 1 - \frac{\rho(x,A)}{\varepsilon} \,\right\}.$$

Then, by Lemma A.1, f is a continuous $\mathbb{R}$-valued function on S. Clearly, $0 \leq f(x) \leq 1$, for each $x \in S$. By Lemma A.1, we have

$$x \;\in\; \overline{A} \quad\iff\quad \rho(x,F) \;=\; 0 \quad\iff\quad f(x) \;=\; 1.$$

This proves $I_{\bar{A}}(x) \leq 1 = f(x)$, for each $x \in \overline{A}$, and hence for each $x \in S$ (since $I_{\bar{A}}(x) = 0$ for $x \in S \setminus \overline{A}$, and the inequality holds trivially). On the other hand,

$$x \;\in\; S \setminus A^\varepsilon \quad\iff\quad \varepsilon \;\leq\; \rho(x,A) \quad\iff\quad 1 - \frac{\rho(x,A)}{\varepsilon} \;\leq\; 0 \quad\implies\quad f(x) \;=\; 0.$$

This proves $f(x) = 0 \leq I_{A^\varepsilon}(x)$, for each $x \in S \setminus A^\varepsilon$, and hence for each $x \in S$ (since $I_{A^\varepsilon}(x) = 1$ for each $x \in A^\varepsilon$ and the inequality holds trivially). This completes the proof of (ii).

□

Lemma A.3

Let $(\,\Omega, \mathcal{A}, P\,)$ be any probability space. Then, for each $p > 0$ and for each non-negative random variable (i.e. measurable function) $f : \Omega \longrightarrow [0,\infty)$, we have:

$$E[\,f^p\,] \;=\; p\int_0^\infty P(\,f > t\,)\cdot t^{p-1}\,\mathrm{d}t \;=\; p\int_0^\infty P(\,f \geq t\,)\cdot t^{p-1}\,\mathrm{d}t\,.$$

Proof

We first prove the first equality: By elementary Calculus (change of variable formula) and Fubini's Theorem, we have

$$\begin{aligned} E[\,f^p\,] \;&:=\; \int_\Omega f(\omega)^p\,\mathrm{d}P(\omega) \;=\; \int_\Omega \left[\int_0^{f(\omega)^p} 1\,\mathrm{d}s\right]\mathrm{d}P(\omega) \;=\; \int_\Omega \left[\int_0^\infty 1_{\{\,0\,<\,s\,<\,f(\omega)^p\}}(s)\,\mathrm{d}s\right]\mathrm{d}P(\omega) \\ &=\; \int_\Omega \left[\int_0^\infty 1_{\left\{\,0\,\leq\,s^{1/p}\,<\,f(\omega)\,\right\}}\,\mathrm{d}s\right]\mathrm{d}P(\omega) \;=\; \int_\Omega \left[\int_0^\infty 1_{\{\,0\,\leq\,t\,<\,f(\omega)\,\}}\cdot p\cdot t^{p-1}\,\mathrm{d}t\right]\mathrm{d}P(\omega) \\ &=\; \int_0^\infty \left[\int_\Omega 1_{\{\,0\,\leq\,t\,<\,f(\omega)\,\}}\cdot p\cdot t^{p-1}\,\mathrm{d}P(\omega)\right]\mathrm{d}t \;=\; p\cdot\int_0^\infty \left[\int_\Omega 1_{\{\,0\,\leq\,t\,<\,f(\omega)\,\}}\,\mathrm{d}P(\omega)\right]\cdot t^{p-1}\,\mathrm{d}t \\ &=\; p\cdot\int_0^\infty P(\,f > t\,)\cdot t^{p-1}\,\mathrm{d}t. \end{aligned}$$

The proof of the second inequality is analogous. □

Lemma A.4

Suppose

- (S, ρ) *is a metric space, and* $\mathcal{B}(S)$ *is its Borel* σ*-algebra.*
- $S = \bigsqcup_{\gamma\in\Gamma} F_\gamma$ *is a partition of* S *into pairwise disjoint* $\mathcal{B}(S)$*-measurable subsets* $F_\gamma \in \mathcal{B}(S)$*.*
  Note that here the index set Γ *may be uncountable.*

Then, for any probability measure $\mu \in \mathcal{M}_1(S, \mathcal{B}(S))$*, we have:*

$$\mu(F_\gamma) \; = \; 0, \; \text{ for all but countably many } \gamma \in \Gamma.$$

PROOF Define $\Gamma_0 := \{\, \gamma \in \Gamma \;\mid\; \mu(F_\gamma) = 0 \,\}$, and for each $n \in \mathbb{N}$, define $\Gamma_n := \left\{ \gamma \in \Gamma \;\middle|\; \mu(F_\gamma) \geq \dfrac{1}{n} \right\}$. Clearly,

$$\Gamma \;=\; \Gamma_0 \bigsqcup \left(\bigcup_{n=1}^{\infty} \Gamma_n \right).$$

Thus, the Lemma follows immediately from the following

Claim: For each $n \geq 1$, Γ_n is a finite set with $|\,\Gamma_n\,| \leq n$.

Proof of Claim: If the Claim were false, there would exist $n \in \mathbb{N}$ such that Γ_n contained at least $n+1$ distinct elements, say $\gamma_1, \gamma_2, \ldots, \gamma_{n+1} \in \Gamma_n$. It would follow that:

$$\mu\left(\bigsqcup_{i=1}^{n+1} F_{\gamma_i} \right) \;=\; \sum_{i=1}^{n+1} \mu(F_{\gamma_i}) \;\geq\; \sum_{i=1}^{n+1} \frac{1}{n} \;=\; \frac{n+1}{n} \;>\; 1,$$

which would contradict that hypothesis that μ is a probability measure. Thus, the Claim must be true. □

References

[1] Billingsley, P. *Convergence of Probability Measures*, second ed. John Wiley & Sons, 1999.

[2] Jacod, J., and Protter, P. *Probability Essentials.* Springer-Verlag, New York, 2004. Universitext.