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* $A \in \mathcal{B}(S)$*, we have*

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

(v) *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 function $f_\varepsilon : S \longrightarrow [0,1]$ such that

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

This implies that, 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

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

(iii) $\Longrightarrow$ (ii)

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

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

(ii) and (iii) $\Longrightarrow$ (iv)

Note first that the middle three inequalities in (iv) are trivially true. On the other hand, the leftmost inequality in (iv) follows immediately from (iii), while the rightmost follows immediately from (ii).

(iv) $\Longrightarrow$ (v)

If $\partial A := \overline{A} \setminus A^{\circ}$ is a P-continuity set, i.e. $P(\partial A) = 0$, then $P(A^{\circ}) = P(\,A\,) = P\big(\overline{A}\big)$, in which case, (v) follows immediately from (iv).

(v) $\Longrightarrow$ (i)

Let $f : S \longrightarrow \mathbb{R}$ be a bounded continuous $\mathbb{R}$-valued function on S. We need to show $\int_S f(s)\,\mathrm{d}P_n(s) \longrightarrow \int_S f(s)\,\mathrm{d}P(s)$. Since f is assumed bounded, there exists some $c > 0$ such that $|\,f\,| < c$. Observe that

$$\int_S f(s)\,\mathrm{d}P_n(s) \longrightarrow \int_S f(s)\,\mathrm{d}P(s) \quad\Longleftrightarrow\quad \int_S \left(\frac{f(s)+c}{2c}\right)\mathrm{d}P_n(s) \longrightarrow \int_S \left(\frac{f(s)+c}{2c}\right)\mathrm{d}P(s).$$

Thus, by replacing f with $(f+c)\,/\,2c$ if necessary, we may assume, without loss of generality, that $0 \leq f \leq 1$. We make this assumption for the remainder of the proof.

Now, the fact that $0 \leq f \leq 1$ and Lemma A.3 together imply:

$$\begin{aligned} \int_S f(s)\,\mathrm{d}P(s) &= \int_0^\infty P(\,f > t\,)\;\mathrm{d}t = \int_0^1 P(\,f > t\,)\;\mathrm{d}t, \text{ and} \\ \int_S f(s)\,\mathrm{d}P_n(s) &= \int_0^\infty P_n(\,f > t\,)\;\mathrm{d}t = \int_0^1 P_n(\,f > t\,)\;\mathrm{d}t, \text{ for each } n \in \mathbb{N}. \end{aligned}$$

Next, by Lemma A.5,

$$\{\,f > t\,\} = \{\, s \in S \;|\; f(s) > t \,\}$$

is a P-continuity set, except for at most countably many $t \in [0, \infty)$. Hence, (v) implies:

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

On the other hand, $0 \leq f \leq 1$ also implies that, for each $n \in \mathbb{N}$,

$$|\, P_n(\, f > t\,)\,| \;\;\leq\;\; I_{[0,1]}(t), \;\; \text{for each } t \in [0, \infty),$$

where the common dominating function $I_{[0,1]}$ is the following characteristic function:

$$I_{[0,1]} \,:\, [0, \infty) \;\longrightarrow\; [\,0, 1\,] \,:\, t \longmapsto \begin{cases} 1, & \text{for } 0 \leq t \leq 1 \\ 0, & \text{for } 1 < t \end{cases}$$

Since $I_{[0,1]}$ is Lebesgue-integrable on $[0, \infty)$ and $P_n(\, f > t\,) \longrightarrow P(\, f > t\,)$ for almost every $t \in [0, \infty)$, the Lebesgue Dominated Convergence Theorem implies:

$$\int_0^1 P_n(\, f > t\,)\; \mathrm{d}t \;\;\longrightarrow\;\; \int_0^1 P(\, f > t\,)\; \mathrm{d}t.$$

Combining all of the preceding observations, 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 (v) $\Longrightarrow$ (i). $\square$

Remark 1.2 *In Theorem 1.1 (the Portmanteau Theorem):*

- *The statements (ii), (iii), and (iv) are essentially restatements of each other. (ii) and (iii) are "complemented" versions of each other. And, the middle three inequalities in (iv) hold trivially, while the leftmost is equivalent to (iii) and the rightmost to (ii).*
- *(v) is an immediate consequence of (iv).*
- *All the implications in the Theorem, except (i) $\Longrightarrow$ (ii), require only the fact that S is a topological space; in other words, their validity does NOT explicitly require the metric space structure of S. This is evident in our proof.*
- *On the other hand, our proof of the implication (i) $\Longrightarrow$ (ii) explicitly uses the metric space properties of S. More precisely, the proof invokes the fact that in a metric space, the characteristic function of a closed subset F can be "doubly enveloped" arbitrarily tightly, by the characteristic function of the open ε-neighbourhood F^ε of F, and additionally by a bounded continuous $\mathbb{R}$-valued function bounded between these two aforementioned characteristic functions. This metric space property allows us to deduce the statement (ii) about closed subsets from the statement (i) about bounded continuous $\mathbb{R}$-valued functions.*

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 $\rho(x_n, A) \longrightarrow \rho(x, A)$. We first make the following two claims:

Claim 1: $\rho(x, A) \,-\, \rho(x_n, A) \,\leq\, \rho(x, x_n)$.

Claim 2: $-\rho(x_n, x) \,\leq\, \rho(x, A) \,-\, \rho(x_n, A)$.

The hypothesis $x_n \longrightarrow x$, Claim 1, and Claim 2 together imply:

$$|\,\rho(x, A) \,-\, \rho(x_n, A)\,| \;\leq\; \rho(x, x_n) \;\;\longrightarrow\;\; 0,$$

which proves (i). We now prove the two Claims.

<u>Proof of Claim 1:</u> By the Triangle Inequality, we have

$$\rho(x, A) \;=\; \inf_{a \in A} \rho(x, a) \;\leq\; \rho(x, y) \;\leq\; \rho(x, x_n) \;+\; \rho(x_n, y), \;\; \text{for each } y \in A,$$

which implies

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

This proves Claim 1.

<u>Proof of Claim 2:</u> By the Triangle Inequality, we have

$$\rho(x_n, A) \;=\; \inf_{a \in A} \rho(x_n, a) \;\leq\; \rho(x_n, y) \;\leq\; \rho(x_n, x) \;+\; \rho(x, y), \;\; \text{for each } y \in A,$$

which implies

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

This proves Claim 2.

(ii)

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

□

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*

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

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

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

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,

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

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

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

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

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

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,

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

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 (iii).

□

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

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

The proof of the second inequality is analogous. □

Lemma A.4

Suppose

- $(X,\mathcal{A})$ *is a measurable space, i.e.* X *is a non-empty set and* $\mathcal{A}$ *is a* σ*-algebra of subsets of* X*.*
- Γ *is an uncountable set and* $\bigsqcup_{\gamma\in\Gamma} F_\gamma$ *is a collection, indexed by* Γ*, of pairwise disjoint* $\mathcal{A}$*-measurable subsets* $F_\gamma \in \mathcal{A}$ *of* X*.*

Then, for any finite probability measure μ *on the measurable space* $(X,\mathcal{A})$*, we have:*

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

Proof Let $M_\mu := \mu(X) < \infty$. Define $\Gamma_0 := \{\,\gamma\in\Gamma \mid \mu(F_\gamma)=0\,\}$, and for each $n\in\mathbb{N}$, define

$$\Gamma_\mu(n) := \left\{\gamma\in\Gamma \;\middle|\; \mu(F_\gamma)\ge\frac{1}{n}\right\}.$$

Clearly,

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

Thus, the Lemma follows immediately from the following

> **Claim:** For each $n\ge 1$, Γ_n is a finite set with $|\,\Gamma_n\,| \le n\cdot M_\mu$.
>
> Proof of Claim: If the Claim were false, then there would exist some $n\in\mathbb{N}$ such that $\Gamma_\mu(n)$ contained at least m distinct elements, say $\gamma_1,\gamma_2,\ldots,\gamma_m\in\Gamma_n$, where $m > n\cdot M_\mu$. It would then lead to the following contradiction:
>
> $$M_\mu = \mu(X) \ge \mu\left(\bigsqcup_{i=1}^m F_{\gamma_i}\right) = \sum_{i=1}^m \mu(F_{\gamma_i}) \ge \sum_{i=1}^m \frac{1}{n} = \frac{m}{n} > M_\mu.$$
>
> Thus, the Claim in fact must be true.

□

Lemma A.5

Suppose:

- *X is a topological space and $\mathcal{B}(X)$ is its Borel σ-algebra.*
- *$f : X \longrightarrow \mathbb{R}$ is a continuous $\mathbb{R}$-valued function defined on X.*
- *$P \in \mathcal{M}_1(X, \mathcal{B}(X))$ is a probability measure defined on the measurable space $(X, \mathcal{B}(X))$.*

Then, $f^{-1}(\,(t,\infty)\,) = \{\, x \in X \mid f(x) > t \,\}$ is a P-continuity set, except for at most countably many $t \in \mathbb{R}$.

Proof

First, note that the continuity of f implies that

$$\partial\{\, x \in X \mid f(x) > t \,\} \;\subset\; \{\, x \in X \mid f(x) = t \,\}, \quad \text{for each } t \in \mathbb{R}.$$

Indeed, for each $t \in \mathbb{R}$, we have

$$\begin{array}{cl}
 & x_0 \in \partial\{\, x \in X \mid f(x) > t \,\} \\
\Longleftrightarrow & \text{every neighbourhood of } x_0 \text{ non-trivially intersects each of } \{\, x \in X \mid f(x) > t \,\} \text{ and } \{\, x \in X \mid f(x) \leq t \,\} \\
\Longrightarrow & \text{for each } \varepsilon > 0, \text{ we have } \left\{ \begin{array}{l} f^{-1}(\,(f(x_0)-\varepsilon, f(x_0)+\varepsilon)\,) \cap \{\, x \in X \mid f(x) > t \,\} \neq \varnothing, \text{ and} \\ f^{-1}(\,(f(x_0)-\varepsilon, f(x_0)+\varepsilon)\,) \cap \{\, x \in X \mid f(x) \leq t \,\} \neq \varnothing \end{array} \right. \\
\Longrightarrow & \text{for each } \varepsilon > 0, \text{ we have } \left\{ \begin{array}{l} \exists\; x_\varepsilon \in X \;\text{ with }\; -\varepsilon < f(x_0) - f(x_\varepsilon) < \varepsilon \;\text{ and }\; t < f(x_\varepsilon), \text{ and} \\ \exists\; x'_\varepsilon \in X \;\text{ with }\; -\varepsilon < f(x_0) - f(x'_\varepsilon) < \varepsilon \;\text{ and }\; f(x'_\varepsilon) \leq t \end{array} \right. \\
\Longrightarrow & \text{for each } \varepsilon > 0, \text{ we have } \; t - \varepsilon < f(x_0) \;\text{ and }\; f(x_0) \leq t + \varepsilon, \;\text{ or equivalently }\; |\, f(x_0) - t \,| \;\leq\; \varepsilon \\
\Longrightarrow & f(x_0) \;=\; t
\end{array}$$

Next, note that, since f is continuous, $f^{-1}(\{\,t\,\})$ is $\mathcal{B}(X)$-measurable for each $t \in \mathbb{R}$. Thus,

$$X \;=\; \bigsqcup_{t\in\mathbb{R}} \{\, x \in X \mid f(x) = t \,\} \;=\; \bigsqcup_{t\in\mathbb{R}} f^{-1}(\{\,t\,\})$$

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

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

which in turn implies

$$P(\,\partial\{\, x \in X \mid f(x) > t \,\}\,) \;=\; 0, \;\text{ for all but countably many } t \in \mathbb{R}.$$

This completes the proof of the Lemma. □

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.