MEASURE THEORETICAL PROBABILITY I HOMEWORK 1

SHAWN LIN; CHENG PENG; YIWEI SHI ; TOMMENIX YU
STAT 38100
DUE THU JAN 12, 2023, 11PM

Discussed with classmates.

Exercise 1. *Prob 1.*

Proof.

Claim: Assume $\{F_n\}_{n\in\mathbb{N}}$ is a collection of compact sets in $\mathbb{R}^d$. If $\displaystyle\bigcap_{n\in\mathbb{N}} F_N = \emptyset$, then there exists a finite k such that $\displaystyle\bigcap_{i=1}^{k} F_k = \emptyset$.

(proof by contradiction) Assume for every $k \in \mathbb{N}$ there exists $x_k \in \cap_{i=1}^k F_i$. We denote this sequence by $\{x_k\}_{k\in\mathbb{N}}$. Moreover, $\{x_k\}_{k\in\mathbb{N}} \subset F_1$, which is a compact subset of $\mathbb{R}^n$. By the compactness of F_1, there exists a subsequence $(x_{k_j})_{j\in\mathbb{N}}$ such that x_{k_j} converges to some $x_* \in F_1$. By the construction, we have $(x_{k_j})_{j\geq i} \subset F_i$ for all $i \in \mathbb{N}$. Since each F_i is closed, it must attain x_*. Hence, $x_* \in \cap_{i\in\mathbb{N}} F_i$.

Fix ε, for each E_n, since it is the finite union of $k(n)$ disjoint boxes, we can for each box in that disjoint collection find a smaller closed box with a difference smaller than $\varepsilon\dfrac{1}{2^n k(n)}$. Now take the union of all such closed boxes, call it F_n, we get that

$$\mu(E_n \backslash F_n) \leq \varepsilon k(n) \frac{1}{2^n k(n)} = \frac{1}{2^n}\varepsilon.$$

But since now F_n is closed and $\displaystyle\bigcap_{n\geq 1} F_n \subset \bigcap_{n\geq 1} E_n = \emptyset$, by the claim, we know that there's a finite collection of F_n such that

$$\bigcap_{n=1}^{N} F_n = \emptyset.$$

But since

$$\mu(E_N) = \mu\left(E_N \Big\backslash \bigcap_{n=1}^{N} F_n\right) = \mu\left[\bigcup_{n=1}^{N}\left(E_N \Big\backslash F_n\right)\right] \leq \sum_{n=1}^{N} \mu(E_n \backslash F_n) = \sum_{i=1}^{N} \frac{1}{2^i}\varepsilon \leq \varepsilon,$$

where the inequality uses the fact that E_n is decreasing. Therefore, for every $\varepsilon > 0$, there exists $N(\varepsilon)$ such that

$$\mu(E_N) \leq \varepsilon, \text{ for all } N \geq N(\varepsilon).$$

By the definition of the limit,

$$\lim_{n\to\infty} \mu(E_n) = 0.$$

□

Exercise 2. *Prob 2.*

Proof. First, consider elementary set $E = \bigcupdot_{i=1}^{N} B_i$ (by prop in class) and $\mu(E) = \sum_{i=1}^{N} |B_i|$.

The definition of outer sums

$$\mu^*(F) := \inf \left\{ \sum_{n\geq 1} |B_n| \middle| F \subset \bigcup_{n=1}^{\infty} B_n \right\}$$

immediately tells us for any subset F, $\mu^*(F) \leq \mu(F)$ since the above set is larger than the finite union version (the one in the definition of μ).

Thus, to prove $\mu(E) = \mu^*(E)$ it suffices to prove $\mu(E) \leq \mu^*(E)$.

For any sequence (B_i') of boxes such that $E \subset \bigcup_{i\geq 1} B_i'$ we have by the construction of E that

$$E = \bigcupdot_{i=1}^{N} B_i \subset \bigcup_{i\geq 1} B_i'$$

which implies by taking the volume on both sides

$$\mu(E) = \sum_{i=1}^{N} |B_i| = V(E) \leq V\left(\bigcup_{i\geq 1} B_i' \right)$$

By taking inf of all (B_i') that covers E on both side of the expression above, we have $\mu(E) \leq \mu^*(E)$.

Now, for arbitrary Jordan measurable sets A, from the lecture we know that $\forall \varepsilon > 0$, $\exists$ elementary set B, C such that $A \subset B$, $C \supset B \backslash A$ and $\mu(C) < \varepsilon$.

Then, for $\varepsilon_i = \dfrac{1}{i}$, $\exists B_i, C_i$ such that $A \subset B_i$, $B_i \backslash A \subset C_i$ and $\mu(C_i) \leq \dfrac{1}{i}$. WLOG we can let $C_i = C_i \cap B_i$ (still elementary) which still has $\mu(C_i) \leq \frac{1}{i}$.

Therefore, $A \supset B_i \backslash C_i$ and

$$\begin{aligned}
\mu^*(A) &\geq \mu^*(B_i) - \mu^*(C_i) \geq \mu^*(B_i) - \frac{1}{i} \\
&= \mu(B_i) - \frac{1}{i} \geq \inf \left\{ \sum_{n=1}^{N} |B_n'| \middle| B_i \subset \bigcup_{n=1}^{\infty} B_n' \right\} - \frac{1}{i} \\
&\geq \inf \left\{ \sum_{n=1}^{N} |B_n'| \middle| A \subset \bigcup_{n=1}^{\infty} B_n' \right\} - \frac{1}{i} = \mu(A) - \frac{1}{i}
\end{aligned}$$

then as $i \to \infty$ we have $\mu^*(A) \geq \mu(A)$.

But again by definition $\mu^*(A) \leq \mu(A)$ and hence $\mu(A) = \mu^*(A)$, which is what we want.

□

Exercise 3. *Prob 3.*

Proof.

(1) To show: $\mathcal{L}$ is stable under countable unions.

Consider any A_0 such that $A_0 = \bigcup_{i=1}^{\infty} A_i$ where $A_i \in \mathcal{L}$. Thus, fix any $\varepsilon > 0$, we can find $C_i := \bigcup_{j=1}^{\infty} B_{i,j}$ such that $A_i \subset C_i$ and $\mu^*(C_i \backslash A_i) < \frac{\varepsilon}{2^{i+1}}$ by definition of elements in $\mathcal{L}$.

Now, we know that $A_0 \subset \bigcup_{i=1}^{\infty} C_i := C_0$ since every point in A_0 is in some A_i, hence in the corresponding C_i. (Note that the countable union of countable sets is countable.) And thus

$$\mu^*(C_0 \backslash A_0) \leq \mu^* \left(\bigcup_{i=1}^{\infty} (C_i \backslash A_i) \right) < \varepsilon \cdot \left(\frac{1}{2} + \frac{1}{4} + \dots \right) = \varepsilon$$

thus $\mathcal{L}$ is stable under countable unions.

(2): To show: $\mathcal{L}$ is stable under complements.

Claim: Let $E \subset \mathbb{R}^d$. If $\mu^*(E) = 0$, then $E \in \mathcal{L}$.

Assume $E \subset \mathbb{R}^d$ and $\mu^*(E) = 0$. Let $\varepsilon > 0$. Then, there exists a collection of boxes $\{B_i\}_{i \in \mathbb{N}}$ such that

$$E \subset \cup_{i=1}^{\infty} B_i \text{ and } \sum_{i=1}^{\infty} |B_i| < \varepsilon.$$

On the other hand,

$$\mu^*(\cup_{i=1}^{\infty} B_i \setminus E) \leq \mu^*(\cup_{i=1}^{\infty} B_i) \leq \sum_{i=1}^{\infty} \mu^*(B_i) < \varepsilon.$$

Therefore, every set with Lebesgue outer measure 0 is Lebesgue measurable.

Claim: $\mathcal{L}$ is closed under taking compliment.

Assume $E \in \mathcal{L} \subset \mathbb{R}^d$. Then, by the definition of Lebesgue measurability, there exists a collection of balls $\{B_{n,j}\}_{j=1}^{\infty} \subset \mathbb{R}^d$ such that

$$E \subset \cup_{j=1}^{\infty} B_{i,j} \text{ and } \mu^*(\cup_{j=1}^{\infty} B_{i,j} \setminus E) \leq \frac{1}{n}.$$

We denote $U_n := \cup_{j=1}^{\infty} B_{n,j}$ for $n \in \mathbb{N}$.

As $E \subset U_n$, $U_n^c \subset E^c$ and

$$\mu^*(E^c \setminus U_n^c) = \mu^*(E^c \cap U_n) = \mu^*(U_n \setminus E) \leq \frac{1}{n}.$$

for all $n \in \mathbb{N}$. Let $U_0^c := \cup_{i=1}^{\infty} U_i^c$, then $U_0^c \subset E^c$.

$$\mu^*(E^c \setminus U_0^c) = \mu^*(E^c \setminus \cup_{i=1}^{\infty} U_i^c) \leq \mu^*(E^c \setminus U_n^c) \leq \frac{1}{n} \text{ for all } n \in \mathbb{N}.$$

Thus, $\mu^*(E^c \setminus U_0^c) = 0$ and $E^c \setminus U_0^c$ is measurable.

$$U_0^c = \cup_{i=1}^{\infty} U_i^c = \cup_{i=1}^{\infty} \cap_{j=1}^{\infty} B_{i,j}^c.$$

If we can show U_0^c is measurable, then $E^c = U_0^c \cup (E^c \setminus U_0^c)$ is measurable by part(1).

Consider $X_k = [-k, k]^d$, then for every $k \in \mathbb{N}$, we have

$$\cap_{j=1}^{\infty} \underbrace{(B_{i,j}^c \cap X_k)}_{\text{elementary set}}, \tag{0.1}$$

which is an infinite intersection of elementary sets. Therefore, (0.1) is Lebesgue measurable.

$$U_0^c = \cup_{i=1}^{\infty} \cup_{k=1}^{\infty} \cap_{j=1}^{\infty} (B_{i,j}^c \cap X_k),$$

which is a countable union of Lebesgue measurable sets. $\square$

Exercise 4. *Prob 4.*

Proof.

[(1) $\Longrightarrow$ (2)] Assume $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$ are pairwise disjoint. Define

$$B_n = (\cup_{i=1}^n A_i)^c = \cap_{i=1}^n A_i^c \text{ for } n \in \mathbb{N},$$

which is then a decreasing sequence.

By part (1) assumption,

$$\begin{aligned}\lim_{n\to\infty} \rho(B_n) = \rho(\cap_{n\in\mathbb{N}} B_n) &= \rho((\cup_{n\in\mathbb{N}} B_n^c)^c) \\ &= \rho(X) - \rho(\cup_{n\in\mathbb{N}} B_n^c) \qquad (\rho(X) < \infty \text{ by assumption}) \\ &= \rho(X) - \rho(\cup_{n\in\mathbb{N}} \cup_{i=1}^n A_i) \\ &= \rho(X) - \rho(\cup_{i\in\mathbb{N}} A_i). \end{aligned} \tag{0.2}$$

One the other hand,

$$\rho(X) - \lim_{n\to\infty} \rho(B_n) = \lim_{n\to\infty} \rho(B_n^c) = \lim_{n\to\infty} \sum_{i=1}^n \rho(A_i).$$

Combining with (0.2), we have

$$\rho(\cup_{i\in\mathbb{N}} A_i) = \sum_{i=1}^\infty \rho(A_i).$$

[(2) $\Longrightarrow$ (1)]

We will prove ρ is continuous from below first.

Assume $\{B_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$ is an increasing sequence. Define $B_1 := B_1$, $B_i' = B_i/B_{i-1}$ for $i = 2, 3, ...$ Then the sequence $\{B_i'\}_{i\in\mathbb{N}}$ is pairwise disjoint.

$$\lim_{n\to\infty} \sum_{i=1}^n \rho(B_i') = \lim_{n\to\infty} \rho(\cup_{i=1}^n B_i') = \lim_{n\to\infty} \rho(B_n).$$

By part (2) result,

$$\lim_{n\to\infty} \rho(B_n) = \lim_{n\to\infty} \sum_{i=1}^n \rho(B_i') = \rho(\cup_{n\in\mathbb{N}} B_n') = \rho(\cup_{n\in\mathbb{N}} B_n).$$

For any decreasing sequence $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$, the sequence defined by $C_i = A_1 - A_i$ for $i \in \mathbb{N}$ is a increasing sequence. By the continuity from below result and note that $\rho(A) < \rho(X) \leq \infty$

$$\rho(\cup_{i\in\mathbb{N}} C_i) = \lim_{i\to\infty} \rho(C_i).$$

$$\rho(\cup_{i\in\mathbb{N}} C_i) = \rho(\cup_{i\in\mathbb{N}} A_1 - A_i) = \rho(A_1 - \cap_{i\in\mathbb{N}} A_i) = \rho(A_1) - \rho(\cap_{i\in\mathbb{N}} A_i). \tag{0.3}$$

$$\lim_{i\to\infty} \rho(C_i) = \lim_{i\to\infty} \rho(A_1 - A_i) = \rho(A_1) - \lim_{i\to\infty} \rho(A_i). \tag{0.4}$$

(0.3) and (0.4) together gives

$$\lim_{i\to\infty} \rho(A_i) = \rho(\cap_{i\in\mathbb{N}} A_i).$$

$\square$

Exercise 5. *Prob 5.*

Proof.

Claim: If $\mathcal{A}$ is infinite, $\mathcal{A}$ can not be countably infinite.

We will prove it by contradiction. Assume $\mathcal{A}$ is countably infinite and denoted by $\{A_i\}_{i\in\mathbb{N}}$. For $x \in X$, define

$$G_x = \cap_{x\in A_i} A_i.$$

Since $\{A_i\}$ is countable, $\cap_{x\in A_i}$ is well-defined.

Subclaim: $G_x = G_y$ or $G_x \cap G_y = \emptyset$ for $x, y \in X$.

Assume $G_x \neq G_y$. Then, either $x \in G_x \cap G_y$ or $x \in G_x \cap G_y^c$. In either way, we can find a proper subset of G_x that contains x, which violates the condition that G_x is the intersection of all A_i containing x.

Denote $G = \{G_x\}_{x\in X}$, which is a disjoint collection of sets.

Subclaim: $A_i = \cup_{j\in I} G_j$ for some index set $I \subset X$.

Denote $A_i' = \cup_{x\in A_i} G_x$. By definition, $A_i' \supset A_i$ as G_x contains x. On the other hand, for each $x \in A_i$, $G_x = \cap_{x\in A_j} A_j \subset A_i$. Therefore, $A_i' = \cup_{x\in A_i} G_x \subset A_i$.

If G is finite, then $\sigma(G)$ will be finite as well. As shown in the previous subclaim, each A_i is in $\sigma(G)$, we must have $\sigma(G)$ be infinite. Moreover, as each element in G is disjoint, any two distinct arbitrary unions of elements from G will be distinct from each other and lie in $\mathcal{A}$.

Therefore, there is an injective map from $\mathcal{P}(G)$ to $\mathcal{A}$. However, as $\mathcal{P}(G)$ is countably infinite, $\mathcal{A}$ is at least countably infinite (contradiction).

Now We can conclude that if $\mathcal{A}$ is infinite, it must be countably infinite/uncountable.

Claim: if $\mathcal{A}$ is finite, it has 2^n sets for some $n \in \mathbb{N}$.

We denote $\mathcal{A} = \{A_i\}_{i=1}^N$ for some $N \in \mathbb{N}$. Similar to previous precedes, we define

$$G_x = \cap_{x\in A_i} A_i,$$

which is well-defined. By the exact same proof, we obtain two subclaims as before. Since $\mathcal{A}$ is finite, $G \subset \mathcal{A}$ and G is finite. Moreover, each element of $\mathcal{A}$ can be represented by unions of elements in G uniquely. This is a directly consequence of the subclaim and the fact that elements in G are pairwise disjoint. Therefore, there is a bijective correspondence between $\mathcal{P}(G)$ and $\mathcal{A}$. By the binomial theorem, $|\mathcal{A}| = |\mathcal{P}(G)| = (1 + 1)^{|G|} = 2^{|G|}$.

Combining these two claims together, if $\mathcal{A}$ is a $\sigma-$algebra on a non-empty set X, then, $\mathcal{A}$ is either uncountable or finite with cardinality 2^n for some $n \in \mathbb{N}$.

$\square$

Exercise 6. *Prob 6*

Proof.

(a): Simply construct

$$\sigma(\mathcal{A}) = \bigcap_{\alpha \in \mathcal{A}} F_\alpha$$

where F_α are all σ-algebras that contains $\mathcal{A}$. Since all F_α are $\sigma-$algebra, it is easy to verify that $\sigma(A)$ is stable under countable union and complement and therefore it is a $\sigma-$ algebra.

We know that $P(X)$, the discrete σ-algebra is one term in the intersection, and $\mathcal{A} \subset F_\alpha$, hence in the subset ordering of sets, $\mathcal{A} \leq \sigma(\mathcal{A}) \leq P(X)$, thus it is well defined.

Uniqueness: If there are 2 distinct such sets we know that $\sigma_1(\mathcal{A}) \subset \sigma_2(\mathcal{A})$ and $\sigma_2(\mathcal{A}) \subset \sigma_1(\mathcal{A})$ by their definition (since $\mathcal{A}$ is in both). Thus $\sigma_1(\mathcal{A}) = \sigma_2(\mathcal{A})$.

(b):

We first show that the intersection of any λ-systems is still a λ-system, i.e. for F_i are λ-systems

$$F := \bigcap_{i \in \mathcal{I}} F_i \text{ is still a } \lambda\text{-system.}$$

The reason is that

- $X \in F_i$ since they are all λ-systems, so $X \in F$.
- $\forall A \in F$ we must have $A \in F_i$ for any i. Thus $A^c \in F_i$ for any i, and hence $A^c \in F$.
- $\forall A_1, A_2, \cdots \in F$ be disjoint, we must have that each is in all F_i, and hence the union of them is in all F_i, so $\biguplus_{n=1}^{\infty} A_n \in F$.

Thus, let

$$l(\mathcal{A}) := \bigcap_{\alpha \in \mathcal{I}} l_\alpha(\mathcal{A})$$

where $l_\alpha(\mathcal{A})$ is any λ-systems containing A. The intersection is well defined since $\mathcal{B}$ is an instance. And, for similar reasons as in (a), as well as the argument above, it is the smallest λ-system that contains $\mathcal{A}$.

Now we prove that $\forall A \in l(\mathcal{A})$

$$G(A) := \{B | B \cap A \in l(\mathcal{A})\}$$

is a λ-system.

The reason is that

- $X \cap A = A \in l(\mathcal{A})$, so $X \in G(A)$.
- $\forall B \in G(A)$ we must have $A \cap B \in l(\mathcal{A})$. Since $l(\mathcal{A})$ is a λ-system we have

  $$(A \cap B^c) = (A^c \uplus (A \cap B))^c \in l(\mathcal{A}),$$

  which implies $B^c \in G(A)$.

- $\forall A_1, A_2, \cdots \in G(A)$ be disjoint, we have that (De Morgan)
$$\left(\bigcupdot_{i=1}^{\infty} A_i\right) \cap A = \bigcupdot_{i=1}^{\infty}(A_i \cap A) \in l(\mathcal{A})$$
and hence
$$\left(\bigcupdot_{i=1}^{\infty} A_i\right) \in G(A).$$

Thus $G(A)$ is a λ-system.

Now we show that $l(\mathcal{A})$ is closed under the intersection.

For any $A \in \mathcal{A}$, we know that all B such that $A \cap B \in l(\mathcal{A})$ is in $G(A)$. But note that at the same time for all $B \in \mathcal{A}$, $B \cap A \in \mathcal{A} \subset l(\mathcal{A})$, which means that $\mathcal{A} \subset G(A)$. That is, $G(A)$ is a λ-system containing $\mathcal{A}$. By definition of $l(\mathcal{A})$, $l(\mathcal{A}) \subset G(A)$ for all $A \in \mathcal{A}$.

Now we look at $\forall A \in l(\mathcal{A})$ and $B \in \mathcal{A}$, we want to show that $B \in G(A)$, i.e. $A \cap B \in l(\mathcal{A})$. But using the argument in last paragraph, $A \in l(\mathcal{A}) \subset G(B)$, which means $A \cap B \in l(\mathcal{A})$. This means that $\mathcal{A} \subset G(A)$, and since $G(B)$ is again a λ-system containing $\mathcal{A}$, we get $l(\mathcal{A}) \subset G(A)$ for all $A \in l(\mathcal{A})$.

This means that $\forall A, B \in l(\mathcal{A})$, $B \in l(\mathcal{A}) \subset G(A)$ and hence $A \cap B \in l(\mathcal{A})$, so $l(\mathcal{A})$ is closed under finite intersection.

Now, at last, we show that any λ-system closed under finite intersection is a σ-algebra. To show this we let (X, L) be such a system. Then we check

- $X \in L$ since it is a λ-system.
- $\forall A \in L$, $A^c \in L$ since it is a λ-system.
- Let $A_1, A_2 \dots$ be a sequence of sets in L. Then, Let $B_1 = A_1$ and
$$B_i = A_i \backslash (A_i \cap B_{i-1}) = A_i \cap (A_i \cap B_{i-1})^c,$$
Since B_1, $(A_2 \cap B_1)^c \in L$, therefore $B_2 \in L$ and all $B_i \in L$ we know that B_i are disjoint and hence $\bigcupdot_{i=1}^{\infty} B_i$ is in L. But note that $\bigcupdot_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$, so we are done.

Now that $l(\mathcal{A})$ is a λ-system under finite intersection, so it is a σ-algebra. Hence $\sigma(\mathcal{A}) \subset l(\mathcal{A})$. To show it is $\sigma(\mathcal{A})$, we simply note that all σ-algebras have to be λ-system. Thus, $l(\mathcal{A}) \subset l(\mathcal{A})$. So $l(\lambda(A)) = \sigma(\mathcal{A})$. This means that $\mathcal{B}$ contains $\sigma(\mathcal{A})$ since it contains $l(\mathcal{A})$.

$\square$

Exercise 7. *Prob 7.*

Proof.

($\Rightarrow$):

Assume that E is Jordan measurable, that is $\forall \varepsilon > 0$, $\exists$ elementary F, G such that $F \subset E \subset G$ and $\mu(G \backslash F) < \varepsilon$, where the measure μ is the Jordan measure. Fix an $\varepsilon > 0$.

Now using the proposition in class we know that F, G are the finite union of disjoint boxes. Moreover, the closeness/openness does not affect the proof below (since it doesn't affect the length of an interval). WLOG, we assume that they are closed. Since we are in $\mathbb{R}$, this just means that

$$F = \dot{\bigcup_{i=1}^{n}} [a_i, b_i] \quad \text{and} \quad G = \dot{\bigcup_{j=1}^{m}} [c_i, d_i].$$

We can partition $[a, b]$ into 3 parts:

WLOG, we may assume $G \subset [a, b]$. Otherwise, we can just take the intersection of G and $[a, b]$. For $x \in [a, b]$, either one of the 3 is true,

(1) $x \in F$ and $x \in G$, i.e. $x \in [a_i, b_i]$ and $x \in [c_j, d_j]$ for some suitable i, j.
(2) $x \in G \backslash F$, i.e. $x \in [c_j, d_j]$ for some j but not in any $[a_i, b_i]$.
(3) $x \notin G$, i.e. $x \notin [a_i, b_i]$ and $x \notin [c_j, d_j]$ for all i, j.

Now, since we have only finite points in the set

$$S := \{x | x = a_i \;\; or \;\; b_i \;\; or \;\; c_j \;\; or \;\; d_j\}$$

and let P be a tagged subdivision of $[a, b]$ that contains an ordered collection of points in S, with arbitrary chosen x_k. All these are pretty legit as S is finite.

Now we plug in the definition of the Jordan measure μ for elementary sets to get (with disjoint addition property)

$$\mu(F) = \sum_{i=1}^{n} \mu([a_i, b_i]) = \sum_{i=1}^{n} (b_i - a_i)$$

$$\mu(G) = \sum_{j=1}^{m} \mu([c_i, d_i]) = \sum_{j=1}^{m} (d_i - c_i)$$

and (disjoint set addition)

$$\sum_{j=1}^{m} (d_i - c_i) - \sum_{i=1}^{n} (b_i - a_i) = \mu(G) - \mu(F) = \mu(G \backslash F) < \varepsilon$$

Then since we have the inequality

$$\mathbb{1}_F \leq \mathbb{1}_E \leq \mathbb{1}_G,$$

which gives us the following inequality

$$RS(\mathbb{1}_F, P) \leq RS(\mathbb{1}_E, P) \leq RS(\mathbb{1}_G, P).$$

Yet as we have discussed, $[a, b]$ can be partitioned into 3 sets. For $x \in F$, $\mathbb{1}_F(x) = \mathbb{1}_G(x) = 1$, for $x \notin G$, $\mathbb{1}_F(x) = \mathbb{1}_G(x) = 0$, and for the other case $\mathbb{1}_F(x) = 0; \mathbb{1}_G(x) = 1$.

Also, we note that for any subdivision in P, all x in that subdivision is in the same category of x as partitioned above.

Thus

$$RS(\mathbb{1}_G, P) - RS(\mathbb{1}_F, P) = \sum_{k=1}^{|S|} (\mathbb{1}_G - \mathbb{1}_F)(x_k)(s_k - s_{k-1}) = \sum_{k \in K} 1 \cdot (s_k - s_{k-1})$$

where $K := \{k | \mathbb{1}_G(x) = 1, \mathbb{1}_F(x) = 0, \forall x \in [s_{k-1}, s_k]\}$.

Now note that all such points are in G but not in F, so they are in $G \backslash F$; also all such points are in $[s_{k-1}, s_k]$ for $k \in K$. Thus we have

$$\sum_{k \in K} (s_k - s_{k-1}) = \sum_{j=1}^{m} (d_i - c_i) - \sum_{i=1}^{n} (b_i - a_i) < \varepsilon.$$

Since $\sum_{i=1}^{n} (b_i - a_i)$ is fixed call it s, we have

$$|RS(\mathbb{1}_E, P) - s| \leq |RS(\mathbb{1}_G, P) - s| < \varepsilon.$$

Now we find a δ such that all P_N with $\delta(P_N) < \delta$ has that $|RS(\mathbb{1}_E, P_N) - s| < \varepsilon$. We simply choose the minimal mesh, m, in the P created above and let $\delta = m/2$.

Now, since all the intervals in P is in one of the 3 categories specified above, if it's contained in both G, F or neither, then the value of $\mathbb{1}_E$ is constant there. If it's in $G \backslash F$, then we note that by properties of the real line (differences of intervals are intervals) there's only 1 discontinuous point in that mesh interval such that the function $\mathbb{1}_E$ has different value on both sides.

That is, for each such interval (with varying value of $\mathbb{1}_E$) in P, there is one and only 1 corresponding interval in $\P_N$ such that it also has varying value of $\mathbb{1}_E$. Yet the mesh of P_N is less than half of the length of the smallest interval in P, so the upper and lower sum on P_N is bounded in the following manner:

$$0 = |RS(\mathbb{1}_F, P) - s| \leq |RS(\mathbb{1}_E, P_N) - s| \leq |RS(\mathbb{1}_G, P) - s| \leq \varepsilon$$

since

$$s = RS(\mathbb{1}_F, P) \leq RS(\mathbb{1}_E, P_N) \leq RS(\mathbb{1}_G, P).$$

($\Leftarrow$):

Assume that E is Riemann integrable, we then have $\forall \varepsilon$ there exists an s'_ε and δ such that for all P_N with mesh smaller than δ we have

$$|RS(\mathbb{1}_E, P_N) - s'_\varepsilon| < \varepsilon.$$

Now fix one such subdivision satisfying the above condition with grid points $s_0, \dots, s_N$.

Note that the choice of a tagged subdivision contains the choice of x_i, and since the tagged subdivision is arbitrary, we can choose x_i as we want.

So we can choose any x_i in the grid points $s_0, \dots, s_N$ as we want. So we pick the upper and lower, i.e. we choose P_N^1, P_N^2 such that the grid points are $s_0, \dots, s_N$ with

$$x_i^1 \in \left\{ x_i \in [s_{i-1}, s_i] \middle| \mathbb{1}_E(x_i) = \sup_{x_i \in [s_{i-1}, s_i]} \mathbb{1}_E(x_i) \right\}$$

and

$$x_i^2 \in \left\{ x_i \in [s_{i-1}, s_i] \middle| \mathbb{1}_E(x_i) = \inf_{x_i \in [s_{i-1}, s_i]} \mathbb{1}_E(x_i) \right\}.$$

They are well-defined since $f(x_i)$ are either 0 or 1.

Now we define F and G based on the construction of P_N^1 and P_N^2. We first note that, let x_i^1 and x_i^2 be the corresponding choice of points in the tagged subdivision P_N^1 and P_N^2, then all the mesh intervals can be partitioned into 3 type of closed intervals:

(1) $\mathbb{1}_E(x_i^1) = \mathbb{1}_E(x_i^2) = 1$;
(2) $\mathbb{1}_E(x_i^1) = 0, \mathbb{1}_E(x_i^2) = 1$;
(3) $\mathbb{1}_E(x_i^1) = \mathbb{1}_E(x_i^2) = 0$.

Let the union of all mesh intervals that is of partition (1) be the set F, then

$$F = \bigcup_{i \in \mathcal{I}_1} [s_{i-1}, s_i] = \bigcupdot_{j \in \mathcal{J}_1} [s_{j-1}, s_j]$$

and let the union of all mesh intervals that is of partition (1) and (2) be the set G, then

$$G = \bigcup_{i \in \mathcal{I}_2} [s_{i-1}, s_i] = \bigcupdot_{j \in \mathcal{J}_2} [s_{j-1}, s_j].$$

Note that the above rearrange into disjoint union is true since if 2 adjacent closed interval are all in the collection, we can just combine them into 1 larger closed interval. Everything is finite so we are good to do so repeatedly until we are done.

But now we have that F and G are all elementary sets and

$$\mu(G \backslash F) = \mu(G) - \mu(F) = \sum_{j \in \mathcal{J}_1} (s_j - s_{j-1}) - \sum_{j \in \mathcal{J}_2} (s_j - s_{j-1}) = |RS(\mathbb{1}_E, P_N^2) - RS(\mathbb{1}_E, P_N^1)| < 2\varepsilon$$

by triangle inequality and the construction of F and G.

Also, since if $x \in F$, $\mathbb{1}_E(x) = 1$ we have $F \subset E$; If $x \notin G$ we have $\mathbb{1}_E(x) = 0$ thus $E \subset G$. Thus $F \subset E \subset G$ and $\mu(G \backslash F) < 2\varepsilon$ for arbitrary ε, we've shown that E is Jordan measurable.

($\boldsymbol{\mu(\mathbf{E}) = \mathbf{R}_{[a,b]}(\mathbb{1}_{\mathbf{E}})}$):

By definition

$$\mu(E) = \inf_{\substack{A \subset G \\ G-elementary}} \mu(G)$$

but this is nothing but the limit of the constant $s_\varepsilon := \sum_{i=1}^n (b_i - a_i)$ for $F = \bigcupdot_{i=1}^n [a_i, b_i]$ as defined in the ($\Rightarrow$) part. Note that F is actually constructed after fixing ε, so it depends on ε. i.e.

$$\mu(E) = \inf_{\substack{A \subset G \\ G-elementary}} \mu(G) = \lim_{\varepsilon \to 0} s_\varepsilon + \varepsilon = \lim_{\varepsilon \to 0} s_\varepsilon .$$

On the other hand

$$R_{[a,b]}(\mathbb{1}_E) = \lim_{\varepsilon \to 0} s'_\varepsilon$$

where s'_ε is such that

$$|RS(\mathbb{1}_E, P_N) - s'_\varepsilon| < \varepsilon .$$

Again, we've shown in the ($\Rightarrow$) part that

$$|RS(1_E, P_N) - s_\varepsilon| < \varepsilon$$

which by triangle inequality yields

$$|s_\varepsilon - s'_\varepsilon| < 2\varepsilon \Rightarrow \lim_{\varepsilon \to 0} s_\varepsilon = \lim_{\varepsilon \to 0} s'_\varepsilon$$

which means

$$\mu(E) = R_{[a,b]}(\mathbb{1}_E).$$

$\square$

Method 2 for backward direction:

Assume E is Riemann integrable. Let $\varepsilon > 0$, then there exists $\delta_\varepsilon > 0$ such that for all tagged partition P_N, we have

$$RS(\mathbb{1}_E, P_N) < \frac{\varepsilon}{2}.$$

Denote $f(x) := \mathbb{1}_E(x)$ and $s := \int_{[a,b]} f(x)dx$. Let $P = \{s_0, \dots, s_N\}$ be any partition with $\delta(P) < \delta_\varepsilon$. Then, we define

$$\overline{x_i} := \arg \max_{x \in [s_{i-1}, s_i]} f(x)$$

$$\underline{x_i} := \arg \min_{x \in [s_{i-1}, s_i]} f(x),$$

for $i = 1, \ldots, N$. The above two expressions are well-defined as there are only two possible values for $f(x)$. Moreover, we have

$$|\sum_{1\leq i\leq N} f(\overline{x_i})(s_i - s_{i-1}) - s| \leq \frac{\varepsilon}{2}$$
$$|\sum_{1\leq i\leq N} f(x_i)(s_i - s_{i-1}) - s| \leq \frac{\varepsilon}{2}.$$

Let $A = \bigcup_{\substack{[s_{i-1},s_i]\subset E \\ i=1,\ldots,N}}[s_{i-1}, s_i]$ and $B = \bigcup_{\substack{[s_{i-1},s_i]\cap E \\ i=1,\ldots,N}}[s_{i-1}, s_i]$.

$$\sum_{1\leq i\leq N} (f(\overline{x_i})(s_i - s_{i-1}) = \sum_{[s_{i-1},s_i]\cap E\neq\emptyset} |s_{i-1} - s_i| = \mu(B),$$

as each interval are almost disjoint.

$$\sum_{1\leq i\leq N} f(\overline{x_i})(s_i - s_{i-1}) = \sum_{[s_{i-1},s_i]\subset E} |s_{i-1} - s_i| = \mu(A),$$

as intervals are almost disjoint.

$$\begin{aligned}\mu(B \setminus A) = \mu(B) - \mu(A) = \mu(B) - s + s - \mu(A) \\ &= \sum_{1\leq i\leq N} f(\overline{x_i})(s_i - s_{i-1}) - s + s - \sum_{1\leq i\leq N} f(x_i)(s_i - s_{i-1}) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon.\end{aligned}$$

Hence, for every $\varepsilon > 0$, there exists A, B elementary such that $A \subset E \subset B$ and $\mu(B\setminus A) < \varepsilon$. We now can conclude that E is Jordan measurable.

In the above proof, we have seen the construction of elementary A and B such that

$$A \subset E \subset B \text{ and } \mu(B) - \mu(A) < \varepsilon.$$

Moreover,

$$\begin{aligned}|\mu(E) - s| &= |\mu(E) - \mu(B) + \mu(B) - s| \\ &\leq |\mu(E) - \mu(B)| + |\sum_{1\leq i\leq N} (f(\overline{x_i})(s_i - s_{i-1}) - s| \\ &\leq \varepsilon + \frac{\varepsilon}{2} \\ &= \frac{3}{2}\varepsilon.\end{aligned}$$

Since the choice of ε is arbitrary, we have $\mu(E) = s = R_{[a,b]}(\mathbb{1}_E)$.