Advanced Probability Theory

Jia Cheng

December 2021

Reference Text: A First Look at Rigorous Probability Theory

Related modules: ST5214

Notation Throughout this document, \mathbb{N} can vary in usage, sometimes $\mathbb{N} = \mathbb{Z}^+$, other times $\mathbb{N} = \mathbb{Z}_0^+$, depending on which is more convenient.

1.1 To see that $\exists z \in \mathbb{R}, P(Z=z) > 0$, we observe that

$$P(Z=0) \ge P(Z=0 \land X=0) = P(Z=0|X=0) \cdot P(X=0) \ge \frac{1}{2} \cdot P(X=0) > 0$$

since P(X=0) > 0.

1.2

Uncountable summation Given an uncountable non-negative set of numbers $\{r_a : a \in I\}$ indexed by I,

$$\sum_{a\in I} r_a := \sup\{\sum_{a\in J} r_a : J\subseteq I \wedge J \text{ finite}\}\$$

R-shift (Equivalent definition) R-shift of $A \subseteq [0,1]$. $A \oplus r = \{(a+r) \mod 1 : a \in A\}$

2.1 Notice that countability is used by 2 constructs. One, probability measure is countably additive (and not uncountably so). Two, the σ -algebra is closed under countable union and intersection (and not uncountably so).

Recall that the reason for disallowing uncountable operations in general is due to the fact that

$$\bigcup_{x \in A} \{x\} = A$$

for any set A, in particular [0,1] when discussing the uniform distribution on the unit interval.

Theorem 2.2.1 We provide a proof for this theorem.

First, we show that \mathcal{F} is a σ -algebra. By definition, $\mathcal{F} = \mathcal{P}(\Omega)$. Hence, the unary complement operation is a mapping $\mathcal{P}(\Omega) \to \mathcal{P}(\Omega)$ whose domain and codomain are just \mathcal{F} as desired. Similarly, for the countable set operations of union and intersection, they are mappings with codomain as $\mathcal{P}(\Omega)$ and are also closed since $\mathcal{F} = \mathcal{P}(\Omega)$.

We also note that both \emptyset , Ω reside in \mathcal{F}

Next, we show that P is a probability measure. By definition of P, P is additive since $A \cap B = \emptyset \implies P(A \sqcup B) = \sum_{\omega \in A \sqcup B} p(\omega) = \sum_{\omega \in A} p(\omega) + \sum_{\omega \in B} p(\omega) = P(A) + P(B)$.

I am not quite sure about showing countable additivity however, perhaps using some form of diagonal summation argument it is possible to prove this.

 $P(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$. Furthermore, p is non-negative, hence P is indeed bounded between 0 and 1.

Ex 2.2.3 First, $\emptyset, \Omega = [0, 1] \in \mathcal{J}$ by definition as they are intervals. Next, to show closure under finite intersection, it suffices to show closure under binary intersection. Consider cases: We only consider one endpoint, since we can "patch" together two endpoints.

- $[a \text{ and } [b \text{ intersect to give } [\max\{a, b\}$
- (a and (b intersect to give $(\max\{a, b\})$
- [a and (b intersect to give [a if a > b and (b otherwise)]

We can do a similar case analysis for right endpoints. Given a stringified left endpoint $l \in \{"[a","(a")]\}$ and a stringified right endpoint $r \in \{"b]","b)"\}$ we can form an interval via concatenation l, r. Hence, \mathcal{J} is closed under finite intersection.

Consider the complement $J = [0, a) \cup (b, 1]$ of an interval $[a, b] \in \mathcal{J}$, where depending on whether the left/right endpoint is closed or open we adjust J accordingly. Regardless, we see that J is a disjoint union of at most 2 intervals in \mathcal{J} .

Hence, \mathcal{J} is a semialgebra of subsets of Ω .

Ex 2.2.5

a $\mathcal{B}_0 \subseteq \mathcal{P}(\Omega)$. Since \mathcal{B}_0 consists of all finite unions of elements of \mathcal{J} , in particular, $\mathcal{J} \subseteq \mathcal{B}_0$, so $\Omega = [0,1] \in \mathcal{J} \subseteq \mathcal{B}_0$.

Next, the finite union and intersection of elements of \mathcal{B}_0 will give finite unions of elements of semialg, so that \mathcal{B}_0 is closed under finite union and intersection. (For intersection, we can argue using distributive law plus observe that the intersection of intervals gives another interval)

Let $B \in \mathcal{B}_0$, so that B is a finite union of the form $\bigcup_{1 \leq i \leq n} I_i$ for some intervals I_i in [0,1]. Then, $B^c = \bigcap_{1 \leq i \leq n} I_i^c$ by DeMorgan's Law, and we have already proven in Ex 2.2.3 that I_i^c is a disjoint union of intervals, i.e. $I_i^c \in \mathcal{B}_0$. Furthermore, we have proven that \mathcal{B}_0 is closed under finite intersection, so $B^c \in \mathcal{B}_0$.

Hence, \mathcal{B}_0 is an algebra.

b The difference between an algebra and a σ -algebra is that σ -algebras are closed under countable union and intersection but algebras are not necessarily so.

We consider Cantor's set C, which is a countable intersection of C_i , where each C_i is formed by removing from each interval in C_{i-1} the middle one-third.

Since each C_i is a union of (disjoint) intervals, by definition, $C_i \in \mathcal{B}_0$. If \mathcal{B}_0 is to be a σ -algebra, then we must have $C \in \mathcal{B}_0$, i.e. C can be formed from a finite union of intervals.

First of all, C does not contain any interval of non-zero length, so our options are reduced to forming C from a finite union of singletons, i.e. intervals of the form $[a,a]=\{a\}$. But we also know C to be uncountable, but a finite union of singletons is finite. Hence we have a contradiction.

Since $C \notin \mathcal{B}_0$, \mathcal{B}_0 is not closed under countable intersection, so it is not a σ -algebra.

We comment that \mathcal{B}_1 is similarly not a σ -algebra using the same counterexample. The countable union of singletons must be at most countable, so they cannot union to form an uncountable set like C.

Ex 2.3.16 Suppose $A \in \mathcal{M} \wedge P^*(A) = 0$ and $B \subseteq A$. To show that $B \in \mathcal{M}$, we show equivalently that:

For each $E \subseteq \Omega$, $P^*(E) \ge P^*(A \cap E) + P^*(A^c \cap E)$ (i.e. superadditivity. But by monotonicity, $P^*(A \cap E) = 0$, so that $P^*(E) \ge P^*(A^c \cap E)$ is automatically true by monotonicity. Hence, we have shown the completeness of the extension $(\Omega, \mathcal{M}, P^*)$.

See https://mathoverflow.net/questions/11554/whats-the-use-of-a-complete-measure

Remark A complete measure is good in the sense that it gives us more measurable sets, the more things we can measure, the better. For instance, one consequence of a complete measure is that when A is measurable, and B differs from A by a subset of a set of zero measure, then B is measurable as well.

Ex 2.4.3

a We make the following manipulations.

$$I \subseteq \bigcup_{1 \le j \le n} I_j = \sqcup_{1 \le j \le n} I'_j$$

where $I'_j = I_j - \bigcup_{1 \le l < j} I_l$. Hence,

$$I = I \cap \sqcup_{1 \le j \le n} I'_j = \sqcup 1 \le j \le n(I \cap I'_j)$$

such that

$$P(I) = P(\sqcup 1 \le j \le n(I \cap I'_j))$$

$$= \sum_{1 \le j \le n} |I \cap I'_j|$$

$$\le \sum_{1 \le j \le n} |I'_j|$$

$$\le \sum_{1 \le j \le n} |I_j|$$

$$= \sum_{1 \le j \le n} P(I_j)$$

where the first equality is proven in proposition 2.4.2 (disjoint union of intervals to form a single larger interval), and the latter inequalities are given by monotonicity of the length function $|\cdot|$.

b In $[0,1] \subseteq \mathbb{R}$, I is a closed and bounded interval, hence compact (Heine-Borel Theorem). Given countable cover $I_i, i \in \mathbb{N}$, there exists a finite subcover of I, i.e. $I \subseteq \bigcup_{1 \le j \le n} I_{i_j}$. By part (a), we know that $P(I) \le \sum_{1 \le j \le n} P(I_{i_j}) \le \sum_j P(I_j)$ since P is non-negative.

c We want to generalize our result in (b), i.e. I can be any interval, not just closed, and I_j can be any interval, not just open.

We extend each $I_j, j \ge 1$ to form $I'_j := (a_j - \epsilon 2^{-j-1}, b_j + \epsilon 2^{-j-1})$, such that $|I'_j| = |I_j| + \epsilon 2^{-j}$.

We compress I to form $I' := [a + \epsilon, b - \epsilon]$. This assumes $I \neq \emptyset$, since if I is empty, we trivially have $P(I) = 0 \leq \sum_{i} P(I_{i})$.

We note that each I'_j may very well exceed the boundaries of [0,1], such that P may not be defined for I'_j , but it doesn't matter, since the length function $|I'_j|$ is still well-defined.

Hence,

$$P(I) = |I| = |I'| + 2\epsilon$$

and

$$|I'| \le \sum_{j} |I'_{j}| = \sum_{j} |I_{j}| + \epsilon = \sum_{j} P(I_{j}) + \epsilon$$

Combining gives

$$P(I) \le \sum_{i} P(I_{i}) + 3\epsilon$$

which proves countable monotonicity in general for all sets $I, I_1, I_2, \dots \in \mathcal{J}$.

An unsuccessful attempt: I tried to use closure $I' := \overline{I} = [a, b]$, but it doesn't seem possible to show that the endpoints lie in some open interval.

Ex 2.4.5 Claim: $\sigma(\{(-\infty, b] : b \in \mathbb{R}\}) = \sigma(\{I \subseteq \mathbb{R} : I \text{ interval}\})$

We first note that for collections \mathcal{A}_{∞} , \mathcal{A}_{\in} , $\mathcal{A}_{\infty} \subseteq \mathcal{A}_{\in} \implies \sigma(\mathcal{A}_{\infty}) \subseteq \sigma(\mathcal{A}_{\in})$, since a Borel algebra that contains all elements of \mathcal{A}_{\in} also contains all elements of \mathcal{A}_{∞} .

Let $\mathcal{A}_{\infty} = \{(-\infty, b] : b \in \mathbb{R}\}$ and \mathcal{A}_{\in} be the set of all intervals in \mathbb{R} . Let $\sigma_1 = \sigma(\mathcal{A}_{\infty}), \sigma_2 = \sigma(\mathcal{A}_{\in})$. Clearly, $\sigma_1 \subseteq \sigma_2$.

Next, we make the following 4 steps

- 1. First, for each $b, (-\infty, b] \in \sigma_1$ by definition
- 2. By closure under complements, for each $a, (-\infty, a]^c = (a, \infty) \in \sigma_1$
- 3. By closure under countable union, $\bigcup_{n\in\mathbb{N}}(-\infty,b-\frac{1}{n}]=(-\infty,b)\in\sigma_1$
- 4. By closure under countable intersection $\bigcup_{n\in\mathbb{N}}(a-\frac{1}{n},\infty)=[a,\infty)\in\sigma_1$

Now we observe that $\forall a, b \in \mathbb{R}, a \leq b$,

- $[a,b] = [a,\infty) \cap (-\infty,b]$
- $(a,b) = (a,\infty) \cap (-\infty,b)$
- $[a,b) = [a,\infty) \cap (-\infty,b)$
- $(a,b] = (a,\infty) \cap (-\infty,b]$

hence by closure, σ_1 contains all intervals, i.e. $A_2 \subseteq \sigma_1$. By minimality of Borel algebra σ_2 , we have $\sigma_2 \subseteq \sigma_1$. Hence $\sigma_1 = \sigma_2$.

(We note that a Borel algebra of a set A is a subset of any σ -algebra containing the same set A)

A Borel algebra $\mathcal{B} = \sigma(A)$ reminds me of the subgroup $\langle S \rangle$ generated by a subset S of a group. $\langle S \rangle$ is equal to the set of all words formed by the (finite) product of elements of S, whereas \mathcal{B} is the set of all countable unions, countable intersections, complements of elements of A.

Ex 2.4.7

- a K is the countable intersection of intervals remaining at each step, whereas K^c is the countable union of the intervals removed at each step of the construction. Hence, by closure of σ -algebras under countable union and intersection, $K, K^c \in \mathcal{B}$.
- **b** Since $\mathcal{B} \subseteq \mathcal{M}, K, K^c \in \mathcal{M}$
- **c** Algebras are closed under countable union, so $K^c \in \mathcal{B}_1$
- **d** K is an uncountable set. Furthermore, K contains no interval of non-zero length. It is not possible to obtain K with the countable union of singletons. Hence $K \notin \mathcal{B}_1$.
- e Since \mathcal{B}_1 is not closed under complement, \mathcal{B}_1 is not a σ -algebra. (We can also say that \mathcal{B}_1 is not closed under countable intersection.)
- **2.5.4** To explicitly show the claim that $D_n = B_n \cap \bigcap_{1 \leq i \leq n-1} B_i^c$ is a disjoint union of elements of \mathcal{J} , we let each $B_n^c = \bigcup_{1 \leq j \leq l_n} F_{n,j}$ where each $F_{n,j} \in \mathcal{J}$.

Then

$$B_n \cap \bigcap_{1 \le i \le n-1} B_i^c = B_n \cap \bigcap_{1 \le i \le n-1} \sqcup_{1 \le j \le l_n} F_{n_j}$$

By distributive law, this resolves to a union of terms of the form $B_n \cap F_{1,\lambda_1} cap F_{2,\lambda_2} \cap \cdots \cap F_{n-1,\lambda_{n-1}}$ where for each $1 \leq i \leq n-1$, $1 \leq \lambda_i \leq l_i$. Since each term is a finite intersection, it resides in \mathcal{J} .

Any 2 distinct terms of the intersection, $B_n \cap F_{1,\lambda_1} \cap F_{2,\lambda_2} \cap \cdots \cap F_{n-1,\lambda_{n-1}}$, and $B_n \cap F_{1,\mu_1} \cap F_{2,\mu_2} \cap \cdots \cap F_{n-1,\mu_{n-1}}$ must have some i for which $\lambda_i \neq \mu_i$. This implies that $F_{i,\lambda_i} \cap F_{i,\mu_i} = \emptyset$ so the 2 terms must also be disjoint.

Ex 2.5.6 Suppose P satisfies finite additivity, and that given a monontonically decreasing set sequence (A_i) , with each A_i being a finite disjoint union of elements of \mathcal{J} and $\bigcap_n A_n = \emptyset$, that $\lim_{n \to \infty} P(A_n) = 0$. (Note that P is extended to sets A_n by finite additivity, mentioned in the errata document)

Suppose we have a countably many pairwise disjoint sets $D_n, n \in \mathbb{N}$. Then, we note that we can write

$$\sqcup_i D_i = D_1 \sqcup D_2 \sqcup \cdots \sqcup D_n \sqcup (\sqcup_{i>n} D_i)$$

We now show some properties of the tail union $\sqcup_{i>n} D_i$. First, letting $A_n = \sqcup_{i>n} D_i = \sqcup_i D_i \cap D_1^c \cap D_2^c \cap \cdots \cap D_n^c$. Similar to **2.5.4**, we can show that A_n is a disjoint union of elements of \mathcal{J} . Furthermore, $A_n \supseteq A_{n+1}$ and $\cap_n A_n = \emptyset$ are easy to verify. Hence, by assumption, $\lim_{n\to\infty} P(A_n) = 0$.

Hence,

$$P(\sqcup_i D_i) = \lim_{n \to \infty} (P(D_1) + \dots + P(D_n) + P(A_n)) = \lim_{n \to \infty} \sum_{1 \le i \le n} P(D_i) = \sum_i P(D_i)$$

as desired.

2.5.9 Let $S = \{(-\infty, x] : x \in \mathbb{R}\}, \mathcal{J}_1 = \{(-\infty, x] : x \in \mathbb{R}\} \cup \{(y, \infty) : y \in \mathbb{R}\} \cup \{(y, x] : y, x \in \mathbb{R}\} \cup \{\emptyset, \mathbb{R}\}, \mathcal{J}_2$ be the set of all intervals on \mathbb{R} . Then $S \subseteq \mathcal{J}_1 \subseteq \mathcal{J}_2$. (Note that S is not a semialgebra but \mathcal{J}_1 is.)

Then $\sigma(S) \subseteq \sigma(\mathcal{J}_1) \subseteq \sigma(\mathcal{J}_2)$. But we have proven in **Ex 2.4.5** that $\sigma(S) = \sigma(\mathcal{J}_2)$, hence $\sigma(\mathcal{J}_1) = \sigma(\mathcal{J}_2)$.

Ex 2.6.1

a First, we check that indeed, \emptyset , $\Omega \in \mathcal{J}$.

Next, we check that \mathcal{J} is closed under finite intersection.

- If \emptyset is part of an intersection, then we just get \emptyset .
- If Ω is part of an intersection, then we can just ignore it.
- Now consider $B_1 \cap B_2$, where $B_1 = A_{a_1,a_2,...,a_m}, B_2 = A_{a'_1,a'_2,...,a'_n}$. We then consider cases. If WLOG, a_1,\ldots,a_m is a prefix of a'_1,\ldots,a'_n , then $B_2 \subseteq B_1$ and $B_1 \cap B_2 = B_1 \in \mathcal{J}$. If neither is a prefix of the other, then $B_1 \cap B_2 = \emptyset \in \mathcal{J}$.

Finally, consider any $B \in \mathcal{J}$. The case where $B = \emptyset \vee B = \Omega$ is trivial. So suppose $B = A_{a_1,...,a_n}$. Then B^c is the finite disjoint union of $A_{a'_1,...,a'_n}$ where $\bigvee_{1 \leq i \leq n} a'_i \neq a_i$. Hence, the complement of any sets of \mathcal{J} can be formed by a finite disjoint union of sets in \mathcal{J} .

Hence \mathcal{J} is a semialgebra.

b Suppose we have N disjoint sets $\{D_1, \ldots, D_n\}$.

Following the hint, letting $k \in \mathbb{N}$ be the largest number such that a_1, a_2, \ldots, a_k is a shared prefix amongst all sets D_1 to D_n .

Let $A = A_{a_1,...,a_k}$. Then, we clearly have $A \supseteq \bigcup_{1 \le i \le n} D_i$. We claim that this is in fact an equality. Note that by assumption, the union of all D_i is a set in the semialgebra \mathcal{J} , so the union must be of the form $A_{a_1,...,a_k,...,a_j}$ for some $j \ge k$. So we just want to show that j = k. Suppose not, such that j > k, then a_1, \ldots, a_j is a shared prefix, but this contradicts the largest-ness of k.

Hence, we have proven $\sqcup_{1 \leq i \leq n} D_i = A$.

Next, we want to show finite additivity, i.e. $P(A) = \sum_{1 \leq i \leq n} P(D_i)$. Let K be the length of the longest sequence specified by D_i , $1 \leq i \leq n$. For e.g. if $D_1 = A_{1,0,1}$, $D_2 = A_{1,1}$, then the longest sequence is of length K = 3 (and k = 1).

We shall prove this by (finite) induction.

Define for each $m \in \{k, k+1, \dots, K\}$ the statement Q(m):

Given any finite collection of disjoint sets D_i whose longest specified sequence is of length K and whose longest disjoint union is $A^m := A_{a_1,...,a_m}$ for some length m sequence, then $\bigsqcup_{1 \leq i \leq n} P(D_i) = P(A^m)$.

The base case, Q(K) is trivial since the finite collection can only consist of 1 set, A^K itself. So we have $P(D_1) = P(A^K)$.

Next, suppose Q(j) holds for some $k < j \le K$. Then, we want to show Q(j-1). We partition the set $D_i, 1 \le i \le n$ into 2 subcollections, B_1, \ldots, B_{n_1} and C_1, \ldots, C_{n_2} , such that each $B_i \subseteq A_{a_1,\ldots,a_j,0}$ and each $C_i \subseteq A_{a_1,\ldots,a_j,1}$. It is not too hard to argue that $\bigsqcup_{1 \le i \le n_1} B_i = A_{a_1,\ldots,a_j,0}$ and $\bigsqcup_{1 \le i \le n_2} C_i = A_{a_1,\ldots,a_j,1}$. Then, applying inductive assumption Q(j), we have $\sum_{1 \le i \le n_1} P(B_i) = P(A_{a_1,\ldots,a_j,0})$ and $\sum_{1 \le i \le n_2} P(C_i) = P(A_{a_1,\ldots,a_j,1})$. Hence,

$$P(A^{j}) = \frac{1}{2^{j}} = \frac{1}{2^{j+1}} + \frac{1}{2^{j+1}} = P(A_{a_{1},\dots,a_{j},0}) + P(A_{a_{1},\dots,a_{j},1}) = \sum_{1 \le i \le n} P(D_{i})$$

This proves Q(j).

By mathematical induction, we have $\forall m \in \{k, k+1, \ldots, K\}, Q(m)$, in particular, Q(k) is true. This completes the proof.

Ex 2.6.4 Clearly $\Omega = \Omega_1 \times \Omega_2 \in \mathcal{J}$ and $\emptyset = \emptyset \times \emptyset \in \mathcal{J}$.

Next, if $A_1 \times B_1$, $A_2 \times B_2 \in \mathcal{J}$, then $(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2) \in \mathcal{F}_1 \times \mathcal{F}_2 = \mathcal{J}$, so \mathcal{J} is closed under finite intersection.

Finally, if $A \times B \in \mathcal{J}$, then $(A \times B)^c = (A^c \times \Omega_2) \cup (\Omega_1 \times B^c) = (A^c \times B^c) \cup (A^c \times B) \cup (A \times B^c)$. Note that $\mathcal{F}_1, \mathcal{F}_2$ are σ -algebra, so $A^c \in \mathcal{F}_1 \wedge B^c \in \mathcal{F}_2$. Hence \mathcal{J} is a semialgebra of $\Omega_1 \times \Omega_2$.

We also verify that $P(\emptyset) = P_1(\emptyset)P_2(\emptyset) = 0$ and $P(\Omega) = P_1(\Omega_1)P_2(\Omega_2) = 1$.

Ex 3.1.4

Here, we use the property that $[a, b] \cap [c, d] = [\max\{a, c\}, \min\{b, d\}].$

We write ω as w. $Z(w) > a \iff 3w + 4 > a \iff w > \frac{a-4}{3}$. Hence, $P(Z > a) = P(\{w \in [0,1]: w > \frac{a-4}{3}\}) = P([0,1] \cap (\frac{a-4}{3},\infty)) = P([0,1] \cap [\frac{a-4}{3},\infty)) = P([\max\{0,\frac{a-4}{3}\},1])$. We note that $0 < \frac{a-4}{3} < 1 \iff 4 < a < 7$, so

$$P([\max\{0, \frac{a-4}{3}\}, 1]) = \begin{cases} 1 \text{ if } a \le 4\\ 1 - \frac{a-4}{3} \text{ if } 4 < a < 7\\ 0 \text{ if } a \ge 7 \end{cases}$$

Next, $\forall w \in \Omega = [0,1], X(w) < a \land Y(w) < b \iff w < a \land w < \frac{b}{2} \iff w < \min\{a, \frac{b}{2}\}.$ Hence, $P(X < a \land Y < b) = P((-\infty, \min\{a, \frac{b}{2}\}) \cap [0,1]) = P((-\infty, \min\{a, \frac{b}{2}\}] \cap [0,1]) = P([0, \min\{1, a, \frac{b}{2}\}])$