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

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

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 $\emptyset, \Omega$ 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.

- $[a$ and $[b$ 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 $\bigcup_{1 \le i \le n} I_i$ for some intervals I_i in $[0, 1]$. Then, $B^c = \bigcap_{1 \le i \le n} I_i^c$ 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.

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

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

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

such that

$$\begin{aligned}
P(I) &= P(\sqcup 1 \leq j \leq n(I \cap I'_j)) \\
&= \sum_{1\leq j\leq n} |I \cap I'_j| \\
&\leq \sum_{1\leq j\leq n} |I'_j| \\
&\leq \sum_{1\leq j\leq n} |I_j| \\
&= \sum_{1\leq j\leq n} P(I_j)
\end{aligned}$$

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. $I \subseteq \bigcup_{1\leq j\leq n} I_{i_j}$. By part (a), we know that $P(I) \leq \sum_{1\leq j\leq n} P(I_{i_j}) \leq \sum_j P(I_j)$ 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 $I_j, j \geq 1$ 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 $P(I) = 0 \leq \sum_j P(I_j)$.

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

$$|I'| \leq \sum_j |I'_j| = \sum_j |I_j| + \epsilon = \sum_j P(I_j) + \epsilon$$

Combining gives

$$P(I) \leq \sum_j P(I_j) + 3\epsilon$$

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

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 $\forall a,b\in\mathbb{R}, a\le b$,

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

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 $D_n = B_n\cap\bigcap_{1\le i\le n-1}B_i^c$ is a disjoint union of elements of $\mathcal{J}$, we let each $B_n^c = \sqcup_{1\le j\le l_n}F_{n,j}$ 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 $1\le i\le n-1$, $1\le\lambda_i\le l_i$. 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 $\lambda_i\ne\mu_i$. 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 $\cap_n A_n = \emptyset$, 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,

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

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 $B_1 = A_{a_1,a_2,\ldots,a_m}, B_2 = A_{a'_1,a'_2,\ldots,a'_n}$. 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 $B = A_{a_1,\ldots,a_n}$. Then B^c is the finite disjoint union of $A_{a'_1,\ldots,a'_n}$ where $\vee_{1\le i\le n} a'_i \neq a_i$. 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 $A = A_{a_1,\ldots,a_k}$. 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 $A_{a_1,\ldots,a_k,\ldots,a_j}$ 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 $\sqcup_{1\le i\le n} D_i = A$.

Next, we want to show finite additivity, i.e. $P(A) = \sum_{1\le i\le n} P(D_i)$. Let K be the length of the longest sequence specified by $D_i, 1 \le i \le n$. 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 $m \in \{k, k+1, \ldots, K\}$ 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 $A^m := A_{a_1,\ldots,a_m}$ for some length m sequence, then $\sqcup_{1\le i\le n} P(D_i) = P(A^m)$.

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 $k < j \leq K$. Then, we want to show $Q(j-1)$. We partition the set $D_i, 1 \leq i \leq n$ 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 $\sqcup_{1\leq i\leq n_1} B_i = A_{a_1,\ldots,a_j,0}$ and $\sqcup_{1\leq i\leq n_2} C_i = A_{a_1,\ldots,a_j,1}$. Then, applying inductive assumption $Q(j)$, we have $\sum_{1\leq i\leq n_1} P(B_i) = P(A_{a_1,\ldots,a_j,0})$ and $\sum_{1\leq i\leq n_2} P(C_i) = P(A_{a_1,\ldots,a_j,1})$. Hence,

$$P(A^j) = \frac{1}{2^j} = \frac{1}{2^{j+1}} + \frac{1}{2^{j+1}} = P(A_{a_1,\ldots,a_j,0}) + P(A_{a_1,\ldots,a_j,1}) = \sum_{1\leq i\leq n} P(D_i)$$

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 $(A \times B)^c = (A^c \times \Omega_2) \cup (\Omega_1 \times B^c) = (A^c \times B^c) \sqcup (A^c \times B) \sqcup (A \times B^c)$.
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

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

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

3.1.5
(ii) We elaborate that $\{X + Y < x\} = \bigcup_{a,b\in\mathbb{R},a+b<x}\{X = a \wedge Y = b\}$, an uncountable union. However, choosing $r \in (a, x-b) \cap \mathbb{Q}$, we see that $\{X = a \wedge Y = b\} \subseteq \{X < r \wedge Y < x - r\}$, so the uncountable union $\bigcup_{a,b\in\mathbb{R},a+b<x}\{X = a \wedge Y = b\}$ is indeed captured by the countable union $\bigcup_{r\in\mathbb{Q}}\{X < r \wedge Y < x - r\}$.

Ex 3.1.7 The goal is to show that

$$\{Z \leq x\} = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{Z_k \leq x + \frac{1}{m}\}$$

For each $w \in \Omega$,

$$\begin{aligned}
& w \in \{Z \leq x\} \iff \\
& Z(w) \leq x \iff \\
& \lim_{n\to\infty} Z_n(w) \leq x \iff \\
& \forall m \in \mathbb{N}, \exists n \in \mathbb{N}, \forall k \geq n, Z_k(w) \leq x + \frac{1}{m} \iff \\
& \forall m \in \mathbb{N}, \exists n \in \mathbb{N}, \forall k \geq n, w \in \{Z_k \leq x + \frac{1}{m}\} \iff \\
& \wedge_{m\in\mathbb{N}} \vee_{n\in\mathbb{N}} \wedge_{k\geq n}, w \in \{Z_k \leq x + \frac{1}{m}\} \iff \\
& w \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{Z_k \leq x + \frac{1}{m}\}
\end{aligned}$$

which proves LHS = RHS.

It is interesting to see the conversion from universal and existential quantifiers (commonly applied in analytical definition of limits) to boolean logic operators, and finally to set operators.

Proposition (Restating the paragraph right below **ex 3.1.7**.) If X is a random variable, and $f : \mathbb{R} \to \mathbb{R}$ is a Borel-measurable function, then the composition $f \circ X$ is also a random variable.

3.1.8
To understand the proof better we consider a few lemmas.

Lemma Characterization of intervals. The following statements about a set I are equivalent in $\mathbb{R}$.

- I is an interval, in particular, I is of one of the following forms $(a, b), (a, b], [a, b), [a, b]$.
- I is a convex set.
- $\forall a, b \in I$, with $a \leq b$, $a \leq c \leq b \implies c \in I$.

Lemma Every open set O in $\mathbb{R}$ is an at most countable union of disjoint open intervals.

The proof involves considering the following equivalence relation: $\forall a, b \in \mathbb{R}, a \equiv b \iff (a, b) \subseteq O$.

- To show that each equivalence class I_j (where J is an arbitrary index set) is an interval, we claim that the left and right endpoints of I_j are respectively $\inf I_j$ and $\sup I_j$.
- To show that each I_j is open, we use the fact that O is open.
- That the $\{I_j : j \in J\}$ are pairwise disjoint is a direct consequence of equivalence classes.
- That there are at most countably many I_j follows from the presence of an injection from index set J to $\mathbb{Q}$ using the axiom of choice and the disjointedness of the I_j.

With this lemma in mind, we see why in the proof of **3.1.8**, $f^{-1}((x, \infty)) \subseteq \mathbb{R}$ open implies $f^{-1}((x, \infty)) \in \mathcal{B}$.

3.1.9 Suppose $(\Omega, \mathcal{F}, P)$ is complete. Let X be a r.v. on Ω. Let $Y : \Omega \to \mathbb{R}$ such that $P(X = Y) = 1$. We claim that Y is also a r.v. The proof is as follows.

We denote the event $E = \{X = Y\}$. Note that $P(E) = 1, P(E^c) = 0$.

For each $b \in \mathbb{R}$, $\{Y \leq b\} = (\{Y \leq b\} \cap E) \cup (\{Y \leq b\} \cap E^c)$, so that $P(Y \leq b) = P(Y \leq b \wedge X = Y) + P(Y \leq b \wedge X \neq Y) = P(Y \leq b \wedge X = Y) = P(X \leq b)$. (To be more precise, we should actually do the derivation backwards, starting with $P(X \leq b)$, since we don't know whether $\{Y \leq b\} \in \mathcal{F}$ or not. But equality is symmetric.)
The last equality is by manipulation of sets. The second last equality is by completeness of $(\Omega, \mathcal{F}, P)$, and that $\{Y \leq b \wedge X \neq Y\} \subseteq E^c$.

3.2.2 Suppose for a collection of events indexed by I is independent, such that for any finite choice $a_1, a_2, \dots, a_j \in I$, $P(A_{a_1} \cap A_{a_2} \cap \dots \cap A_{a_j}) = P(A_{a_1}) \dots P(A_{a_j})$.

a $P(A_{a_2} \cap \dots \cap A_{a_j}) = P(subseqconnAaj\cap) + P(A^c_{a_1} \cap A_{a_2} \cap \dots \cap A_{a_j})$ by additivity.

So we have $P(A^c_{a_1} \cap A_{a_2} \cap \dots \cap A_{a_j}) = P(A_{a_2} \cap \dots \cap A_{a_j}) - P(A_{a_1} \cap A_{a_2} \cap \dots \cap A_{a_j}) = P(A_{a_2}) \dots P(A_{a_j}) - P(A_{a_1}) \dots P(A_{a_j}) = (1 - P(A_{a_1})P(A_{a_2}) \dots P(A_{a_j}) = P(A^c_{a_1})P(A_{a_2}) \dots P(A_{a_j})$.

b Proven by induction, plus associative and commutative properties of $\cap$ and $\cdot$ (product). The idea is to replace each $A_{a_i}, 1 \le i \le j$ by $A^c_{a_i}$ one by one.

c A direct consequence of part (b).

3.2.4

- We make a note that if X, Y are random variables, then their joint distribution is always defined, since for Borel set A, B, $\{X \in A, Y \in B\} = \{X \in A\} \cap \{Y \in B\} \in alg$, of which $X^{-1}(A), Y^{-1}(B)$ both reside in $\mathcal{F}$, and $\mathcal{F}$ is closed under countable intersection.
- Another detail is that $P' : \mathcal{B} \to \mathbb{R}, S \mapsto P(Y \in S)$ is a probability measure sample space $\mathbb{R}$ and the borel algebra on $\mathbb{R}$, $\mathcal{B}$. To prove this, we first verify that indeed, for each borel set $A \in \mathcal{B}$, by definition of a r.v., $P'(S) = P(Y^{-1}(S))$ is defined. So, the domain of P' is correct. Next, we show countable additivity. Given disjoint set $S_1, S_2, \dots$, we have $P'(\sqcup_i S_i) = P(Y \in \sqcup_i S_i) = P(\sqcup_i \{Y \in S_i\} = \sum_i P(Y \in S_i) = \sum_i P'(S_i)$, where the second last equality holds because P is a probability measure.
  Consequently, in the proof of **proposition 3.2.4**, Q, P' are both probability measures defined on $\mathcal{B}$, and Q agrees with P' on sets $(-\infty, y]$ and hence the semialgebra $\mathcal{J}$ defined in **2.5.9**, so $Q = P'$ on $\mathcal{B}$.

3.3.1 Set theory exercise. We show that with $\{A_n : n \in 1, 2, \dots\} \nearrow A, B_n := A_n - A_{n-1}$, we have the properties below. For convenience, we also define $A_0 := \emptyset$ so that B_1 is defined.

1. $\cup_n B_n = \cup_n A_n$
   To show this, $B_n \subseteq A_n$ implies LHS $\subseteq$ RHS. Furthermore, with $x \in$ RHS, let $n = \min\{n : x \in A_n\} > 0$, then $x \in A_n - A_{n-1} = B_n$, so RHS $\subseteq$ LHS. The minimum exists by well-ordering principle.
2. and $i \neq j \implies B_i \cap B_j = \emptyset$.
   WLOG, $i < j$, so $B_i B_j = (A_i A^c_{i-1})(A_j A^c_{j-1}) \subseteq A_i A^c_{j-1} \subseteq A_i A^c_i = \emptyset$, hence $B_i B_j = \emptyset$.

Notation Here, we introduce a notation from Ross's A First course in probability. Given a monotonic sequence of events, be it $\{A_n\} \nearrow A$ or $\{A_n\} \searrow A$, we denote $\lim_{n\to\infty} A_n := A$.

3.4.2 Let (B_n) be a sequence of events, and (B_{n_k}) is a subsequence. Then $\{B_n i.o.\} \supseteq \{B_{n-k} i.o.\}$. In particular, if $P(B_{n_k} i.o.) = 1$, then $P(B_n i.o.) = 1$. Hence, with this proposition, to find $P(B_n i.o.)$ we can attempt to find an independent subsequence and use the Borel-Cantelli lemma.

3.5 It is defined that the tail field of a sequence of events $A_1, A_2, \dots$

$$\tau = \bigcap_{n=1}^{\infty} \sigma(A_n, A_{n+1}, \dots)$$

In general, we $\{A_n i.o.\} \in \tau$ since for each $N \in \mathbb{N}$,

$$\{A_n i.o.\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \bigcap_{n=N}^{\infty} \bigcup_{k=n}^{\infty} A_k \in \sigma(A_N, A_{N+1}, \dots)$$

Hence $\{A_n i.o.\} \in \tau$.

Similarly for any subsequence (A_{n_k}) of (A_n), we have $\{A_{n_k} i.o.\} \in \tau$.

In the special case of $A_n = H_n$, where H_n is the event that in infinite fair coin tossing, the nth coin comes up heads. Then the event $B_n = \{r_n = r_{n+1} = r_{n+2}\} = H_n H_{n+1} H_{n+2} \cup H_n^c H_{n+1}^c H_{n+2}^c$ and $P(B_n) = \frac{1}{4}$. Additionally, $\{B_n i.o.\} \in \tau$ since each B_n can be expressed in terms of $H_k, k \geq n$. To elaborate, we observe that for each $N \in \mathbb{N}$

$$\{B_n i.o.\} \in \sigma(B_N, B_{N+1}, \dots) \subseteq \sigma(H_N, H_{N+1}, \dots)$$

3.5.1 Some notes on the proof of **theorem 3.5.1**.

- By independence of $A_1, A_2, \dots$, for each n, $\sigma(\{A_1, A_2, \dots, A_{n-1}\}), \sigma(\{A_n, \dots\})$ are independent classes. Since $A \in \sigma(\{A_n, \dots\})$ and $A, A_1, A_2, \dots, A_{n-1}$ are independent.
- For any finite subcollection S of $A, A_1, A_2, \dots$, there is a largest k such that $A_k \in S$. Hence, by independence of $A, A_1, A_2, \dots, A_k$, $S \subseteq \{A, A_1, A_2, \dots, A_k\}$ is also independent.

4.1.1 Alternatively, in Iverson's notation, for a simple r.v. X with $range(X) = \{x_1, x_2, \dots, x_n\}$,

$$X = \omega \mapsto \sum_{1 \leq i \leq n} x_i [X(\omega) = x_i]$$

Questions for further research

Borel algebra How is a borel algrebra $\mathcal{B} = \sigma(A)$ rigorously defined? Intuitively, I would say "countable combinations of elements of A" involving set operations. However, according to wikipedia, it appears some form of advanced induction (transfinite induction?) is needed.

Follow up questions:

- Prove rigorously that if $A' \subseteq \mathcal{B}$, then $\sigma(A') \subseteq \mathcal{B}$. This may require some form of advanced induction.