Math 418 Notes

Henry Xia

November 7, 2018

Contents

1 Introduction

Text: R. Durrett. Probability Theory adn Examples. Version 5a.

Definition 1.1. A *sample space* Ω is a non-empty set.
The set of subsets $\mathcal{F} \in 2^\Omega$ is a *field* if and only if

(i) $\emptyset \in \mathcal{F}$.

(ii) Closed under complement. $\forall A \in \mathcal{F}$, $A^C \in \mathcal{F}$.

(iii) Closed under finite union. $\forall A, B \in \mathcal{F}$, $A \cup B \in \mathcal{F}$.

Definition 1.2. The set of subsets $\mathcal{F} \in 2^\Omega$ is a *σ-field* iff

(i) $\emptyset \in \mathcal{F}$.

(ii) Closed under complement. $\forall A \in \mathcal{F}$, $A^C \in \mathcal{F}$.

(iii) Closed under countable union.

Definition 1.3. Call $(\Omega, \mathcal{F})$ a *measureable space* if $\mathcal{F}$ is a σ-field.

Definition 1.4. We write $A = \dot{\bigcup}_{i \in I} A_i$ if the sets A_i are disjoint.

Definition 1.5. Let $(\Omega, \mathcal{F})$ be a measurable space. The *measure* μ on $(\Omega, \mathcal{F})$ is a countably additive, non-negative set function $\mu : \mathcal{F} \to [0, \infty]$ such that

(i) $\mu(\emptyset) = 0$

(ii) $\mu(\dot{\bigcup}_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$

where $I = \mathbb{N}$ or $I = \{1, \ldots, n\}$.
$(\Omega, \mathcal{F}, \mu)$ is a *measure space*. If μ is a probability, that is $\mu(\Omega) = 1$, then the said measure space is a *probability space*.

Definition 1.6. A *finite additive measure* on $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a field, is a function $\mu : \mathcal{F} \to [0, \infty]$ such that

(i) $\mu(\emptyset) = 0$

(ii) $\mu(\dot{\bigcup}_{i\in I} A_i) = \sum_{i\in I} \mu(A_i)$

where (ii) only holds for I finite.

Example 1.1. An example of a finite additive measure is

$$\mu(A) = \begin{cases} 0 \text{ if A is finite} \\ \infty \text{ if A is infinite} \end{cases}$$

Lemma 1.1. *Let $\mathcal{F}$ be a field and μ be a finite additive measure, then*

(a) Monotonicity. $A, B \in \mathcal{F}$ with $A \subset B$ implies $\mu(A) \leq \mu(B)$.

(b) Subadditivity. $A, B \in \mathcal{F}$ means $\mu(A \cup B) \leq \mu(A) + \mu(B)$.

Proof of (a). Elementary. □

Proof of (b). Use (a) and disjoint union. □

Definition 1.7. We write $A_n \uparrow A$ if $A_n \subset A_{n+1}$ and $\bigcup A_n = A$. Defined similarly for $\downarrow$.

Theorem 1.2. *Let μ be a measure on $(\Omega, \mathcal{F})$, and $\{A_n : n \in \mathbb{N}\}$ be a sequence in $\mathcal{F}$.*

(a) Continuity from below. $A_n \uparrow A \implies \mu(A_n) \uparrow \mu(A)$

(b) Continuity from above. $\mu(A_1) < \infty$ and $A_n \downarrow A \implies \mu(A_n) \downarrow \mu(A)$

(c) Countable subadditivity. $\mu(\bigcup_n^\infty A_n) \leq \sum_n^\infty \mu(A_n)$

Proof of (a). Observe that $A_n = \dot{\bigcup}_{k=1}^n (A_k \setminus A_{k-1})$ where we define $A_0 = \emptyset$. Then

$$\mu(A) = \sum_{k=1}^\infty \mu(A_k \setminus A_{k-1}) = \lim_{n\to\infty} \mu(A_n).$$

Also, $\mu(A_n) \leq \mu(A_{n+1})$, so $\mu(A_n) \uparrow \mu(A)$. □

Proof of (b). We can take complements with respect to A_1 and use (a) on $(A_1 \setminus A_n) \uparrow (A_1 \setminus A)$. Then $\mu(A_1 \setminus A_n) \uparrow \mu(A_1 \setminus A)$. □

Proof of (c). We use finite subadditivity as follows

$$\mu(\cup_k^n A_k) \leq \sum_k^n \mu(A_k) \leq \lim_{n\to\infty} \sum_k^n \mu(A_k).$$ □

Partial converse to Theorem 1.2 (Continuity from below): If μ is finitely additive on $(\Omega, \mathcal{F})$ and continuous from below, then μ is a measure (countably additive).

Example 1.2. Consider the counting measure. Let $A_n = \{n, n+1, \ldots\}$, then $A_n \downarrow \emptyset$, but $\mu(A_n) = \infty$ for all n and $\mu(\emptyset) = 0$.

Definition 1.8. A $\mathcal{A} \subset 2^\Omega$ is a π-system if it is closed under intersection.
A $\mathcal{A} \subset 2^\Omega$ is a λ-system if

(i) $\Omega \in \mathcal{A}$

(ii) If $A, B \in \mathcal{A}$ and $B \subset A$, then $A \setminus B \in \mathcal{A}$

(iii) If $A_n \in \mathcal{A}$ and $A_n \uparrow A$, then $A \in \mathcal{A}$

Theorem 1.3 (π-λ Theorem)**.** *Let $\mathcal{C}$ be a π-system in 2^Ω, and $\mathcal{L}$ be a λ-system in 2^Ω. If $\mathcal{C} \subset \mathcal{L}$, then $\sigma(\mathcal{C}) \subset \mathcal{L}$.*

Proof. Draw a venn diagram, or show that the smallest λ-system containing $\mathcal{C}$ is a σ-field. $\square$

Corollary 1.3.1. *Let P_1 and P_2 be probabilities on $(\Omega, \mathcal{F})$, and $\mathcal{C} \subset \mathcal{F}$ be a π-system. If $P_1 = P_2$ on $\mathcal{C}$, then $P_1 = P_2$ on $\sigma(\mathcal{C})$.*

Proof. Let $\mathcal{L} = \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$, then $\mathcal{C} \subset \mathcal{L}$. We want to show that $\mathcal{L}$ is a λ-system.

(i) Clearly $\Omega \in \mathcal{L}$.

(ii) If $A, B \in \mathcal{L}$ and $A \subset B$, then $B \setminus A \in \mathcal{L}$ by finite additivity.

(iii) If $A_n \in \mathcal{L}$ and $A_n \uparrow A$, then we can use continuity from below to get $P_1(A) = P_2(A)$. $\square$

Example 1.3. Let $p_1, \dots, p_6 \in [0,1]$ and $\sum_{i=1}^6 p_i = 1$. We can define $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = 2^\Omega$, and $P(\{i\}) = p_i$. Then by finite additivity, we get $P(A) = P(\dot\bigcup_{i \in A} \{i\}) = \sum_{i \in A} p_i$ for any $A \in \mathcal{F}$.

2 The Law of Large Numbers

2.1 The Law of Large Numbers for Coin Tossing

Experiment: Toss a fair coin an infinite number of times.
Outcome: An infinite sequence of 0's and 1's. $w = (w_1, w_2, \dots)$.
Let $\Omega = \{0,1\}^{\mathbb{N}}$ be the sample space. Let $x_n : \Omega \to \{0,1\}$ with $x_n(w) = w_n$, and $s_n : \Omega \to \{0, 1, \dots, n\}$ with $s_n(w) = \sum_{i=1}^n x_i(w)$.

Definition 2.1. The event $A \subset \Omega$ is a finite dimensional event if $\exists n, B \subset \{0,1\}^n$ such that $A = \{w : (w_1, w_2, \dots, w_n) \in B\}$. (ie. we only need to look at the first n things.)
Let $\mathcal{F}_0$ be the set of finite dimensional events.

We want to show that $\lim_{n \to \infty} \frac{s_n(n)}{n} = \frac{1}{2}$.

Lemma 2.1. *$\mathcal{F}_0$ is a field.*

Proof. Observe that $\emptyset = \{w \in \Omega : w_1 \in \emptyset\} \in \mathcal{F}_0$.
Let $A_1, A_@ \in \mathcal{F}$, then there exist n_1, n_2, B_1, B_2 such that $A_i = \{w : (w_1, \dots, w_{n_i}) \in B_i\}$ for $i \in \{1, 2\}$. Now we may assume without loss of generality that $n_1 = n_2$, because we can extend the smaller B_i so that it has the same dimension as the larger one. This is also closed under complement. $\square$

Definition 2.2. Define $P : \mathcal{F}_0 \to [0,1]$ such that $P(\{w : (w_1, \dots, w_n) \in B\}) = \frac{\#B}{2^n}$, where $B \subset \{0,1\}^n$. We can check that P is well defined.

Lemma 2.2. *P is a probability.*

Proof. We can check that $P(\emptyset) = 0$ and $P(\Omega) = 1$.
Let disjoint $A_1, A_2 \in \mathcal{F}$. Then $\emptyset = A_1 \cap A_2 = \{w : (w_1, \dots, w_n) \in B_1 \cap B_2\}$, implying that $B_1 \cap B_2 = \emptyset$. Then, we can write the disjoint union in the same way. $\square$
We can compute $P(\{s_n(w) = k\}) = \binom{n}{k} 2^{-n}$, and say "$s_n$ has binomial distribution with parameters $(n, \frac{1}{2})$".

Let $C = \left\{w : \lim_{n\to\infty} \frac{s_n(w)}{n} = \frac{1}{2}\right\}$. Observe that $C \notin \mathcal{F}_0$. However, we will extend P from $\mathcal{F}_0$ to $\sigma(\mathcal{F}_0)$.

Lemma 2.3. *$C \in \sigma(\mathcal{F}_0)$.*

Proof. Observe that

$$\begin{aligned} w \in C &\iff \lim_{n\to\infty} \frac{S_n(w)}{n} = \frac{1}{2} \\ &\iff \forall M \in \mathbb{N},\ \exists N \text{ such that } \forall n \geq N,\quad \left|\frac{S_n(w)}{n} - \frac{1}{2}\right| < \frac{1}{M} \\ &\iff w \in \bigcap_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{w : \left|\frac{S_n(w)}{n} - \frac{1}{2}\right| < \frac{1}{M}\right\} \end{aligned}$$

Let $C_{n,M}$ be the latter set, which is finite dimensional, hence $C_{n,M} \in \mathcal{F}_0$. It follows that $C \in \sigma(\mathcal{F}_0)$ as σ-fields are closed under finite unions and intersections. □

Suppose that we can extend P to $\sigma(\mathcal{F}_0)$ (by Carathéodory). Now we need to prove that $P(C) = 1$, that is the event happens *almost surely.*

Lemma 2.4. *$P(C_{n,M}^c) \leq \frac{M^2}{4n}$.*

Proof. We observe that S_n is a binomial random variable, and $\frac{n}{2}$ is the mean. Then we can use Chebychev's inequality.

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \frac{1}{M}\right) = P\left(\left|S_n - \frac{n}{2}\right| \geq \frac{n}{M}\right) \leq \frac{\text{var}(S_n)}{(n/M)^2} = \frac{M^2}{4n}.$$

There is also a direct calculation proof given on Perkins's webpage. □

Theorem 2.5 (The Strong Law of Large Numbers for Coin Tossing)**.** *$P(C) = 1$.*

Proof. Let $\hat{C} = \left\{w : \lim_{m\to\infty} \frac{S_{m^2}}{m^2} = \frac{1}{2}\right\}$. This set is equal to C as shown in homework 1. Similarly, we can write

$$\hat{C} = \bigcap_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \left\{\left|\frac{S_{m^2}}{m^2} - \frac{1}{2}\right| < \frac{1}{M}\right\}.$$

We want to show that $P\left(\bigcap_{m=N}^{\infty} C_{m^2,M}\right) = 1$, after which continuity from below and continuity from above will yield the result.

$$P\left(\bigcap_{m=N}^{\infty} C_{m^2,M}\right) = 1 - P\left(\bigcup_{m=N}^{\infty} C_{m^2,M}^c\right) \geq 1 - \sum_{m=N}^{\infty} P(C_{m^2,M}^c) \geq 1 - \sum_{m=N}^{\infty} \frac{M^2}{4m^2} \to 1,$$

where the convergence holds because $\frac{1}{m^2}$ is summable.

Increasing N means we take the intersection of less sets, which increases. Continuity from below gives $P(\bigcup\bigcap C_{m^2,M}) = 1$.

Increasing M makes the epsilon smaller, so we are decreasing the size of each set. This will decrease the intersection. Continuity from above gives $P(\bigcap\bigcup\bigcap C_{m^2,M}) = 1$.

Hence $P(C) = P(\hat{C}) = 1$. □

Now we prove that we can indeed extend P to $\sigma(\mathcal{F}_0)$.

Theorem 2.6. *Let $\mathcal{C}$ be a π-system, μ_1, μ_2 be measures on $\sigma(\mathcal{C})$. If $\mu_1 = \mu_2$ on $\mathcal{C}$ and both are σ-finite on $\mathcal{C}$, then $\mu_1 = \mu_2$ on $\sigma(\mathcal{C})$.*

Proof. Let $\{A_n\} \in \mathcal{C}$ and $A_n \uparrow \Omega$ where $\mu_1(A_n) = \mu_2(A_n) < \infty$. For all $\mu_i(A_n) > 0$, define $P_{i,n}(\mathcal{E}) = \frac{\mu_i(E \cap A_n)}{\mu_i(A_n)}$. Then $P_{i,n}$ is a probability. We can check that $\mu_i(A_n) > 0$ for n sufficiently large. We can also show that $P_{1,n} = P_{2,n}$ using continuity from below. Now we can use π-λ theorem to finish the proof. $\square$

Lemma 2.7. *Let $P : \mathcal{F}_0 \to [0, 1]$ be a finitely additive probability, then P is a countably additive probability, where $\mathcal{F}_0$ is defined as above.*

Proof. Give $\Omega = \{0, 1\}^{\mathbb{N}}$ the product topology. That is for $\{w_n\} \subset \Omega$, then $w^n \to w$ as $n \to \infty$ if and only if $\forall i \in \mathbb{N}$, we have $\lim_{n\to\infty} w_{n,i} = w_i$. Equivalently, we can define the metric $d(w, w') = \sum_{i=1}^{\infty} \frac{|w_i - w'_i|}{2^i}$. Observe that (Ω, d) is compact by Cantor diagonalization.
Let $\{A_n\}$ be disjoint sets in $\mathcal{F}_0$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_0$. Let $K_n = \bigcup_{j=1}^{\infty} A_j \setminus \bigcup_{j=1}^{n} A_j$. Observe that $K_n \downarrow \emptyset$. If each K_n is compact, then $\exists N$ such that $\forall n \geq N$, we have $K_n = \emptyset$. Hence finite additivity implies countable additivity.
Now it suffices to show that each $K_n = \bigcup_{j=1}^{\infty} A_j \setminus \bigcup_{j=1}^{n} A_j$ is compact. Suppose that K_n is nonempty for all n. $\square$

3 Some stuff missing here