STATS310A - Lecture 6

Persi Diaconis
Scribed by Michael Howes

10/07/21

Contents

1 An IOU

All lot of what we have done has been done assuming that length is countably subadditive. That is if $(a, b] \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]$, then $b - a \leq \sum_{i=1}^{\infty} b_i - a_i$. Today we prove this fact.

Proof. First we will show that if $(a, b] \subseteq \bigcup_{i=1}^{n}(a_i, b_i]$, then $b - a \leq \sum_{i=1}^{n} b_i - a_i$. We will proceed by induction, if $(a, b] \subseteq (a_1, b_1]$, then $b - a \leq b_1 - a \leq b_1 - a_1$ and we are done.

Thus suppose the result holds for $n-1$ and that $(a, b] \subseteq \bigcup_{i=1}^{n}(a_i, b_i]$. Then we must have $b \in (a_i, b_i]$ for some i. By relabelling we can assume without loss of generality that $b \in (a_n, b_n]$. If $a_n \leq a$, then we are done since then $(a, b] \subseteq (a_n, b_n]$. Now suppose $a_n > a$. We thus have $(a, a_n] \subseteq \bigcup_{i=1}^{n-1}(a_i, b_i]$ and by the inductive hypothesis $a - a_n \leq \sum_{i=1}^{n-1} b_i - a_i$. Thus

$$b - a \leq b_n - a = b_n - a_n + a_n - a \leq \sum_{i=1}^{n} b_i - a_i,$$

as required. Now consider the infinite case when $(a, b] \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]$. Let $\varepsilon > 0$ and note that

$$[a + \varepsilon, b] \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i + 2^{-i}\varepsilon).$$

Since the set $[a + \varepsilon, b]$ is compact, the open cover $\{(a_i, b_i + 2^{-i}\varepsilon)\}_{i=1}^{\infty}$ must contain a finite subcover. Thus

$$(a + \varepsilon, b] \subseteq [a + \varepsilon, b] \subseteq \bigcup_{i=1}^{n}(a_i, b_i + 2^{-i}\varepsilon) \subseteq \bigcup_{i=1}^{n}(a_i, b_i + 2^{-i}\varepsilon].$$

Thus by the finite case we have

$$b - a - \varepsilon \leq \sum_{i=1}^{n} b_i - a_i + 2^{-i}\varepsilon \leq \varepsilon + \sum_{i=1}^{\infty} b_i - a_i.$$

Thus, by letting ε go to zero we are done and Persi has paid up. □

2 Semi-rings

Definition 1. A class, $\mathcal{A}$, of subsets of a set Ω, is a *semi-ring* if

(a) $\emptyset \in \mathcal{A}$.

(b) $\mathcal{A}$ is closed under finite intersections.

(c) If $A, B \in \mathcal{A}$ and $A \subseteq B$, then there exist disjoint sets $C_1, \ldots, C_k \in \mathcal{A}$, such that

$$B \setminus A = \bigcup_{i=1}^{k} C_i.$$

Some examples are

- Intervals in $\mathbb{R}$.
- Rectangles in $\mathbb{R}^2$.
- Rectangular prisms in $\mathbb{R}^3$.
- Hypercubes in $\mathbb{R}^k$.

Note that semi-rings are not necessarily closed under complements and need not contain Ω. Semi-rings are π-systems.

Theorem 1. *Let Ω be a set and $\mathcal{A}$ a semi-ring of subsets of $\mathcal{A}$. Let $\mu : \mathcal{A} \to \mathbb{R} \cup \{\infty\}$ be a function such that*

(a) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$.

(b) μ is finitely additive on $\mathcal{A}$.

(c) μ is countably subadditive on $\mathcal{A}$.

Then μ has an extension to $\sigma(A)$ and the extension is unique if μ is σ-finite with respect to $\mathcal{A}$.

Proof. If $A \subseteq B$ are both in $\mathcal{A}$, then

$$\mu(B) = \mu(A) + \sum_{i=1}^{k} \mu(C_i) \geq \mu(A).$$

Thus μ is monotone on $\mathcal{A}$. Define

$$\mu^*(B) = \inf\left\{\sum_{i=1}^{\infty} \mu(A_i) : B \subseteq \bigcup_{i=1}^{\infty} A_i, A_i \in \mathcal{A}\right\}.$$

We know that μ^* is a measure on the σ-algebra $\mathcal{M}(\mu^*)$. We need to show that

(a) $\mathcal{A} \subseteq \mathcal{M}(\mu^*)$, and

(b) $\mu^*(A) = \mu(A)$ if $A \in \mathcal{A}$.

For (a), suppose that $A \in \mathcal{A}$ and $E \subseteq \Omega$. If $\mu^*(E) = \infty$, then $\mu^*(E) \geq \mu^*(A \cap E) + \mu^*(A^c \cap E)$ and we are done.

Thus suppose $\mu^*(E) < \infty$ and fix $\varepsilon > 0$. There exists $A_n \in \mathcal{A}$ such that $E \subseteq \bigcup_{n=1}^\infty A_n$ and

$$\sum_{n=1}^\infty \mu(A_n) \leq \mu^*(E) + \varepsilon.$$

Define $B_n = A \cap A_n$. Since $\mathcal{A}$ is closed under intersections we have $B_n \in \mathcal{A}$. For each n we have $B_n \subseteq A_n$ and hence by the third semi-ring property there exist disjoint sets $C_{n,i} \in \mathcal{A}$ for $i = 1, \ldots, k_n$ such that

$$A^c \cap A_n = A_n \setminus B_n = \bigcup_{i=1}^{k_n} C_{n,i}.$$

We know that $\bigcup_{n=1}^\infty B_n$ covers $A \cap E$ and that $\bigcup_{n=1}^\infty \bigcup_{i=1}^{k_n} C_i$ covers $A^c \cap E$. Thus, since μ is finitely additive on $\mathcal{A}$ we have

$$\begin{aligned} \mu^*(A \cap E) + \mu^(A^c \cap E) &\leq \sum_{n=1}^\infty \mu(B_n) + \sum_{n=1}^\infty \sum_{i=1}^{k_n} \mu(C_{n,i}) \\ &= \sum_{n=1}^\infty \left(\mu(B_n) + \sum_{i=1}^{k_n} \mu(C_{n,i}) \right) \\ &= \sum_{n=1}^\infty \mu(A_n) \\ &\leq \mu^*(E) + \varepsilon. \end{aligned}$$

Thus we can conclude that $\mu^*(E) \geq \mu^*(A \cap E) + \mu^*(A^c \cap E)$ and so $A \in \mathcal{M}(\mu^*)$.

Now we will prove (b). Since $A \subseteq A$ for all $A \in \mathcal{A}$, we immediately have $\mu^*(A) \leq \mu(A)$. Also if $A \subseteq \bigcup_{i=1}^\infty A_i$, then $\mu(A) \leq \sum_{i=1}^\infty \mu(A_i)$ by countable subadditivity so $\mu(A) \leq \mu^*(A)$. □

3 Distribution functions

Let μ be a probability on $\mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}$. Define

$$F(x) = \mu((-\infty, x]),$$

then F satisfies

(a) $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$.

(b) $x \leq y$ implies $F(x) \leq F(y)$.

(c) If $x_n \searrow x$, then $F(x_n) \searrow F(x)$. Since $\bigcap_{n=1}^\infty (-\infty, x_n] = (-\infty, x]$.

Conversely if F satisfies (a)-(c), then $\mu((a, b]) = F(b) - F(a)$ extends to a measure on $(\mathbb{R}, \mathcal{B})$ and $\mu((-\infty, x]) = F(x)$ for all x.

For example

- The normal distribution: $F(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{1}{2}t^2\right) dt$.

- The exponential distribution:

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-x} & \text{if } x \geq 0. \end{cases}$$

- A point mass at x_0

$$F(x) = \begin{cases} 0 & \text{if } x < x_0, \\ 1 & \text{if } x \geq x_0. \end{cases}$$

3.1 Distribution functions in higher dimensions

If $x \in \mathbb{R}^k$, let $A_x = \{y \in \mathbb{R}^k : y_i \leq x_i, 1 \leq i \leq k\}$. If μ is a probability on $\mathbb{R}^k$, define

$$F(x) := \mu(A_x).$$

The function F satisfies (a)-(c) from before we also have the following extra condition! Suppose A is a rectangle

$$A = \{y \in \mathbb{R}^k : a_i < x \leq b_i, i = 1, \ldots, k\},$$

for some $a, b \in \mathbb{R}^k$. Then define

$$\Delta_F(A) = \sum_{v \text{ a vertex of } A} \operatorname{sgn}(v) F(v).$$

where

$$\operatorname{sgn}(v) = \begin{cases} 1 & \text{if } v_i = a_i \text{ for an even number of } i's, \\ -1 & \text{if } v_i = a_i \text{ for an odd number of } i's. \end{cases}$$

For example if $k = 1$, then $A = (a, b]$ and $\Delta_F(A) = F(b) - F(a)$. If $k = 2$, then

$$\Delta_F(A) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2).$$

Note that by the inclusion exclusion principle

$$\Delta_F(A) = \mu(A) \geq 0,$$

if $F(x) = \mu(A_x)$.

Theorem 2. *If F satisfies (a)-(c) and $\Delta_F(A) \geq 0$ for all rectangles A, then there exists a unique probability μ on $\mathbb{R}^k$ such that $\mu(A_x) = F(x)$ for all $x \in \mathbb{R}^k$*

Proof. The set of all rectangles is a semi-ring and $\mu(A) := \Delta_F(A)$ is a non-negative, finitely additive, countably subadditive function. □

4 The basic problem in probability

We can now restate our "basic problem in probability". Let Ω be a set, $\mathcal{A}$ a class of subsets of Ω and μ a probability measure on $\sigma(\mathcal{A})$ which we can compute on $\mathcal{A}$. Given $B \in \sigma(\mathcal{A})$, can we compute/approximate $\mu(B)$?

5 Random variables

Definition 2. Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be measurable spaces (that is, $\mathcal{F}, \mathcal{F}'$ are σ-algebras). A function $T : \Omega \to \Omega'$ is *measurable* if for all $A' \in \mathcal{F}'$, $T^{-1}(A') \in \mathcal{F}$.

For example suppose $\Omega = S_n$ the set of all permutations of $\{1, \ldots, n\}$ and $\mathcal{F}$ is the set of all subsets of Ω. Suppose also that $\Omega' = \{1, \ldots, n\}$ and again $\mathcal{F}'$ is the collection of all subset of Ω'. Define $T(\pi) = \pi(i)$. Then $T^{-1}(\{j\}) = \{\pi \in \Omega : \pi(i) = j\}$ and $T^{-1}(A) = \{\pi \in \Omega : \pi(i) \in A\}$. The function T is thus measurable.

Lemma 1. *Let $T : \Omega \to \Omega'$ be a function and let $\{A_i\}_{i \in I}$ be a collection of subsets of Ω', then*

(a) $T^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} T^{-1}(A_i)$.

(b) $T^{-1}\left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} T^{-1}(A_i)$.

(c) For all $A \subseteq \Omega$, $T^{-1}(A^c) = T^{-1}(A)^c$.

Proposition 1. *Suppose $\mathcal{F}' = \sigma(\mathcal{A}')$ for some collection of subsets $\mathcal{A}'$. Let $T : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ be a function such that $T^{-1}(A') \in \mathcal{F}$ for all $A' \in \mathcal{A}'$, then T is a measurable function.*

Proof. Define $\mathcal{G}' = \{A' \in \mathcal{F}' : T^{-1}(A') \in \mathcal{F}\}$, by assumption $\mathcal{G}'$ contains $\mathcal{A}'$ and by the above lemma $\mathcal{G}'$ is a σ-algebra. Thus $\mathcal{F}' = \sigma(\mathcal{A}') \subseteq \mathcal{G}'$, as required. □

Proposition 2. *Suppose that $T : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ and $S : (\Omega', \mathcal{F}') \to (\Omega'', \mathcal{F}'')$ are both measurable functions, then $S \circ T : (\Omega, \mathcal{F}) \to (\Omega'', \mathcal{F}'')$ is also a measurable function.*

Definition 3. A *random variable* is a measurable function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$.