Probability and Stochastic Processes I - Lecture 3

Michael Evans
University of Toronto
http://www.utstat.utoronto.ca/mikevans/stac62/STAC622023.html

2023

I.5 Continuity of P

- consider why P is required to be countably additive rather than just finitely additive

- for this we need to define what it means for a sequence of sets $A_n \subseteq \Omega$ to converge to a set $A \subseteq \Omega$

Definition 1.5.1 For a sequence $A_n \subseteq \Omega$ define

$$\begin{aligned} \liminf A_n &= \cup_{n=1}^{\infty} \cap_{i=n}^{\infty} A_i = \{\omega : \omega \text{ is in all but finitely many } A_i\}, \\ \limsup A_n &= \cap_{n=1}^{\infty} \cup_{i=n}^{\infty} A_i = \{\omega : \omega \text{ is in infinitely many } A_i\}. \end{aligned}$$

Then A_n *converges* to the set A, and write $\lim_{n\to\infty} A_n = A$ or $A_n \to A$, whenever $A = \liminf A_n = \limsup A_n$. ■

- $\liminf A_n \subseteq \limsup A_n$

- $\cap_{i=1}^{\infty} A_i \subseteq \cap_{i=2}^{\infty} A_i \subseteq \cdots$ and so $\cap_{i=n}^{\infty} A_i$ is an increasing sequence of sets

- $\cup_{i=1}^{\infty} A_i \supseteq \cup_{i=2}^{\infty} A_i \supseteq \cdots$ and so $\cup_{i=n}^{\infty} A_i$ is a decreasing sequence of sets

- if $A_n \in \mathcal{A}$ for every n, then $\cap_{i=n}^{\infty} A_i, \cup_{i=n}^{\infty} A_i \in \mathcal{A}$ for every n (they are "events") and this implies $\liminf A_n, \limsup A_n \in \mathcal{A}$ and also, when $A_n \to A$, then $A \in \mathcal{A}$ (are all events)

Proposition 1.5.1 If $A_n \in \mathcal{A}$ for every n and $A_1 \supseteq A_2 \supseteq \cdots$ (a monotone decreasing sequence of sets), then $A_n \to A = \cap_{i=1}^{\infty} A_i$.

Proof: Now let $\omega \in \cap_{i=n}^{\infty} A_i$ so $\omega \in A_n \subseteq A_{n-1} \subseteq \cdots \subseteq A_1$ which implies $\omega \in \cap_{i=1}^{\infty} A_i$ and therefore $\cap_{i=n}^{\infty} A_i \subseteq \cap_{i=1}^{\infty} A_i$ while it is clear that $\cap_{i=1}^{\infty} A_i \subseteq \cap_{i=n}^{\infty} A_i$ for every n. Therefore, $\cap_{i=n}^{\infty} A_i = \cap_{i=1}^{\infty} A_i$ for every n which implies $\liminf A_n = \cap_{i=1}^{\infty} A_i$. Also $\cup_{i=n}^{\infty} A_i = A_n$ by the monotonicity and so $\limsup A_n = \cap_{i=1}^{\infty} A_i$. Therefore, $A_n \to A = \cap_{i=1}^{\infty} A_i$. ■

Exercise 1.5.1 If $A_1 \subseteq A_2 \subseteq \cdots$ (a monotone increasing sequence of sets), then prove $A_n \to A = \cup_{i=1}^{\infty} A_i$.

Proposition 1.5.2 (*Continuity of* P) If $A_n \in \mathcal{A}$ for every n and $A_n \to A$, then $P(A_n) \to P(A)$ as $n \to \infty$.

Proof: As noted $\cup_{i=n}^{\infty} A_i$ is a monotone decreasing sequence and so (Prop. 1.4.1) $\cup_{i=n}^{\infty} A_i \to \cap_{n=1}^{\infty} \cup_{i=n}^{\infty} A_i = \limsup A_n$ and similarly (Ex. 1.4.1) $\cap_{i=n}^{\infty} A_i \to \liminf A_n$. So, if we prove the result for monotone sequences, then

$$\begin{aligned} P(\cup_{i=n}^{\infty} A_i) &\to P(\limsup A_n), \\ P(\cap_{i=n}^{\infty} A_i) &\to P(\liminf A_n). \end{aligned}$$

Now note

$$P(\cap_{i=n}^{\infty} A_i) \le P(A_n) \le P(\cup_{i=n}^{\infty} A_i),$$

and since $A = \liminf A_n = \limsup A_n$ we would have $P(A_n) \to P(A)$.

Now suppose A_n is a monotone increasing sequence of sets, so (Ex. 1.4.1) $A_n \to A = \cup_{i=1}^{\infty} A_i$. Put

$$B_1 = A_1, B_2 = A_2 \cap A_1^c, B_3 = A_3 \cap A_2^c, \ldots$$

and note that the $B_n \in \mathcal{A}$, are mutually disjoint with $A_n = \cup_{i=1}^n B_i$. Therefore, $P(A_n) = \sum_{i=1}^n P(B_i)$ and

$$\begin{aligned} \lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} \sum_{i=1}^n P(B_i) = \sum_{i=1}^\infty P(B_i) \\ &= P(\cup_{i=1}^\infty B_i) = P(\cup_{i=1}^\infty A_i) = P(\lim_{n\to\infty} A_n). \end{aligned}$$

A similar argument establishes this result when A_n is a monotone decreasing sequence of sets. ■

Exercise 1.5.2 Prove that if A_n is a monotone decreasing sequence of sets, then $\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n)$. (Hint: A_n^c is a monotone increasing sequence of sets.)

- the converse (under finite additivity) of Prop. 1.4.2 is also true

Proposition 1.5.3 If $P : \mathcal{A} \rightarrow [0, 1]$ satisfies (i) $P(\Omega) = 1$, (ii) P is additive ($A, B \in \mathcal{A}$ mutually disjoint then $P(A \cup B) = P(A) + P(B)$) and (iii) $P(A_n) \rightarrow P(A)$ as $n \rightarrow \infty$ whenever $A_n \in \mathcal{A}$ for every n and $A_n \rightarrow A$, then P is a probability measure on $\mathcal{A}$.
Proof: **Exercise 1.5.3**.

- so countable additivity is equivalent to continuity of P which is only really needed when $\#(\Omega) = \infty$ and in that case (in practice) we are approximating something that is essentially finite

Proposition 1.5.4 (*Boole's inequality*) If $A_n \in \mathcal{A}$ for every n, then (i) $P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$ and (ii) $P(\cup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty P(A_i)$.

Proof: Recall from Exercise I.1.4,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq P(A_1) + P(A_2)$$

and now assume (i) holds for a specified $n \geq 2$. Then

$$P(\cup_{i=1}^{n+1} A_i) = P((\cup_{i=1}^n A_i) \cup A_{n+1}) \leq P(\cup_{i=1}^n A_i) + P(A_{n+1}) \leq \sum_{i=1}^{n+1} P(A_i)$$

and by induction the result (i) holds for every n. Also, since

$$P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$$

the LHS converges to $P(\cup_{i=1}^\infty A_i)$ (since $\cup_{i=1}^n A_i$ is monotone increasing and P is continuous) and the RHS converges to $\sum_{i=1}^\infty P(A_i)$ proving (ii). ■

Proposition 1.5.5 (*Borel-Cantelli lemma*) If $A_n \in \mathcal{A}$ for every n and $\sum_{i=1}^{\infty} P(A_i) < \infty$, then $P(\limsup A_n) = 0$.

Proof: We have that $P(\limsup A_n) = P(\cap_{n=1}^{\infty} \cup_{i=n}^{\infty} A_i) \leq P(\cup_{i=n}^{\infty} A_i)$ for every n and, using Boole's inequality, $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i) \to 0$ as $n \to \infty$ which establishes the result. ■

- Borel-Cantelli says if the sum of all the probabilities $P(A_i)$ is finite, then it is impossible that infinitely many of the events are true

Example 1.5.1

- consider a sequence of experiments where a fair coin is tossed n times and let $A_n =$ "n heads are obtained in the n-th experiment" so $P(A_n) = 1/2^n$ and (summing a geometric series)

$$\sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} \frac{1}{2^i} = \sum_{i=0}^{\infty} \frac{1}{2^i} - 1 = \frac{1}{1 - 1/2} - 1 = 2 - 1 = 1 < \infty$$

and so by Borel-Cantelli the event that all heads occurs can only happen finitely many times as $n \to \infty$ ■

1.6 Conditional Probability

- the most important concept in probability (relationships among variables, measuring evidence, etc.)

Definition 1.6.1 When $(\Omega, \mathcal{A}, P)$ is a probability model and $C \in \mathcal{A}$ satisfies $P(C) > 0$, then the *conditional probability model given* C is $(\Omega, \mathcal{A}, P(\cdot \mid C))$ where $P(\cdot \mid C) : \mathcal{A} \to [0, 1]$ is given by

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)}. \blacksquare$$

Exercise 1.6.1 Prove that $(\Omega, \mathcal{A}, P(\cdot \mid C))$ is a probability model.

- application: initially the measure of belief that A is true is given by $P(A)$ but then the information is provided that C is true and so the belief measure is modified to $P(A \mid C)$

- principle of conditional probability: you **must** modify beliefs in this way

- note: if $A \cap C = \phi$, then $P(A \mid C) = 0$ while $P(C \mid C) = 1$ so really the probability model can be taken to be $(C, \mathcal{A} \cap C, P(\cdot \mid C))$ where

$$\mathcal{A} \cap C = \{A \cap C : A \in \mathcal{A}\}$$

is a σ-sigma (closure under complementation means complements wrt C)

Proposition 1.6.1 (*Theorem of Total Probability*) Suppose $C_1, C_2, \ldots \in \mathcal{A}$ with $P(C_i) > 0$ for all i and $\Omega = \cup_{i=1}^{\infty} C_i$ with $C_i \cap C_j = \phi$ for all i, j, then for any $A \in \mathcal{A}$

$$P(A) = \sum_{i=1}^{\infty} P(C_i) P(A \mid C_i).$$

Proof: Clearly $A = \cup_{i=1}^{\infty} A \cap C_i$ and the sets $C_i \cap A$ are mutually disjoint. Therefore,

$$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i) = \sum_{i=1}^{\infty} \frac{P(A \cap C_i)}{P(C_i)} P(C_i) = \sum_{i=1}^{\infty} P(C_i) P(A \mid C_i). \blacksquare$$

- we call $\{C_i : i = 1, 2, \ldots\}$, as described in Prop. 1.5.1, a *partition* of Ω

Exercise 1.6.2 Prove that $A = \cup_{i=1}^{\infty}(A \cap C_i)$ and the sets $C_i \cap A$ are mutually disjoint when $\{C_i : i = 1, 2, \ldots\}$ is a partition of Ω.

- the value of the theorem lies in simplifying calculations

Example 1.6.1

- suppose there are three urns, with the following contents

urn 1 contains 50 white balls and 50 black balls
urn 2 contains 60 white balls and 80 black balls
urn 3 contains 20 white balls and 30 black balls

- an urn is selected according to a probabilistic mechanism where

$P(\text{"urn 1 is selected"}) = 2/3,$
$P(\text{"urn 2 is selected"}) = 1/5,$
$P(\text{"urn 3 is selected"}) = 2/15$

and then a ball is drawn from the selected urn after thorough mixing

- question: what is the probability that a white ball is selected?

- $\Omega = \{(1,B),(1,W),(2,B),(2,W),(3,B),(3,W)\}, \mathcal{A} = 2^\Omega$
- partition Ω via $C_i =$ "urn i is selected" $= \{(i,B),(i,W)\}$
- by the Theorem of Total Probability

$$\begin{aligned} & P(\text{"white ball is selected"}) \\ = & \sum P\left(\text{"white ball is selected"} \,\middle|\, \begin{array}{c}\text{"urn i is} \\ \text{selected"}\end{array}\right) P\left(\begin{array}{c}\text{"urn i is} \\ \text{selected"}\end{array}\right) \\ = & \sum P(\{(i,W)\} \mid \{(i,B),(i,W)\}) P(\{(i,B),(i,W)\}) \\ = & \frac{50}{100}\frac{2}{3} + \frac{60}{140}\frac{1}{5} + \frac{20}{50}\frac{2}{15} = \frac{248}{525} = 0.47238 \end{aligned}$$