STAB52, An Introduction to Probability - Tutorial 2

Chong Chen

University of Toronto, Scarborough

We begin by discussing independence and conditional probability. Two events A and B are said to be independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$. Recall from section 1.4 (Evans & Rosenthal) that this amounts to saying that the combinatorial multiplication rule" is preserved. In general, a set of events $A_1, ..., A_n$ are said to be independent if

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

For a concrete example, consider flipping a coin 10 times and let $A_i, i = 1, ..., 10$ be the event that the i^{th} flip is a head. Without any prior education in probability, the lay person would surmise that the A_i are independent events since one of the coins coming up heads has no influence on any other coin coming up heads. The definition of independence in probability theory is motivated by capturing this notion. One can easily verify that the definition of independence is satisfied here for the A_i (this essentially boils down to saying that we can apply the multiplication rule". Let us now consider the following problem:

Problem 1. (Independence of finite compliment events)
Suppose that $A_1, A_2, ..., A_n$ are independent events. Show that independence still holds if we replace any number of the A_i's with A_i^c. That is, $B_1, B_2 ..., B_n, B_i \in \{A_i, A_i^c\}$ are independent. *(Hint: Argue that it suffices to show $A_1^c, A_2, ..., A_n$ are independent).*

Solution: Following the hint, if we can show that $A_1^c, A_2, ..., A_n$ are independent, then without loss of generality the events $A_1, ..., A_j^c, ..., A_n$ are independent for $1 \leq j \leq n$. Using this result iteratively, it follows that $B_1, B_2 ..., B_n, B_i \in \{A_i, A_i^c\}$ are independent and we are done.

So it suffices to prove $A_1^c, A_2, ..., A_n$ are independent. To do this, we compute

$$\begin{aligned}
\mathbb{P}(A_1^c \cap A_2 \cap ... \cap A_n) &= \mathbb{P}(A_2 \cap ... \cap A_n) - P(A_1 \cap ... \cap A_n) \\
&= \mathbb{P}(A_2) \cdots \mathbb{P}(A_n) - \mathbb{P}(A_1) \cdots \mathbb{P}(A_n) \\
&= \mathbb{P}(A_2) \cdots \mathbb{P}(A_n) \cdot [1 - \mathbb{P}(A_1)] \\
&= \mathbb{P}(A_1^c) \cdot \mathbb{P}(A_2) \cdots \mathbb{P}(A_n),
\end{aligned}$$

where the first equality uses the law of total probability and the second line is simply the definition of the events $A_1, A_2, ..., A_n$ being independent.

□

Often, we wish to consider the probability of events given some information. Given events A and B such that $\mathbb{P}(B) > 0$, we define

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

as the "probability of A given B". Intuitively, this is the probabilistic proportion of of the event B which is also contained in the events A. It is natural to conjecture that if A and B are independent events, then $\mathbb{P}(A|B) = \mathbb{P}(A)$ (i.e. conditioning on an independent event does not influence the probability). This is indeed true and follows immediately from definitions. We now consider a famous problem which demonstrates the power of conditional probability.

Problem 2. (Monty Hall Problem)

First see `http://www.youtube.com/watch?v=Zr_xWfThjJ0`. Though this link does not explain the mathematical reasoning behind solving the Monty Hall Problem (an exercise left to the reader), it does present the problem nicely.

We now motivate "continuity of probability". In calculus, the idea of a limit is the number L that a sequence of numbers $(a_n)_{n=1}^{\infty}$ will eventually approach (this is essentially what the δ, ϵ definition is trying to capture). In set theory, though it makes no sense to talk about limits of sets (this would have to be a number after all), we do have the analogous ideas of

$$\bigcup_{n=1}^{\infty} A_n \quad \text{and} \quad \bigcap_{n=1}^{\infty} A_n.$$

That is, given infinitely many sets $\{A_n\}_{n=1}^{\infty}$, we can take the union and intersection of all of them to yield a new set (or event in the probability context). To calculate the probabilities of such events requires the aforementioned "continuity of probability". This is, if $\{A_n\}_{n=1}^{\infty}$ are events such that either $A_n \subseteq A_{n+1}$ or $A_{n+1} \subseteq A_n$, then

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right)$$

respectively. So the probability of an infinite union or intersection of events is the limit of their probabilities provided they are "growing" or "shrinking". This is a rather abstract idea so let us examine two concrete examples.

Problem 3. (Evans & Rosenthal, Exercise 1.6.10)
Let $\mathbb{P}$ be some probability measure on the sample space $\mathcal{S} = [0, 1]$.
(a) Prove that we must have $\lim_{n\to\infty} \mathbb{P}((0, 1/n)) = 0$.
(b) Show by example that we might have $\lim_{n\to\infty} \mathbb{P}([0, 1/n)) > 0$.

Solution: (a) Let $A_n = (0, 1/n)$. Then clearly $A_{n+1} \subseteq A_n$ and

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} (0, 1/n) = \emptyset$$

for if $\omega \in \cap_{n=1}^{\infty}(0, 1/n)$, then $\omega < 1/n$ for all n (i.e. $\omega \leq 0$). But $\omega \in (0, 1/n)$ implies $\omega > 0$. Hence, no such ω can exist. By continuity of probability, we conclude

$$\lim_{n\to\infty} \mathbb{P}((0, 1/n)) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbb{P}(\emptyset) = 0$$

as required.

$\square$

(b) Given $\mathcal{S} = [0, 1]$, let $\mathbb{P}(\{0\}) = 1$ (i.e. all the probability is concentrated at $\{0\}$). Let $A_n = [0, 1/n)$. Then $\{A_n\} \searrow \{0\}$ and so by continuity of probability, we have

$$\lim_{n\to\infty} \mathbb{P}([0, 1/n)) = \mathbb{P}(\{0\}) = 1 > 0.$$

$\square$

Problem 4. (Evans & Rosenthal, Exercise 1.6.1)
Suppose we know that $\mathbb{P}$ is *finitely additive*, but we do not know that it is *countably additive*. In other words, that $\mathbb{P}(A_1 \cup ... \cup A_n) = \mathbb{P}(A_1) + ... + \mathbb{P}(A_n)$ for any finite collection of disjoint events $\{A_1, ..., A_n\}$, but we do not know about $\mathbb{P}(A_1 \cup A_2 \cup ...)$ for infinite collection of disjoint events. Suppose further that we know that $\mathbb{P}$ is continuous in the sense of Theorem 1.6.1. Using this, give a proof that $\mathbb{P}$ must be countable additive.

(In effect, you are proving that continuity of $\mathbb{P}$ is equivalent to countable additivity of $\mathbb{P}$, at least one we know that $\mathbb{P}$ is finite additive).

Solution: Let $\{A_n\}_{n=1}^{\infty}$ be disjoint events. Define $B_n = \bigcup_{i=1}^{n} A_i$. Then $B_n \subseteq B_{n+1}$ and hence by finite additivity we have

$$\mathbb{P}\left(\bigcup_{n=1}^{n} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{n} B_n\right) = \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} \mathbb{P}(A_1 \cup ...A_n) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

as required.

$\square$