STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-09-14 12:52

a non-everyday application of this axiomatic approach to P

towards showing the "continuity" of P

The culmination of our axiomatic approach will be to define the notion of "continuity" for P and prove that the defined property holds.

Recall from the prerequiste the notion of a *continuous* function. There are several equivalent definitions, one of which uses left- and right-continuity.

A function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is left-continuous at x if for any non-decreasing sequence $x_1 \leq x_2 \leq x_3 \leq \dots$ that converges to x, then $f(x_i) \rightarrow f(x)$.

(Right continuity is the same but with a non-increasing sequence.)

f is continuous at x if it is left- and right-continuous at x.

f is continuous if it is continuous at every point in its domain.

increasing sequences of events that converge

The domain of P is a collection of events $\mathcal{A}$. We need a notion of the following for events:

$$A_1, A_2, \dots \text{ increases to } A$$

Definition: $A_n \nearrow A$ means $A_i \subseteq A_{i+1}$ and $\bigcup_{i=1}^{\infty} A_i = A$

Example: Consider $S = (0, 1)$. Let $A_n = \left(0, \frac{1}{2} - \frac{1}{2^{n+1}}\right)$ for $n \geq 1$ and $A = \left(0, \frac{1}{2}\right)$

What about the probabilities of these events under the uniform model?

the continuity theorem

Theorem 6 (The Continuity Theorem): If A_n and A are events and $A_n \nearrow A$, then $P(A_n) \longrightarrow P(A)$.

Proof: $\ldots$

This is analogous to left-continuity. There is also a right-continuity:

Corallary: Suppose A_n and A are events in $\mathcal{A}$ with $A_i \supseteq A_{i+1}$ and $\bigcap_{i=1}^{\infty} A_i = A$. Then $P(A_n) \longrightarrow P(A)$.

Proof: The Continuity Theorem, a de Morgan's Law, and "Theorem 3".

Something to try if you like: finite additivity together with The Continuity Theorem implies $\sigma-$additivity.

application to the continuous sample space example

Reconsider the uniform pick on $S = (0, 1)$, where the probability of choosing a number in any $0 < a < b < 1$ is $b - a$.

What is the probability of choosing exactly $\frac{1}{2}$?

Let A be the event that the number chosen is *rational*. What is P(A)?

some computations for finite and countable sample spaces

finite and countable S in general

Starting with:

$$S = \{\omega_1, \ldots, \omega_n\} \qquad \text{(finite), or,}$$
$$S = \{\omega_1, \omega_2, \omega_3, \ldots\} \qquad \text{(countable)}$$

then a valid probability can always be based on $P(\{\omega_i\}) = p_i$ with $0 \leq p_i \leq 1$ and $\sum p_i = 1$.

An important special case for finite S is the uniform case: $p_i = \frac{1}{n}$.

In this case many problems can be solved by counting the number of outcomes in S and counting the number of outcomes in an event.

Some people enjoy these problems. Others don't. Fortunately for you, I do not!

permutations and combinations

At the very least we'll need to recall (or learn!) these.

Number of ways to choose k items out of n where order matters:

$$ {}_nP_k = \begin{cases} 0 & \text{if k > n,} \\ \frac{n!}{(n-k)!} & \text{otherwise.} \end{cases} $$

and when order doesn't matter:

$$ {}_nC_r = \binom{n}{k} = \frac{n!}{k!(n-k)!} $$

Two classic examples: "The Birthday Problem" and "Lotto"

conditional probability

partial information

I'll role a six-sided die. $S = \{1, 2, 3, 4, 5, 6\}$. Consider these events:

$$A = \{2, 5\},$$
$$B = \{2, 4, 6\},$$
$$C = \{1, 2\}.$$

So $P(A) = \frac{2}{6} = \frac{1}{3}$.

What if I peek and tell you "Actually, B occurred". What is the probabality of A given this partial information? It is $\frac{1}{3}$.

I roll the die again, peek, and tell you "Actually, C occurred". Now the probability of A is $\frac{1}{2}$.

Intuitively we used a "sample space restriction" approach.

elementary definition of conditional probability

Given B with $P(B) > 0$,

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

"The conditional probability of A given B"

The answers for the previous example coincide with the intuitive approach.

Theorem 7: For a fixed B with $P(B) > 0$, the function $P_B(A) = P(A|B)$ is a probability measure.

Proof: exercise.

useful expressions for calculation - I

$P(A \cap B) = P(A|B)P(B)$ often comes in handy.

Consider the testing for, and prevalence of, a viral infection such as HIV.

Denote by A the event "tests positive for HIV", and by B the event "is HIV positive."

For the ELISA screening test, $P(A|B)$ is about 0.995. The prevalence of HIV in Canada is about $P(B) = 0.00212$.

useful expressions for calculation - II

Take some event B. The sample space can be divided in two into B and B^C.

This is an example of a *partition* of S, which is generally a collection $B_1, B_2, \dots$ of disjoint events (could be infinite) such that $\bigcup_i B_i = S$.

Theorem 8: If $B_1, B_2, \dots$ is a partition of S with all $P(B_i) > 0$, then

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

Proof: …

Continuing with the HIV example, suppose we also know $P(A|B^c) = 0.005$ ("false positive").

We can now calculate $P(A)$.

useful expressions for calculation - III

Much to my amusement, Theorem 8 gets a grandiose title: ***"THE! LAW! OF! TOTAL! PROBABILITY!!!"***

Now, in the HIV example, we also might be interested in $P(B|A)$, the chance of an HIV+ person testing positive.

A little algebra:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

In our example this is $\frac{0.0021094}{0.0070988} = 0.2971$.

Bayes' rule

Theorem 9: If $B_1, B_2, \dots$ is a partition of S with all $P(B_i) > 0$, then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_i P(A|B_i)P(B_i)}$$