Math 340: Lec 3 (Conditional probability)

Asa Royal (ajr74)

January 18, 2024

Conditional probability

Definition 1 (conditional probability)**.** Given two events $A, B \subset \Omega$ with $\mathbb{P}(B) > 0$,

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

This is the probability that an outcome in A and B occurs divide by the probability that an outcome in B occurs. In a manner of thinking, B has become the updated sample space, and the event we're calculating the probability of is the set of outcomes in A that overlap with B.

Note: conditional probability requires that $\mathbb{P}(B) \neq 0$, because it doesn't make much sense to think about the probability of A given an event that cannot happen.

Conditional probabilities let us update our models with new information. For example, imagine we have some prior expectation about $\mathbb{P}(A)$. If we know B has occurred, an updated forecast for A is $\mathbb{P}(A|B)$.

Bayes' Theorem

Derivation

A corollary of the formula for conditional probability is

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

Since $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$,

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)} \tag{1}$$

Equation (1) is called Bayes' Theorem: it allows us to express the conditional probability of A on B in terms of the conditional probability of B on A, once again giving us a way to update models.

Partition rule

Definition 2 (partition)**.** Events $B_1, \ldots, B_n$ are a partition of Ω if the B_k'sQ are mutually disjoint and their union is equal to Ω.

Theorem 3 (Partition rule)**.** The partition rule states that for any event $A \subset \Omega$, the sets $\{(A \cap B_K)\}^n$ are a partition of A. i.e. the set of those intersections is disjoint and their union is A. Mathematically, by the additivity proposition,

$$\mathbb{P}(A) = \sum_{k=1}^{n} \mathbb{P}(A|B_k)\mathbb{P}(B_k)$$

p>

Remark. Note 2: We can think think about the calculations involved in the partition formula as cascading probabilities in a decision tree.

Remark. The partition rule is used when a problem's setup includes randomization in the first stage, then selection. The randomization forms partitions! Problem examples: Boxes and marbles.

Theorem 4 (Baye's theorem w/ partition rule). We can express the denominator in Baye's theorem as a sum of disjoint partition intersections. I.e.

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\sum_{k=1}^{n} \mathbb{P}(A|B_k)\mathbb{P}(B_k)}$$

Note 3: We can use the partition rule to calculate the denominator in the Baye's theorem formula by summing up the weighted conditional probability of the denominator event. See e.g., medical diagnosis problem.