Elements of Probability and Statistics

# Lecture 05: Conditional Probability, Bayes Rule, Independence of Events

IAI, TCG-CREST August 22 & 25, 2023

## 4.1 Conditional Probabilities

The tool of conditioning allows us to consider events that are observed to occur, and update our beliefs towards the subsequent events. This ability to update our belief on events is a very useful tool to determine probabilities associated with complex random experiments. Conditional probabilities are a ubiquitous concept that we'll familiarize with before discussing probability distributions. One can even argue that there is never an absolute probability, i.e., all probabilities are conditional.

**Defn.** Conditional probability: If $A$ and $B$ are events with $P(B) > 0$, then the conditional probability of $A$ given $B$, denoted by $P(A|B)$, is defined as,

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

From the Bayesian perspective, conditional probabilities are conceptualized in the following manner. The event $A$ initially has a *prior* probability $P(A)$. On observation of evidence $B$, the probability of event $A$ gets updated to the *posterior* probability $P(A|B)$.

From the frequentist perspective, we are observing two events occurring several times, from which we only consider those events where $B$ occurs. From these selected events, we identify the events where $A$ and $B$ both occur.

Note that '$A|B$' is not a new event, it is still the event $A$, whose probability is being evaluated given the observation of event $B$.

The conditional probability of any event given an event $E$, i.e., $P(.|E)$, is a valid probability, which can be verified from the axioms of probability:

$$P(S|E) = 1,\; P(\phi|E) = 0,\; \text{and for disjoint } A_1, A_2, ...,\; P(\cup_{i=1}^{\infty} A_i|E) = \sum_{i=1}^{\infty} P(A_i|E).$$

**Q:** There exists two coins, one is a fair coin whereas the other coin gives heads 70% of the time. One of these two coins is selected at random, and tossed twice. Two heads, i.e., 'HH' is observed to occur. What is the probability that the first coin was selected?

**Q:** From a shuffled deck of playing cards, two cards are drawn randomly one after the other without replacement. Let $A$ be the event that the first card is a diamond, and $B$ be the event that the second card is red. Find $P(A|B)$ and $P(B|A)$.

From the definition of conditional probability, the following identities can be derived:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$

These identities can be generalized to $n$ events $A_1, ..., A_n$:

$$P(A_1, ..., A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1, A_2)...P(A_n|A_1, ..., A_{n-1}).$$

Here commas (,) represent intersection ($\cap$). The above equation actually defines $n!$ number identities. As an example for 3 events,

$$\begin{aligned}P(A_1, A_2, A_3) &= P(A_1)P(A_2|A_1)P(A_3|A_1, A_2)\\ &= P(A_2)P(A_3|A_2)P(A_1|A_2, A_3)\\ &= \text{... and so on.}\end{aligned}$$

### 4.2 Bayes Rule and LOTP

In a wide variety of problems, we may be interested in finding $P(A|B)$, however calculating $P(B|A)$ may be relatively easier. In such situations, we can use $P(B|A)$ to calculate $P(A|B)$. This is done using Bayes Rule, which can be derived from the previously defined identities.

**Defn.** Bayes Rule:

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

Event $B$ is observed, however that does not mean that the denominator $P(B)$ is 1. It is instead usually calculated using the Law of Total Probability.

**Defn.** Law of Total Probability (LOTP): If the $n$ events $A_1, ..., A_n$ are disjoint, with $P(A_i) > 0\ \forall i$, and they partition the sample space $S$, then the probability of any event $B$ can be calculated as,

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

As a corollary of the above, if we are interested in finding $P(A|B)$ using Bayes Rule, then the denominator $P(B)$ can always be calculated as $P(B|A)P(A) + P(B|A^c)P(A^c)$, since $A$ and $A^c$ are disjoint and partition the sample space.

**Q.** A rare disease exists that affects 1% of the population. There exists a test for the disease, which is 95% accurate. For a patient, the test has returned positive, then what is the probability that the patient has the disease?

Both Bayes Rule and LOTP can be extended to incorporate more than one event as evidence.

$$P(A_1|B, E) = \frac{P(B|A_1, E)P(A_1|E)}{P(B|E)}\ ,$$

$$\text{with, } P(B|E) = \sum_{i=1}^{n} P(B|A_i, E)P(A_i|E)\ .$$

### 4.3 Independence of Events

Conditional probabilities are defined to be able to update our belief of an event, given certain observed evidence. However there can also be situations where the observation of some evidence does not change our belief of an event. This situation motivates the definition of the independence of events.

**Defn.** Independence of two events: Two events $A$ and $B$ are independent if $P(A \cap B) = P(A)P(B)$.

This implies that even after observing event $B$, the posterior remains equal to the prior, i.e., $P(A|B) = P(A)$. Equivalently, independence also implies that $P(B|A) = P(B)$.

Note that in general, *disjoint events do not imply independence*. Therefore if $P(A \cap B) = 0$, then $A$ and $B$ can be independent for the trivial case where either $P(A) = 0$ or $P(B) = 0$. Intuitively for non-trivial cases, disjoint events should not be independent: If we know two events $A$ and $B$ are disjoint, then the occurrence of $A$ necessarily implies that $B$ cannot occur, hence the probability of observing $B$ will be affected by the observation of $A$ and vice-versa.

**Defn.** Independence of three events: Three events $A, B, C$ are independent if,

$$\begin{aligned} P(A \cap B) &= P(A)P(B), \\ P(B \cap C) &= P(B)P(C), \\ P(C \cap A) &= P(C)P(A), \\ P(A \cap B \cap C) &= P(A)P(B)P(C)\ . \end{aligned}$$

The first three equations from the above are conditions for *pairwise-independence* between the three events $A$, $B$ and $C$. Only when all four equations are true do we state that the three events are *independent*. Pairwise independence in general does not imply independence.

The independence conditions above can be generalized: $n$ events are independent if all finite subsets of the events are independent.

**Defn.** Conditional Independence: Events $A$ and $B$ are said to be conditionally independent given event $E$ if $P(A \cap B|E) = P(A|E)P(B|E)$.

Note the following:

- The conditional independence of $A$ and $B$ given $E$ does not imply their independence with respect to $E^c$.
- Conditional independence of $A$ and $B$ given $E$ does not imply the independence of $A$ and $B$.
- The independence of $A$ and $B$ does not imply the conditional independence of $A$ and $B$ given $E$.

**Q.** Simpson's Paradox: Two doctors, Dr. Hilbert and Dr. Nick, each perform two types of surgeries that have very different levels of difficulty: heart surgery and band-aid application. The outcome of each surgery can be a success or a failure. The records of the two doctors are given in the following tables:

Dr. Hilbert

| | Heart | Band-Aid |
|---|---|---|
| Success | 70 | 10 |
| Failure | 20 | 0 |

Dr. Nick

| | Heart | Band-Aid |
|---|---|---|
| Success | 2 | 81 |
| Failure | 8 | 9 |

Dr. Hilbert has a significantly higher success rate than Dr. Nick in surgeries, and also a higher rate of success in band-aid application. But when aggregated, Dr. Hilbert has a success rate of 80% overall whereas Dr. Nick has a success rate of 83%. How can LOTP help explain this paradox?

**Q.** The Marilyn vos Savant / Monty Hall Problem: In a game show hosted by Monty Hall, a participant is presented initially with 3 doors. Behind one of the doors is a car, whereas a goat is present behind each of the other two doors.

The participant is first asked to choose a door. Once the participant selects a door, Monty Hall opens one of the two remaining doors that does not have a car behind it, to reveal the goat behind it.

Thereafter, the participant is now left with two remaining doors. The participant has two choices. Either they can stay on the door they initially chose. Or they can switch to the other remaining door.

Between these two strategies: (i) staying on the initial door, and (ii) switching to the other remaining door, which of these two strategies is more likely to let the participant win the car?

Avisek Gupta
Postdoctoral Fellow
IAI, TCG CREST.