CMPT 210: Probability and Computing

Lecture 9

Sharan Vaswani

February 2, 2023

Recap

For events E and F, we wish to compute $\Pr[E|F]$, the probability of event E conditioned on F.

Approach 1: With conditioning, F can be interpreted as the *new sample space* such that for $\omega \notin F$, $\Pr[\omega|F] = 0$.

Approach 2: $\Pr[E|F] = \frac{\Pr[E \cap F]}{\Pr[F]}$.

Multiplication Rule: For events $E_1, E_2, \ldots, E_n$,
$\Pr[E_1 \cap E_2 \ldots \cap E_n] = \Pr[E_1] \Pr[E_2|E_1] \Pr[E_3|E_1 \cap E_2] \ldots \Pr[E_n|E_1 \cap E_2 \cap \ldots E_{n-1}]$.

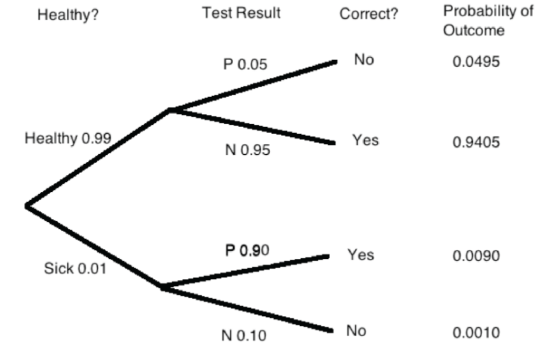
Conditional Probability - Examples

Q: A test for detecting cancer has the following accuracy – (i) If a person has cancer, there is a 10% chance that the test will say that the person does not have it. This is called a "false negative" and (ii) If a person does not have cancer, there is a 5% chance that the test will say that the person does have it. This is called a "false positive". For patients that have no family history of cancer, the incidence of cancer is 1%. Person X does not have any family history of cancer, but is detected to have cancer. What is the probability that the Person X does have cancer?

Conditional Probability - Examples

$\mathcal{S} = \{(Healthy, Positive), (Healthy, Negative), (Sick, Positive), (Sick, Negative)\}$.

A is the event that Person X has cancer. B is the event that the test is positive.



$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{\Pr[\{(S,P)\}]}{\Pr[\{(S,P),(H,P)\}]} = \frac{0.0090}{0.0090+0.0495} \approx 15.4\%.$$

Questions?

Conditional Probability

Conditional probability for complement events: For events E, F, $\Pr[E^c|F] = 1 - \Pr[E|F]$.

Proof: Since $E \cup E^c = \mathcal{S}$, for an event F such that $\Pr[F] \neq 0$,

$$
\begin{aligned}
(E \cup E^c) \cap F &= \mathcal{S} \cap F = F \\
(E \cup E^c) \cap F &= (E \cap F) \cup (E^c \cap F) && \text{(Distributive Law)} \\
\implies \Pr[(E \cap F) \cup (E^c \cap F)] &= \Pr[(E \cup E^c) \cap F]
\end{aligned}
$$

Since $E \cap F$ and $E^c \cap F$ are mutually exclusive events,

$$
\begin{aligned}
\Pr[E \cap F] + \Pr[E^c \cap F] = \Pr[F] &\implies \frac{\Pr[E^c \cap F]}{\Pr[F]} = 1 - \frac{\Pr[E \cap F]}{\Pr[F]} \\
&\implies \Pr[E^c|F] = 1 - \Pr[E|F] && \text{(By def. of conditional probability)}
\end{aligned}
$$

Bayes Rule

Bayes Rule: For events E and F if $\Pr[E] \neq 0$ and $\Pr[F] \neq 0$, then, $\Pr[F|E] = \frac{\Pr[E|F]\Pr[F]}{\Pr[E]}$.
Proof: Using the formula for conditional probability,

$$\Pr[E|F] = \frac{\Pr[E \cap F]}{\Pr[F]} \quad ; \quad \Pr[F|E] = \frac{\Pr[F \cap E]}{\Pr[E]}$$

$$\implies \Pr[E \cap F] = \Pr[E|F]\Pr[F] \quad ; \quad \Pr[F \cap E] = \Pr[F|E]\Pr[E]$$

$$\implies \Pr[E|F]\Pr[F] = \Pr[F|E]\Pr[E]$$

$$\implies \Pr[F|E] = \frac{\Pr[E|F]\Pr[F]}{\Pr[E]}$$

Allows us to compute $\Pr[F|E]$ using $\Pr[E|F]$. Later in the course, we will see an application of the Bayes rule to machine learning.

Law of Total Probability and Bayes rule

Law of Total Probability: For events E and F, $\Pr[E] = \Pr[E|F]\Pr[F] + \Pr[E|F^c]\Pr[F^c]$.

Proof:

$$E = (E \cap F) \cup (E \cap F^c)$$

$$\implies \Pr[E] = \Pr[(E \cap F) \cup (E \cap F^c)] = \Pr[E \cap F] + \Pr[E \cap F^c]$$

(By union-rule for disjoint events)

$$\Pr[E] = \Pr[E|F]\Pr[F] + \Pr[E|F^c]\Pr[F^c]$$

(By definition of conditional probability)

Combining Bayes rule and Law of total probability

$$\Pr[F|E] = \frac{\Pr[F \cap E]}{\Pr[E]} = \frac{\Pr[E|F]\ \Pr[F]}{\Pr[E]}$$

(By definition of conditional probability)

$$\Pr[F|E] = \frac{\Pr[E|F]\ \Pr[F]}{\Pr[E|F]\Pr[F] + \Pr[E|F^c]\Pr[F^c]}$$

(By law of total probability)

Questions?

Total Probability - Examples

Q: In answering a question on a multiple-choice test, a student either knows the answer or she guesses. Let p be the probability that she knows the answer and $1 - p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $\frac{1}{m}$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

Let C be the event that the student answers the question correctly. Let K be the event that the student knows the answer. We wish to compute $\Pr[K|C]$.

We know that $\Pr[K] = p$ and $\Pr[C|K^c] = 1/m$, $\Pr[C|K] = 1$. Hence,
$\Pr[C] = \Pr[C|K]\Pr[K] + \Pr[C|K^c]\Pr[K^c] = (1)(p) + \frac{1}{m}(1-p)$.

$\Pr[K|C] = \frac{\Pr[C|K]\ \Pr[K]}{\Pr[C]} = \frac{mp}{1+(m-1)p}$.

Total Probability - Examples

Q: An insurance company believes that people can be divided into two classes — those that are accident prone and those that are not. Their statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability decreases to 0.2 for a non-accident-prone person. If we assume that 30% of the population is accident prone, what is the probability that a new policy holder will have an accident within a year of purchasing a policy?

Let A = event that a new policy holder will have an accident within a year of purchasing a policy. Let B = event that the new policy holder is accident prone. We know that $\Pr[B] = 0.3$, $\Pr[A|B] = 0.4$, $\Pr[A|B^c] = 0.2$. By the law of total probability,
$\Pr[A] = \Pr[A|B]\Pr[B] + \Pr[A|B^c]\Pr[B^c] = (0.4)(0.3) + (0.2)(0.7) = 0.26$.

Q: Suppose that a new policy holder has an accident within a year of purchasing their policy. What is the probability that they are accident prone?

Compute $\Pr[B|A] = \frac{\Pr[A|B]\ \Pr[B]}{\Pr[A]} = \frac{0.12}{0.26} = 0.4615$.

Total Probability - Examples

Q: Alice is taking a probability class and at the end of each week she can be either up-to-date or she may have fallen behind. If she is up-to-date in a given week, the probability that she will be up-to-date (or behind) in the next week is 0.8 (or 0.2, respectively). If she is behind in a given week, the probability that she will be up-to-date (or behind) in the next week is 0.6 (or 0.4, respectively). Alice is (by default) up-to-date when she starts the class. What is the probability that she is up-to-date after three weeks?

Let U_i and B_i be the events that Alice is up-to-date or behind respectively after i weeks. Since Alice starts the class up-to-date, $\Pr[U_1] = 0.8$ and $\Pr[B_1] = 0.2$. We also know that $\Pr[U_2|U_1] = 0.8$, $\Pr[U_3|U_2] = 0.8$ and $\Pr[B_2|U_1] = 0.2$, $\Pr[B_3|U_2] = 0.2$. Similarly, $\Pr[U_2|B_1] = 0.6$, $\Pr[U_3|B_2] = 0.6$ and $\Pr[B_2|B_1] = 0.4$, $\Pr[B_3|B_2] = 0.4$.

We wish to compute $\Pr[U_3]$. By the law of total probability,
$\Pr[U_3] = \Pr[U_3|U_2]\Pr[U_2] + \Pr[U_3|B_2]\Pr[B_2]$ and
$\Pr[U_2] = \Pr[U_2|U_1]\Pr[U_1] + \Pr[U_2|B_1]\Pr[B_1]$.

Hence, $\Pr[U_2] = (0.8)(0.8) + (0.6)(0.2) = 0.76$, and $\Pr[U_3] = (0.8)(0.76) + (0.6)(0.24) = 0.752$.

Simpson's Paradox

In 1973, there was a lawsuit against a university with the claim that a male candidate is more likely to be admitted to the university than a female.

Let us consider a simplified case – there are two departments, EE and CS, and men and women apply to the program of their choice. Let us define the following events: A is the event that the candidate is admitted to the program of their choice, F_E is the event that the candidate is a woman applying to EE, F_C is the event that the candidate is a woman applying to CS. Similarly, we can define M_E and M_C. Assumption: Candidates are either men or women, and that no candidate is allowed to be part of both EE and CS.

Lawsuit claim: Male candidate is more likely to be admitted to the university than a female i.e. $\Pr[A|M_E \cup M_C] > \Pr[A|F_E \cup F_C]$.

University response: In any given department, a male applicant is less likely to be admitted than a female i.e. $\Pr[A|F_E] > \Pr[A|M_E]$ and $\Pr[A|F_C] > \Pr[A|M_C]$.

Simpson's Paradox: Both the above statements can be simultaneously true.

Simpson's Paradox

CS	2 men admitted out of 5 candidates	40%
	50 women admitted out of 100 candidates	50%
EE	70 men admitted out of 100 candidates	70%
	4 women admitted out of 5 candidates	80%
Overall	72 men admitted, 105 candidates	$\approx$ 69%
	54 women admitted, 105 candidates	$\approx$ 51%

In the above example, $\Pr[A|F_E] = 0.8 > 0.7 = \Pr[A|M_E]$ and $\Pr[A|F_C] = 0.5 > 0.4 = \Pr[A|M_C]$. $\Pr[A|F_E \cup F_C] \approx 0.51$. Similarly, $\Pr[A|M_E \cup M_C] \approx 0.69$.

In general, Simpson's Paradox occurs when multiple small groups of data all exhibit a similar trend, but that trend reverses when those groups are aggregated.

Questions?