STAT 516: Basic Probability and its Applications

Lecture 3: Conditional Probability and Independence

Prof. Michael Levine

September 29, 2015

Motivating Example

- Experiment ξ consists of rolling a fair die twice;
  $A = \{$ the first roll is 6 $\}$ amd
  $B = \{$ the sum of the two rolls is 12$\}$.
- Under the assumption of the equally likely sample points, $P(B) = \frac{1}{36}$
- If we know that A already happened, intuitively we feel that the probability of the event B, given A, should be $\frac{1}{6}$
- We say that $P(B|A) = \frac{1}{6}$ and call it the **conditional probability** of A
- The conditional probability tells us *among the times that A already happened how often B also happens*

Formal definition of the conditional probability

- Conditional probability results from restricting the outcomes to only those that are *inside* C_1 where $P(C_1 > 0)$
- Therefore, instead of

$$P(C_2) = \frac{N(C_2)}{N(\mathcal{S})}$$

we have

$$\frac{N(C_2 \cap C_1)}{N(\mathcal{S} \cap C_1)} = \frac{P(C_2 \cap C_1)}{P(\mathcal{S} \cap C_1)} = \frac{P(C_2 \cap C_1)}{P(C_1)}$$

- Thus, the definition:

$$P(C_2|C_1) = \frac{P(C_2 \cap C_1)}{P(C_1)}$$

Properties of Conditional Probabilities

- $P(C_2|C_1) \geq 0$
- $P(C_1|C_1) = 1$
- $P(\cup_{j=2}^{\infty} C_j|C_1) = \sum_{j=2}^{\infty} P(C_j|C_1)$

Example I

- Consider the situation where of all individuals buying a digital camera 60% include an optional memory card, 40% - an extra battery and 30% - both. What is the probability that a person who buys an extra battery also buys a memory card?
- Let $A = \{\text{memory card purchased }\}$ and $B = \{\text{battery purchased }\}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.4} = 0.75$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.6} = 0.50$$

Multiplication Rule I

- A simple **multiplication rule** follows from the definition of conditional probability
- For any two events C_1 and C_2 such that $P(C_1) > 0$

$$P(C_1 \cap C_2) = P(C_1)P(C_2|C_1)$$

- For example, take a bowl with eight chips. 3 of the chips are red and the remaining 5 are blue. 2 chips are drawn without replacement.
- What is the probability that the 1st draw results in a red chip (C_1) and the second draw results in a blue chip (C_2)

Multiplication rule II

- It is reasonable to assume that $P(C_1) = \frac{3}{8}$ and $P(C_2|C_1) = \frac{5}{7}$
- Then, $P(C_1 \cap C_2) = \frac{3}{8}\frac{5}{7} = \frac{15}{56} = 0.2679$

Multiplication rule III

- The multiplication rule can be extended to three or more events
- Using the basic multiplication rule iteratively, and assuming $P(C_1 \cap C_2) > 0$, we have

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2|C_1)P(C_3|C_1 \cap C_2)$$

Multiplication rule IV

- Four cards are dealt successively from an ordinary deck of playing cards
- Dealing is *at random* and *without replacement*
- The probability of receiving a spade, a heart, a diamond, and a club, in that order, is

$$\frac{13}{52}\frac{13}{51}\frac{13}{50}\frac{13}{49} = 0.0044$$

The Law of Total Probability

- Let $C_1, \ldots, C_k$ form a **partition** of the sample space S
- This means that $C_i \cap C_j = \varnothing$ for all $i \neq j$ and $\cup_{i=1}^{k} A_i = S$
- Moreover, let $0 < P(C_i) < 1$, $i = 1, 2, \ldots, k$
- Then, the **Law of Total Probability** is

$$P(C) = \sum_{i=1}^{k} P(C_i) P(C|C_i)$$

Example

- One of the cards from a deck of 52 cards is missing from the deck but we don't know which one
- One card is chosen at random from the remaining 51 cards. What is the probability that it is a spade?
- Let $A =$ the missing card is a spade and $B=$ the card chosen from the remaining ones is a spade
- Then,

$$P(B) = P(B|A)P(A) + P(B|A^{'})P(A^{'}) = \frac{12}{51}\frac{1}{4} + \frac{13}{51}\frac{3}{4} = \frac{1}{4}$$

- Note that it is impossible to have 12.5 spade cards in the remaining deck of 51 cards...but $P(B) = \frac{1}{4}$ nevertheless

Example

- A certain item is produced in a factory on one of the three machines A, B, or C. The percentages of items produced on A, B, and C, are 50%, 30%, and 20%, respectively
- 4%, 2%, and 4% of their products, respectively, are defective. What is the percentage of all copies of this item that are defective?
- Let the event that an item is produced by A, B, or C, respectively, be A_1, A_2, and A_3; let D be the event that it is a defective item
- $$P(D) = \sum_{i=1}^{3} P(D|A_i)P(A_i) = .04\times.5+.02\times.3+.04\times.2 = .034$$

A more complicated example

- A lottery with n tickets, out of which m prespecified tickets will win a prize
- There are n players who choose one ticket at random successively from available tickets
- What is a probability that the ith player wins a prize?

A more complicated example

- For the first player, the probability is $p_1 = \frac{m}{n}$
- By total probability formula,

$$p_2 = \frac{m-1}{n-1} \times \frac{m}{n} + \frac{m}{n-1} \times \left(1 - \frac{m}{n}\right) = \frac{m}{n}$$

- Again, by total probability formula

$$\begin{aligned} p_3 &= \frac{m}{n} \times \frac{m-1}{n-1} \times \frac{m-2}{n-2} + \left(1 - \frac{m}{n}\right) \times \frac{m}{n-1} \times \frac{m-1}{n-2} \\ &+ \frac{m}{n} \times \left(1 - \frac{m-1}{n-1}\right) \times \frac{m-1}{n-2} \\ &+ \left(1 - \frac{m}{n}\right) \times \left(1 - \frac{m}{n-1}\right) \times \frac{m}{n-2} \\ &= \frac{m}{n} \end{aligned}$$

- It can be shown by induction that for any $i \geq 1$ we have $p_i = \frac{m}{n}$ - no early entry advantage for this lottery!

Bayes Theorem

- Also, the **Bayes theorem** follows directly:

$$P(C_j|C) = \frac{P(C \cap C_j)}{P(C)} = \frac{P(C_j)P(C|C_j)}{\sum_{i=1}^{k} P(C_i)P(C|C_i)}$$

Example III: setting up

-
- The following problem was published in a syndicated column *Ask Marilyn* that was run by Marilyn vos Savant
- Consider three fair coins that are HH, HT and TT
- A randomly selected coin is tossed; the face-up side is H
- What is the probability that the face-down side is H too?
- Face-up H $\rightarrow$ HH or HT
- If HH was selected, the face - down is H; otherwise, it is T
- Thus, the probability of H facing down is 50%... Is that correct?

Example III: the truth will set you free!

- $$P(\text{down} = H|\text{up} = H) = \frac{P(\text{down} = H \cap \text{up} = H)}{P(\text{up} = H)}$$
- Need to compute

  $$P(\text{up} = H)$$
  $$= P(HT \cap \text{up} = H \cap down = T) + P(HH \cap \text{up} = H \cap \text{down} = H)$$

- $$P(HT \cap \text{up} = H \cap \text{down} = T)$$
  $$= P(HT \cap \text{up} = H) \times P(\text{down} = T|HT \cap \text{up} = H)$$
  $$= P(HT) \times P(up = H|HT) \times 1 = \frac{1}{3}\frac{1}{2} = \frac{1}{6}$$

A simple example I

- A store stocks light bulbs from three suppliers. Suppliers A, B, and C supply 10%, 20%, and 70% of the bulbs respectively.
- It has been determined that company A's bulbs are 1% defective while company B's are 3% defective and company C's are 4% defective.
- If a bulb is selected at random and found to be defective, what is the probability that it came from supplier B?

A simple example II

- Let D = defective lightbulb

$$P(B|D) = \frac{P(B)P(D|B)}{P(A)P(D|A)+P(B)P(D|B)+P(C)P(D|C)} = \frac{0.2(0.03)}{0.1(0.01)+0.2(0.03)+0.7(0.04)} \approx 0.1714$$

An Example of Bayes Theorem in action I

- HIV testing: four outcomes are $D \cap +$, $D \cap -$, $D^{'} \cap +$ and $D^{'} \cap -$
- The **prevalence** of the disease $P(D)$ is commonly known; say, $P(D) = 0.001$
- $+|D^{'}$is a **false positive** and $-|D$ is a **false negative**.
- Diagnostic procedures undergo extensive evaluation and, therefore, probabilities of false positives and false negatives are commonly known
- We assume that $P(+|D^{'}) = 0.015$ and $P(-|D) = 0.003$; for more details, see E.M. Sloan et al. (1991) "HIV testing: state of the art", JAMA, 266:2861-2866

An Example of Bayes Theorem in action II

- The quantity of most interest is usually the **predictive value** of the test $P(D|+)$
- By definition of conditional probability and multiplication rule, we have

$$P(D|+) = \frac{P(D \cap +)}{P(+)} = \frac{P(D \cap +)}{P(D \cap +) + P(D' \cap +)}$$

$$= \frac{P(D) * P(+|D)}{P(D) * P(+|D) + P(D') * P(+|D')}$$

$$= \frac{P(D) * [1 - P(-|D)]}{P(D) * [1 - P(-|D)] + (1 - P(D)) * P(+|D')}$$

- The resulting probability will be small because false positives are much more common then false negatives in the general population

Independence

- The first definition

$$P(C_2 \cap C_1) = P(C_1)P(C_2)$$

- An alternative definition follows from multiplication formula:

$$P(C_1|C_2) = P(C_1)$$

- Note that the first definition also allows us to consider the case where $P(C_1) = P(C_2) = 0$

Some immediate consequences of the definition of independence

- The independence of A and B implies that
  1. A and $B^{'}$ are independent
  2. $A^{'}$ and B are independent
  3. $A^{'}$ and $B^{'}$ are independent

Example I

- Consider the fair six-sided die. Define $A = \{2, 4, 6\}$, $B = \{1, 2, 3\}$ and $C = \{1, 2, 3, 4\}$.
- Clearly, $P(A) = \frac{1}{2}$, $P(A|B) = \frac{1}{3}$ and $P(A|C) = \frac{1}{2}$.
- Therefore, A and C are independent but A and B are NOT!

Example II

- A red and a white die are cast independently
- Let $C_1 = \{4 \text{ red die}\}$ and $C_2 = \{3 \text{ white die}\}$
- $P(C_1) = \frac{1}{6}$ and $P(C_2) = \frac{1}{6}$
- By independence,
$$P(C_1 \cap C_2) = \frac{1}{36}$$
- The probability that the sum of the up spots is equal to 7 is
$$P[(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)] = 6 * \left(\frac{1}{6}\right)^2 = \frac{1}{6}$$

Some additional consequences of the definition of independence

- If A and B are mutually exclusive events, they **cannot** be independent.
- If either $P(A) = 0$ or $P(B) = 0$, A and B are always independent:

$$0 \leq P(A \cap B) \leq min(P(A), P(B)) = 0,$$

therefore $P(A \cap B) = 0$

Independence: practical considerations

- In practice, we do not verify independence of events. Instead, we ask ourselves whether independence is a property that we wish to incorporate into a mathematical model of an experiment, based on the common sense
- Thus, independence is commonly *assumed*
- Example: let C_2=(A student is a female) and C_1=(A student is concentrating in elementary education); clearly, $P(C_2|C_1) \neq P(C_2)$

Mutual independence

- $C_1, C_2, \ldots, C_n$ are **mutually independent** if for every $2 \leq k \leq n$, and every possible subset of indices $i_1, i_2, \ldots, i_k$

$$P(C_{i_1} \cap C_{i_2} \cap \ldots \cap C_{i_k}) = P(C_{i_1}) \cdot P(C_{i_2}) \cdot \ldots \cdot P(C_{i_k})$$

- In particular, this means that

$$P(C_1 \cap C_2 \cap C_n) = P(C_1)P(C_2)\ldots P(C_n)$$

The relationship between mutual independence and pairwise independence

- Pairwise independence DOES NOT imply mutual independence
- Let a single outcome be ω_1 and the entire sample space is $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
- Then let $C_1 = \{\omega_1, \omega_2\}$, $C_2 = \{\omega_1, \omega_3\}$ and $C_3 = \{\omega_1, \omega_4\}$
- If each outcome is equally likely with $p = \frac{1}{4}$, it is easy to check that pairwise independence is valid
- However,

$$P(C_1 \cap C_2 \cap C_3) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3 = P(C_1)P(C_2)P(C_3)$$

Example

- Consider sample space S of 36 ordered pairs (i, j) with $i, j = 1, \ldots, 6$
- We assume that for each pair the probability of occurring $p = \frac{1}{36}$
- Let $C_1 = \{(i, j) : j = 1, 2 \text{ or } 5\}$, $C_2 = \{(i, j) : j = 4, 5 \text{ or } 6\}$ and $C_3 = \{(i, j) : i + j = 9\}$
- Then,

$$P(C_1 \cap C_2) = \frac{1}{6} \neq \frac{1}{4} = P(C_1)P(C_2)$$

$$P(C_1 \cap C_3) = \frac{1}{36} \neq \frac{1}{18} = P(C_1)P(C_3)$$

and

$$P(C_2 \cap C_3) = \frac{1}{12} \neq \frac{1}{18} = P(C_2)P(C_3)$$

- However,

$$P(C_1 \cap C_2 \cap C_3) = \frac{1}{36} = P(C_1)P(C_2)P(C_3)$$

Independent experiments

- Events as outcome of **independent experiments**
- A computer system is built with K_1, K_2 and K_3
- Each successive component is a backup for the previous one
- If probabilities of failure are $P(K_1) = 0.01$, $P(K_2) = 0.03$, and $P(K_3) = 0.08$, the system fails with probability

$$(0.01)(0.03)(0.08) = 24 \times 10^{-6}$$

- The system functions successfully with the probability

$$1 - 24 \times 10^{-6} = 0.999976$$