Lecture 2: Probability

MATH 697

Omidali Jazi

September 6, 2018

McGill University

Conditional Probability and Independence

- The idea of conditional probability is that the probability of an event depends on our knowledge of other events that have occurred.

- Example:
  Imagine I am holding a card I picked at random out of a standard deck of playing. What is the probability it is the ace of hearts?

- It is $1/52$ or just under 0.02.
  Now, if I tell you the card I am holding is an ace, what is the chance I'm holding the ace of hearts?

- It's $1/4$ (or 25%) since we know it has to be one of only 4 aces in the deck.

Conditional Probability

Definition 1

The conditional probability of an event A, given that an event B has occurred, is equal to

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

provided that $P(B) > 0$.

Back to the Ace of Hearts Example

- Define the events: $A =$ choose the ace of hearts and B $=$ choose an ace (randomly out of 52 cards)

- Calculate the probabilities: Let's start with $P(B)$

- $P(B) = 4/52$ (there are 4 aces in 52 cards)
  Now, $P(A \cap B)$ means "the probability of choosing the ace of hearts and choosing an ace"

- In this case, that is the same as "the probability of choosing the ace of hearts"

- So, $P(A \cap B) = P(A) = 1/52$. Then
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/52}{4/52} = 1/4$$

Independence

- The idea of independent events is that the probability of occuring event A is unaffected by whether B occurs or not.
  Another way to think about it is that knowledge about event B does not affect your probabilistic assessment of A.
  This is the intuitive notion that A and B are independent.
  But, if knowledge of B tells you something about the probability of A, then the two events are dependent.

Independent Events

Definition 2

Two events A and B are said to be independent if any one of the following holds:

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

Otherwise, the events are dependent.
Note that each condition follows from either of the other two conditions and the definition of conditional probability.

Exmaple

Consider the following events in the toss of a single balanced six-sided die:

- A=Observe an odd number
- B= Observe an even number
- C= Observe a 1 or a 2.
- Question 1: Are A and B independent events?
  Solution: No, because

$$P(A \cap B) = 0 \neq P(A)P(B) = 9/36$$

  Conclusion: Two disjoint events with positive probabilities are always dependent.

Example

- Question 2: Are A and C independent events?
  Solution: Yes, because

$$P(A \cap C) = 1/6 = P(A)P(C) = (3/6)(2/6) = 1/6$$

  Conclusion: Disjoint events and independent events are different notions and should not be mixed up. In this case, two events share a sample point, but they are independent.

Two Laws of Probability

Theorem 3

(Multiplicative Law of Probability): The probability of the intersection of two events A and B is

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

If A and B are independent, then

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

Extending the Multiplicative Law

By twice applying the Theorem

$$\begin{aligned} P(A \cap B \cap C) &= P[(A \cap B) \cap C] \\ &= P(A \cap B) P(C|A \cap B) \\ &= P(A) P(B|A) P(C|A \cap B) \end{aligned}$$

And, this can be extended to any number, say k, events:

$$\begin{aligned} &P(A_1 \cap A_2 \cap .... \cap A_k) = \\ &\qquad P(A_1) P(A_2|A_1) ... P(A_k|A_1 \cap A_2 \cap ... \cap A_{k-1}) \end{aligned}$$

Application Example

- The Multiplicative Law is very useful for calculating complicated probabilities when conditional probabilities are easier to determine.
- Example: What is the probability of choosing a female local undrgraduate stuent when the university consists of 60% female and 40% male students. If a student is female, 70% and 30% are the chances of being local and international respectively. Further, half of of the female local students are undergraduate, while 30% and 20% are master and PhD respectively.
  Solution: $0.6 \times 0.7 \times 0.5 = 0.21$

The Addition and Complement Rules

Theorem 4

(The Additive Law of Probability): The probability of the union of two events A *and* B *is*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A *and* B *are mutually exclusive, then*

$$P(A \cup B) = P(A) + P(B)$$

Theorem 5

For event A*,*

$$P(A) = 1 - P(\overline{A})$$

Example Application: Gene Expression Profiling

Example 6

Gene expression profiling is a state-of-the-art method for determining the biology of cells. In Briefings in Functional Genomics and Proteomics (Dec. 2006), biologists at Pacific Northwest National Laboratory reviewed several gene expression profiling methods. The biologists applied two of the methods (A and B) to data collected on proteins in human mammary cells. The probability that the protein is cross-referenced (i.e., identified) by method A is .41, the probability that the protein is cross-referenced by method B is .42, and the probability that the protein is cross-referenced by both methods is .40.

Example Application

- Find the probability that the protein is cross-referenced by either method A or method B.
  Solution: $0.41 + 0.42 - 0.40 = 0.43$
- On the basis of your answer to part a, find the probability that the protein is not cross-referenced by either method.
  Solution: $1 - 0.43 = 0.57$

Example Application

Example 7

It is known that a patient with a disease will respond to treatment with probability 0.9. If three patients with the disease are treated and respond independently, find the probability that at least one will respond.

Solution: $1 - P(No\ Response) = 1 - 0.1^3 = 0.999$

Extension: Iclusion-Exlusion Formula

The Additive Law of Probability can be extended to k events, $A_1, A_2, ..., A_k$ as

$$P\left(\bigcup_{i=1}^{k} A_i\right)$$
$$= \sum_{i=1}^{k} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + ... + (-1)^{k-1} P(A_1 \cap ... \cap A_k)$$

Calculating Probabilities Via the Event-Composition Method

- The event-composition method uses the various laws of probability that we have just learned to calculate the probability of some event A.
- The idea is to express A as the union, intersection, or compliment of other sets.
- Then, calculate the probabilities of those events and appropriately combine the probabilities to obtain $P(A)$.

Law of Total Probability and Bayes' Rule

It is often useful to view a sample space as the union of mutually exclusive subsets.

Definition 8

For some positive integer k, let the sets $\{B_1, B_2, ..., B_k\}$ be such that

1. $S = B_1 \cup B_2 \cup ... \cup B_k$
2. $B_i \cap B_j = \phi$ for $i \neq j$.

Then the collection of sets $\{B_1, B_2, ..., B_k\}$ is said to be a partition of S.

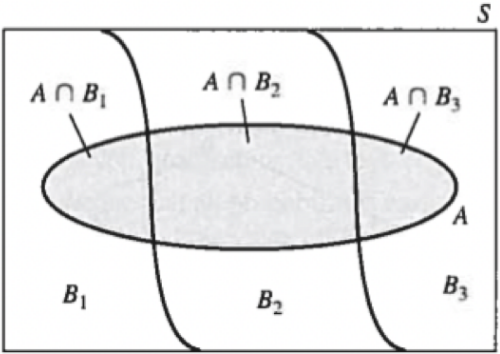
Using the Partition to Decompose a Set

Now, if A is any subset of S and $\{B_1, B_2, ..., B_k\}$ is a partition of S, then A can be decomposed as follows:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup ... \cup (A \cap B_k)$$

An illustration for $k = 3$

FIGURE **2.12** Decomposition of event A



Law of Total Probability

Theorem 9

Assume that that $\{B_1, B_2, ..., B_k\}$ *is a partition of* S *such that* $P(B_i) > 0$ *for* $i = 1, ..., k$. *Then for any event* A,

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

Example Application

Example 10

Of the voters in a city, 40% are Republican and 60% are Democrats. Among the Republicans, 70% are in favor of a bond issue, while 80% of the Democrats favor this issue. If a voter is selected at random, what is the probability that he/she will favor the bond issue?

Solution: $P(F) = 0.7 \times 0.4 + 0.8 \times 0.6 = 0.76$

Bayes' Rule

Theorem 11

Assume that $\{B_1, B_2, ..., B_k\}$ *is a partition of* S *such that* $P(B_i) > 0$ *for* $i = 1, ..., k$. *Then,*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

Applications of the Bayes' Rule

- Invert conditional probabilities
  You know the various conditional probabilities $P(A|B_1)$, $P(A|B_2)$,..., $P(A|B_k)$ but you want to know $P(B_j|A)$ for some B_j

- Update probabilities $P(B_j|A)$
  You know or obtain $P(B_1)$, $P(B_2)$,...,$P(B_k)$ then subsequently observe/collect some information $P(A|B_1)$, $P(A|B_2)$,..., $P(A|B_k)$.
  Then update your probabilities conditioned on what you've observed by calculating $P(B_1|A)$, $P(B_2|A)$, ...,$P(B_k|A)$

Application: A Drug Testing Example

Example 12

Based on an investigation, it is believed that 1% of the sailors use illegal drugs. A test detects drug use as follows:

- If someone has used illegal drugs, there is a 99% chance it will detect it (true positive).
- If someone does not use illegal drugs, there is a 0.5% chance the test will indicate it (false positive).

What is the chance that a positive drug test signals a drug user?

Solution

- U_+ is the event an individual used illegal drugs, U_- is the event he/she did not
- T_+ is the event the drug test is positive for the individual
- We know $P(T_+|U_+) = 0.99$ and we want to calculate $P(U_+|T_+)$.

$$\begin{aligned} P(U_+|T_+) &= \frac{P(T_+|U_+)P(U_+)}{P(T_+|U_+)P(U_+) + P(T_+|U_-)P(U_-)} \\ &= \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.005 \times (1 - 0.01)} \\ &= \frac{2}{3} \end{aligned}$$

Interpretation

- Say, there are 10,000 sailors: 100 are drug users (1%) and 9,900 are not.
  99 will be detected (99% of 100).
  49.5 will be false positives (0.5% of 9,900)
- Total of 148.5 positives. Of those, 99/148.5 (2/3) are drug users while 49.5/148.5 (1/3) are clean.
- Conclusions:
  Don't use drug tests with high false positives.
  No point in testing clean population (all you get are false positives).

Example Application (Maize seeds)

Example 13

The genetic origin and properties of maize (modern-day corn) were investigated in Economic Botany. Seeds from maize ears carry either single spikelets or paired spikelets, but not both. Progeny tests on approximately 600 maize ears revealed the following information: Forty percent of all seeds carry single spikelets, while 60% carry paired spikelets. A seed with single spikelets will produce maize ears with single spikelets 29% of the time and paired spikelets 71% of the time. A seed with paired spikelets will produce maize ears with single spikelets 26% of the time and paired spikelets 74% of the time.

Example Application (Maize seeds)

- Find the probability that a randomly selected maize ear seed carries a single spikelet and produces ears with single spikelets.
  Solution: Exercise
- Find the probability that a randomly selected maize ear seed produces ears with paired spikelets.
  Solution: Exercise

Random Sampling

One application of probability is in selecting random samples from a population

- Important when executing various types of experiments as well as in collecting survey data.
- Random selection of a sample from a population is critical for being able to do valid inference. Lots of ways to draw "random samples". Two broad categories include:
    - Sampling without replacement: Population elements cannot be sampled more than once.
    - Sampling with replacement: Population elements can be sampled multiple times.

(Simple) Random Sampling

Definition 14

Let N be the number of elements in the population and n be the number in the sample. A (simple) random sample is a sample such that each of the $\binom{N}{n}$ possible samples has an equal probability of being selected.

This is just one possible type of sample. Others include stratified sampling, cluster sampling, and complex sampling that combines two or more of these methods.

What We Have Just Learned

- Introduction to Probability
- Methods for Calculating Probabilities (Sample-point Method, Event-composition Method)
- Counting Tools (Combinations, Permutations, etc)
- Conditional Probability and Independence
- Probability Laws and Rules (Multiplicative and Additive Laws, Law of Total Probability and Bayes' Rule)
- Random sampling