STA 111: Probability & Statistical Inference

Lecture Two – Counting Methods and Conditional Probability

Instructor: Olanrewaju Michael Akande

Department of Statistical Science, Duke University

Outline

- Questions?
- Introduction
- Counting Methods
- Conditional Probability
- Independent Events
- Bayes' Theorem
- Recap

Introduction

– In the last lecture we discussed the different interpretations of probability.

– We also talked about the mathematical definition of probability and how to calculate probability of events in simple experiments.

– Today we will look at how to calculate the probability of one event given what we know about another event that has already occurred.

– This would lead us to the concept of “independent events”.

– Lastly, we will look at one of the core theorems of probability, Bayes' Theorem.

Counting Methods

– First we need to review important methods for counting the number of outcomes in an event/set.

– This is particularly useful when the process of enumerating all possible outcomes is simply inefficient.

– Multiplication Rule (D.S. Theorem 1.7.2): Suppose an experiment has k parts $(k \geq 2)$, such that the ith part of the experiment can have n_i possible outcomes $(i = 1, \dots, k)$. If all of the outcomes in each part can occur regardless of which specific outcomes have occurred in the other parts, then the total number of outcomes will be equal to the product: $n_1 n_2 \dots n_k$.

Examples

– *Example 1:* Back to our toy example of tossing a coin twice. Since the outcome of each toss doesn't affect the other, the total number of outcomes, that is $|\Omega|$, is $2 \times 2 = 4$

– *Example 2:* Suppose you roll a die twice. Again, since each roll is "independent" of the other (we will look at the mathematical definition of independence soon), $|\Omega| = 6 \times 6 = 36$

– *Example 3:* Suppose you are at an ice cream shop with three different cup/cone options, 12 different options for flavor, and five different options for toppings/mix-ins. Then you have $3 \times 12 \times 5 = 90$ cup-flavor-topping options in all; as long as your choices for each one of cup, flavor or mix-in doesn't affect the choices for the others!

Counting Methods (Cont'd)

– Permutation: The number of ways to arrange n distinct objects in a line is

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

This is the number of permutations of n distinct objects. There are n choices for the the first position, then $n-1$ for the second and so forth; multiplication gives the number of distinct elements.

– Then the number of permutations of n elements taken k at a time is

$$P_{n,k} = n \times (n-1) \times (n-2) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$$

– By convention, $0! = 1$. Thus, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, ...

– $P_{n,k}$ is also the number of distinct "orderings" of k items selected without replacement from a collection of n different items.

Examples

– *Example 3:* In how many ways can 8 people line up? $8! = 40320$.

– *Example 4:* In how many ways can 4 married couples stand in a police line-up , if couples must stand together?

There are $4! = 24$ ways that the couples can be arranged, and each couple can be arranged in $2! = 2$ ways. Thus the answer is

$$4! \times 2! \times 2! \times 2! \times 2! = 384$$

– *Example 5 (D.S. Example 1.7.8):* In how many ways can you select a president and a secretary from a club consisting of 25 members?

Counting Methods (Cont'd)

– Notice that ordering is important in permutations. What if we wish to select k items from a collection of n different items without regard for specific orderings? In other words, we care about selecting subsets of size k.

– Combination: The number of distinct subsets of size k that can be chosen from a set of size n is:

$$C_{n,k} = \binom{n}{k} = \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!}$$

– That is, once we have the number of possible permutations, we need to get rid of repeated subsets. Since there are $k!$ ways to arrange each of the k repeated elements, we divide by $k!$.

– $\binom{n}{k}$ is also called the binomial coefficient but we will get to that later in the course.

Examples

– *Example 6:* If you have 20 different courses to pick from in a school year, in how many ways can you choose 7 courses?

$$\binom{20}{7} = \frac{20!}{7!13!} = 77520$$

Aren't you glad there is a more "systematic" way to select courses?

– *Example 7 (D.S. Example 1.8.2):* The number of different groups of people that might be on the committee composed of eight people from a group of 20 people is?

Conditional Probability

– In the last class, we saw how to calculate the probability of events in a simple experiment. We also saw how to calculate the probability that two events occur together. That is, for two events A and B, P(A and B) = P(A ∩ B).

– What if we are instead interested in the probability that A happens given that B has already happened? This is called conditional probability. A quick example to illustrate this:

Example 8: Back to our toy example of tossing a fair coin twice. Let A be the event that only one head is observed, and B the event that at least one head is observed. Then A = $\{HT, TH\}$ and B = $\{HH, HT, TH\}$. What is the probability that A occurs given that we know B already occurred?

First notice that if B already happened, $P(\{TT\}) = 0$, so that B is our new sample space. If B is the new sample space, and all outcomes are equally likely, then the probability we seek to calculate is simply $\frac{2}{3}$.

Conditional Probability

- Formally, the probability of A given B is:

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

which is not defined if $P(B) = 0$. That is, it doesn't make sense to condition on an event that cannot happen. Also, $P(A|B) + P(A^c|B) = 1$.

Example 8 again: Using this formula instead of what we did before,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/4}{3/4} = \frac{2}{3} \quad \text{same as before!}$$

- A useful implication of conditional probability is:

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

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

Examples

Example 9: Suppose you have a standard deck of 52 cards which includes 13 cards of each of the four suits: clubs, diamonds, hearts and spades. Clubs and spades are black while diamonds and hearts are red. Let A be the event that you draw a diamond card and B the event that you draw a red card. Then,

- $P(A) = \frac{13}{52} = \frac{1}{4}$
- $P(B) = \frac{26}{52} = \frac{1}{2}$
- $P(A \text{ and } B) = P(A \cap B) = \frac{13}{52} = \frac{1}{4}$
- $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{4} + \frac{1}{2} - \frac{1}{4} = \frac{1}{2}$
- $P(A|B) = \frac{1/4}{1/2} = \frac{1}{2}$
- $P(B|A) = \frac{1/4}{1/4} = 1$. So if A already happened, B has to happen.

Independent Events

– Two events A and B are said to be independent if and only if

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

Notice already that if A and B are independent,

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

so that A and B are independent if the occurrence of one doesn't affect the other in any way.

– Can you see why this is true for our toy example of tossing a fair coin twice? Define any event based on the first toss and another event based on the second toss. Then use conditional probability to show independence between them.

Examples

– *Example 9 again:* Are A and B independent?

$$P(A \cap B) = \frac{1}{4} \neq \frac{1}{4} \times \frac{1}{2} = P(A)P(B)$$

Of course not! If I already picked a red card, I know the card has to be a diamond or a heart, which changes my uncertainty about picking a diamond card.

– *Example 10 (D.S. Example 2.2.3) :* Suppose that a balanced/fair die is rolled. Let A be the event that an even number is obtained, and let B be the event that one of the numbers 1, 2, 3, or 4 is obtained. Are they independent?

– These examples help present the idea of independence as information gain. If I know an event A has happened, do I gain any information from A about another event B beyond what I already know about B? If no, they are independent.

Bayes' Rule/Theorem

– Partition: The events $A_1, \ldots, A_k$ form a finite partition of the sample space Ω if they are all disjoint and $\bigcup_{j=1}^{k} A_j = \Omega$. Note that $\sum_{j=1}^{k} P(A_j) = 1$ and $\sum_{j=1}^{k} P(A_j|B) = 1$ for any other event B.

– Law of Total Probability: Suppose the events $A_1, \ldots, A_k$ form a finite partition of Ω and $P(A_i) > 0$ for all $i = 1, \ldots, k$. Then for every event B in Ω,

$$P(B) = \sum_{j=1}^{k} P(A_j)P(B|A_j)$$

Draw a Venn Diagram!

– Bayes' Rule/Theorem: Suppose the events $A_1, \ldots, A_k$ form a finite partition of Ω and $P(A_i) > 0$ for all $i = 1, \ldots, k$. Then

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{k} P(A_j)P(B|A_j)}$$

Examples

Example 11: Suppose Duke has only three possible majors: economics, statistics, and biology. Also, suppose 25% of the students are in economics, 45% are in stats, and the rest are in bio (with no double majors).

Campus folklore says that 90% of econ majors want to go on dates, compared with 10% of stats majors and 50% of bio majors.

Your room-mate meets someone at Duke gardens and gets a date. What is the probability that he/she is dating an economist?

Note that the majors define a finite partition, and the campus folklore gives us the conditional probabilities $P(B|A_i)$

The point of Bayes' rule is to reverse the conditioning to get $P(A_i|B)$.

Examples

Let A_1 = being an econ major, A_2 = being a biology major, A_3 = being a stats major, and B = going on a date. Then we want to find $P(A_1|B)$. From the question, we have:

$$P(A_1) = 0.25, \; P(A_2) = 0.45, \; P(A_3) = 0.30.$$

$$P(B|A_1) = 0.90, \; P(B|A_2) = 0.10, \; P(B|A_3) = 0.50.$$

Using Bayes' rule,

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{\sum_{j=1}^{3} P(A_j)P(B|A_j)}$$

$$= \frac{0.25 \times 0.90}{(0.25 \times 0.90) + (0.45 \times 0.10) + (0.30 \times 0.50)} = \frac{0.225}{0.42} = 0.5357$$

In-class Exercise

To be done in-class with one teammate: ELISA is a test for HIV. Like all technology, it is not completely reliable, but the Food And Drug Administration (FDA) has collected extensive information on its error rates.

- If a person has HIV, ELISA has probability 0.997 of signaling.
- If a person does not have HIV, then ELISA does not signal with probability 0.985.
- About 0.32% of the U.S. population has HIV.

Suppose someone walks into a clinic to get an HIV test at random and the test comes back positive. What is the chance that the person has HIV?

Recap

Today, we talked about

- Some useful counting methods
- Conditional probability
- The concept of independent events
- How to apply Bayes' theorem