Quantitative Tools - Level 2 - Fall 2017
Louis de Charsonville

Lecture 5

Probability (1/2)

Introduction

Numerous natural events occurs at a "regular frequency". For instance, whatever the country of time period, the sex ratio, the ratio of males to females in a population, tends to be 1. Heights in a population tends to the follow the same pattern with a lot of people centered around the mean and a few people to the ends. Hence, the idea that real life events might be approximated by statistical laws. These laws touches on the notion of probability. Probability is a large and complex field of study. This course covers the basic issues of probability.

1 Discrete Probability

1.1 Definitions

If you roll a six-sided die, there are six possible outcomes, and each of these outcomes is equally likely. A six is as likely to come up as a three, and likewise for the other four sides of the die. What, then, is the probability that a one will come up? Since there are six possible outcomes, the probability is 1/6.

- Rolling the dice is an **experiment**, denoted as $\mathcal{E}$.
- The result of this experiment is called an **event**, denoted E.
- The set of all possible results is called the **sample space** of the experiment and denoted $\mathcal{F}$.
- The **power set** of the sample space is formed by considering all different collections of possible results and denoted Ω. For instance, one possible collection of the experiment of rolling a dice is 2,4,6, which is the event that the dice falls on an even number.
- A singleton ω with $\omega \in \Omega$ is called an **elementary event**.
- A **probability** is a way of assigning every event to a value between 0 and 1. We denote the probability of E as $P(E)$ with the following properties:
  1. $0 \leq P(E) \leq 1$
  2. $\emptyset$ is the impossible event (empty set) with $P(\emptyset) = 0$.
  3. Ω is the certain event with $P(\Omega) = 1$
  4. Two events, A and B, are said to be **mutually exclusive** if they can't happen simultaneously. A dice can't fall on 1 and on 2. The sum of the probabilities of all mutually exclusive events equals 1.
  5. Two events are said to be **complementary** if they are mutually exclusive and their union is the sample space. We denote $\bar{A}$ the complementary event of A. *Example*: if A is the event "the dice rolls on 1", $\bar{A}$ is "the dice rolls on 2,3,4,5 or 6".

1.2 Properties

Definition 1.1. Finite sample space
The sample space is finite when Ω is finite, i.e. the power set of events the sample space is finite.

Properties 1.1. For any events, A and B:

1. $P(\bar{A}) = 1 - P(A)$
2. $P(B \setminus A) = P(B) - P(A \cap B)$
3. if $A \subset B$ then $P(A) \leq P(B)$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. if A and B are mutually exclusive then $P(A \cup B) = P(A) + P(B)$. More generally, if $A_1, ..., A_n$ are mutually exclusive then $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

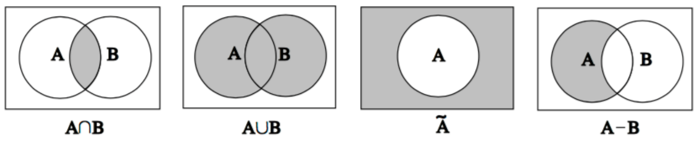

Figure 1: Union, intersection of sets

Properties 1.2. In the case where each elementary event has the same probability, we have for any elementary event A:

$$P(A) = \frac{Number\ of\ favorables\ outcomes\ to\ A}{Number\ of\ possible\ outcomes}$$
$$P(A) = \frac{Card(A)}{Card(\Omega)}$$

2 Combinatorics

Many problems in probability theory require that we count the number of ways that a particular event can occur. For this, we study *permutations* and *combinations*.

2.1 Tree diagrams

It will often be useful to use a tree diagram when studying probabilities of events relating to experiments that take place in stages and for which we are given the probabilities for the outcomes at each stage. A task is to be carried out in a sequence of s stages. There are n_1 ways to carry out the first stage; for each of these n_1 ways, there are n_2 ways to carry out the second stage, and so forth. Then the total number of ways in which the entire task can be accomplished is given by the product $N = n_1 * n_2 * ... * n_r$.

Example 2.1. For instance, assume that in a restaurant, 30% of the customers choose to have a starter, among those who choose a starter, 40% choose the soup and 60% choose the salad. Among those who choose the soup 50% choose the pumpkin soup, 20% choose the cauliflower soup, 30% choose the tomato soup. How many customers choose the pumpkin soup ?

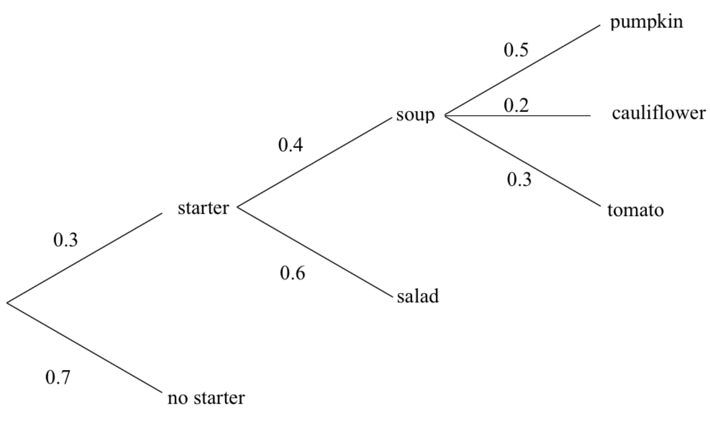

Figure 2: Three-stages probability tree diagram

2.2 Permutations

Definition 2.1. Let E be any finite set. A **permutation** of E is an *ordered arrangement* of the elements of E. We also say that a *permutation* is a one-to-one mapping of E into itself, or a *bijection* from E to itself.

Example 2.2. Let E the set of possible outcomes of rolling dices $\{1, 2, 3, 4, 5, 6\}$. A permutation σ of E can be :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \end{pmatrix}$$

Theorem 2.1. The total number of permutations of a set E of n elements is given by $n*(n-1)*(n-2)*...*1$.

Example 2.3. Let E be a set of 3 elements $\{a, b, c\}$. There are $3 * 2 = 6$ permutations of E:

$$abc\, acb\, bac\, bca\, cab\, cba.$$

Definition 2.2. Let E be an $n-element$ set, and let k be an integer between 0 and n. Then a **k-permutation** of E is an ordered listing of a subset of E of size k.

Theorem 2.2. The total number of *k-permutations* of a set E of n elements is given by $n * (n-1) * (n-2) * ... * (n-k+1)$.

Example 2.4. Let E be a set of 4 elements $\{a, b, c, d\}$. There are $4 * 3 = 12$ *2-permutations* of E:

$$ab\, ac\, ad\, ba\, bc\, bd\, ca\, cb\, cd\, da\, db\, dc$$

Definition 2.3. We called **factorial** of n, denoted $n!$, the number:

$$n! = n * (n-1) * (n-2) * ... * 1$$

The expression 0! is defined to be 1.

2.3 Combinations

We now consider **combinations**. Let E be a set with n elements; we want to count the number of distinct subsets of E that have exactly p elements. The empty set, $\emptyset$, and the set E are considered to be subsets of E.

Example 2.5. Let E be a set of 3 elements $\{a, b, c\}$. The subsets of E are:

$$\emptyset\, , \{a\}\, , \{b\}\, , \{c\}\, , \{a, b\}\, , \{b, c\}\, , \{a, b, c\}$$

Definition 2.4. The number of distinct subsets with p elements that can be chosen from a set that contains n elements is called a **binomial coefficient** and is denoted $\binom{n}{p}$ and is pronounced "n choose p".

Properties 2.1. We can derive from the definition the following properties:

- $\binom{n}{p} = \binom{n}{n-p}$
- $\binom{n}{0} = \binom{n}{n} = 1$
- $\binom{n}{1} = \binom{n}{n-1} = n$

Theorem 2.3. Pascal's triangle
For any n, p we have:

$$\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$$

Theorem 2.4. Binomial theorem
For any $\{x, y\} \in \mathbb{R}^2$, $n \in \mathbb{N}$ we have:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

3 Conditional Probability and Independence

Definition 3.1. Let A be an event with a non-zero probability. For any event B, we have:

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

We called $P(B|A)$ the conditional probability of B given A.

Example 3.1. We roll a dice once. Let X be the outcome, A be the event $\{X \leq 2\}$ and B the event $\{X = 1\}$. We have:

$$\begin{aligned} P(A \cap B) = P(B) &= \frac{1}{6} \\ P(A) &= \frac{1}{3} \\ P(B|A) &= \frac{1}{2} \end{aligned}$$

Theorem 3.1. Bayes theorem
Let A, B be two events with a non-zero probability. We have:

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

Example 3.2. Probability of getting AIDS
Pierre is a French young man who wants to know if he is HIV-infected. He does a HIV-test which happens to be positive. What is the probability that Pierre is HIV-infected ?
Let A be the probability of having AIDS, and B the probability that the test is positive. In France, the number of HIV-infected people was estimated at 152,000. The current population is 66 millions, so the proportion of HIV-infected people is 0.02%. A HIV test returns a positive or a negative result. There are four cases:

- True positive: the test is positive and the patient is HIV-infected.
- False positive: the test is positive but the patient is not HIV-infected.
- True negative: the test is negative and the patient is not HIV-infected
- False negative: the test is negative but the patient is HIV-infected

The sensitivity of a test is the probability of obtaining a positive result given that the patient is HIV-infected. The specifity of a test is the propability of obtaining a negative result given that the patient is not HIV-infected. A good test has both a high sensitivity and high specificity. Lets say that both specificity and sensitivity are 99.9%. Let's recap. We know that:

$$\begin{aligned} P(A) &= 0.02\% \\ P(B|A) &= 99.9\% \\ P(\neg B|\neg A) &= 99.9\% \end{aligned}$$

We are looking for $P(A|B)$ which is the probability that Pierre is HIV-infected given that the test is positive. We use the Bayes theorem:

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

We need to compute $P(B)$, the probability that the test is *positive*.

$$\begin{aligned} &P(B) = P(B|A) * P(A) + P(B|\neg A)P(\neg A) \\ &P(B) = P(B|A) * P(A) + (1 - P(\neg B|\neg A) * (1 - P(\neg A)) \\ &P(B) = 0.999 * 0.02 + (1 - 0.999) * (1 - 0.02) \\ &P(B) = 0.12\% \end{aligned}$$

We then use the Bayes theorem:

$$\begin{aligned} P(B|A) &= \frac{P(B)}{P(A)} P(A|B) \\ &= 0.999 * \frac{0.02}{0.12} \\ &= 16.7\% \end{aligned}$$

The probability that Pierre is HIV-infected given that the test is positive is 16.7%.

Definition 3.2. Independent events It often happens that the knowledge that a certain event A has occurred has no effect on the probability that some other event B has occurred, that is, that $P(B|A) = P(B)$, which implies that $P(A|B) = P(A)$. If theses equations are true, we say that B is independent of A.

Properties 3.1. Let A, B two events. If A and B are independent, then we have:

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

Definition 3.3. A set of events $\{A_1, A_2, ..., An\}$ is said to be **mutually independent** if for any subset $\{A_i, A_j, ..., Am\}$ of these events we have:

$$P(A_i \cap A_j \cap ... \cap A_m) = P(A_i) * P(A_j) * ... * P(A_m)$$

* * *

Bibliography

These notes draw heavily on the following books that are worth reading by the interested reader:

- Introduction to Statistics, David M. Lane, 2013.
- De l'analyse à la prévision, Volume 3, Didier Schlacther, 2009.
- Analysis of Economic Data, Third Edition, Gary Koop, 2009.