Probability Theory for Econometricians

Sven Otto

December 27, 2023

Table of contents

Welcome

This tutorial gives a short introduction of the most important basic concepts from probability theory and statistics for econometricians. The sections are complemented with short R sessions.

To learn R or refresh your skills, please check out my tutorial **Getting Started With R**.

1 Probability theory

1.1 Random experiment

In econometrics, **observed data** is treated as an **outcome** of a **random experiment**. It is convenient to describe random outcomes numerically by a **random variable**. We distinguish between **discrete** and **continuous** random variables. Discrete random variables have only countably many potential outcomes (e.g., integers) whereas continuous random variables may produce any real number.

Table 1.1: Examples of random variables

Variable	Potential outcomes	Type
Coin flip	$C = 1$ (heads), $C = 0$ (tails)	discrete
Dice	$D \in \{1, ..., 12\}$	discrete
Years of schooling	$Y \in \{0, 1, 2, ...\}$	discrete
Stock returns	$R \in \mathbb{R}$	continuous
Wages	$W \geq 0$	continuous

1.2 Events

The coin flip random experiment has two potential outcomes, which are 0 and 1. The set S of all potential outcome values is called the **sample space**, i.e., $S = \{0, 1\}$. An **event** is a collection of outcomes and therefore a subset of the sample space. Any outcome is an event, the collection of all outcomes is an event, any union of multiple events is an event, and the complement of an event is an event. Thus, $\{0\}$ and $\{1\}$ are events, the union $\{0\} \cup \{1\} = \{0, 1\} = S$ is an event, and the complement $\{0, 1\}^c = \{\} = \varnothing$ (empty set) is an event. The collection of all events is called **sigma algebra**, which in our case is

$$\mathcal{B} = \{\{0\}, \{1\}, \{0, 1\}, \varnothing\}.$$

The sigma algebra of a discrete random variable with n different potential outcomes contains 2^n different events.

A continuous random variable has the sample space $S = \mathbb{R}$ (stock returns example) or a subset of the real line, e.g. $S = [0, \infty)$ as in the wage example. There are an uncountable infinite number of events, which are collected in the so called Borel sigma algebra. It contains all open, half-open, and closed intervals of outcomes, and their unions, intersections, and complements. Here are some examples of events: $[-1, 1]$, $(2, \infty)$, $(0, 2/3] \cup [3 \cup 4)$, $\{0.5\}$.

1.3 Probability distribution

The value of a random variable is determined by chance. It can be described by a probability distribution P, which assigns a probability to each event in $\mathcal{B}$. A fair coin has the following probability distribution:

$$P(C = 0) = \frac{1}{2}, \quad P(C = 1) = \frac{1}{2}, \quad P(C \in \{0, 1\}) = 1, \quad P(C \in \varnothing) = 0.$$

For the concept of a probability distribution to be coherent, Andrey Kolmogorov formulated the axioms of probability. Let X be a random variable with sample space S and sigma algebra $\mathcal{B}$. Then,

(i) $P(X \in A) \geq 0$ for all events $A \in \mathcal{B}$,
(ii) $P(X \in S) = 1$,
(iii) $P(X \in \bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(X \in A_i)$ for any disjoint sequence $A_1, A_2, \ldots \in \mathcal{B}$

The three axioms of probability imply the following rules of calculation:

1.4 Distribution function