Introduction: Week 1

Econ 361

Amherst College

There are known knowns.

There are things we know we know.

We also know there are known unknowns.

That is to say, we know there are some things we do not know.

But there are also unknown unknowns,

the ones we don’t know we don’t know.

Donald Rumsfeld, February 12th 2002

What do we mean by *uncertainty*?

- Are there events that are intrinsically uncertain?

- Does uncertainty simply reflect our ignorance, our inability to observe all the relevant “state variables” underlying the event?

- Is there a meaningful difference between these two views?

Thought Experiment …

- Consider a **black box** which we will dub an **experiment**
- The **experiment** takes inputs which we will dub **conditioning variables** to realize output which we will dub **outcomes**
  - A collection of these outcomes from past repeated applications of the experiment we will dub as **data** (later, **sample**)
- Suppose knowledge of the conditioning variables does **not** allow one to predict accurately the outcome of the experiment but it does inform one of the set of possible outcomes that might be realized
  - In this scenario, we will dub outcomes as being **uncertain**
- If even the set of possible outcomes cannot be known …

Relationship between Conditioning Variables and Outcomes

- What if there were some stable relationship between conditioning variables and realized outcomes for the experiment?
  - Such stable relationship we dub the **data generating process (DGP)**
- What if we knew this DGP?
  - Could use our knowledge of the DGP to arrive at more informed **prediction**
  - Likely still not perfectly accurate, but would be “more accurate”
- What if we only knew this DGP imperfectly/incompletely?
  - Could we use observations about past outcomes (data) to infer more about the imperfectly understood DGP?
  - **Estimation** (point and interval) and **Hypothesis Testing**

Statistical Inference

- Prediction

- Estimation
  - Point
  - Interval

- Hypothesis Testing

How to characterize the DGP?

- Let us consider, for now, only the scenario where knowledge of the conditioning variables enable us to know the set of possible outcomes (i.e. outcomes are uncertain)
- We want a mapping from each of these possible outcomes to some measure of how “likely” the outcome is to be realized given the conditioning variables
- Consider the following normalization for one such measure (Kolmogorov)
    - If an outcome is certain to be realized (only possible outcome given conditioning variables), assign a measure of one (1)
    - If an outcome is certain not to be realized (not within the set of possible outcomes given conditioning variables) assign a measure of zero (0)
    - Allow all other outcomes to be assigned a measure within (0, 1) with higher values indicating a more likely outcome

Probability

- A mapping from some space that defines the set of possible outcomes to [0, 1]

- Such a mapping would effectively characterize the DGP

- But how to arrive at such mapping?
  - What about outcomes that are related to each other?
  - What if there is an infinite number of possible outcomes? Worse, an uncountably large number of outcomes?

Classic Notion of Probability

- Define set of possible outcomes in terms of atomistic, mutually exclusive events (“simple” events)
  - For a die roll: {roll “1,” roll “2,” … roll “6”}
- Assume that the possible outcomes are equally likely
  - For a fair die roll, assign the measure of 1/6 for each of the possible atomistic, mutually exclusive events
  - Why 1/6?
- Allows us to consider the measure for composite events (some combination of the atomistic, mutually exclusive events)
  - For the composite outcome “roll an even number,” a measure of 1/6 + 1/6 + 1/6 = ½ is assigned … why?

Modern Frequentist Notion of Probability

- Do not need the space of possible outcomes to be defined by atomistic, mutually exclusive, equally likely events

- But, instead, requires the notion that the experiment may be repeated (under identical conditions) an arbitrary number of times

- The assigned measure for any outcome is, then, the limiting relative frequency of that experiment under infinite repetition

Random Variables and Probability Function

- Set theory is a pain. Calculus is much easier.
- Arrive at another mapping, this time between the space of possible outcomes and the real line: **random variables**
- So, first map from possible outcomes to real line via random variables and then from real line to [0, 1] via **probability function**

Events → Real line → [0, 1]

Distributions and Probability Mass/Density Functions

- Probability function may be expressed in many ways
- Primarily through what we dub **distribution –** more formally, the **cumulative distribution function (cdf)**
- Sometimes we will refer to a partial characterization of the distribution function known as **probability mass function (pmf)** or **probability density function (pdf)**
  - pmf or pdf depends on whether the set of possible outcomes is countable

Change of Variables and Distribution Theory

- What about transformations of random variables?

- Transformations of random variables are, generally, random variables as well

- Can we use our knowledge about the distribution of some random variables to derive the distribution of transformations of those random variables?

Moments and Expectations

- Sometimes, we will work with a partial characterization of the pmf/pdf, most notably **moments**

- Moments are usually expressed as **expected values** of the random variables underlying the experiment
  - Think of expected values as “weighted sums” (or weighted averages)

- Famous moments: mean and variance

Statistical Inference, Revisited

- Prediction

- Estimation
  - Point
  - Interval
- Hypothesis Testing