Probability Review I

Harvard Math Camp - Econometrics

Ashesh Rambachan

Summer 2018

Outline

Outline

Outline

The sample space and events

We wish to model a random experiment - an experiment/process whose outcome cannot be predicted beforehand. What are the building blocks?

- The **sample space** Ω is the set of all possible outcomes of a random experiment. We denote an outcome as $\omega \in \Omega$.
- An **event** A is a subset of the sample space, $A \subseteq \Omega$. Let $\mathcal{A}$ denote the family of all events.

Simple examples

Example: Suppose we survey 10 randomly selected people on their employment status and count how many are unemployed.

$$\Omega = \{0, 1, 2, \ldots, 10\}$$

A is the event that more than 30% of those surveyed are unemployed.

$$A = \{4, 5, 6, \ldots, 10\}$$

Example: Suppose we ask a random person what is their income.

$$\Omega = \mathbb{R}_+$$

A is the event that the person earns between \$30,000 and \$40,000.

$$A = [30,000, 40,000]$$

Outline

Putting structure on the set of events

To be able to sensibly define probabilities, we need to place some additional structure on the set of events, $\mathcal{A}$.

Let Ω be a set and $\mathcal{A} \subseteq 2^{\Omega}$ be a family of its subsets. $\mathcal{A}$ is a σ-**algebra** if and only if it satisfies the following

1. $\Omega \in \mathcal{A}$.
2. $\mathcal{A}$ is closed under complements: $A \in \mathcal{A}$ implies that $A^C = \Omega - A \in \mathcal{A}$.
3. $\mathcal{A}$ is closed under countable union: If $A_n \in \mathcal{A}$ for $n = 1, 2, \ldots$, then $\cup_n A_n \in \mathcal{A}$.

$\Longrightarrow$ We assume that $\mathcal{A}$ is a σ-**algebra**. $(\Omega, \mathcal{A})$ is a **measurable space** and $A \in \mathcal{A}$ is **measurable** with respect to $\mathcal{A}$.

Properties of a σ-algebra

If $\mathcal{A}$ is a σ-algebra, then ...

1. $\emptyset \in \mathcal{A}$.
2. $\mathcal{A}$ is closed under countable intersection i.e, if $A_n \in \mathcal{A}$ for $n = 1, 2, \ldots$, then $\cap_n A_n \in \mathcal{A}$.

Why?

1. This one's simple.
2. Hint: DeMorgan's Law - $(A \cup B)^C = A^C \cap B^C$.

What is probability?

We're now ready to finally define what is probability! We will provide the "mathematical" definition.

- Not defined directly as a "long-run frequency" or 'subjective beliefs." But it will capture all of the properties associated with these.

Let $(\Omega, \mathcal{A})$ be a measurable space. A **measure** is a function, $\mu : \mathcal{A} \rightarrow \mathbb{R}$ such that

1. $\mu(\emptyset) = 0$.
2. $\mu(A) \geq 0$ for all $A \in \mathcal{A}$.
3. If $A_n \in \mathcal{A}$ for $n = 1, 2, \ldots$ with $A_i \cap A_j = \emptyset$ for $i \neq j$, then

$$\mu(U_n A_n) = \sum_n \mu(A_n)$$

If $\mu(\Omega) = 1$, μ is a **probability measure**, denoted as $P : \mathcal{A} \rightarrow [0, 1]$.

Putting it all together

So, we model a random experiment as a **probability space**, $(\Omega, \mathcal{A}, P)$.

1. Ω - set of outcomes.
2. $\mathcal{A}$ - σ-algebra on the set of outcomes.
3. P - a probability measure defined on the σ-algebra.

Outline

Basic probability rules

We can prove all of the usual probability rules from this.

Consider a probability space $(\Omega, \mathcal{A}, P)$. The following hold:

1. For all $A \in \mathcal{A}$, $P(A^C) = 1 - P(A)$.
2. $P(\Omega) = 1$.
3. If $A_1, A_2 \in \mathcal{A}$ with $A_1 \subseteq A_2$, then $P(A_1) \leq P(A_2)$.
4. For all $A \in \mathcal{A}$, $0 \leq P(A) \leq P(1)$.
5. If $A_1, A_2 \in \mathcal{A}$, then

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Outline

Outline

Conditional Probability

Given a random experiment and the information that event B has occurred, what is the probability that the outcome also belongs to event A?

Let $A, B \in \mathcal{A}$ with $P(B) > 0$. The **conditional probability of A given B** is

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

- $P(A|B)$ is a probability measure so all the usual probability rules apply!
- We use conditioning to describe the partial information that an event B gives about another event A.

Implies that

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

Outline

Multiplication Rule

$$P(\cap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1) \ldots P(A_n| \cap_{i=1}^{n-1} A_i)$$

Proof?

The Law of Total Probability

Consider K disjoint events C_k that partition Ω. That is, $C_i \cap C_j = \emptyset$ for all $i \neq j$ and $\cup_{i=1}^{K} C_i = \Omega$. Let C be some event.

$$P(C) = \sum_{i=1}^{K} P(C|C_i)P(C_i)$$

Proof?

Bayes' Rule

Bayes' Rule:

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

- Proof?

Definitely the most important probability rule out there...

Outline

Independence

What if event B has no information about event A?

Two events A, B are **independent** if

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

Equivalently,

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

or

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

Independence

Let $E_1, \ldots, E_n$ be events. $E_1, \ldots, E_n$ are **jointly independent** if for any $i_1, \ldots, i_k$

$$P(E_{i_1} | E_{i_2} \cap \ldots \cap E_{i_k}) = E_{i_1}$$

Given an event C, events A, B are **conditionally independent** if

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