Econometrics – Probability

William M Volckmann II

January 3, 2017

This set of notes will be mostly a reference. It will not have many proofs or derivations, and it will lack exposition.

1 Introduction

1.1 Probability

Definition 1. A **random experiment** has three properties:

(a) All possible outcomes are known a priori;

(b) The outcome is unknown a priori in any trial;

(c) It can be repeated many times under identical conditions.

An example would be tossing a coin and recording the outcome on each toss. We know the possible outcomes – heads or tails – but we do not know which outcome will be realized on each toss. And we can toss the coin under (essentially) identical conditions many times.

Definition 2. A **probability space** is composed of the following three components:

(a) The **sample space** S of all outcomes of the experiment;

(b) The set of all events of interest $\mathcal{B}$ (formally a σ-algebra);

(c) A **probability function** P that assigns probabilities to events.

The elements of S are called **elementary events**. Events of interest may be combinations of elementary events. For instance, suppose the experiment is to toss a coin twice. Then

$$S = \{TT, TH, HT, HH.\}$$

Each of the four elements in S are elementary events. Now consider the event that at least one toss is a head. Then $S_{H\geq 1} = \{TH, HT, HH\}$.

Definition 3. A **sigma algebra** $\mathcal{B}$ is a collection of the subsets of S that satisfies

(a) *Closure under complements*: $x \in \mathcal{B}$ implies $x^c \in \mathcal{B}$.

(b) *Closure under countable unions*: $x_j \in \mathcal{B}$ for all $j \in \mathbb{N}$ implies $\cup_{j=1}^{\infty} A_j \in \mathcal{B}$.

(c) $\emptyset \in \mathcal{B}$.

DeMorgan's law implies closure under countable intersections as well. The smallest possible σ-algebra is $\mathcal{B} = \{\emptyset, S\}$. The largest possible σ-algebra is the powerset $\mathcal{P}(S)$. That said, it satisfices to use the smallest possible relevant σ-algebra. For example, if we want to restrict our attention to the number of heads in two coin tosses, then the events of interest are $\{TT\}, \{HH\}, \{TH, HT\}$. To construct the smallest relevant σ-algebra, we need to include the complements of each of these events as well as their unions, as well as the empty set and the whole set S, giving

$$\mathcal{B} = \big\{\emptyset, S, \{TT\}, \{TH, HT, HH\}, \{HH\}, \{TT, TH, HT\}, \{TH, HT\}, \{TT, HH\}.\big\}$$

Definition 4. A **probability function** P is a function defined on the σ-algebra $\mathcal{B}$ and sample spaces S that satisfies

(a) *The probability of an event is nonnegative*: $P(C) \geq 0$ for all $C \in \mathcal{B}$.

(b) *Probabilities sum to one*: $P(S) = 1$.

(c) *Countable additivity of disjoint events is additive in probability*:

$$P\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} P(C_j) \quad \text{where } C_m \cap C_n = \emptyset \text{ for all } m \neq n.$$

Definition 5. A **probability space** $(S, \mathcal{B}, P)$ is a triple constituting a sample space, σ-algebra, and probability function.

1.2 Probability Functions

Probability functions exhibit the following properties:

(a) $P(C) = 1 - P(C^c)$

(b) $P(\emptyset) = 0$

(c) $P(C_1) \leq P(C_2)$ if $C_1 \subseteq C_2$

(d) $0 \leq P(C) \leq 1$ for $C \in \mathcal{B}$

(e) $P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2)$

Furthermore, **Bonferroni's inequality** states that

$$P(C_1 \cup C_2) \geq P(C_1) + P(C_2) - 1.$$

This follows from property (e) because $P(C_1 \cap C_2) \leq 1$. **Boole's inequality** states that

$$P\left(\bigcup_{j=1}^{\infty} C_j\right) \leq \sum_{j=1}^{\infty} P(C_j).$$

Theorem 1. *Suppose C_n be a nondecreasing sequence of events, i.e. $C_n \subseteq C_{n+1}$ for all n. Furthermore suppose*

$$\lim_{n\to\infty} C_n = \bigcup_{n=1}^{\infty} C_n.$$

Then we can interchange the probability and the limit,

$$\lim_{n\to\infty} P(C_n) = P\left(\lim_{n\to\infty} C_n\right) = P\left(\bigcup_{n=1}^{\infty} C_n\right).$$

Theorem 2. *Suppose C_n be a nonincreasing sequence of events, i.e. $C_{n+1} \subseteq C_n$ for all n. Furthermore suppose*

$$\lim_{n\to\infty} C_n = \bigcap_{n=1}^{\infty} C_n.$$

Then we can interchange the probability and the limit,

$$\lim_{n\to\infty} P(C_n) = P\left(\lim_{n\to\infty} C_n\right) = P\left(\bigcap_{n=1}^{\infty} C_n\right).$$

For example, the sequence $\{1 + 1/n\}$ is monotonically decreasing, and the limit of the sequence 1. Thus, via Theorem 2, we can say that

$$\lim_{n\to\infty} P\left(X \geq 1 + \frac{1}{n}\right) = P(X \geq 1).$$

Suppose we want to restrict our attention to outcomes that are a subset B of the sample space S.

2 Permutations and Combinations

Permutations and combinations are the devil. Anyway, the **Fundamental Theorem of Counting** says that if experiment A has m outcomes and experiment B has n outcomes, then the combined experiment has $m \times n$ outcomes.

Now suppose we have n objects and we want to sample k of them. We can do so either with or without replacement; and the order of sampling might or might not matter. Thus, there are four possible cases.

(a) *Sampling with replacement when order matters.* There are n^k different ways of sampling k out of n objects. For instance, if you sample three times from $\{a, b\}$ without replacement, then there are $2^3 = 8$ different possible outcomes.

(b) *Sampling without replacement when order matters.* This is called a **permutation**. The first draw has n possibilities; the second draw has $n-1$ possibilities; the kth draw has $n-k+1$ possibilities. Thus there are

$$n \times (n-1) \times \ldots \times (n-k+1) = \frac{n!}{(n-k)!}$$

different possible outcomes.

(c) *Sampling without replacement when order does not matter.* This is called a **combination**. It is similar to a permutation except we have to divide out all "like" outcomes. For example, we do not want to include both $\{a, b\}$ and $\{b, a\}$ since order doesn't matter and thus these constitute the same outcome. Each subset of k elements has $k!$ different orderings, so we divide the permutation by $k!$. Thus, the total number of combinations we could draw of size k given n elements is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

(d) *Sampling with replacement when order does not matter.* This isn't used very often, but in any case the formula is

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}.$$