Advanced Probability and Statistics for Economists

Yoosoon Chang
Department of Economics
Rice University

and

Joon Y. Park
Division of Economics
Seoul National University
and
Department of Economics
Rice University

September 1997

Part I

Probability

1. Introduction to Probability

1. Probability Space

A *probability space* is the triple $(\Omega, \mathcal{F}, \mathbf{P})$, each of which will be introduced below.

First, Ω is a set called the *sample space*. It is the set consisting of all the possible outcomes of a random experiment. An element ω of Ω is called an *outcome*. Second, $\mathcal{F}$ is a collection of subsets of Ω, called a σ-field. More precisely,

Definition 1 (σ-field) *A class $\mathcal{F}$ of subsets of Ω is called a σ-*field *if it satisfies the following three properties:*

(a) $\Omega \in \mathcal{F}$.

(b) $E \in \mathcal{F}$ *implies* $E^c \in \mathcal{F}$.

(c) $E_1, E_2, \ldots \in \mathcal{F}$ *implies* $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$.

An element E of $\mathcal{F}$ is called an *event*. The σ-field $\mathcal{F}$ contains Ω itself and is closed under the formation of complements and countable unions. If a collection of subsets of Ω satisfying (a) and (b) is closed only under the formation of finite unions, then it is called a *field*. Notice that

Remarks It follows immediately that

(a) $\emptyset \in \mathcal{F}$ since $\emptyset = \Omega^c$.

(b) $E_1, E_2, \ldots \in \mathcal{F}$ implies $\bigcap_n E_n \in \mathcal{F}$, because $\bigcap_n E_n = (\bigcup_n E_n^c)^c$.

(c) A σ-field is a field; a field is a σ-field only when Ω is finite.

Third, $\mathbf{P}$ is a real-valued function defined on the σ-field $\mathcal{F}$. We define

Definition 2 (Probability) *A set function $\mathbf{P}$ on the σ-field $\mathcal{F}$ is a probability or probability measure if it satisfies the following conditions:*

(a) $\mathbf{P}(E) \geq 0$ *for all* $E \in \mathcal{F}$.

(b) $\mathbf{P}(\Omega) = 1$.

(c) *If* $E_1, E_2, \ldots \in \mathcal{F}$ *are disjoint, then*

$$\mathbf{P}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathbf{P}(E_n)$$

The three properties given above are often referred to as the *axioms of probability.* Modern probability theory views a probability simply as any function satisfying the axioms, following an *axiomatic approach.*

Properties Many 'already-well-known' properties of probability follow directly from the above axioms.

(a) For $E, F \in \mathcal{F}$ such that $E \subset F$, it follows from the axiom (c) that $\mathbf{P}(E) + \mathbf{P}(F - E) = \mathbf{P}(F)$, since $F = E \cup F = E \cup (F - E)$. We therefore have by the axiom (a) that

$$E \subset F \ \text{ implies } \ \mathbf{P}(E) \leq \mathbf{P}(F)$$

i.e., probability is *monotone*. Moreover, $\mathbf{P}(F - E) = \mathbf{P}(F) - \mathbf{P}(E)$, and we have as special cases

$$\mathbf{P}(\emptyset) = 0 \quad \text{and} \quad \mathbf{P}(E^c) = 1 - \mathbf{P}(E)$$

by the axiom (b).

(b) It can also be easily deduced from the axiom (c) that

$$\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)$$

since $E \cup F = E \cup (F \cap E^c)$ and $F = (F \cap E) \cup (F \cap E^c)$. Clearly,

$$\mathbf{P}(E \cup F) \leq \mathbf{P}(E) + \mathbf{P}(F)$$

since $\mathbf{P}(E \cap F) \geq 0$ by the axiom (a).

2. Limit Concepts in Probability

Let $\{E_n\}$ be a sequence of events. We say that $\{E_n\}$ is *monotone* when

$$E_1 \subset E_2 \subset \cdots \quad \text{or} \quad E_1 \supset E_2 \supset \cdots$$

For a monotone sequence $\{E_n\}$ of events, we define

$$\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n \quad \text{or} \quad \lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$$

depending upon whether the sequence is increasing or decreasing.

Theorem 1 *Let E_n be a monotone sequence of events. Then*

$$\mathbf{P}\left(\lim_{n\to\infty} E_n\right) = \lim_{n\to\infty} \mathbf{P}(E_n)$$

Proof Assume $\{E_n\}$ is monotonically increasing. Define

$$F_n = E_n - E_{n-1}$$

for $n = 1, 2, \ldots$, with the convention that $E_0 = \emptyset$. It follows that

$$\begin{array}{ccc} \mathbf{P}\left(\bigcup_{n=1}^{\infty} F_n\right) & = & \sum_{n=1}^{\infty} \mathbf{P}(F_n) \\ \| & & \| \\ \mathbf{P}\left(\lim_{n\to\infty} E_n\right) & & \lim_{n\to\infty} \mathbf{P}(E_n) \end{array}$$

as was to be shown. For a decreasing sequence E_n of events, let $F_n = E_1 - E_n$ and apply the above result. ■

For a sequence $\{E_n\}$ of events, we generally define

$$\begin{aligned} \limsup_{n\to\infty} E_n &= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \\ \liminf_{n\to\infty} E_n &= \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k \end{aligned}$$

Note that

$$\omega \in \limsup E_n \quad \text{iff} \quad \text{for any } n, \exists\, k \geq n \text{ such that } \omega \in E_k$$

$$\omega \in \liminf E_n \quad \text{iff} \quad \exists\, n \text{ such that } \omega \in E_k \text{ for all } k \geq n$$

For obvious reasons, we often write

$$\begin{aligned} \limsup E_n &= \{E_n \ \text{ i.o.}\} \\ \liminf E_n &= \{E_n \ \text{ ev}\} \end{aligned}$$

where "i.o." stands for "infinitely often" and "ev" is the abbreviation for "eventually". When $\limsup E_n = \liminf E_n$, we say that E_n has limit $E = \limsup E_n = \liminf E_n$.

Corollary 2 *Let $\{E_n\}$ be a sequence of events. We have*

$$\begin{aligned}\mathbf{P}(\liminf E_n) &\leq \liminf \mathbf{P}(E_n) \\ &\leq \limsup \mathbf{P}(E_n) \leq \mathbf{P}(\limsup E_n).\end{aligned}$$

Proof Since

$$\begin{aligned}\bigcap_{k=n}^{\infty} E_k &\uparrow \bigcup_n \bigcap_{k=n}^{\infty} E_k \\ \bigcup_{k=n}^{\infty} E_k &\downarrow \bigcap_n \bigcup_{k=n}^{\infty} E_k\end{aligned}$$

it follows that

$$\begin{aligned}\mathbf{P}(E_n) &\geq \mathbf{P}\left(\bigcap_{k=n}^{\infty} E_k\right) \to \mathbf{P}(\liminf E_n) \\ \mathbf{P}(E_n) &\leq \mathbf{P}\left(\bigcup_{k=n}^{\infty} E_k\right) \to \mathbf{P}(\limsup E_n)\end{aligned}$$

which completes the proof. ■

Theorem 3 (Borel-Cantelli) *Let $\{E_n\}$ be a sequence of events. We have*

$$\sum_{n=1}^{\infty} \mathbf{P}(E_n) < \infty \quad \textit{implies} \quad \mathbf{P}\left(\limsup_{n\to\infty} E_n\right) = 0$$

Proof Notice that

$$\begin{aligned}\mathbf{P}(\limsup E_n) &= \mathbf{P}\left(\bigcap_n \bigcup_{k\geq n} E_k\right) \\ &\leq \mathbf{P}\left(\bigcup_{k\geq n} E_k\right) \\ &\leq \sum_{k\geq n} \mathbf{P}(E_k) \to 0\end{aligned}$$

as was to be shown. ■

3. π- and λ-systems

It is in general hard to deal with a σ-field, since it has 'too many' elements. We may often consider simpler classes such as π- and λ-systems. The systems will be introduced below.

Definition 3 (π-system) *A class $\mathcal{P}$ of subsets of Ω is a π-system if $E, F \in \mathcal{P}$ implies $E \cap F \in \mathcal{P}$.*

It can easily be shown by induction that for events $E_1, \ldots, E_n \in \mathcal{P}$ we have $\bigcap_{k=1}^n E_k \in \mathcal{P}$. That is, π-system is closed under finite intersection.

Definition 4 (λ-system) *A class $\mathcal{L}$ of subsets of Ω is a λ-system if $\mathcal{L}$ satisfies the following three properties:*

(a) $\Omega \in \mathcal{L}$.

(b) *If $E, F \in \mathcal{L}$ and $E \subset F$, then $F - E \in \mathcal{L}$.*

(c) *If $E_1, E_2, \ldots \in \mathcal{L}$ and $E_n \uparrow E$, then $E \in \mathcal{L}$.*

It follows from (b), as a special case, that $E \in \mathcal{L}$ implies $E^c \in \mathcal{L}$ for a λ-system $\mathcal{L}$, i.e., the system is closed under the formation of complements. For union, however, (c) implies that it is closed only for monotone increasing sequences of events. This is in contrast with a σ-field, for which the formation of union is closed for arbitrary sequences of events.

Lemma 1 *A class $\mathcal{F}$ of subsets of Ω is a σ-field if and only if $\mathcal{F}$ is both a π-system and a λ-system.*

Proof The *only if* part is trivial. To prove the *if* part, let $E_1, E_2, \ldots \in \mathcal{F}$ which is assumed to be both a π-system and λ-system. Since $\mathcal{F}$ is a λ-system, it is closed under complementation. Moreover, it is a π-system and closed under finite intersection. Hence, it follows directly that $\bigcup_n E_n \in \mathcal{F}$, since

$$\left(\bigcap_{k=1}^n E_k^c\right)^c = \bigcup_{k=1}^n E_k \uparrow \bigcup_n E_n$$

The rest of the proof is obvious. ■

Let $\mathcal{S}$ be a class of subsets of Ω. The σ-field generated by $\mathcal{S}$, $\sigma(\mathcal{S})$, is defined to be the smallest σ-field containing $\mathcal{S}$. We may easily show that $\sigma(\mathcal{S})$ is the intersection of all the σ-fields containing $\mathcal{S}$, since an arbitrary intersection of σ-fields is a σ-field. We may similarly define the π- and λ-systems generated by $\mathcal{S}$, denoted by $\pi(\mathcal{S})$ and $\lambda(\mathcal{S})$ respectively, to be the smallest π- and λ-systems containing $\mathcal{S}$. Arbitrary intersections of π- and λ-systems are again π- and λ-systems, as for σ-fields. Therefore, $\pi(\mathcal{S})$ and $\lambda(\mathcal{S})$ may also be equivalently defined to be the intersections of all the π- and λ-systems containing $\mathcal{S}$. It is immediate that

$$\pi(\mathcal{S}) \subset \sigma(\mathcal{S}) \quad \text{and} \quad \lambda(\mathcal{S}) \subset \sigma(\mathcal{S})$$

since $\sigma(\mathcal{S})$ is a π- and λ-system containing $\mathcal{S}$.

Lemma 2 (Dynkin) *Let $\mathcal{P}$ be a π-system. Then*

$$\lambda(\mathcal{P}) = \sigma(\mathcal{P})$$

Proof It suffices to show that $\lambda(\mathcal{P})$ is a π-system. Define

$$\mathcal{G}_H = \{G \mid G \cap H \in \lambda(\mathcal{P})\}$$

Fix an arbitrary $E \in \mathcal{P}$. As we may easily show, $\mathcal{G}_E$ is a λ-system containing $\mathcal{P}$, and $\lambda(\mathcal{P}) \subset \mathcal{G}_E$. Therefore, for $E \in \mathcal{P}$, $F \in \lambda(\mathcal{P})$ implies that $E \cap F \in \lambda(\mathcal{P})$. Since this is true for any $E \in \mathcal{P}$, it follows that for any $F \in \lambda(\mathcal{P})$, $\mathcal{G}_F$ is a λ-system containing $\mathcal{P}$ and $\lambda(\mathcal{P}) \subset \mathcal{G}_F$. We thus have $E \in \lambda(\mathcal{P})$ implies $E \cap F \in \lambda(\mathcal{P})$, for any $F \in \lambda(\mathcal{P})$. This was to be shown. ■

One immediate consequence of the lemma is that if $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system such that $\mathcal{P} \subset \mathcal{L}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$.

Theorem 4 (Uniqueness of Extension) *Let $\mathcal{P}$ be a π-system on Ω, and $\mathbf{P}_1$ and $\mathbf{P}_2$ be probability measures on $\sigma(\mathcal{P})$. If $\mathbf{P}_1$ and $\mathbf{P}_2$ agree on $\mathcal{P}$, then they agree on $\sigma(\mathcal{P})$.*

Proof Due to the Dynkin's lemma, it suffices to show that

$$\mathcal{L} = \{E \,|\, \mathbf{P}_1(E) = \mathbf{P}_2(E)\}$$

is a λ-system. Clearly, $\Omega \in \mathcal{L}$. Moreover, if $E, F \in \mathcal{L}$ and $E \subset F$, then

$$\mathbf{P}_1(F - E) = \mathbf{P}_1(F) - \mathbf{P}_1(E) \;=\; \mathbf{P}_2(F) - \mathbf{P}_2(E) = \mathbf{P}_2(F - E)$$

so that $F - E \in \mathcal{L}$. Finally, if $E_n \in \mathcal{L}$ and $E_n \uparrow E$, then by Theorem 1,

$$\mathbf{P}_1(E) = \lim \mathbf{P}_1(E_n) \;=\; \lim \mathbf{P}_2(E_n) = \mathbf{P}_2(E)$$

so that $E \in \mathcal{L}$. This was to be shown. ■

An Example The *Borel σ-field* on a topological space is the σ-field generated by the family of open subsets. The most important of all is the σ-field $\mathcal{B}(\mathbf{R})$ of the set $\mathbf{R}$ of real numbers with the usual topology.

Though not all, virtually every subset of $\mathbf{R}$ is in $\mathcal{B}(\mathbf{R})$. Indeed, it is not easy to construct a subset of $\mathbf{R}$ which is *not* in $\mathcal{B}(\mathbf{R})$. Elements of $\mathcal{B}(\mathbf{R})$ can therefore be quite complicated. Instead of $\mathcal{B}(\mathbf{R})$, we may often consider a π-system defined by

$$\mathcal{P} = \{(-\infty, x]\}$$

for $x \in \mathbf{R}$, which is much simpler to deal with. It is not difficult to show that $\mathcal{P}$ generates $\mathcal{B}(\mathbf{R})$. It is immediate that $\sigma(\mathcal{P}) \subset \mathcal{B}(\mathbf{R})$, since for all $x \in \mathbf{R}$

$$(-\infty, x] = \bigcap_n \left(-\infty, x + \tfrac{1}{n}\right)$$

To show the other direction, we first note that every open set of $\mathbf{R}$ is a *countable* union of open intervals. It therefore suffices to show that the open intervals of the form (a, b) are in $\sigma(\mathcal{P})$, which follows directly from

$$(a, b) = (-\infty, a]^c \cap \left(\bigcup_n \left(-\infty, b - \tfrac{1}{n}\right]\right)$$

Here we made the convention that intervals with starting points bigger than ending points are null sets.

Thus we have seen that the collection $\mathcal{P}$ generates $\mathcal{B}(\mathbf{R})$. Theorem 6 indicates that the two probabilities $\mathbf{P}_1$ and $\mathbf{P}_2$ on $\mathcal{B}(\mathbf{R})$ are identical, if they agree on $\mathcal{P}$, a π-system defined above.

4. Conditional Probability and Independence

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space.

Definition 5 *For an event F such that $\mathbf{P}(F) > 0$, we define the conditional probability of E given F by*

$$\mathbf{P}(E \,|\, F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)}$$

Properties

(a) For a fixed F the function $\mathbf{Q}(\cdot) = \mathbf{P}(\cdot \,|\, F)$ is a set function which satisfies the three axioms of probability. We may indeed easily show that $\mathbf{Q}(E) \geq 0$, $\mathbf{Q}(\Omega) = 1$, and, for disjoint events $E_1, E_2, \ldots$, $\mathbf{Q}(\bigcup E_n) = \sum \mathbf{Q}(E_n)$. Being a probability, all the properties of probabilities introduced earlier hold for $\mathbf{Q}$. For instance, it is immediate that $\mathbf{Q}(\emptyset) = 0$ and $\mathbf{Q}(E^c) = 1 - \mathbf{Q}(E)$, which implies that $\mathbf{P}(\emptyset \,|\, F) = 0$ and $\mathbf{P}(E^c \,|\, F) = 1 - \mathbf{P}(E \,|\, F)$.

(b) It is often convenient to define a probability of intersection via a conditional probability, i.e.,

$$\mathbf{P}(E \cap F) = \mathbf{P}(E \,|\, F)\, \mathbf{P}(F)$$

which can easily be extended to

$$\mathbf{P}(E \cap F \cap G) = \mathbf{P}(E \,|\, F \cap G)\, \mathbf{P}(F \,|\, G)\, \mathbf{P}(G)$$

or more general cases.

(c) It is said that $\{F_n\}$ is a *partition* of Ω when F_n's are mutually exclusive and exhaustive, i.e., Ω is their disjoint union. If a sequence $\{F_n\}$ of events is a partition of Ω, then

$$\mathbf{P}(E) = \sum_n \mathbf{P}(E \,|\, F_n)\mathbf{P}(F_n)$$

for any event E. This property is often referred to as the *theorem of total probability.* The *Bayes' Formula*

$$\mathbf{P}(F_k \,|\, E) = \frac{\mathbf{P}(E \,|\, F_k)\mathbf{P}(F_k)}{\sum_n \mathbf{P}(E \,|F_n)\mathbf{P}(F_n)}$$

is also immediate.

Definition 6 *Events E and F are called independent if*

$$\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$$

Remarks

(a) We may equivalently reformulate the definition of independent events as

$$\mathbf{P}(E \,|\, F) = \mathbf{P}(E) \quad \text{or} \quad \mathbf{P}(F \,|\, E) = \mathbf{P}(F)$$

when $\mathbf{P}(F) > 0$ or $\mathbf{P}(E) > 0$, respectively.

(b) In general, $E_1, E_2, \ldots$ are said to be independent if

$$\mathbf{P}(E_{i_1} \cap E_{i_2} \cap \cdots E_{i_k}) = \prod_{j=1}^{k} \mathbf{P}(E_{i_j})$$

for any $k \geq 2$, where $(i_1, \ldots, i_k)$ is an arbitrary collection of k numbers from the index set $\{1, 2, 3, \ldots\}$.

5. Exercises

1. Show that an arbitrary intersection of σ-field, π-system and λ-system is a σ-field, π-system and λ-system, respectively.

2. Answer the following:

(a) Suppose that $\mathcal{F}_n$ are fields satisfying $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. Show that $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ is a field.

(b) Suppose that $\mathcal{F}_n$ are σ-fields satisfying $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. Show by providing a counter-example that $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ need not be a σ-field.

3. Answer the following:

(a) In a topological space, closed sets are defined to be the sets whose compliments are open. Show that the Borel σ-field is also generated by the family of closed sets.

(b) Give a collection of the subsets of $\mathbf{R}$ other than $\mathcal{P}$, which generates $\mathcal{B}(\mathbf{R})$.

4. Let $\{E_n\}$ be a monotone decreasing sequence of events, and let $E = \lim E_n$. Prove that $\lim \mathbf{P}(E_n) = \mathbf{P}(E)$.

5. Show that $\lim_{n\to\infty} \left(-1 + \frac{1}{n}, 1 - \frac{1}{n}\right] = (-1, 1)$.

6. Let $\mathbf{R}$ be the sample space, and the probability be given by the Lebesque measure on the unit interval $[0, 1]$. We define a sequence E_n of subsets of $\mathbf{R}$ by

$$E_n = \begin{cases} \left(-\frac{1}{n}, \frac{1}{2} - \frac{1}{n}\right] & \text{if } n \text{ is odd} \\ \left[\frac{1}{3} - \frac{1}{n}, \frac{2}{3} + \frac{1}{n}\right) & \text{if } n \text{ is even} \end{cases}$$

Find $\liminf E_n$ and $\limsup E_n$. Also, compare $\mathbf{P}(\liminf E_n)$ with $\liminf \mathbf{P}(E_n)$ and $\mathbf{P}(\limsup E_n)$ with $\limsup \mathbf{P}(E_n)$.

7. Answer the following:

(a) If the events E and F are independent, then so are E^c and F^c.

(b) The events Ω and $\emptyset$ are, respectively, independent of any event E.

(c) Is there an event which is independent of itself?