Chapter 5

MEASURE THEORY WITH APPLICATIONS TO ECONOMICS

ALAN P. KIRMAN

Université d'Aix-Marseille

This chapter will first present problems arising from economic theory, the modelling of which has required, in an essential way, measure theory. Having explained why measure theory is needed, we will give, for reference, some basic measure theoretical concepts and results, and this will be followed by a development and discussion of some particularly useful and less accessible results.[1]

1. The use of measure theory in economics

1.1. Perfect competition: Large economies

The idea of "perfect" or "pure" competition is a very old one in economics.[2] Any economist will have an intuitive idea as to what is meant by it, though the definitions may vary. The underlying principle may be captured by saying that a situation in which a group of individuals together are involved in economic activity, exchange for example, and in which no individual can, and therefore no individual will try to, affect the outcome is one of perfect competition. The first and most obvious requirement for such a situation is that there should be "many" individuals. This is clearly not enough; we also need that none of these individuals should be "important". The conceptual difficulty arises if we really insist that each individual shall have no influence whatsoever. As an example, think of the productive sector of an economy, and then what we require for it to be "perfectly competitive" would be, for example, that if a producer stopped production completely, the price of the commodity he produces should not

[1] The presentation is aimed at the "informed reader", that is, someone acquainted with the basic ideas of measure theory as presented in a course on probability theory. The better informed mathematician will see from the first and third sections where notions familiar to him are used in economics. The reader who is not sure where he stands should read quickly through Section 2. If it has a familiar ring, he should find the chapter profitable. If not, he would be well advised to first consult any standard text on measure theory, the classic reference being Halmos (1961).

[2] Adam Smith is clearly aware in the *Wealth of Nations* (Book 1, ch. 7, for example) that the existence of large numbers of agents, that is of a situation approaching perfect competition, diminishes the power of an individual agent to influence prices in a market process.

change. If prices can vary, then for this to be strictly true for every producer, there must be an infinite number of producers, or, in other words, his "influence as a producer must be genuinely negligible". As the term "perfect competition" suggests, this is an idealisation, not a description of reality; but the examination of such an ideal case, as in other sciences, provides us with useful insights into the working of economics.[3]

The idea that individuals should have no weight but that collectively they have positive weight is a familiar one in mathematics, and it is the basis of measure theory. If we wished merely to describe a perfectly competitive economy, it would be enough to consider any infinite set as the set of agents, or agents' names, and to specify the "characteristics" of each agent. Thus, in an exchange economy, agents are characterized by two things: preferences and an initial bundle of goods. Without entering into any details, consider the set of possible preferences as $\mathcal{P}$ and the set of bundles of goods on which these preferences are defined as R^{ℓ}_{+}, that is, there are ℓ goods. Then an exchange economy $\mathcal{E}$ is given by

$$\mathcal{E}: A \to \mathcal{P} \times R^{\ell}_{+},$$

where A is some arbitrary set of agents. Now, if we require that there are an infinite number of agents in A and that no agent has too much of any goods, e.g. that we restrict attention to some bounded set of R^{ℓ}_{+} for initial endowments, then we have a description of a perfectly competitive exchange economy. Provided that all we require is a definition, this would indeed suffice. However, if we wish to work with this model, we will need more than this. Suppose that we are concerned with the problem of competitive or Walrasian equilibrium. We need now some way of expressing the idea that for some *allocation of goods f* "supply equals demand".[4] In a finite economy, we simply add the demands of all the individuals and check that this is equal to the sum of all the initial resources. Thus we require that, referring to the initial bundle of an individual as $e(a)$,[5] that

$$\sum_{a \in A} f(a) = \sum_{a \in A} e(a), \tag{1.1}$$

and that $f(a) \in \varphi(p, a)$ where φ, the demand of an individual a at prices p, is defined in the normal way. Now, in our infinite economy, we can no longer add

[3]A discussion of the notion of perfect competition and its relation to recent theoretical developments is to be found in Mas-Colell (1979).

[4]An allocation of goods is here $f: A \to R^{\ell}_{+}$.

[5]$e(.)$ is the projection of the mapping $\mathcal{E}$ onto R^{ℓ}_{+}.

supply and demand. Instead we can substitute the idea that *mean supply is equal to mean demand*. In a finite economy, we could write

$$\frac{1}{\#A}\sum_{a\in A} f(a) = \frac{1}{\#A}\sum_{a\in A} e(a), \tag{1.2}$$

which is clearly identical to (1.1). However, in the infinite case, we can resort to the equivalent idea, one which will be familiar to all those who know a little probability theory, and write

$$\int_A f(a)\,\mathrm{d}\nu = \int_A e(a)\,\mathrm{d}\nu. \tag{1.3}$$

In writing (1.3), although its intuitive meaning is clear, we have introduced a number of technical complications. We have *integrated* but for this to be well defined, we have to integrate "with respect to some measure", that is, we must define a function ν which attributes a certain weight to each set of individuals. Intuitively, we can think of this as a "counting measure", i.e., one which says what proportion of individuals are in each set. The only important thing for us, for the moment, is that for an infinite economy such a measure should give zero weight to individuals, and for convenience, that it should give weight one to the whole set. Such a function is an *atomless probability measure*.[6] The machinery of measure theory provides a convenient way of resolving many economic problems in the context of such ideal economies. To use the standard tools of this theory imposes some technical restrictions, which will be specified in the next two sections, but suffice it to say that to construct an idealised or *perfectly competitive economy*, we take the set of agents A to be represented by an *atomless measure space* $(A, \mathscr{Q}, \nu)$, the three components being the set A, the collection $\mathscr{Q}$ of subsets on which the measure ν is defined,[7] and the measure ν itself.

The notion of an ideal economy, in the context discussed, was introduced by Aumann (1964), but a continuum of agents had already been used in economics by Allen and Bowley (1935), and in game theory by Shapley (1953), and in a number of other papers in the early 1900's. The idea of perfect competition in the sense that individuals believe that prices are given and beyond their control has a long history, accounts of which can be found in Schumpeter (1954) and Blaug (1968) for example, but it is only with the introduction of the "continuum theory" that such a behaviour is strictly justified. Indeed in the work of Torrens, Cournot, and Edgeworth is to be found a discussion as to whether it is rational for individuals to behave in this way. As Viner remarked, the fact that it is not has remained a "skeleton in the cupboard of free trade". The use of a measure

[6]We will in Section 2 come back to the precise definition of "atomless", and the reader will hopefully pardon a slight looseness in the statements above.

[7]Unfortunately, this is not $P(A)$, the set of all subsets of A, but more of this later.

space of agents thus enables us to formalise the notion of perfect competition,[8] but the next question is obviously: "Does it enable us to develop stronger results?" A first result showing how the assumption that an economy is large, in the sense described, leads us to drop assumptions necessary in the finite case, concerns the existence of equilibrium.[9] In a finite economy, we typically need to make an assumption about the convexity of the preferences of individuals to prove the existence of equilibrium. If we make a strong assumption that preferences are strictly convex, then the bundle demanded by an individual a at prices p, denoted $\varphi(a, p)$, will be unique, that is, φ will be a function. If we weaken the assumption to making preferences convex, then $\varphi(a, p)$ will be a set but a convex one, and we will have

$$\sum_{a \in A} \varphi(a, p) \quad \text{is convex for all } p.$$

Provided that total demand, or equivalently mean demand, is a convex set, we can prove that equilibrium exists.[10] If however individuals have nonconvex preferences, their demand may not be a convex set for some prices and the proof of existence no longer goes through. To look at this question in the continuum case, we must first be able to define the integral of individuals' demands, which are set-valued functions or correspondences. The *integration of correspondences* is discussed in Section 3 of this chapter. The important result is that even if we do not assume individual's preferences to be convex, nevertheless $\int_A \varphi(a, p) \mathrm{d}\nu$ is convex.

Thus, what might be thought of as irregular behaviour in individuals becomes "well-behaved" in large economics. This fact enables one to prove the existence of equilibria in large economies under weaker assumptions than in the finite case. See Aumann (1966) and Hildenbrand (1970).

1.2. Different solutions for the market problem

A further important result proved by Aumann (1964) was the equivalence between two different solution concepts in a continuum economy. One based on

[8] The approach adopted here is not by any means the only possible one. We use σ additive measures, and it may be possible to work with only finitely additive measures, but this slight conceptual weakening of assumptions leads to other complications in the definition of "atomless" for example. Another different approach is to consider agents as infinitesimally small but not null. To do this involves using non-standard analysis developed by Robinson (1965) and used in economics by Brown–Robinson (1974) and Khan (1973). The disadvantage of this approach is that the mathematical apparatus employed is familiar to a very limited audience.

[9] The essential result in this connection is Liapunov's theorem which will be given later.

[10] The standard discussion of this problem is given in Debreu's *Theory of Value* (1959), and a complete survey of the work in this area is given in Chapter 15 of this Handbook.

the price mechanism gives us the set of allocations which are equilibria denoted $W(\mathcal{E})$ and the other, the *core*, is the set of allocations upon which no coalition S of individuals can improve. "Improve upon" in this sense means that a coalition S of agents could reallocate its initial resources to make its members better off. Thus in a continuum economy $\mathcal{E}$, for example, where the set of agents A is the closed unit interval of the real line, an allocation f would be improved upon by S if the members can find g with

$$g(a) \succ_a f(a) \quad \text{for all members of } S,^{11}$$

and (1.4)

$$\int_S g(a)\,\mathrm{d}\nu = \int_S e(a)\,\mathrm{d}\nu.^{12}$$

Allocations which can be improved by no coalition form the core of the economy $\mathcal{E}$ denoted $C(\mathcal{E})$.

Aumann's result is that for "continuum economies",

$$W(\mathcal{E}) = C(\mathcal{E}).$$

This exact equality for an ideal economy confirmed in a more general setting an old asymptotic result of Edgeworth (1881) and a later result of Debreu and Scarf (1963) and gave rise in turn to a whole series of very general asymptotic results which are treated in detail in Chapter 18 of this Handbook on the core, and to which we will return shortly.

In discussing perfect competition, we have given an idea as to why atomless measure spaces provide a useful formalisation of the idea of a large economy in which each agent is insignificant. If this were indeed the only value of such tools, then it would be difficult to persuade economic theorists of the virtue of acquiring them. In fact, measure theory provides extremely useful insights at a conceptual level.

1.3. Distributions of characteristics

In a large economy, listing the characteristics of all the individual agents would be both a tedious and an elaborate task. Indeed, economists often make the simplifying step of describing an economy by the *distribution of its agents' characteristics*. The idea of the income distribution and describing it by some

[11] $x \succ_a y$ denotes that agent a strictly prefers x to y.

[12] The informed reader will note that (1.4) is not, for technical reasons, defined for all coalitions S; details will be forthcoming in Section 2.

such function as the Pareto distribution is well established in economics. The use of such functions relies implicitly on the idea that a large economy may be represented as a continuum, and the measure space of agents approach leads naturally to the development of the distribution as a fundamental concept.

If we consider a mapping from a probability space into the space of characteristics, then it is clear that a natural probability measure is induced on the latter. If we take a subset B of the characteristics space then consider the set C in the original space whose image lies in that subset, that is $C=\mathfrak{E}^{-1}(B)$. Now, let the measure of B be the measure of C; this gives us a measure on the characteristics space itself. Thus, instead of asking which agent has which characteristic in an economy, we might ask what proportion of agents have certain characteristics? Instead of thinking of an economy as a detailed listing of all the characteristics of the agents in that economy, we can think of it as a distribution of characteristics. Indeed as we have said, economists are in the habit of viewing economies as characterized by their income distribution, for example. We might, indeed, reasonably say that two economies for which the distributions of agents characteristics are the same are effectively the same economy. For this to be acceptable, we would have to show that the equilibria of these economies are the same. A full treatment of this sort of problem may be found in Hildenbrand (1975).

A little more formally, consider $(A, \mathfrak{A}, \nu)$ a probability space, M the space of characteristics, and f a mapping of A into M. The distribution ν of f denoted by $\mu \circ f^{-1}$ is defined by

$$\nu(B)=\mu\{a\in A \mid f(a)\in B\} \quad \text{for every subset } B \text{ of } M.$$

The reader will already be familiar with this idea from probability theory and will recognise f as a random element and, in particular if M were the real line, would recognise f as a random variable.

Now, since for many purposes we take some arbitrary basic measure space as a starting point, it is frequently the distribution that conveys the real information with which we are concerned. For example, in studying "large economies", the choice of the unit interval [0,1] where it is used as the space of agents is purely for convenience and has no particular significance itself. In fact, given a suitable distribution on the space of characteristics of agents, we could always construct an associated economy with the unit interval as the space of agents.

1.4. Limit theorems

The idea of using distributions as the description of the essential features of an economy proved extremely useful in translating results from ideal or continuum economies to large but finite economies. For results on ideal economies to be of any interest, they must also hold, at least approximately, for large enough finite economies.

Thus, rather than make a statement that such and such a result is true for a continuum economy say $\mathcal{E}_\infty$, we would like to construct an increasing sequence of economies $\mathcal{E}_n$ converging in some sense to $\mathcal{E}_\infty$ and then make the assertion that our result is approximately true for large enough n. The problem is that if we think of our economies as being listings of all the characteristics of the agents, the dimension of this description changes as more agents are added and as the economies of the sequence increase in size. How then can we construct an increasing sequence of economies and in what sense can that sequence be said to converge to the limit, atomless, economy?

An important key to solving this problem is that we can construct a sequence of parallel "equivalent" economies each with a continuum of agents and establish our results via this "equivalent" sequence. However, we will need to establish the meaning of the "equivalence" between the original sequence of finite economies and the sequence of artificially constructed economies. To handle these problems, we will need a number of mathematical tools, in particular, we will need to study the *convergence of measures* or more exactly *weak convergence of measures*. We will need subsequently to develop the idea of *convergence in distribution* so that we can give precision to the requirement that for a given sequence of economies "the distribution of agents' characteristics" should be "close", for n large, to that of the limit economy.

1.5. Many but different agents

As must by now be evident, much of the value of measure theoretical tools is to handle situations in which there are "many" agents. We have discussed the weakening of assumptions possible in "ideal" economies to achieve standard results. Thus the assumption of large numbers may be seen to be a substitute for restrictive hypotheses at the individual level. Sometimes however we need more than simply "many" agents. We will need that the agents are, in some sense, different, thus not only numbers but variety will be important. If we think of the distribution of agents characteristics, then we could require for example that the *support* of that measure should not be "too small", the support of a measure being the smallest set that has full measure. Thus we would require that peoples' characteristics in an economy are not too similar.

For what sort of economic problem is this of interest? A well-known difficulty in economics is that associated with the assumption of strict convexity of preferences. Although every elementary text in micro-economics has diagrams of preferences which inevitably satisfy this hypothesis, only the most hardened economic theorist feels completely at ease with it. Plausible counter-examples are so easy to find that one would be happy to dispense with it. However, the formal difficulties that arise when it is removed are far from trivial to overcome. In particular, as we have already remarked, since at given prices p the bundle of goods demanded by agent a, that is $\varphi(a, p)$, is not necessarily unique, one can no longer work with demand functions. However, intuitively it is clear that if there is a large number of agents and the number of these who have more than one element in their demand set, is "negligible", then we have essentially what we require. For this idea to make perfect sense, we must have an infinite number of agents. Now, if we have an infinite number of agents, what we need is that "mean demand" should be unique. For this we will have to integrate over our agents,[13] and hence what we must show is that the "bad" set of agents have *measure zero*. For this we obviously must require that the preferences are sufficiently "dispersed". Results in this direction using assumptions of differentiability have been obtained by Sondermann (1975), Dierker, Dierker and Trockel (1978) and Araujo and Mas Colell (1978). Hildenbrand (1979) has shown with a suitable assumption about dispersion of preferences that the almost sure uniqueness of maximisers and hence the continuity of mean demand functions can be obtained without any differentiability assumptions. Again the usefulness of measure-theoretic tools in making precise an intuitive idea should be emphasised. For the use of continuous demand functions to be strictly justified in a context of non-strictly convex preferences, an infinite number of agents is essential, and to use the natural notion of the mean demand function, the measure theoretical approach is necessary.

Before leaving this topic, an important observation should be made. How are the above results obtained? They depend on showing that a certain phenomenon is "exceptional" or "unusual". The significance of this is that for a long time, unless we made extremely restrictive assumptions in economics, we were unable to rule out intractable situations even though it seemed unlikely that they might occur. One approach to this is topological. Thus rather than make strong assumptions to rule out certain phenomena one can show that the set of economies that exhibits these phenomena is "negligible", that is, that the set of well-behaved economies is *open and dense* in the set of all the economies under consideration.[14]

Thus, one can in a certain sense ignore such phenomena. One might also like to say that, in a probabilistic sense, certain things are unlikely, or more precisely, that the set of objects, economies, for example, exhibiting certain phenomena has measure *zero*, or that such a phenomena is "almost sure" not to occur.[15] In

[13]Agents are, in fact, identified by preferences, and A is an open subset of R^n. Thus preferences or utility functions can be classified by n parameters.

[14]The pitfalls of too facile a use of this approach are alluded to in Grandmont, Kirman and Neuefeind (1974), and the same strictures of course apply to the measure-theoretic approach.

[15]A fundamental paper which shows that economies with an infinite number of equilibria are unlikely in both the topological and probabilistic sense is that of Debreu (1970).

the papers mentioned above on the uniqueness of maximising elements, it is precisely this notion that allows the passage from individual demand correspondences to mean demand functions.

1.6. *Price forecasting, tight measures, and compactness*

In many situations we are led to introduce restrictions of the opposite sort of those mentioned earlier. When, for example, we want to establish existence of an equilibrium, we will need certain "compactness" properties. In particular, if we define measures on a space which is not itself "compact", we will need to restrict ourselves to families of measures which are, in a technical sense, concentrated essentially on a compact set. This technical requirement arises naturally in work on temporary equilibria.[16]

Consider traders who base their forecasts of future prices on today's prices. Thus any price vector today generates a measure on the space of tomorrow's prices. In a model of this sort, to ensure the existence of an equilibrium, one is led to assume that tomorrow's prices do not depend "too strongly" on today's prices. In other words, if some prices today become very high, then individuals attach a low probability to their being exceeded tomorrow. This rules out, for example, the simple-minded forecast that tomorrow's prices will, with probability one, be equal to today's prices.

The underlying stabilising assumption is clear; what we want is that if prices become very high today, for example, traders will attach a high probability to their diminishing tomorrow, and it is this that prevents prices exploding. The formal requirement is that the family of measures or forecasts should be *tight*. This requirement also plays an important role in work on large economies.[17]

1.7. *Social choice with many agents*

Arrow (1963) proved a theorem which is widely regarded as the most important in the field of social choice.[18] What he showed was that there is no rule for aggregating individual preferences, which respects certain apparently reasonable conditions. It was later shown by Fishburn (1970) that Arrow's theorem is not true if there is an infinite number of individuals in the society in question. This has been interpreted as meaning that in large societies Arrow's result loses its significance and its importance is thus diminished. However with many agents, we may obtain a measure theoretic equivalent of Arrow's theorem; see Kirman and Sondermann (1972).

[16] See Chapter 19 by Grandmont.
[17] See Chapter 18 by Hildenbrand.
[18] See Chapter 22 by Sen.

To sketch the problem briefly, consider A the set of individuals, X a set of alternatives, and P the set of preorders (preferences) on X. What we are looking for is a rule that will associate with a given distribution of preferences among the individuals (we will call this a "situation"), preferences for the society.

Let $f: A \to P$ be a situation, then $\mathcal{F}$ is the set of all possible situations. Then a social preference rule is $\sigma: \mathcal{F} \to P$. What Arrow shows is that given certain reasonable restrictions on σ the only rule that exists is the following "dictatorial" one:

Choose one individual a and, no matter what the preferences of the other individuals, if a prefers x to y, then x is socially preferred to y. Written with the obvious notation

$$xf(a)y \quad \text{implies} \quad x\sigma(f)y.$$

Since Arrow rules out such a dictatorial function, no social welfare function σ is possible. The mathematical structure of this problem is now well-known. The Arrovian axioms impose a very specific structure on the sets of individuals who are "socially decisive". That is the set B is decisive if, when all the members of B prefer x to y, then x is socially preferred to y. In the case where the set A of all individuals is finite, these socially decisive sets consist of all the sets that contain a given individual a and, in particular, the set $\{a\}$ consisting of just a himself. Now, suppose that the set A is infinite, for example, the interval $[0,1]$, then we could, from Arrow's axioms, define a measure which could give weight 1 to the decisive sets and 0 to the others. If Arrow's theorem translated directly to this case, the measure μ would necessarily have the form

$$\mu(C)=1 \quad \text{if and only if} \quad a \in C.$$

Thus a would be the dictator. In particular, note that such a measure is *not atomless* and that, for this reason, unlike the other measures with which we shall work, it is defined on every subset of A.

However, we know that Arrow's result does not hold in this case, but we also know that to discuss single individuals in such a case does not make much sense. What we can show is a different sort of result. If A is $[0,1]$ then, given Arrow's axioms, any social rule σ has the following property:

Given any ε, there is a socially decisive set C with $\mu(C) < \varepsilon$,

where μ is the natural Lesbegue measure, i.e., the "length" of the set C. Thus, though no single individual determines society's preferences, arbitrarily small coalitions do so. Thus, the measure theoretic approach enables us to show that Arrow's result remains essentially true even in the infinite case.

1.8. How to cut a cake fairly

An old problem which has intrigued mathematicians is that of how to divide up some object "fairly" in some sense, among n individuals. The object to be interesting, of course, must be differently appreciated by different individuals. One could think of a block of ice cream with different flavours. Thus one could think of each individual i assigning a "measure" μ_i to the parts of the ice cream, each attributing 1 to the whole for example. Thus, what we would like is to find a way of dividing the ice cream U, i.e. a partition of U, $\{U_1,\ldots,U_n\}$, such that

$$\mu_i(U_i) \geqq 1/n \quad \text{for} \quad i=1,\ldots,n.$$

This would be fair in the sense that each individual receives in his own eyes at least $1/n$ of the value of the ice cream. In the case of two individuals, all those who have children will know that the method of "divide and choose" solves the problem. However, much better results in the n person case have been proved by Steinhaus, Banach and Knaster, and references are given and very general theorems proved in an elegant paper by Dubins and Spanier (1961). A very striking result shows that one can partition the ice cream or cake in question in such a way that each individual n believes that *all* the pieces of the cake are worth $1/n$. That is, one can find a partition $\{U_1,\ldots,U_n\}$ such that

$$\mu_i(U_j)=1/n \quad \text{for} \quad i=1,\ldots,n \quad \text{and} \quad j=1,\ldots,n.$$

This rules out an individual getting $1/n$ of the cake but being jealous of another individual. This is in fact equivalent to the old problem of the agricultural land of an Egyptian village which is flooded by the Nile to different heights each year. How should the land be divided so that each of the n landowners always has $1/n$ of the land remaining above water?

In addition, Dubins and Spanier show that there are "optimal" partitions in different senses. For example, there are partitions $\{U_1,\ldots,U_n\}$ which maximise

$$\sum_{i=1}^{n} \mu_i(U_i),$$

thus which are optimal in a utilitarian sense. Connections with other mathematical results are shown in their paper and the central role played by *Liapunov's theorem* mentioned above is clear.

Here again, the measure-theoretical approach has solved a number of interesting problems arising in an economic context. Having given a number of examples to motivate the use of measure theory in economics, we now turn to the mathematical tools themselves.

2. Some basic measure theory

The area covered by measure theory may be thought of as that concerned with attributing numbers to the parts of an object or set in such a way that these numbers correspond intuitively to the "size" or "measure" of those parts. Physically one might think of the weight of some object and its component parts, or if one takes an interval of the real line, one might be interested in the "length" of some subset of that interval. Again, from the point of view of intuition, it is important that if the numbers assigned are to be meaningful, they should have certain properties of additivity. Thus, if one takes two disjoint parts of an object, one would naturally require that the weights of these two parts taken together should equal the sum of their separate weights. Indeed, we would require that this be true not just for any two sets, but for arbitrary collections of subsets. The passage from the simple idea of adding the weights of a finite collection of subsets to find the weight of their union to the problem of adding the weights of an arbitrary collection of subsets is not even in general possible, and we will have to restrict ourselves to a less ambitious task. The specification of the functions that designate the measure of each subset of some set, the collection of subsets on which they are defined, and the properties of those functions will be the concern of the second part of this chapter.

2.1. Classes of subsets and algebras

Before developing the theory of set functions and measures in particular, we must first study the classes of subsets on which they are defined. If we consider any set E then we will denote $\mathcal{P}(E)$ the set of all subsets of E.

Definition 1

An *algebra* or *Boolean algebra* of sets $\mathcal{Q}$ is a non-empty class of subsets of E such that if

$$A \in \mathcal{Q} \quad \text{and} \quad B \in \mathcal{Q},$$

then

$$A \cup B \in \mathcal{Q} \quad \text{and} \quad A \backslash B \in \mathcal{Q},$$

and

$$E \in \mathcal{Q}.$$

It follows obviously that if $\mathcal{Q}$ is an algebra, then

$$A \in \mathcal{Q} \quad \text{implies} \quad A^c \in \mathcal{Q}.$$

Examples

It is clear that for any set E, $\mathcal{P}(E)$ is a Boolean algebra.

The set of all intervals on the real line does not form an algebra since it is closed neither under the operation of difference nor that of union. However, the reader will be able to show that the set of all finite unions of intervals is an algebra.

We will need to consider classes of sets where there are members which cannot only be formed by the finite union of other members but also by countable unions. That is, if we consider some set E then we will need to be able to talk of the "weight", "size", or "measure" of some set which can be made up of a countable number of "pieces" of E. Thus we have:

Definition 2

A *σ algebra* is an algebra with the property that if

$$A_i \in \mathcal{Q} \quad \text{then} \quad \bigcup_{i=1}^{\infty} A_i \in \mathcal{Q}, \qquad i=1,2,\ldots$$

It is clear that the countable intersection of sets in a σ algebra belongs to that σ algebra.

2.2. *Generated algebras and σ algebras*

If we could always work with the set of all subsets of some set E, that is with $\mathcal{P}(E)$, and could define a measure on such subsets, things would be very simple. However, this is not possible and we have to restrict ourselves to subclasses of $\mathcal{P}(E)$. It is for this reason that we have introduced notions of algebras and σ algebras. In particular, it will often be useful to start with some simple class of subsets and to construct from it a larger class. To this end, we give the following:

Theorem 1

If $\mathcal{Q}$ is any class of sets then there exists a unique algebra (resp. σ algebra) such that $R \supset \mathcal{Q}$, and if R' is also an algebra (resp. σ algebra) such that $R' \supset \mathcal{Q}$, then

$$R \subset R'.$$

The class R is referred to as the algebra (resp. σ algebra) generated by the class $\mathcal{Q}$.

A particularly important class of sets is given by the smallest σ algebra containing the open sets of some topological space. Formally, we have:

Definition 3

For a topological space X the *class of Borel sets* is the σ algebra $\mathcal{B}$ generated by the open sets of X. The reader will have no difficulty in showing that the Borel sets are also generated by the closed sets of X.

It will be useful later to work with the σ algebra generated by a class of sets. Although it is unfortunately impossible to give a constructive procedure for obtaining this σ algebra, this will not, at the level of presentation here, present any difficulty.

With these simple set structures in mind, we now pass on to consider set functions, and in particular those set functions which are called measures.

2.3. *Set functions*

We will confine our attention to set functions which will be defined on a non-empty class $\mathcal{C}$ of subsets of some set E. Thus μ associates with a set $A\in\mathcal{C}$ a real number or $\pm\infty$. The empty set $\varnothing$ is always a member of $\mathcal{C}$. If we denote by R^* the compactification of the real line by the addition of the two points $+\infty$ and $-\infty$, then the operations represented by $+$ and $\times$ are extended in the conventional way, for example,

$$0\times\pm\infty=0.$$

The purpose of this chapter is not to consider arbitrary functions of abstract interest, but to tie ourselves to those which will be of use in economic theory. A first condition that the functions must satisfy if they are to correspond to the intuitive idea of assigning "weights" or "lengths" is that the "weight" of two disjoint sets taken together should be equal to the sum of their individual weights.

Definition 4

A set function $\mu:\mathcal{C}\to R^*$ is said to be *(finitely) additive* if

(i) $\mu(\varnothing)=0$,

(ii) for every finite collection $E_1, E_2,\dots, E_n$ of disjoint sets of $\mathcal{C}$ such that $\bigcup_{i=1}^n E_i\in\mathcal{C}$,

then

$$\mu\left(\bigcup_{i=1}^{n} E_i\right)=\sum_{i=1}^{n}\mu(E_i). \tag{2.1}$$

In fact, condition (i) is superfluous provided that for some set E in $\mathscr{Q}$, $\mu(E)$ is finite. It is natural to define an additive set function on an algebra since we have

$$E_i \in \mathscr{Q} \quad \text{implies} \quad \bigcup_{i=1}^{n} E_i \in \mathscr{Q}, \qquad i=1,\ldots,n.$$

The reader will note that if $\mathscr{Q}$ is an algebra, we cannot have sets E and $F \in \mathscr{Q}$ with $E \cap F = \varnothing$ and $\mu(E) = +\infty$ and $\mu(F) = -\infty$. Thus, although it is not always sufficient to confine our attention to finite valued set functions, they will not take on both the values $+\infty$ and $-\infty$.

We will now give several examples of set functions which are additive which will aid in understanding the nature of measure.

Example 1

Consider X any set with infinitely many points and the set of all subsets of X. Define μ by

$$\begin{aligned} \mu(E) &= \#E \quad \text{if } E \text{ is finite,} \\ &= +\infty \quad \text{if } E \text{ is infinite,} \end{aligned} \qquad \text{for} \quad E \in \mathscr{P}(X).$$

Thus if X represents the individuals in a large economy, this measure simply "counts" the agents in any coalition. We will encounter a more useful "counting measure" later in the chapter.

Example 2

Consider X any set and define $\mathscr{P}(X)$ as before.

For $\hat{x}$ a point of X, let

$$\begin{aligned} \mu(A) &= 1 \quad \text{if} \quad \hat{x} \in A, \\ &= 0 \quad \text{if} \quad \hat{x} \notin A, \end{aligned} \qquad \text{for} \quad A \in \mathscr{P}(X).$$

Example 3

Let $X = R$ and let $\mathscr{Q}$ be the set of all finite intervals of R. Any $E \in \mathscr{Q}$ is then defined by its end points a, b, and let

$$\mu(E) = b - a.$$

Thus we simply take the value of an interval to be its length.

Example 4

Let X be the half open interval $(0,1]$ and let $\mathscr{Q}$ be the class of half open intervals $(a,b]$ with $0 \leqslant a \leqslant b \leqslant 1$, and let

$$\mu(a,b) = b - a \quad \text{if} \quad a \neq 0,$$

and

$$\mu(0,b) = +\infty.$$

All these examples are of additive set functions, but we will come back to see whether they satisfy the additional criteria that we will impose.

Having defined finite additivity, we will now give a stronger requirement—that of σ additivity—that is we will ask that our set function should be additive not only on a finite union of sets but on countable unions as well. Why is this necessary? The following example from probability theory gives a clear answer.

Definition 5

Consider a set X and $\mathscr{Q}$ an algebra of subsets of X. Define a finitely additive function,

$$\mu : \mathscr{Q} \to [0,1],$$

with $\mu(X) = 1$. Such a function is called a *probability distribution*. If one thinks of an experiment with a number of possible outcomes then $\mu(S)$ expresses the intuitive idea of the probability that the outcome of the experiment will be in the set S.

Now consider a map $f: X \to R$. Such a map is called a *simple random variable*.[19] Thus, it associates a real number to each possible outcome of an experiment.

Next consider an infinite sequence of independent trials of the experiment. That is, from the population X is drawn each time, according to the probability distribution μ, an element $x \in X$. An outcome then may be represented as $(x_1, x_2, \ldots)$. Let X_∞ be the space of all such outcomes.

If we wish to be able to make such statements as "the 'sample mean' of n observations converges to some number α as $n \to \infty$", we will need to construct sets which can only be obtained in a countable and not a finite number of operations. That is, to construct the set of all sequences in X_∞ whose sample mean converges to α is not possible in a finite number of operations.

[19] In fact this map must satisfy a regularity condition, that of measurability, which we will shortly define but for the purpose of the example we will ignore this requirement.

σ algebra

Once again, we assume we are interested in collections of subsets of some set X which have X itself as a member. We can thus define:

Definition 6

A collection $\mathcal{C}$ of subsets of a set X is called a *σ algebra*, if

(i) $\varnothing \in \mathcal{C}$,
(ii) $A \in \mathcal{C}$ then $A^c \in \mathcal{C}$,
(iii) $A_1, A_2, \ldots \in \mathcal{C}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.

Now we may extend out definition of an additive set function to the following:

Definition 7

A set function $\mu : \mathcal{C} \to R^*$ is *σ additive*, if

(i) $\mu(\varnothing) = 0$,
(ii) for any sequence $E_1, E_2, \ldots$ of sets of $\mathcal{C}$, where

$$E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{C},$$

then

$$\mu(E) = \sum_{i=1}^{\infty} \mu(E_i).$$

Obviously any σ additive set function is also additive but the converse does not hold. Consider Example 4 given earlier. Let in that example

$$E = (0,1] \quad \text{and} \quad E_n = \left(\frac{1}{n+1}, \frac{1}{n}\right], \qquad n = 1, 2, \ldots$$

Now the sequence (E_n) has each of its elements in A, and E itself is in $\mathcal{C}$, but clearly

$$\mu(E) = +\infty \quad \text{and} \quad \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1.$$

Having discussed various types of set functions we can now restrict the class of such functions to those which have particular interest for us; that is, those to which we will refer as measures.

Consider a set X and a σ algebra $\mathcal{Q}$ of its subsets. We will refer to the couple $(X, \mathcal{Q})$ as a *measurable space*.

Definition 8[20]

If $(X, \mathcal{Q})$ is a measurable space then any function $:\mathcal{Q} \to R_+$ (where $R_+ = \{x| x \in R^*$ and $x \geqslant 0)$ which is σ additive is called a *measure*.

Definition 9

For $(X, \mathcal{Q})$ a measurable space, if μ is a measure on $\mathcal{Q}$ and $\mu(X) = 1$, then μ is called a *probability space*.

It will, in general, be enough for us to restrict our attention to probability measures but it is useful to have the more general definition.

The reader should now return to the examples given and check which are measures.

Before proceeding we need a further definition.

Definition 10

A set function $\mu: \mathcal{Q} \to R^*$ is called *σ finite* if for each $E \in \mathcal{Q}$ there exists a sequence of sets E_i $(i = 1,2,\ldots)$ with $E_i \in \mathcal{Q}$ such that $E \subset \bigcup_{i=1}^{\infty} E_i$ and $\mu(E_i) < \infty$ for all i.

We will now show that starting with a measure on an algebra $\mathcal{Q}$ we can extend it uniquely to a measure on the σ algebra generated by $\mathcal{Q}$.

If we start with a measure μ defined on an algebra $\mathcal{Q}$ of subsets of a set X consider the function defined by

$$\mu^*(E) = \inf \sum_{i=1}^{\infty} \mu(F_i),$$

where the infimum is taken over all sequences of sets (F_i) such that $E \subset \bigcup_{i=1}^{\infty} F_i$. We can now state the following:

Theorem 2

Let $\mathcal{Q}$ be an algebra of subsets of a set X and $\mu: \mathcal{Q} \to R^+$ a measure on $\mathcal{Q}$. Then

[20] It is worth noting that, although we start here with the natural domain of definition, a σ algebra, there is no need to do so and we could have started with some other class of subsets. In addition, we have required that a measure be positive, a restriction not generally imposed.

there is an extension of μ to a measure r where $r: S(\mathcal{Q}) \rightarrow R^+$ and $S(\mathcal{Q})$ is the σ algebra generated by $\mathcal{Q}$. The extension is unique and σ finite on $S(\mathcal{Q})$ if μ is σ finite. r is the restriction of μ^* to $S(\mathcal{Q})$ where μ^* is defined as above by

$$\mu^*(E) = \inf \sum_{i=1}^{\infty} \mu(F_i).$$

We have then arrived at the point where given some arbitrary set X and a measure defined on a simple structure (an algebra) of its subsets we can extend this measure uniquely to the σ algebra generated by that algebra.

We can now give the following:

Definition 11

A *measure space* $(X, \mathcal{Q}, \mu)$ is a triple where X is a set, $\mathcal{Q}$ is a σ algebra of subsets of X, and μ a measure defined on $\mathcal{Q}$.

An example: Lebesgue measure

In Euclidean space the notion of measure corresponds intuitively to the ideas of length, area, or volume depending upon the dimension in question. How is the measure of a set defined in this case? In R we consider the class $\mathcal{P}$ of half open intervals $(a, b]$; these generate the class $\mathcal{R}$ of all elementary figures, i.e., sets of the form

$$E = \bigcup_{i=1}^{n} (a_i, b_i] \quad \text{with} \quad b_i < a_{i+1}, i = 1,2,\ldots, n-1.$$

In other words, the elementary figures consist of all sets which are expressible as a finite union of disjoint sets of $\mathcal{P}$.

In R^ℓ the half open intervals are given by

$$\{(x_1, x_2, \ldots, x_\ell)\}, \quad a_i < x_i \leqslant b_i, \qquad i = 1,2,\ldots,\ell;$$

these generate analogously the elementary figures $\mathcal{R}^\ell$.

Define now the natural set function to express length, i.e.,

$$\mu(a, b] = b - a,$$

or area or volume,

$$\mu\{(x_1, \ldots, x_\ell): a_i < x \leqslant b_i, \quad i = 1,2,\ldots,\ell\} = \prod_{i=1}^{\ell} (b_i - a_i).$$

Such a μ is a measure, and it can be uniquely extended by our previous results to the σ algebra generated by the elementary figures, i.e., the Borel sets $\mathcal{B}^\ell$, and is referred to as *Lebesgue measure*.

In fact, the class of Lebesgue measurable sets $\mathcal{L}^\ell$ is larger than $\mathcal{B}^\ell$, but this is not of great importance for the present discussion.

The interpretation of measures in probabilistic terms is clear and the measure of a subset is the probability that the outcome of some "experiment" will fall in that subset. In economics, we often need to formalise the idea that people forecast future prices. Such a forecast by an individual of n prices would be given by a measure on the unit simplex R^n. A discussion of such forecasts and requirements imposed on them is to be found in Chapter 19 on temporary general equilibrium theory.

A wholly different approach to the use of measure theory in economics, as has been mentioned, is the idea of representing a purely competitive economy by a continuum of agents or more generally by a measure space. Having discussed the nature of measure space and the nature of measures at length, we can now look at the economic interpretation given to them.

A *measure space of economic agents* $(A, \mathcal{Q}, \mu)$ can be viewed as follows: A is the set of individual names or labels. For example, in a finite economy A could be a set of integers, while in a continuum economy we could, as Aumann (1964) did, use the closed unit interval $[0,1]$ as the underlying set of labels.

$\mathcal{Q}$ is the σ algebra of subsets of A which can be thought of as corresponding to the possible coalitions of A. As we have observed, we cannot in general define the measure on all the subsets of A, but the reader can consider for practical purposes all subsets as possible coalitions. In other words those coalitions which are eliminated by confining our attention to a σ algebra, rather than all subsets, are of no special interest. In probability theory, for example, we may well be interested in the probability that a number drawn from some set falls into a certain interval or collection of intervals, but we are probably less interested in the probability of its being rational or irrational. Even this is manageable, but there are sets to which we cannot assign probabilities. However, such sets are not constructed as collections of intervals, and it is in these that our interest generally lies. μ the measure on A simply conveys the idea of the proportion of agents belonging to any subset. Thus we will have for a finite economy a measure space given by the following:

Definition 12

A measure space $(A, \mathcal{Q}, \mu)$ is *simple* if A is finite, $\mathcal{Q}$ is the set of all subsets of A and $\mu(E) = \#E/\#A$.

We should also note in our discussion of the continuum economy as a representation of a perfectly competitive situation we suggested that each individual had no weight. If we use Lebesgue measure on the unit interval this is clearly the case but in general we need to make an assumption that the measure space satisfies the following:

Definition 13

A measure space $(A, \mathscr{Q}, \mu)$ is *atomless* if for every $E \in \mathscr{Q}$ with $\mu(E) > 0$ there exists $F \in \mathscr{Q}$ and $F \subset E$ such that $\mu(E) > \mu(F) > 0$.

This rules out some individual having positive "weight" or "influence". Indeed the idea of a measure space with atoms had already been used to designate situations which are not perfectly competitive that is to convey the idea of monopoly. See e.g. Shitovitz (1974).

An alternative use of the notion of an an atom would be as mentioned earlier when we wish to define the notion of what Arrow described as a "dictator" in social choice with an arbitrarily large number of members.

That is if a^* an agent is "decisive" for A in that his preferences determine those of the society as a whole then we define the Dirac measure as follows:

$$\begin{aligned} \mu(E) &= 1 \quad \text{if} \quad a^* \in E, \\ &= 0 \quad \text{otherwise,} \end{aligned} \qquad \forall E \in \mathscr{P}(A).$$

Clearly μ defines a measure and a^* may be thought of as a dictator in the Arrovian sense, if we make the rule that if for some coalition E, $x \succ_a y$, $\forall a \in E$, then x is socially preferred to y if $\mu(E) = 1$.

Incidentally, one can see from the above example that in general it is the requirement that the measure be atomless which prevents us from defining it on all subsets of A. In the example it is clear that μ is a measure defined on all subsets of A.

The notion of a measure on a set allows us, as we mentioned in the introduction, to make statements about which subsets are of no importance, that is, which are "negligible".

If $(A, \mathscr{Q}, \mu)$ is a measure space then a set $B \subset A$ is said to be *negligible* (for μ) if there exists a set $E \in \mathscr{Q}$ such that $\mu(E) = 0$, and $B \subset E$.

If a certain property holds for all points of A except for a set B where $\mu(B) = 0$ then we say that that property holds "*almost everywhere*". In economics such a description is useful as a way of characterising particular phenomena as exceptional or rare.

If μ is a probability measure then the term "*almost surely*" replaces almost everywhere.

Situations in economics which occur only for configurations of parameters which together have measure zero in the space of all such configurations may be thought of as "unlikely" or "rare". This is a useful idea which enables us to avoid making strong or unnatural assumptions to rule out cases which are "exceptional" and which enables us to give a precise interpretation of the word "exceptional".

Liapunov's theorem

We now give a result of considerable importance in applying measure theory to economics and one which has played a central role in the formalisation of, and the equivalence between, solution concepts for large economies.

Theorem 3 (Liapunov)[21]

Let $\mu_1,\dots,\mu_m$ be atomless measures on $(A,\mathcal{Q})$, then the set

$$\{(\mu_1(E),\mu_2(E),\dots,\mu_m(E))\in R^m,\quad E\in\mathcal{Q}\}$$

is a closed and convex subset of R^m.

This theorem is particularly useful since when considering a continuum economy we can always find a "scaled down" version of that economy and the reader will find a discussion in Chapter 18 by Hildenbrand. We also note in passing that this theorem is fundamental in the article by Dubins and Şpanier (1961) to which we referred earlier.

Measurable mappings

In economics we will frequently be concerned with mappings from one measurable space to another. Indeed, when defining an exchange economy, for example, we will be concerned with identifying with each agent his endowments and preferences. We will need a certain regularity property of such a mapping, in particular that the pre-image of every set in the σ algebra of the range shall be a set in the σ algebra of the domain. This is inconvenient but necessary for technical reasons.

Definition 14

For two measurable spaces $(A_1,\mathcal{Q}_1)$ and $(A_2,\mathcal{Q}_2)$ a mapping $f\colon A_1\to A_2$ is *measurable* if $f^{-1}(E)=\{a\in A_1\,|\,f(a)\in E\}\in\mathcal{Q}_1$ for each $E\in\mathcal{Q}_2$.

[21]A proof is given in Lindenstrauss (1969).

Note that the measurability of a function depends upon the σ algebras, and thus for the same underlying sets A_1 and A_2 changing the σ algebra associated with each can change whether a function is measurable or not. When A_1 and A_2 are metric spaces we will generally take $\mathcal{Q}_1$ and $\mathcal{Q}_2$ to be the respective Borel σ algebras.

It would seem at first sight that it might be difficult to determine whether a given function is, in fact, measurable, but in fact it is sufficient to check for any class of subsets of A_2 which generates $\mathcal{Q}_2$. More formally, we have:

Remark

If for a class $\mathcal{C}$ of subsets of A_2 which generates $\mathcal{Q}_2$, and a mapping f from a measurable space $(A_1, \mathcal{Q}_1)$ into a measurable space $(A_2, \mathcal{Q}_2)$, $f^{-1}(C) \in \mathcal{Q}_1$, for every $C \in \mathcal{C}$; then f is measurable.

It is also important to note that composing two measurable mappings preserves measurability. Thus we have:

Proposition 1

Let f and g be two measurable mappings from $(A_1, \mathcal{Q}_1)$ to $(A_2, \mathcal{Q}_2)$ and from $(A_2, \mathcal{Q}_2)$ to $(A_3, \mathcal{Q}_3)$, respectively, then the composition $g \circ f$ is a measurable mapping.

In addition, the following result is frequently useful:

Proposition 2

Let g be a measurable mapping from a measurable space $(A_1, \mathcal{Q}_1)$ into a measurable space $(A_2, \mathcal{Q}_2)$ and f a function from A into R^m, then f is measurable with respect to the σ algebra $g^{-1}(\mathcal{Q}_2)$ if and only if there exists a measurable function h of $(A_2, \mathcal{Q}_2)$ into R^m such that $f = h \circ g$.

Real-valued measurable functions

In particular if we consider a mapping f from a measurable space $(A, \mathcal{Q})$ to the extended real line R^*, then any of the following conditions are necessary and sufficient for f to be measurable:

(i) $\{x \mid f(x) \leqslant c) \in \mathcal{Q}$ for all $c \in R$,
(ii) $\{x \mid f(x) > c\} \in \mathcal{Q}$ for all $c \in R$,
(iii) $\{x \mid f(x) < c\} \in \mathcal{Q}$ for all $c \in R$,
(iv) $\{x \mid f(x) \geqslant c\} \in \mathcal{Q}$ for all $c \in R$.

Other useful properties of real-valued or extended real-valued measurable functions are given by the following:

Proposition 3

If $(A, \mathcal{Q})$ is a measurable space and f and g two measurable functions into R. (resp. into R^*), then the functions

(i) $f+g$. (resp. $f+g$ if the function if defined),
(ii) $\sup(f, g)$,
(iii) $\inf(f, g)$,
(iv) $f.g$,
(v) αf, $\forall \alpha \in R$,

are measurable.

Examples and further properties

Consider now a generalisation of the special mapping mentioned earlier often referred to as the "indicator variable" that is $\mathcal{X}_c: A \in R$ such that

$$\begin{aligned} \mathcal{X}_c &= 1 \quad \text{if} \quad a \in C, \\ &= 0 \quad \text{if} \quad a \notin C, \end{aligned} \quad \text{for every} \quad C \in \mathcal{Q},$$

then the mapping is measurable.

If we wish to confine our attention to a restricted class of a σ algebra then it is useful to know that, if $(A, \mathcal{Q})$ is a measurable space and $\mathcal{Q}'$ a sub σ algebra of $\mathcal{Q}$ then the *identity map*,

$$\text{id}.(A, \mathcal{Q}) \to (A, \mathcal{Q}') \quad \text{where} \quad \text{id}.(a) = a,$$

is measurable.

When we consider functions from a metric space M into R^* it is important to observe that:

Proposition 4

Every lower or upper semi-continuous function from a metric space M into R^* (and thus in particular every continuous function) is measurable.

Finally we give a result which will be used in the next section:

Proposition 5

Let the sequence $(f_n)(A, \mathcal{Q})$ into R be such that:

(i) f_i is measurable $(i = 1, 2, \ldots)$.

Then (a) the functions $\sup_n f_n$ and $\inf_n f_n$ are measurable, and (b) the functions $\limsup_n f_n$ and $\liminf_n f_n$ are measurable.

Furthermore if the following condition is also satisfied:

(ii) $\lim f_n(a)$ exists for every $a \in A$.

Then the function g defined by $g(a) = \lim f_n(a)$ is measurable.

Note that we cannot extend these results to include non-countable operations. To see this consider the following:

Example 5

Let A be a subset of $[0,1]$ which is not Lebesgue measurable. Let

$$f_\alpha(x) = 1 \quad \text{if} \quad x = \alpha,$$

$$f_n(x) = 0 \quad \text{if} \quad x \neq \alpha.$$

For each $\alpha \in A$ the function f is clearly measurable, but

$$\chi_A(x) = \sup_{\alpha \in A} f_\alpha$$

is obviously not Lebesgue measurable.

This creates particular problems, for example, when considering stochastic processes with a continuous time parameter.

Integration

The idea of the integral of a function plays a very important role whether we are considering the probabilistic aspect of measure theory of whether we are considering the application of measure theory directly to "idealised", "perfectly competitive" or "limit" economies. In the former case the reader will be aware that the integral gives the "mean" or "expectation" of a given function f with respect to a particular probability distribution. In this case the function is a "random variable with distribution μ" and the integral gives the familiar idea of the expected value of the random variable.

Recall that in economies with a measure space of agents we are faced with a simple definitional problem. How, with an infinite number of agents each possessing a positive bundle of goods, can we talk of an equilibrium in which the demand for these goods equals the supply of them? Since the sum is of no interest, the appropriate notion is that average, or "per capita", demand equals supply. Here again the integral will be the appropriate concept. The integral $I(f)$ will be a real number associated with a particular function f and we will require that for suitable functions f the operator $I(f)$ should satisfy certain properties. Let $\mathfrak{F}$ be a class of functions $f: A \to R^*$ and let $I: \mathfrak{F} \to R$ define a real number for each $f \in \mathfrak{F}$, then the following properties would seem intuitively, to be required of I, particularly if one thinks of the interpretation of "the area under a curve" as the integral of a function from R into R.

(i) If for all $f \in \mathfrak{F}$ we have $f(a) \geqslant 0$ for all $a \in A$, then we have $I(f) \geqslant 0$; that is, I preserves non-negativity.
(ii) For f and $g \in \mathfrak{F}$ and α and $\beta \in R$, it holds that $\alpha f + \beta g \in \mathfrak{F}$ and $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$; in other words I is linear on $\mathfrak{F}$.
(iii) I is continuous on $\mathfrak{F}$, in that, if (f_n) is an increasing sequence of functions in $\mathfrak{F}$ and

$$f_n(a) \to f(a) \quad \text{for all} \quad a \in A,$$

then $f \in \mathfrak{F}$ and $\lim_{n\to\infty} I(f_n) = I(f)$.[22]

Our procedure for obtaining an integral which satisfies these three conditions is, first, to restrict our attention to a particular class of functions for which the integral has an obvious intuitive definition, and then to extend this class of functions to as large a class as possible.

To do this we need first the idea of a "simple function" from a set A to R which is one which takes on a finite number of values, one for each set of a *partition* of A, and is constant on each set of the *partition*. The idea is illustrated in Figure 2.1 for a function from [0, 1] into R.

Definition 15

A finite collection of sets $E_1, \ldots, E_n$ such that

$$E_i \cap E_j = \varnothing, \qquad i = 1, \ldots, n, \quad j = 1, \ldots, n,$$

and

$$\bigcup_{i=1}^{n} E_i = A,$$

is said to form a *finite partition* of A. In particular, if $E_i \in \mathfrak{A}$ $(i = 1, \ldots, n)$ then

[22] The Riemann integral with which the reader will be familiar from the integral calculus does not satisfy this property but does satisfy the following weakened version of it:

(iii*) Let (f_n) be a monotone decreasing sequence of functions with $\lim_{n\to\infty} f_n(a) = 0$ for all $a \in A$, then $\lim_{n\to\infty} I(f_n) = 0$.

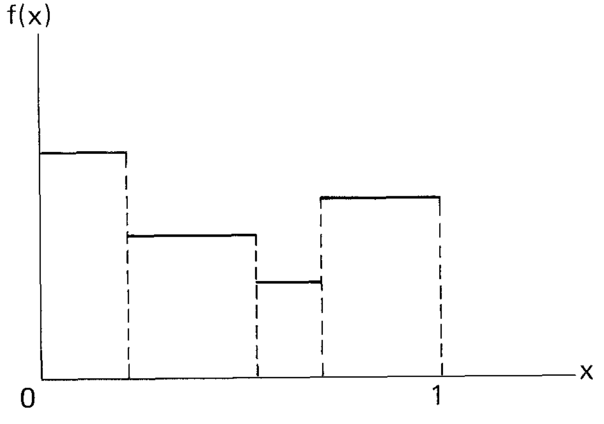

Figure 2.1

these sets form an $\mathscr{A}$ *partition* of A. With this we can proceed to the following:

Definition 16

A function $f: A \to R$ is called $\mathscr{A}$ *simple* if it can be expressed as $f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$, where $E_1, E_2, \ldots, E_n$ form an $\mathscr{A}$ partition of A and $c_i \in R$ ($i = 1, 2, \ldots, n$).

Remark

If f and g are two simple functions from A to R then the functions

$$f+g, \quad f-g, \quad fg,$$

are also simple functions.

Note also that an $\mathscr{A}$ simple function is $\mathscr{A}$ measurable.[23] Now we can start to extend our attention to measurable functions by considering the following:

Theorem 4

If a function $f: A \to R_+$ is measurable then it is the limit of a monotone increasing sequence of non-negative simple functions.

Now to move towards the desired results, we must show that any measurable function is the limit of a sequence of simple functions.

[23] We will frequently speak of measurable functions rather than $\mathscr{A}$ measurable functions when only one σ algebra is under consideration.

First for a function $f: A \to R^*$ define

$$f_+(x) = \max[0, f(x)], \qquad f_-(x) = -\min[0, f(x)],$$

Clearly,

$$f(x) = f_+(x) - f_-(x).$$

Now from a previous remark, if f is measurable so are f_+ and f_-, and since both are non-negative each is the limit of a sequence of non-negative simple functions. Applying the remark again we then have the following important theorem which provides the basis for the definition of the integral:

Theorem 5

Any measurable function $f: A \to R^*$ is the limit of a sequence of simple functions.

This link between simple and measurable functions enables us to proceed to the definition of the integral for simple functions and to extend it to measurable functions.

Thinking of measure on a set A as the distribution of mass in physical terms or as a probability distribution over a set of outcomes, it is clear that the natural notion of the integral for the particularly convenient case of a non-negative simple function is given by:

Definition 17

Given a measure space $(A, \mathscr{Q}, \mu)$ and a non-negative simple function,

$$f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x) \quad \text{with} \quad c_i \geqslant 0, \qquad i = 1, \ldots, n,$$

(with respect to μ), the *integral* $\int f \mathrm{d}\mu$ is defined by

$$\int f \mathrm{d}\mu = \sum_{i=1}^{n} c_i \mu(E_i).$$

Referring back to Figure 2.1, it is clear that the integral of such a function consists of the sum of the area of the rectangles under each step of the function. This sum is always defined since the individual terms are non-negative,[24] and it is independent of which of the possible representations of f is chosen.

[24]If we are treating general measures it is possible that $\mu(E_i) = \infty$ and $c_i = 0$; in this case we take $\mu(E_i)c_i = 0$.

Remark

It is easily shown that the integral is linear on the class of non-negative simple functions S_+, that is, if $f, g \in S_+$ and $\alpha, \beta \geqslant 0$, then

$$\int (\alpha f + \beta g)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu + \beta \int g\,\mathrm{d}\mu.$$

Furthermore the integral is order preserving on the same class, i.e., if $f, g \in S_+$ and $f \geqslant g$, then

$$\int f\,\mathrm{d}\mu \geqslant \int g\,\mathrm{d}\mu.$$

Now we can proceed to the second step—that of extending the definition of the integral to the class of non-negative measurable functions M_+.

For f in M_+ there exists by Theorem 6 a monotone increasing sequence (f_n) of simple functions with $f_n \to f$. Now for each f_n in the sequence $\int f_n\,\mathrm{d}\mu$ is defined, and by our previous observations the sequence $(\int f_n\,\mathrm{d}\mu)$ is monotone increasing and has a limit.[25] Hence we define for $f \in M_+$,

$$\int f\,\mathrm{d}\mu = \lim_{n \to \infty} \int f_n\,\mathrm{d}\mu.$$

Clearly the monotone sequence (f_n) which converges to a given f is not unique, but the integral, as defined, is independent of the choice of sequence. Note that it follows directly from our earlier observation for functions in the class S_+ that the integral operator is linear on the class M_+, i.e., for $f, g \in M_+$ and $\alpha, \beta \geqslant 0$,

$$\int (\alpha f + \beta g)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu + \beta \int g\,\mathrm{d}\mu.$$

Definition 18

A non-negative measurable function f is said to be *integrable* if

$$\int f\,\mathrm{d}\mu \quad \text{is finite.}$$

Thus far we have been concerned with measurable functions in M_+. We now extend our definition of the integral to the class of integrable measurable functions.

[25]The limit may, of course, be $+\infty$.

First, observe again that if $f\colon A\to R^*$ is measurable, then so are f_+ and f_-. In particular if the two non-negative measurable functions f_+ and f_- are integrable that we say that f is integrable. More precisely, we have the following:

Definition 19

If $f\colon A\to R^*$ is such that f_+ and f_- are integrable, then f *is called integrable* and the integral of f with respect to μ is given by

$$\int f_+\,\mathrm{d}\mu-\int f_-\,\mathrm{d}\mu.$$

Often we will be concerned with the integral of a function f over only a subset $E\in\mathcal{A}$, and in this case we define

$$\int_E f\mathrm{d}\mu=\int f\cdot\chi_E\mathrm{d}\mu,$$

provided that $f\cdot\chi_E$ is defined. Then there are two conditions each of which will ensure the integrability of a function over a given set. Either:

(i) $f\cdot\chi_E$ is non-negative and measurable, or
(ii) $f\cdot\chi_E$ is measurable and integrable.

f is then integrable over A if $f\cdot\chi_A$ is integrable. We denote the set of all integrable functions from $(A,\mathcal{A},\mu)$ into R^* by $\mathcal{L}(A,\mathcal{A},\mu)$.

Confirmation of the properties we demanded of the integral at the outset is given by the following:

Theorem 6

If $(A,\mathcal{A},\mu)$ is a measure space, E,F are two disjoint sets in $\mathcal{A}$, and f,g are two functions belonging to $\mathcal{L}(A,\mathcal{A},\mu)$, then

(i) f,g are integrable over E and F;
(ii) $f+g,|f|,|g|$ belong to $\mathcal{L}(A,\mathcal{A},\mu)$;
(iii) $\int_{E\cup F}f\mathrm{d}\mu=\int_E f\mathrm{d}\mu+\int_F f\mathrm{d}\mu$;
(iv) f,g are finite μ a.e.;
(v) $\int(f+g)\mathrm{d}\mu=\int f\mathrm{d}\mu+\int g\mathrm{d}\mu$;
(vi) $|\int f\mathrm{d}\mu|\leqslant\int|f|\mathrm{d}\mu$;
(vii) for $c\in R$, $c\cdot f$ is μ integrable and $c\int f\mathrm{d}\mu=\int cf\mathrm{d}\mu$;
(viii) $f\geqslant 0\Rightarrow\int f\mathrm{d}\mu\geqslant 0: f\geqslant g\Rightarrow\int f\mathrm{d}\mu\geqslant\int g\mathrm{d}\mu$;
(ix) if $f\geqslant 0$ and $\int f\mathrm{d}\mu=0$, then $f=0$ μ a.e.;
(x) $f=g\mu$ a.e.$\Rightarrow\int f\mathrm{d}\mu=\int g\mathrm{d}\mu$;
(xi) If $h: A\to R^*$ is A measurable and $|h|\leqslant f$ then $h\in\mathcal{L}(A,\mathcal{A},\mu)$.

From these results follows:

Corollary to Theorem 6

If a function $f: A \to R^*$ is bounded, $\mathcal{A}$ measurable and if $f(x)=0$ when $x \notin E$ for some $E \in \mathcal{A}$ with $\mu(E) < \infty$ then f is μ integrable.

As we will see in what follows, an exchange economy will be defined by a measurable mapping from the underlying measure space of agents to the space of agents' characteristics. In other words, defining an economy consists of specifying for each agent his preferences and his initial endowments. Now for many of the results in Chapter 18 of this Handbook, it will be important to show that properties of very large economies — that is economies with a measure space of agents — are, in some sense, also true for large finite economies. To do this we will need to consider sequences of economies and sequences of allocations, i.e., sequences of mappings from the space of agents to R^{ℓ}_{+}. The following three results will prove to be particularly useful, and we will later investigate in more detail different notions of convergence of measurable functions.

Proposition 6

If the sequence (f_n) in $\mathcal{L}(A, \mathcal{A}, \mu)$ is increasing (decreasing), $\lim f_n(a)$ is finite for every $a \in A$, and if $\lim_n f_n$ is finite, then

$$\lim_n f_n \in \mathcal{L}(A, \mathcal{A}, \mu) \quad \text{and} \quad \lim_n \int f_n = \int \lim_n f_n.$$

Lemma 1 (Fatou)

If (f_n) *is a sequence in* $\mathcal{L}(A, \mathcal{A}, \mu)$ *and if* $f_n \leqslant g$ *where* $g \in \mathcal{L}(A, \mathcal{A}, \mu)$ *then*

$$\int \limsup_n f_n \geqslant \limsup_n \int f_n.$$

Furthermore, if $h \leqslant f_n$ where $h \in \mathcal{L}(A, \mathcal{A}, \mu)$, then

$$\int \liminf_n f_n \leqslant \liminf_n \int f_n.$$

Theorem 7 (Lebesgue)

If (f_n) is a sequence in $\mathcal{L}(A, \mathcal{A}, \mu)$, and if $\lim_n f_n(a)$ exists for every $a \in A$ and $|f_n| \leqslant g$ $(n = 1, 2, \ldots)$, where $g \in \mathcal{L}(A, \mathcal{A}, \mu)$, then

$$\lim_n f_n \in \mathcal{L}(A, \mathcal{A}, \mu) \quad \text{and} \quad \int \lim_n f_n = \lim_n \int f_n.$$

2.4. Product spaces and product measures

Before proceeding to our discussion of convergence of measurable functions we will need to discuss the idea of product spaces and product measures. To see why we need these notions consider again the example of the exchange economy mentioned earlier. It will be defined by a mapping from the measure space of agents to the space of agents characteristics. The latter, however, is the product of two spaces, the space of preferences $\mathcal{P}$ and the space of initial endowments R_+^ℓ. Now the natural procedure is to use the structure of each space to define the product space and product measure since, in particular, this allows us to use the natural idea of "projection", for example, when we wish to concentrate on the distribution of initial endowments. Recall that the Cartesian product $A\times B$ of two spaces A and B is the set of all ordered pairs (a, b) where $a\in A$ and $b\in B$.

Definition 20

A set in $A\times B$ of the form $E\times F$ with $E\subset A$ and $F\subset B$ is called a *rectangle.*

Definition 21

Let $\mathcal{A}$ and $\mathcal{B}$ be classes of subsets in A and B, respectively, then $\mathcal{A}\times\mathcal{B}$ denotes the class of all rectangles $E\times F$ with $E\in\mathcal{A}$ and $F\in\mathcal{B}$, i.e., *the product of the classes* $\mathcal{A}$ *and* $\mathcal{B}$.

Definition 22

Let $\mathcal{A}$ and $\mathcal{B}$ be algebras (resp. σ algebras) in A and B, respectively, then the *product algebra* (resp. σ algebra) is the algebra (resp. σ algebra) generated by $\mathcal{A}\times\mathcal{B}$, and is denoted $\mathcal{A}\otimes\mathcal{B}$.

It is important to note that if $\mathcal{A}$ and $\mathcal{B}$ are σ algebras, $\mathcal{A}\times\mathcal{B}$ will not be a σ algebra.

As we mentioned above we will sometimes be interested in restricting our attention to one of the components of the product space and for this we need two definitions:

Definition 23

For any set $E\subset A\times B$ and any point $a\in A$, the set

$$E_a=\{b|(a,b)\in E\}$$

is called the *section* of E at a. Similarly for any $b\in B$ the subset $E^b=\{a|(a,b)\in E\}$ is called the *section* of E at b.

Definition 24

For any set $E \subset A \times B$ the sets $\{x \mid \text{there exists } y \text{ with } (x,y) \in E\}$ and $\{y \mid \text{there exists } x \text{ with } (x,y) \in E\}$ are called the *projections* of E into the respective spaces A and B, and are denoted $\text{proj}_A E$ and $\text{proj}_B E$.

It is again important to note that $E \in \mathcal{A} \otimes \mathcal{B}$ does not mean that $\text{proj}_A E \in \mathcal{A}$.

Now with these definitions we may proceed to consider product measures. Consider two measure spaces $(A_1, \mathcal{A}_1, \mu_1)$ and $(A_2, \mathcal{A}_2, \mu_2)$ where μ_1 and μ_2 are both σ finite. Define for any rectangle set $E_1 \times E_2$,

$$\mu(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2).$$

It is easy to show that μ is finitely additive on $\mathcal{A}_1 \times \mathcal{A}_2$; indeed μ is a measure on the semi algebra $\mathcal{A}_1 \times \mathcal{A}_2$ which can be extended uniquely to the generated algebra, and thence of course, by previous results, to the generated σ algebra which is $\mathcal{A}_1 \otimes \mathcal{A}_2$. The resulting λ^* is called the *product measure* on $\mathcal{A}_1 \otimes \mathcal{A}_2$. This may be summarised in the following:

Theorem 8

Given two measure spaces $(A_1, \mathcal{A}_1, \mu_1)$, $(A_2, \mathcal{A}_2, \mu_2)$ such that μ_1 and μ_2 are σ finite, there is an unique measure λ defined on the product σ algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ on $A_1 \times A_2$ such that

$$\lambda(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2) \quad \text{for} \quad E_1 \in \mathcal{A}_1 \quad \text{and} \quad E_2 \in \mathcal{A}_2.$$

Definition 25

If μ is a measure on the product space $A \times B$, then the *marginal distribution* of μ on A is given by

$$\mu(C) = \mu(C \times B) \quad \text{for } C \text{ a subset of } A.$$

Although this result extends immediately to any finite Cartesian product of σ finite measure spaces, more care has to be taken in defining product measures for countable products of measure since one has to make sure that the infinite products of real numbers involved converge. The problem may be avoided by sticking to probability measure spaces where the natural generalisation holds.

If we have two measure spaces $(A_1, \mathcal{A}_1, \mu_1)$ and $(A_2, \mathcal{A}_2, \mu_2)$, we can in fact show quite simply the link between the integral of a function f from $A_1 \times A_2 \to R^*$ with respect to the product measure and a two-step procedure including integrating with respect to μ_1 and μ_2. The idea is clear: we fix $x \in A_1$ then integrate f with respect to μ_2. The resulting function from A_1 to R^* is then integrated with respect to μ_1, and the result turns out, with certain restrictions, to be equivalent to having integrated with respect to the product measure λ directly.

We first state a result which shows how to obtain the product measure λ by integrating the measure of the section for each fixed x over all x with respect to μ_1.

Theorem 9

For two σ finite measure spaces $(A_1, \mathcal{A}_1, \mu_1)$, $(A_2, \mathcal{A}_2, \mu_2)$ define the product measure λ on the σ algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$. Then for each $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $\mu_2(E_x)$ is $\mathcal{A}_1$ measurable and $\mu_1(E_y)$ is $\mathcal{A}_2$ measurable, and

$$\lambda(E) = \int \mu_1(E_y)\,\mathrm{d}\mu_2 = \int \mu_2(E_x)\,\mathrm{d}\mu_1.$$

Incidentally, it now follows from our previous discussion that we have:

Corollary to Theorem 9

For $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $\lambda(E) = 0$ if and only if $\mu_2(E_x) = 0$ for almost all x, and if and only if $\mu_1(E_y) = 0$ for almost all y.

We now state the following important:

Theorem 10

Under the conditions of Theorem 11 denote by $\mathcal{A}$ the σ algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$. Then if h is any $\mathcal{A}$ measurable function from $A_1 \times A_2 \to R^+$, then

$$\int h\,\mathrm{d}\lambda = \int \left(\int h_x\,\mathrm{d}\mu_2 \right) \mathrm{d}\mu_1 = \int \left(\int h_y\,\mathrm{d}\mu_1 \right) \mathrm{d}\mu_2.$$

We now move on to develop the idea of a derivative of a set function, but to do so we will need two definitions:

Definition 26

Given a measure space $(A, \mathcal{A}, \mu)$, the function $v: \mathcal{A} \to \mathfrak{R}^*$ is *absolutely continuous* with respect to μ, if for any $E \in \mathcal{A}$, $\mu(E) = 0$ implies $v(E) = 0$.

In particular, if $f: A \to R^*$ is μ integrable, then

$$v(E) = \int_E f\,\mathrm{d}\mu \quad \text{for} \quad E \in \mathcal{A}$$

is a finite valued absolutely continuous set function.

In order to define our derivative, we will take a general σ additive set function and decompose it into an absolutely continuous part and a remainder which is concentrated on a set which is μ null. To make this last remark more precise, we give the following:

Definition 27

For a measure space $(A, \mathscr{A}, \mu)$, a set function $v: \mathscr{A} \to R^*$ is *singular* with respect to μ, if there exists a set $E_0 \in \mathscr{A}$ with $\mu(E_0)=0$ and

$$v(E)=v(E \cap E_0) \quad \text{for all } E.$$

We can now give the important:

Theorem 11

For a σ finite measure space $(A, \mathscr{A}, \mu)$ and a σ additive, σ finite set function $v: \mathscr{A} \to R^*$ there is a unique decomposition

$$v=v_1+v_2,$$

where v_1 and v_2 are σ additive and σ finite, such that v_1 is singular with respect to μ and $v_2 < \mu$.

In addition, there is a finite valued measurable $f: A \to R$ such that

$$v_2(E)=\int_E f \,\mathrm{d}\mu \quad \text{for all} \quad E \in \mathscr{A};$$

f is unique in that if there is a function g such that

$$v_2(E)=\int_E g \,\mathrm{d}\mu \quad \text{for all} \quad E \in \mathscr{A},$$

then $f=g$ except on a set of zero measure.

This last observation is important, for it means that when we define the derivative of a set function this is not defined uniquely at any given point but as a function must coincide with any other function representing the same derivative except on a set of measure zero. With this in mind we give the following:

Definition 28

For a σ finite measure space $(A, \mathscr{A}, \mu)$, if $v(E)=\int_E f \mathrm{d}\mu$ for all $E \in \mathscr{A}$, f is called the *Radon–Nikodyn derivative* of v with respect to μ and is denoted $\mathrm{d}v/\mathrm{d}\mu$.

2.5. *Convergence of measurable functions*

As we have said before it will be important for later economic applications such as those found in Chapter 18 of this Handbook, to study the convergence of measurable functions. Several different types of convergence can be defined, and we will always be considering a sequence (f_n) of measurable functions from a measure space $(A, \mathscr{A}, \mu)$ to R^*.

Definition 29

If (f_n) is a sequence of measurable functions from $(A, \mathscr{A}, \mu)$ to R^*, (f_n) is said to *converge point-wise* to a measurable function f on E if for every $x \in E$ $\lim_{n\to\infty} f_n(x) = f(x)$. If $\mu(E) = \mu(A)$ then we say (f_n) *converges to* f *almost everywhere* (a.e.).

Furthermore if (f_n) and f are finite-valued, then we add the following:

Definition 30

If a sequence (f_n) and f are finite-valued functions from E to R then (f_n) is said to *converge uniformly* to f if for each $\varepsilon > 0$ there exists an integer N such that

$$x \in E \quad \text{and} \quad n \geqslant N \quad \text{implies} \quad |f_n(x) - f(x)| < \varepsilon.$$

The idea of convergence uniformly a.e. is self-evident.

A slightly weaker notion of convergence is given by the following:

Definition 31

Let $f_n : E \to R^*$ $(n = 1, 2, \ldots)$ and $f: E \to R^*$ be functions which are a.e. finite on E. Then f_n *converges almost uniformly* to f on E if for each $\varepsilon > 0$ there is a set $F_\varepsilon \subset E$, $F_\varepsilon \in \mathscr{A}$, $\mu(F_\varepsilon) < \varepsilon$, such that $f_n \to f$ uniformly on $(E - F_\varepsilon)$.

If μ is the Lebesgue measure on $E = [0, 1]$ it is clear that the sequence $f_n(x) = x^n$ converges almost uniformly but not uniformly a.e. From the definition it should be evident that convergence uniformly a.e. implies almost uniform convergence.

However, a less obvious implication is given by the following:

Theorem 12 (Egoroff)

Let $E \in \mathscr{A}$ with $\mu(E) < \infty$, and let (f_n) be a sequence of measurable functions from E to R^* which are finite a.e. and converge a.e. to a function $f: E \to R^*$ which is finite a.e.. Then $f_n \to f$ almost uniformly in E.

We consider next a rather different idea of proximity in which we look at the measure of the set on which two functions differ by some given number.

Definition 32

Let $f_n : A \to R^*$ and $f: A \to R^*$ be $\mathscr{A}$ measurable functions. Then f_n *converges in measure* (μ) to f if for each $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\{x : |f_n(x) - f(x)| \geqslant \varepsilon\} = 0.$$

It should be clear that the functions in question must be finite a.e. for the definition to be meaningful.

Our final notion of convergence makes use of the fact that the set $\mathfrak{L}_m$ of μ integrable functions is a linear space in which the idea of mean is defined. We have then:

Definition 33

Let (f_n) be a sequence of functions in $\mathfrak{L}_m$. Then (f_n) *converges to f in mean* if

$$\int |f_n - f| \, \mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} 0.$$

The different notions of convergence are, of course, related and, as they are used very generally, in particular in studying large economies, we give the following basic results:

Theorem 13

If a sequence (f_n) of measurable functions converges almost everywhere to f, then (f_n) converges in measure to f.

For a more limited class of functions we have the following:

Theorem 14

Let (f_n) be a sequence of positive integrable functions. The sequence (f_n) converges in the mean to the integrable function f if and only if (f_n) converges to f in measure, and

$$\lim_n \int f_n = \int f.$$

One further result will complete the basic results we need on the convergence of measurable functions.

Theorem 15 (Scheffé)

If (f_n) is a sequence of positive integrable functions with

$$\int \liminf_n f_n = \lim_n \int f_n < \infty,$$

then (f_n) converges in the mean to $\lim_n \inf f_n$.

We will need these ideas of convergence for many purposes and, in particular, when discussing the notion of a sequence of economies which converges to a limit economy. In order to give a simple and concise description of such a notion we take a sequence of finite economies for which we wish to show that certain properties true for "continuum economies" are approximately true for "large enough" economies. To avoid the problem that the space of agents and hence the associated mapping changes dimension as the number of agents increases we construct a series of parallel "equivalent" economies each with a continuum of agents and show that our results hold via the equivalent sequence. However, before discussing this problem in more detail we have to establish the meaning of this equivalence and for this we will need to discuss the idea of the convergence of measures and of a distribution.

2.6. On metric spaces: Weak convergence

We will focus our attention here on "weak convergence" which has been extensively used by Hildenbrand (1974) in particular.[26] This convergence may be characterised in several different ways two of which are fairly intuitive. If we consider any "well behaved" function f from a metric space T into the real line and a sequence (μ_n) of measures[27] on T then a requirement that (μ_n) converge to μ would be that the integral of f with respect to μ_n should converge to the integral of f with respect to μ. Alternatively, and perhaps more naturally, for any "convenient" subset B of T we should have $\lim_n \mu_n(B)=\mu(B)$. The basic idea is clearly that we require in a certain sense that the "weights" attached by μ_n to the various subsets should be very little different from those given by μ for n large enough. In particular, we might require for a sequence of economics that the "distribution of agents characteristics" should be "close" for n large to that of the limit economy. These ideas will be developed in detail in Chapter 18 of this Handbook. To make our previous remarks precise and to add two other equivalent definitions of weak convergence of measures we give the following:

Proposition 7

If T is a metric space and (μ_n) a sequence of probability measures on T then the following are equivalent:

(i) (μ_n) converges weakly to μ;
(ii) $\int f\mathrm{d}\mu_n \to \int f\mathrm{d}\mu$ for every bounded and uniformly continuous function f: $T \to R$;

[26]The reader is referred to Billingsley (1968) for a complete treatment of the problems mentioned here.

[27]We will be dealing exclusively with probability measures in this section; hence measure should be read as probability measure.

(iii) $\lim_n \mu_n(B)=\mu(B)$ for every subset $B \subset T$ for which the μ measure of the boundary of B is zero;
(iv) $\lim_n \sup \mu_n(C) \leqslant \mu(C)$ for every closed subset C in T;
(v) $\lim_n \inf \mu_n(D) \geqslant \mu(D)$ for every open subset D in T.

The following example cited by Hildenbrand (1974) may aid the reader's intuition.

Example 6

Let $T=R^m$. Define for a measure μ on R^m the distribution function,

$$F_\mu: R^m \to R,$$

by

$$F_\mu(x)=\mu\{z \in R^m \mid z \leqslant x\}.$$

The sequence (μ_n) of measures on R^m converges weakly to the measure μ on R^m if and only if the sequence (F_{μ_n}) of distribution functions converges to F_μ at every point x where F_μ is continuous.

Now suppose that we are concerned with a measure on a product space $A \times B$, for example, when we consider the space of agents' characteristics $\mathcal{P} \times R^\ell_+$. Then we will be concerned with marginal distributions.

The following result shows the relationship between the convergence of marginal distributions and the convergence of measures on a product space:

Theorem 16

If the sequence (μ_n) of probability measures on the separable measure space $A \times B$ converges weakly to the measure μ, then the sequences of marginal distributions (μ_n^A) and (μ_n^B) converge weakly to the marginal distributions μ^A and μ^B, respectively.

Let (μ_n) and (v_n) be sequences of measures on the separable metric spaces A and B, respectively. Then the sequence $(\mu_n \times v_n)$ of product measures on $A \times B$ converges weakly to the product measure $\mu \times v$ on $A \times B$ if and only if (μ_n) converges weakly to μ and (v_n) converges weakly to v.

If in particular A is a separable metric space and we denote by $\mathcal{M}(A)$ the set of all probability measures on A, then we have the following:

Proposition 8

There exists a metric ρ on $\mathcal{M}(A)$ such that the space $(\mathcal{M}(A), \rho)$ is separable and a sequence (μ_n) converges to μ in $(\mathcal{M}(A), \rho)$ if and only if it converges weakly to μ.

Such a distance between measures enables us to endow $\mathfrak{M}(A)$ with a structure similar to that of A. An explicit example of such a metric is given by the Prohorov-metric defined as follows:[28]

$$\rho(\mu, v)=\inf\{\varepsilon>0 \mid v(E)\leqslant\mu(B_\varepsilon(E))+\varepsilon \text{ and } \mu(E)\leqslant v(B_\varepsilon(E))+\varepsilon, \text{ for any } E\in\mathfrak{B}(A)\}.$$

One more notion that is important for many applications to economics is that of the support of a probability measure. Frequently we will be concerned with knowing that a measure concentrates all of its weight on a compact set, that is, that only isolated exceptions lie outside this set. For example, as we indicated in the introduction, we might require of somebody forecasting prices that with probability one he expects prices to fall within some compact set, or, more generally, we might require the following:

If μ is a probability measure on a separable metric space A then there is a closed subset B of A such that $\mu(B)=1$ and if $F\subset A$ is closed and $\mu(F)=1$ then $B\subset F$. Now consider:

Definition 34

The *support* of a probability measure μ denoted supp(μ), on a separable metric space A, is the smallest closed subset of A with measure one.

Then a very useful result which takes us in the right direction is the following:

Proposition 9

The set of probability measures with finite support is dense in $(\mathfrak{M}(A), \rho)$.

Now recalling an earlier discussion of price forecasting, we need to be sure that if the underlying space of outcomes is not compact that forecasts are "essentially" concentrated on some compact subset.

What is needed to formalise this requirement is the following:

Definition 35

A family of probability measures M on the metric space A is called *tight* if for every $\varepsilon>0$ there exists a compact set $K\subset A$ such that $\mu(K)>1-\varepsilon$ for every $\mu\in M$.

In the light of this definition the reader should consider the example mentioned previously for the case of one good in which the family μ_n is such that the forecast of tomorrow's price attaches probability one to $p^{t+1}=n$.

[28] Here $B_\varepsilon(E)$ denotes, as usual, the ε-neighbourhood of E, i.e., $B_\varepsilon(E)=\{x\in A \mid \text{dist}(x, E)<\varepsilon\}$.

Alternatively consider the family of measures (μ_n) where μ_n is the uniform probability distribution on $[0, n]$. The problems such examples pose will be evident in the chapter on temporary general equilibrium theory (Chapter 19).

Two results of particular interest are given by:

Theorem 17

If the family of probability measures M on a metric space A is tight, then every sequence (μ_n) of probability measures contains a weakly converging subsequence.

Specialising to families with only one member we have:

Theorem 18

Every probability measure on a complete separable measure space is tight.

Further results may be found in Hildenbrand (1974); and for a more complete mathematical development, see Billingsley (1968).

2.7. *Distributions*

We return now to a concept which shows, in particular, as we observed earlier the true value of using measure spaces as a description of an economy. This is the idea of a distribution, and, as we suggested, it is frequently useful to work with the distribution of characteristics, for example, as the basic description of an economy.

We give now the formal version of the definition given in the introduction:

Definition 36

Let $(A, \mathscr{Q}, \mu)$ be a probability space, M a metric space, and f a measurable mapping of A into m. The *distribution* v of f denoted by $\mu \circ f^{-1}$ is defined by

$$v(B) = \mu\{a \in A \mid f(a) \in B\} \quad \text{for every} \quad B \in \mathscr{B}(M).$$

As already observed, the choice of the measure space is arbitrary, and it is frequently the distribution that conveys the real information with which we are concerned. In particular, in studying "large economies" the frequent choice of the unit interval [0, 1] as the space of agents is purely for convenience and has no particular significance in itself. Indeed, we know that every measure on a metric space M is the distribution of some measurable mapping on some measure space. More particularly, if M is complete and separable then for every probability measure on M there exists a measurable mapping f of the closed unit interval

$[0,1]$ into M such that $\mu=\lambda\circ f^{-1}$, where λ denotes the Lebesgue measure on $[0,1]$. Thus we could, given a suitable distribution on the space of characteristics of agents, always construct an associated economy with the unit interval as the space of agents.

We now come to a result which proved crucial in establishing general limit theorems concerning the equivalence of different solutions to the problem of allocating resources in a market. This result indicates clearly how we may overcome the problem that if in a sequence of economies the number of agents changes then so does the space of agents and the notion of convergence to a limit is unclear.

Theorem 19 (Skorokhod)

Let T be a separable metric space and (μ_n) a weakly converging sequence of measures on T with limit μ. Then there exists a measure space $(A,\mathcal{Q},\nu)$ and measurable mappings f and f_n $(n=1,2,\ldots)$ of A into T such that $\mu=v\circ f^{-1}$, $\mu_n=v\circ f_n^{-1}$ and $\lim_n f_n=f$ a.e. in A.

Furthermore if T is complete then the measure space $(A,\mathcal{Q},v)$ can be chosen to be the unit interval with Lebesgue measure.

In Chapter 18 by Hildenbrand the reader will encounter an extensive discussion of sequences of finite economies which converge to limit economies. What is important is that the preference endowment distributions of the finite economies are always defined on the same space. Thus, although each economy has a different space of agents we can by Skorokhod's theorem construct an analogous space of agents which is the same for each of the economies in the sequence. In other words, the distribution of agents characteristics will give us the information required and we can replace the original agent space by a more convenient artifact without changing any of the economic features of the model. This discussion anticipates our next section.

2.8. *Convergence in distribution*

Given a sequence of measurable mappings (f_n), each from a measure space $(A_n,\mathcal{Q}_n,\mu_n)$ into a metric space T, we will want to define a sense in which these mappings converge. This leads us to:

Definition 37

A sequence (f_n) of measurable mappings with values in a metric space T *converges in distribution* to a measurable mapping f with values in T if the sequence (ν_n) of distributions of (f_n) converges weakly to the distribution ν of f.

Consider the special case in which $T=R$ and all the functions f_n and f are defined on the same measure space. In this case convergence in measure, and hence convergence almost everywhere, implies convergence in distribution. The converse is true only if f is a constant function.

We now give three results on convergence in distribution which will be of particular use when studying limit theorems for increasing sequences of economics.[29]

Proposition 10

Let (f_n) and f be a sequence of functions and a function all from a measure space $(A, \mathcal{A}, \mu)$ into a separable metric space (T, d) (where d is the metric). Then the function $w\mapsto d(f_n(w), f(w))$ from $A\to R$ is measurable and if the sequence $(d(f_n(\cdot), f(\cdot)))$, $n=1,2,\ldots,$ converges in measure to zero then the sequence (f_n) converges in distribution to f.

Before proceeding to the next result we will need to extend the notion of integrability to a sequence of functions:

Definition 38

Let (f_n) be a sequence of measurable functions and $(A_n, \mathcal{A}_n, \mu_n)$ a sequence of measure spaces with $f_n\colon A_n\to R$. Then (f_n) is said to be *uniformly integrable* if (i)

$$\lim_{q\to\infty}\left(\sup_n \int_{|f_n|>q} |f_n|\,\mathrm{d}\mu_n\right)=0;$$

or, equivalently, (ii)

$$\sup_n \int |f_n|\,\mathrm{d}\mu_n<\infty;$$

or (iii)

$$\lim_n \int_{E_n} |f_n|\,\mathrm{d}\mu_n\to 0 \quad \text{for every sequence } (E_n) \text{ for which } \mu_n(E_n)\to 0.$$

We note that:

Proposition 11

If the sequence (f_n) is uniformly integrable then the sequence of distributions of f_n is tight.

[29] These results, together with much of this section, are taken directly from Hildenbrand (1974).

Now we state a result of fundamental importance in studying sequences of economies:

Theorem 20 (Generalisation of Lebesgue's Theorem)

Let the sequence (f_n) of measurable functions converge in distribution to the measurable function f. If the sequence (f_n) is uniformly integrable then f is integrable, and furthermore,

$$\lim_n \int f_n \mathrm{d}\mu_n = \int f \mathrm{d}\mu .$$

If f and all f_n are positive and integrable, then the above equation implies that the sequence (f_n) is uniformly integrable.

The last result of this section is of interest since it shows how we may deal with the situation typically found in establishing results for growing sequences of economies. At each step we deal with a finite economy, and it is only in the limit that we are concerned with an infinite economy. To see how we may think of the infinite economy as representing the limit of the sequence of finite economies, we consider the following idea: At each step we draw from some fixed hypothetical distribution a finite sample and as these samples increase in size we would want the sample distributions to approximate more and more closely that of the underlying infinite population. With this in mind we state the following:

Theorem 21 (Glivenko–Cantelli)

Let $(A, \mathscr{Q}, \mu)$ be a measure space and (x_n) an independent sequence of identically distributed measurable mappings x_n of A into a separable metric space T. For every $a \in A$, let $\nu_n(a,.)$ be the distribution of the sample $\{x_1(a), \ldots, x_n(a)\}$ of size n $(n = 1, 2, \ldots)$, i.e.,

$$\nu_n(a, B) = \frac{1}{n}\{i \mid x_i(a) \in B, i = 1, \ldots, n\}.$$

Then for almost all $a \in A$ the sequence $(\nu_n(a,.))$, $n = 1, 2, \ldots$, of sample distributions converges weakly to the distribution of x_n.

3. Some results

We will now look at some examples in which many of the preceding concepts are used.

3.1. A large economy

Example 7[30]

Consider $\mathcal{P}$, the set of irreflexive and continuous binary relations on R_+^ℓ with the property that for every price vector $p \gg 0$ the set $\phi(\succ, e, p)$ of maximal elements for $\succ$ in the consumer's budget set $\{x \in R_+^\ell \mid p \cdot x \leqslant p \cdot e\}$ is non-empty. If $\mathcal{P}$ is endowed with Hausdorff's topology of closed convergence it is a separable metric space.

First we look at the case of a finite number of economic agents and define, as mentioned earlier, an exchange economy as a mapping from the space of agents to the space of characteristics, i.e., preferences and endowments,

$$\mathcal{E}: A \to P \times R_+^\ell .$$

Now the distribution μ_ε of agents' characteristics is given by

$$\mu_\varepsilon(B) = \frac{\#\mathcal{E}^{-1}(B)}{\#A} \quad \text{for every} \quad B \text{ of } \mathcal{P} \times R_+^\ell .$$

Again we emphasise that in the case of a large economy this second description may well be considered as more appropriate since we are not really concerned with specifying the characteristics of each individual, but are more interested in knowing what proportion of individuals fall within any given subset of characteristics.

Now for the infinite case we define the space of agents as a measure space $(A, \mathcal{Q}, v)$ with $v(A) = 1$, i.e., v is a *probability measure* and an economy is a *measurable mapping* $\mathcal{E}: (A, \mathcal{Q}, v) \to \mathcal{P} \times R_+^\ell$ such that $\int e \, dv < \infty$.

As will be described in detail in Chapter 18 by Hildenbrand, we may show the equivalence of two different solution concepts for the problem of allocating goods in an infinite economy. For this result to be interesting we must show it to be essentially true for large economies. To do this we must be able to describe a sequence of economies converging to a limit (infinite) economy, where, in particular, that limit economy is *atomless*. To this end we introduce the following:

Definition 39

The sequence $(\mathcal{E}_n)$, $\mathcal{E}_n: A_n \to \mathcal{P} \times R_+^\ell$, is called *purely competitive* if and only if

(i) the number $\#A_n \underset{n \to \infty}{\to} \infty$;

[30] For a full discussion of this example, see Hildenbrand (1975) from which it is taken, and Hildenbrand (1974).

(ii) the sequence (μ_{ε_n}) of preference endowment distributions *converges weakly* to a limit μ;

(iii) $\int e\,\mathrm{d}\mu_\varepsilon \to \int e\,\mathrm{d}\mu \gg 0$.

Note that the important idea here is that the distributions converge and that the limit economy is characterised by its distribution. This is again because the underlying measure space is arbitrary and, since it can be shown that two economies with the same preference endowment distribution are essentially the same, we can forego the micro distribution of an economy.

With Definition 39 a number of important limit theorems can be proved and details are given in Chapter 18.

3.2. *Fair division: Some results*

We return now to a problem mentioned in the introduction which has interested mathematicians for some period of time and which, with the increasing interest of economists in equitable distributions, shows clearly how measure-theoretic tools may be useful. Recall that the simplest expression of the problem, as we saw it, is that of dividing some object U amongst a finite number n of individuals such that each individual i receives in his own estimation (expressed by a measure μ_i on U) at least $1/n$ of the total "value" of U. More generally we might want to assign parts of the object to individuals such that the ith individual receives α_i and the others receive α_j, in his opinion, of the total where $\sum_{i=1}^{n}\alpha_i = 1$. The answer to this problem is to be found in Dubins and Spanier (1961). Their first result is:

Theorem 22

Let $(U, \mathfrak{U})$ be a measurable space $\mu_1,\ldots,\mu_n$ atomless probability measures on $(U, \mathfrak{U})$. Then given k and $\alpha_1,\ldots,\alpha_k > 0$ with $\sum_{i=1}^{k}\alpha_i = 1$, there exists a partition $A_1,\ldots,A_k$ of U such that $\mu_i(A_j) = \alpha_j$ for all $i = 1,\ldots,n$ and $j = 1,\ldots,k$.

The reader will observe that this result answers the question posed for $k = n$ and in particular setting $\alpha_i = 1/n$ for all i shows that a partition can be found that not only gives the ith individual his fair share but one which everybody believes gives the others their fair share too.

Provided that at least two individuals have different measures then there exist partitions which give strictly more than α_i to the ith individual, i.e., such that $\mu_i(A_i) > \alpha_i$, $i = 1,\ldots,n$.

This result involves an extension of Liapunov's theorem (Theorem 3), and Dubins and Spanier give a proof of that Theorem in proving the proposition from which Theorem 24 is derived.

The authors go on to prove the existence of partitions which are optimal in some sense. For example, one might wish to adopt the utilitarian criterion and find a partition $A_1, \ldots, A_n$ to maximise

$$\sum_{i=1}^{n} \mu_i(A_i),$$

and indeed the maximum is shown to exist.

Perhaps of more interest is the anticipation of Rawl's criterion of maximising the welfare of the least well-off individual.

Of all partitions consider those which maximise the amount received by the person who gets least. From these select those which give the most to the person who gets next to the least and so forth. More precisely if P is a partition then arrange the members $\mu_i(A_i)$ in non-decreasing order to construct the sequence

$$a_1(P), a_2(P), \ldots, a_n(P).$$

Now construct the lexicographic ordering on partitions P. Thus P is maximal in that ordering if for any other partition P' either $a_i(P) = a_i(P')$ for all i or if j is the smallest i such that $a_j(P) \neq a_j(P')$; then $a_j(P) > a_j(P')$. Such a maximal element we will call an *optimal partition*. Dubins and Spanier prove that such partitions exist and furthermore that if each μ_i is absolutely continuous with respect to every other then every optimal partition is equitable in the sense that

$$\mu_i(A_i) = \mu_j(A_j) \quad \text{for all } i \text{ and } j.$$

3.3. Integration of correspondences[31]

It is often the case that we are concerned with set-valued mappings, or correspondences, in economics. For example the demand of a given individual may be a set of bundles rather than a particular bundle for some given prices. Again we may wish to associate with an individual a production technology which would be a set of possible combinations of inputs and outputs. If we wish to talk about perfect competition in such circumstances and to use a continuum economy to do so, then we will need to be able to integrate such correspondences. If, for example, we want to be able to talk about mean demand for the

[31] This brief discussion is intended to give an indication of the sort of problem encountered in an area that demands greater sophistication than the other topics mentioned. A full treatment and references may be found in Hildenbrand (1974, pp. 53–79). A standard treatment of this problem is one where the integral is considered as the expectation of a set-valued random variable.

whole continuum economy we will be forced to integrate the demand correspondences of individuals.

Consider first a function f from a measure space $(A, \mathcal{Q}, \mu)$ into R^m, that is, $f=(f^1,\ldots,f^m)$ where each $f^i: A \to R$ $(i=1,\ldots,m)$. The function f is said to be *integrable* if each coordinate function f^i is integrable and the integral $\int f \mathrm{d}\mu$ is defined by

$$\left(\int f^1 \mathrm{d}\mu,\ldots,\int f^m \mathrm{d}\mu\right).$$

Now consider ϕ a set-valued mapping or correspondence of A into R^m. Denote by $\mathcal{L}_\phi$ the set of all μ integrable functions that have the property

$$f(a)\in\phi(a) \quad \text{a.e. in } A.$$

The functions in the set $\mathcal{L}_\phi$ are called integrable selections of ϕ.

Definition 40[32]

The set $\{\int f \mathrm{d}\mu \in R^m \mid f\in\mathcal{L}_\phi\}$ is called the *integral* of the correspondence ϕ and is denoted by $\int\phi \mathrm{d}\mu$ or by $\int\phi$.

Although the meaning of this definition is clear we have yet to show that there is a large class of correspondences for which the integral is non-empty, that is for which there exists an integrable selection. If a correspondence admits an integrable selection we say that the correspondence itself is integrable.

To ensure that the integral is non-empty we need first to establish that there exists a measurable selection. First, then, we give the following:

Definition 41[33]

A *correspondence* ϕ from a measure space $(A, \mathcal{Q}, \mu)$ into a complete separable metric space S is *measurable* if the graph of that correspondence belongs to $\mathcal{Q}\otimes\mathcal{B}(S)$.

Now the basic result is the following:

Theorem 23 (Measurable Selection)

Let ϕ be a measurable correspondence of a measure space $(A, \mathcal{Q}, \mu)$ into a complete separable metric space S. Then there exists a measurable function f of A into S such that $f(a)\in\phi(a)$ a.e. in A.

[32] This is the definition given by Aumann (1965).
[33] Note that for many purposes in economics $S=R^\ell$.

We need now, of course, to make sure that the correspondence is bounded in order to ensure that it is integrable:

Definition 42

A correspondence ϕ of $(A, \mathcal{A}, \mu)$ into R^ℓ is *integrably bounded* if there exists an integrable function h of A into R^ℓ_+ such that for every $a \in A$ and for every $x \in \phi(a)$ we have

$$|x| = (|x_1|, \ldots, |x_\ell|) \leqslant h(a).$$

Now we can give the following:

Theorem 24

A measurable, integrably bounded correspondence from $(A, \mathcal{A}, \mu)$ into R^ℓ is integrable.

Now we turn to a particularly interesting aspect of large economies, that is, the "convexifying effect" of large numbers. Thus, at an individual or micro level we are obliged to make assumptions of convexity of individual preferences to guarantee that the demand correspondence is convex valued. This property then carries over to the aggregate excess demand correspondence, and one can make use of Kakutani's fixed point theorem to prove the existence of equilibrium. This problem is fully dealt with in the chapter on the existence of competitive equilibrium, Chapter 15 by Debreu.

However, if we have an atomless measure space of agents we may dispense with the requirement of convexity of preferences, since even though individual demand correspondences will not be convex valued, mean demand will necessarily be so. This follows directly from:

Theorem 25

Let ρ be a correspondence from an atomless measure space $(A, \mathcal{A}, \mu)$ into R^m then the set

$$\mathfrak{X} = \left\{ \int_E f \mathrm{d}\mu \,|\, f \in \mathfrak{L}_\rho,\ E \in \mathcal{A},\ \mu(E) > 0 \right\}$$

is a convex set in R^m.

The proof of this theorem depends on Liapunov's theorem, emphasising the importance of the latter. One can also proceed to develop approximation results for large finite economies.

The frequent use of correspondences in economic theory has led to the development of a substantial literature on the integration of correspondences and the interested reader will find the appropriate references in Hildenbrand (1974).

In passing, it is interesting to note that measure theory and a measure-theoretical approach to the description of large economies has in a sense diminished some of the problems posed by the use of correspondences.

In particular, we discussed earlier a standard problem, namely that for individuals with convex preferences the appropriate demand concept is a correspondence. Yet it can be shown that if we have a measure space of consumers and the support of that measure is sufficiently rich, then the aggregate or mean demand may, in fact, be considered as a function if we add some more structure to the space of preferences. That is, for any given prices, only a negligible subset of consumers will have a set rather than a single bundle as their demand; hence the demand will be a function. Thus, in a sense, the importance of correspondences as such is somewhat diminished in large economies.

4. Conclusion

In this chapter we have presented a number of results which should be useful for the reader interested in economic theory in two ways: they should provide an underpinning for many of the probabilistic approaches to problems in economic theory, and they should also provide a basis for an understanding of that part of the literature which adopts the measure-theoretic approach as a description of an economy.

No attempt has been made to give a comprehensive survey of the general literature in which measure theory is applied to problems in economic theory since this literature is already extensive and is involving an increasing number of different areas. However, with the selection of results presented here the reader should find much of this literature readily accessible.

References

Allen, R. G. D. and A. L. Bowley (1935), Family expenditure: A study of its variation. London: Staples Press.

Araujo, A. and A. Mas-Colell (1978), "Notes on the smoothing of aggregate demand", Journal of Mathematical Economics, 5:113–129.

Arrow, K. J. (1963), Social choice and individual values, 2nd ed. New York: Wiley.

Aumann, R. J. (1964), "Markets with a continuum of traders", Econometrica, 32:39–50.

Aumann, R. J. (1965), "Integrals of set-valued functions", Journal of Mathematical Analysis and Applications, 12:1–12.

Aumann, R. J. (1966), "Existence of competitive equilibria in markets with a continuum of traders", Econometrica, 32:1–17.

Billingsley, P. (1963), Convergence of probability measures. New York: Wiley.

Blaug, M. (1963), Economic theory in retrospect. Homewood, IL: Richard D. Irwin.

Brown, D. J. and A. Robinson (1972), "A limit theorem on the core of large standard exchange economies", Proceedings of the National Academy of Sciences, USA, 69:1258–1260. Correction, 69:3068.

Debreu, G. (1959), Theory of value. New York: Wiley.

Debreu, G. (1970), "Economies with a finite set of equilibria", Econometrica, 38:387–392.

Debreu, G. and H. Scarf (1963), "A limit theorem on the core of an economy", International Economic Review, 4:235–246.

Dierker, E., H. Dierker and W. Trockel (1978), "Continuous mean demand function", Discussion paper. Bonn: University of Bonn.

Dubins, L. E. and E. H. Spanier (1961), "How to cut a cake fairly", American Mathematical Monthly, 1:1–17.

Edgeworth, F. Y. (1881), Mathematical psychics. London: P. Kegan.

Fishburn, P. C. (1970), "Arrow's impossibility theorem, concise proof, and infinite voters", Journal of Economic Theory, 2:103–106.

Grandmont, J. M., A. P. Kirman and W. Neuefeind (1974), "A new approach to the uniqueness of equilibrium", Review of Economic Studies, 41:289–292.

Halmos, P. R. (1961), Measure theory, 7th ed. Princeton, NJ: Van Nostrand.

Hildenbrand, W. (1970), "Existence of equilibria for economies with production and a measure space of consumers", Econometrica, 38:608–623.

Hildenbrand, W. (1974), Core and equilibria of a large economy. Princeton, NJ: Princeton University Press.

Hildenbrand, W. (1975), "Distributions of agents' characteristics", Journal of Mathematical Economics, 2:129–138.

Hildenbrand, W. (1979), "On the uniqueness of mean demand for dispersed families of preferences", Discussion Paper. Bonn: University of Bonn.

Ichiishi, T. (1976), "Economies with a mean demand function", Journal of Mathematical Economics, 3:167–171.

Khan, M. A. (1974), "Some equivalence theorems", Review of Economic Studies, 41:549–565.

Kirman, A. P. and D. Sondermann (1972), "Arrow's theorem, many agents, and invisible dictators", Journal of Economic Theory, 4:267–277.

Lindenstrauss, J. (1966), "A short proof of Liapunov's convexity theorem", Journal of Mathematics and Mechanics, 15:971–972.

Mas-Colell, A. (1979), "Perfect competition and the core", Working paper no. 1P280. Berkeley, CA: Center for Research in Management Science, University of California.

Neveu, J. (1965), Mathematical foundations of the calculus of probability. San Francisco, CA: Holden-Day.

Robinson, A. (1965), Non-standard analysis. Amsterdam: North-Holland.

Schumpeter, J. (1954), History of economic analysis. Oxford: Oxford University Press.

Shapley, L. S. (1953), "Stochastic games", Proceedings of the National Academy of Sciences, USA, 39:327–332.

Shitovitz, B. (1974), "Oligopoly in markets with a continuum of traders", Econometrica, 41:467–501.

Sondermann, D. (1975), "Smoothing demand by aggregation", Journal of Mathematical Economics, 2:201–224.