CHAPTER 1

Topics from Probability Theory

In this preliminary chapter we shall give an exposition of certain topics in probability theory which are necessary to understand and interpret the definition and properties of entropy. We have tried to write the chapter in such a way that a reader with a knowledge of measure theory as given in Ash [15], Halmos [55], or any other basic measure theory text can follow the arguments and understand the examples. We introduce just those parts of probability theory which are necessary for the subsequent chapters and attempt to make them meaningful by use of very simple examples. We also restrict the discussion to "nice" probability spaces, so that conditional expectation and conditional probability are more intuitive and hopefully easier to understand. These "nice" spaces also make it possible to use partitions as models for random experiments, even those experiments which are limits of sequences of experiments.

1.1 Probability Spaces

Entropy is a quantitative measurement of uncertainty associated with random phenomena. In order to define this quantity precisely, it is necessary to have a mathematical model for random phenomena which is general enough to include many different physical situations and which has enough structure to allow us to use mathematical reasoning to answer questions about the phenomena.

Such a model is given by a mathematical structure called a probability space, which is nothing more than a measure space in which the measure of the universe set is 1. Thus, a probability space is a triple $(\Omega, \mathfrak{F}, P)$ where Ω is a set, $\mathfrak{F}$ is a collection of subsets of Ω, and P is a nonnegative real valued function defined on $\mathfrak{F}$ such that

ENCYCLOPEDIA OF MATHEMATICS and Its Applications, Gian-Carlo Rota (ed.). Vol. 12: Nathaniel F. Martin and James W. England, Mathematical Theory of Entropy. ISBN 0-201-13511-6

C1. $\mathcal{F}$ is a σ-field, i.e., it is closed under countable unions, complements, and contains Ω;

C2. P is a measure, i.e., if $\{E_n\}$ is a countable pairwise disjoint collection of sets from $\mathcal{F}$, then

$$P\left(\bigcup E_n\right)=\sum P(E_n);$$

3.C $$P(\Omega)=1.$$

In the triple $(\Omega, \mathcal{F}, P)$, Ω is called the *sample* space, or *outcome* space; the points of Ω are *outcomes*, sets in $\mathcal{F}$ are *events*, and P is the probability.

A random experiment such as tossing a coin or drawing a ball from an urn can be represented mathematically as a probability space. For example, consider the experiment which consists of drawing a ball from an urn which contains 3 red balls, 2 white balls, and 5 blue balls, where the only distinguishing feature of the balls is their color. The only possible outcomes of the experiment will be a ball of a certain color, so the outcomes can be represented by the set $\Omega=\{r,w,b\}$ where r will stand for "a red ball was drawn" and similarly for w and b.

If we actually perform this experiment a great many times, each time replacing the ball after noting its color, the ratio of the number of times a ball of a given color is drawn to the total number of draws will seem to approach a certain limiting value. This value is taken to be the probability or likelihood of obtaining a ball of that color on any draw. In the experiment described in the previous paragraph these values will be .3 for red, .2 for white, and .5 for blue. What is being demonstrated is that in the long run, one will get a red ball 30% of the time, a white ball 20% of the time, and a blue ball 50% of the time. This should be expected, since 30% of the balls in the urn are red, etc.

Using these numbers, we get a function f, called the distribution, on the space $\Omega=\{r,w,b\}$, which we use to obtain the probability of events.

The probability measure is to be defined on subsets of the outcome space Ω, and these subsets are to represent events associated with the experiment. In this example, an event is any meaningful statement which can be made concerning the occurrence or nonoccurrence of red, white, or blue balls. Such a statement can be represented by a subset of Ω. For example, the statement "either a red or a blue ball is drawn" is represented by the set $\{r,b\}$. An event $E\subset\Omega$ will occur provided the outcome of the draw is a member of this set. In this finite case the collection of events will be the collection of all unions of outcomes. This is the collection of all subsets of Ω. Thus, $\mathcal{F}=\{\{r\},\{w\},\{b\},\{r,w\},\{r,b\},\{w,b\},\Omega,\varnothing\}$.

To obtain a probability measure of an event $E\in\mathcal{F}$ from the distribution function f given by $f(r)=.3$, $f(w)=.2$, $f(b)=.5$, all we need to do is sum the values of f over the outcomes in the event. Thus $P(\{r,b\})=f(r)+f(b)=.8$. The value of $P(\{r,b\})$ should represent the relative frequency of obtaining either a red or a blue ball in many independent repetitions of the experiment. In actual experiments this relative frequency does come close to .8. We should expect this, since 80% of the balls in the urn are either red or blue.

This same random experiment can be represented by another measure space in which the outcomes are not points but sets themselves. Consider the space $(I,\mathfrak{L},\lambda)$, where I is the unit interval $[0,1]$, $\mathfrak{L}$ is the collection of Lebesgue measurable subsets of I, and λ is Lebesgue measure. Let R, W, B be any three measurable subsets of $[0,1]$ which are pairwise disjoint and such that $\lambda(R)=.3$, $\lambda(W)=.2$, $\lambda(B)=.5$. The experiment can be modeled by the probability space $(\Omega', \mathfrak{F}', P')$ where Ω' is the set $\{R,W,B\}$, $\mathfrak{F}'$ is the collection of subsets of $\mathfrak{L}$ which are unions of elements from Ω', and P' is the restriction of λ to $\mathfrak{F}'$.

It is easy to see that the two measure spaces $(\Omega,\mathfrak{F},P)$ and $(\Omega',\mathfrak{F}',P')$ are the same in the sense that there is a bijection ϕ of Ω to Ω' with the property that $\phi(\mathfrak{F})=\mathfrak{F}'$ and for any set $F\in\mathfrak{F}$, $P'(\phi(F))=P(F)$. The second representation of the experiment has some advantages over the first because it is embedded in a space with a rich mathematical structure which is well understood. This second characterization is a *factor space* of a Lebesgue space, as we shall see in the next section.

1.2 Measurable Partitions and Lebesgue Spaces

A partition of a probability space $(\Omega,\mathfrak{F},P)$ is a collection of subsets of Ω which are disjoint and whose union is Ω. The sets in the partition are called *atoms* of the partition. There are two partitions which we will use quite often. They are the point partition ε whose atoms are the singleton sets of Ω, i.e.,

$$\varepsilon=\{\{\omega\}:\omega\in\Omega\}$$

and the trivial partition ν whose atoms are the empty set and Ω.

If ξ is a partition of $(\Omega,\mathfrak{F},P)$, any subset of Ω which is a union of atoms of ξ is called a *ξ-set*. For example, the collection of ε-sets of Ω is the collection of all subsets of Ω and the collection of ν-sets consists of $\varnothing$ and Ω. A partition ξ is *measurable* if there exists a countable family $\{B_n : n=1,2,\dots\}$ of ξ-sets which are $\mathfrak{F}$-measurable (i.e., members of $\mathfrak{F}$) and have the following separation property:

S1. *For any pair C_1, C_2 of atoms from ξ with C_1 not equal to C_2, there exists a set B_n such that either $C_1\subset B_n$ and $C_2\subset\Omega-B_n$ or vice versa.*

It can be shown that the atoms of a measurable partition are measurable, but it is not necessarily true that any partition of an arbitrary measure space into measurable sets is a measurable partition.

Measurable partitions are also models of random experiments, as we saw in the example $(\Omega', \mathfrak{F}', P')$ in the last section. In this example, the Lebesgue space being partitioned is the unit interval, and the atoms of the partition $\{R, W, B\}$ represent the outcomes of the experiment.

As another example consider the experiments of tossing a fair coin and drawing a ball from an urn which contains red, white, and blue balls in the proportion 3:2:5. The latter experiment is modeled as the space $(\Omega_1, \mathfrak{F}_1, P_1)$ where $\Omega_1 = \{r, w, b\}$ and $\mathfrak{F}_1$ and P_1 are as described in Section 1.1. The coin tossing experiment can be modeled as the space $(\Omega_2, \mathfrak{F}_2, P_2)$ where $\Omega_2 = \{h, t\}$, $\mathfrak{F}_2 = \{\{h\}, \{t\}, \Omega_2, \varnothing\}$, and P_2 is obtained from a distribution on Ω_2 which assigns $\frac{1}{2}$ to h and $\frac{1}{2}$ to t. All of the probablistic structure in the coin tossing experiment can be obtained as a measurable partition ζ of the urn space as follows.

Take $\zeta = \{\{r, w\}, \{b\}\}$ and identify $\{r, w\}$ with h and $\{b\}$ with t. Since $P_1\{r, w\} = P_2\{h\}$ and $P_1\{b\} = P_2\{t\}$, it is very easy to see that if we define an outcome space to be the atoms of ζ, the field of events to be the $\mathfrak{F}_1$-measurable ζ-sets, and the probability to be P_1 restricted to these ζ-sets, the resulting space is the coin tossing space. This construction makes the coin tossing experiment a *factor space* associated with the partition ζ of the urn experiment.

Let $(\Omega, \mathfrak{F}, P)$ be a probability space and ξ a measurable partition of Ω. The factor space of Ω associated with the partition ξ is the probability space $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$ where Ω_ξ consists of the atoms of ξ, $\mathfrak{F}_\xi$ consists of $\mathfrak{F}$-measurable ξ-sets and P_ξ is the restriction of P to $\mathfrak{F}_\xi$.

The representation $(\Omega_2, \mathfrak{F}_2, P_2)$ of the coin tossing experiment given above is not strictly the factor space of $(\Omega_1, \mathfrak{F}_1, P_1)$ associated with the partition $\{\{r, w\}, \{h\}\}$. However, it is *isomorphic* to this factor space. Isomorphism is a way of identifying different mathematical models of the same observed phenomenon. Two probability spaces $(\Omega_1, \mathfrak{F}_1, P_1)$ and $(\Omega_2, \mathfrak{F}_2, P_2)$ are isomorphic if there exists a bijection $\mathbf{T}$ of Ω_1 onto Ω_2 such that both $\mathbf{T}$ and $\mathbf{T}^{-1}$ are measurable and $P_2(\mathbf{T}E_1) = P_1(E_1)$ for all $E_1 \in \mathfrak{F}_1$ and $P_1(\mathbf{T}^{-1}E_2) = P_2(E_2)$ for all $E_2 \in \mathfrak{F}_2$. The bijection $\mathbf{T}$ is called an isomorphism between the spaces.

The definition of isomorphism just given is accurate enough for the urn space and coin tossing spaces, since the only set of measure zero in either space is the empty set. In more complex spaces, which are needed for more complex phenomena such as the action of a particle under Brownian motion, there are many sets of probability zero which are not empty. Since these events do not affect the probabilities, the definition of isomorphism between probability spaces should only require that the function $\mathbf{T}$ be a bijection after a set of probability zero is removed from both Ω_1 and Ω_2.

Sometimes this is called isomorphism (mod 0) or isomorphism almost everywhere. In this book, an isomorphism will be an isomorphism (mod 0).

Returning to the example of the coin tossing and urn experiments, it is easy to check that $(\Omega_2, \mathfrak{F}_2, P_2)$ is isomorphic to $(\Omega_{1\zeta}, \mathfrak{F}_{1\zeta}, P_{1\zeta})$ and the urn experiment $(\Omega_1, \mathfrak{F}_1, P_1)$ is isomorphic to the factor space $(I_\xi, \mathfrak{F}_\xi, \lambda_\xi)$ of $(I, \mathfrak{L}, \lambda)$ defined in Section 1.1, where $\xi = \{R, W, B\}$.

There is a class of probability spaces which is well understood mathematically and is such that most interesting random phenomena can be modeled as a factor space of them. Such a space, called a *Lebesgue space*, is a measure space which is isomorphic to a segment of the unit interval with Lebesgue measure together with a countable number of point masses such that the total measure of the segment and the point masses is one. Thus a nonatomic Lebesgue space is isomorphic to $(I, \mathfrak{L}, \lambda)$ described in Section 1.1. A totally atomic Lebesgue space is isomorphic to a countable collection of point masses. (Countable includes finite as well as countably infinite sets.)

The axiomatic definition of a Lebesgue space is rather complicated, but it is worth giving, since it will allow us to identify many common probability spaces as Lebesgue spaces with a minimum of effort.

The first objects we must define in order to understand a Lebesgue space are basis and complete basis for a measure space. (Keep in mind that we have the uniform assumption that all measure spaces are complete in the measure theory sense, i.e., all subsets of sets of measure zero are measurable. Unfortunately, the word complete associated with basis has a different meaning.)

A countable collection $\Gamma = \{B_n\}$ of measurable sets is said to be a *basis* for the probability space $(\Omega, \mathfrak{F}, P)$ if it satisfies the following two conditions:

B1. Γ *separates points of* Ω, *i.e., for any two points* ω, ω' *of* Ω *there exists a set* $B \in \Gamma$ *such that either* $\omega \in B$ *and* $\omega' \notin B$ *or vice versa.*

B2. *The (measure theoretic) completion of the σ-field generated by* Γ *is* $\mathfrak{F}$.

Notice that if a space $(\Omega, \mathfrak{F}, \mathrm{P})$ has a basis, then the point partition in this space is a measurable partition.

Now suppose that $\Gamma = \{B_n : n \in N\}$ is a basis for $(\Omega, \mathfrak{F}, P)$, where N is either Z^+ or $\{1, 2, \ldots, N\}$. For the moment let B^0 denote B, and B^1 denote $\Omega - B$ for any set $B \in \Gamma$. Suppose that a is a member of $\{0,1\}^N$, the set of all functions from N to $\{0,1\}$. Consider the set

$$A = \bigcap_{n \in N} B_n^{a(n)}.$$

[For example, if $N = Z^+$ and $a = \{1, 0, 0, \ldots\}$, then $A = (\Omega - B_1) \cap B_2 \cap B_3 \cap \cdots$.] Because Γ is a basis, property B1 implies that A contains at most one point. Using this procedure we can associate with each sequence $a \in \{0,1\}^N$ a set A which is either the empty set or a set containing one point. It is also true that if a_1 and a_2 are distinct functions with associated sets A_1 and A_2, then $A_1 \cap A_2 = \varnothing$. Therefore the collection γ of all sets A obtained from the sequences in $\{0,1\}^N$ is a partition of Ω. Even more is true. This partition is a measurable partition (since the sets in Γ are measurable γ-sets which separate atoms of γ) and is either the point partition $\varepsilon = \{\{\omega\} : \omega \in \Omega\}$ or the point partition with the empty set adjoined.

The basis Γ is said to be *complete* if the associated partition γ just described is the point partition, or equivalently if each sequence $a \in \{0,1\}^N$ gives a nonempty set $\bigcap B_n^{a(n)}$. Notice that a basis is complete if and only if the map which sends $\omega \in \Omega$ to the sequence $a \in \{0,1\}^N$, where $\bigcap_{n \in N} B_n^{a(n)} = \{\omega\}$, is a bijection.

As an example of a space with a complete basis, take the unit interval with Lebesgue measure, $(I, \mathfrak{L}, \lambda)$, and define B_n by

$$B_n = \bigcup \left[\frac{2j}{2^n}, \frac{2j+1}{2^n} \right],$$

where the union is extended over all j between 0 and $2^{n-1}-1$. Then $\Gamma = \{B_n : n \in Z^+\}$ is a complete basis.

Since we are usually concerned just with the measure on the space, and sets of measure zero can be neglected, a weaker form of completeness is adequate. Essentially, this type of completeness for a basis means that the space of the basis can be embedded in another space with a basis in such a way that the universe sets have the same measure and the basis elements in the space are the restrictions of the basis elements of the larger space to the original set. More precisely, a measure space $(\Omega, \mathfrak{F}, \mu)$ is *complete* (mod 0) with respect to a basis $\{B_n\}$ [or the basis is complete (mod 0)] if and only if there exists a measure space $(\Omega', \mathfrak{F}', \mu')$ with a complete basis $\{B_n'\}$ and an injection $\mathbf{T}$ of Ω into Ω' such that the image of Ω in Ω' is measurable, $\mu'(\Omega' - \mathbf{T}\Omega) = 0$, and $B_n = \mathbf{T}^{-1}B_n'$. It is clear that $(\mathbf{T}\Omega, \mathfrak{F}' \cap \mathbf{T}\Omega, \mu')$ is (strictly) isomorphic to $(\Omega, \mathfrak{F}, \mu)$ and $\{B_n' \cap \mathbf{T}\Omega\}$ is a basis for the space $(\mathbf{T}\Omega, \mathfrak{F}' \cap \mathbf{T}\Omega, \mu')$ which corresponds to $\{B_n\}$. The isomorphism is given by $\mathbf{T}$.

The ideas in the above paragraphs are due to Rohlin, and in his development [122] of them he shows that if a space is complete (mod 0) with respect to one basis, then it is complete (mod 0) with respect to every basis. A *Lebesgue space* is defined to be any totally finite measure space which is complete (mod 0) with respect to some basis. We shall always assume that our Lebesgue spaces are probability spaces. Thus a Lebesgue space for the remainder of this book will mean a probability space with a basis which is complete (mod 0) with respect to this basis.

In the next few paragraphs, we list several examples of Lebesgue spaces which arise from natural phenomena. Some of these models are associated with various stochastic processes which we will discuss more fully in Section 1.7.

Example 1.1 (*Discrete Lebesgue space* $(S, \mathbb{S}, \mu_f)$). Let S denote a finite or countably infinite set, and let f denote a distribution on S. (Recall that a distribution is any function on S such that $f(s) \geqslant 0$ and $\Sigma_{s \in S} f(s) = 1$.) Let $\mathbb{S}$ denote the collection of all subsets of S, and define μ_f on $\mathbb{S}$ by $\mu_f(E) = \Sigma_{s \in E} f(s)$. The resulting space is a totally atomic probability space which is a Lebesgue space.

A complete (mod 0) basis for $(S, \mathbb{S}, \mu_f)$ may be constructed as follows: If S is finite, enlarge it to a set S' whose cardinality is 2^k, where k is the integer such that $2^{k-1} < |S| \leqslant 2^k$. ($|S|$ denotes the cardinality of S.) Without loss of generality we may assume $S = \{1, 2, \ldots, |S|\}$. Extend f to f' on S' by defining $f(j) = 0$ for $|S| < j \leqslant 2^k$. A complete basis for $(S', \mathbb{S}', \mu_{f'})$ can be obtained by taking B'_1 to be the odd integers; B'_2 is obtained by taking the first two integers, skipping the next two, taking the next two, etc; B'_3 is obtained by taking the first four integers, skipping the next four, etc. In general B'_n is obtained by taking the first 2^{n-1} integers, skipping the next 2^{n-1}, etc. It is not difficult to see that $(S, \mathbb{S}, \mu_f)$ can be embedded in $(S', \mathbb{S}', \mu_{f'})$ in such a way that the basis $\{B'_n\}$ gives a complete (mod 0) basis.

In case S is countably infinite, the same construction will work directly on S to obtain a complete (mod 0) basis.

Example 1.2 (*The unit interval* $(I, \mathfrak{L}, \lambda)$). Let I denote the closed unit interval $[0, 1]$, and $\mathfrak{L}$ the collection of all Lebesgue measurable subsets of I. (Recall that the Lebesgue sets are the completion of the Borel sets with respect to Lebesgue measure.) The measure λ denotes Lebesgue measure on $\mathfrak{L}$.

A complete basis for $(I, \mathfrak{L}, \lambda)$ is given by $\{B_n\}$ for $B_n = \cup_j [j/2^n, (j+1)/2^n]$, where $j = 1, 3, 5, \ldots, 2^n - 1$. Since $(I, \mathfrak{L}, \lambda)$ has a complete basis, every basis is complete (mod 0), and hence a basis such as the set of open intervals with rational end points is also a complete (mod 0) basis.

Example 1.3 (*A product sequence space over* S, $(\Sigma(S), \mathfrak{F}, \mu)$). This space consists of all doubly infinite sequences of elements from a discrete space $(S, \mathbb{S}, \mu_f)$ with product measure.

The construction of this space is obtained as follows: The set $\Sigma(S)$ consists of all functions on the integers Z to S. For any finite set $G \subset Z$ and collection $\{s_i : i \in G\}$ let

$$C\{s_i : i \in G\} = \{\omega \in \Sigma(S) : \omega(i) = s_i, i \in G\}.$$

This set is called a cylinder set with base G. Define a set function μ on the collection of all cylinder sets by the equation

$$\mu=(C\{s_i: i\in G\})=\prod_{i\in G}\mu_f(\{s_i\})$$

$$=\prod_{i\in G}f(s_i).$$

The collection of all finite pairwise disjoint unions of cylinder sets forms a field, and the usual extension of μ to this field is countably additive. By the Caratheodory extension theorem there exists a unique extension, also denoted by μ, to the σ-field generated by this field. Let $\mathfrak{F}$ denote the completion of this σ-field with respect to μ.

A complete (mod 0) basis for the sequence space is given by the collection of all cylinder sets based on a single integer. That is, the basis consists of sets of the form $\{\omega\in\Sigma(S):\omega(i)=s\}$ where s is selected from S, and i ranges over the integers.

Example 1.4 (*A product sequence space over I*, $(\Sigma(I),\mathfrak{F},\mu)$). This space consists of all doubly infinite sequences of elements from the unit interval with product measure. The construction is similar to Example 1.3 and is as follows:

The set $\Sigma(I)$ consists of all functions from Z to I. For any finite set $G\subset Z$ and collection $\{E_i: i\in G\}$ of sets E_i from $\mathfrak{L}$, the set

$$C\{E_i: i\in G\}=\{\omega\in\Sigma(I):\omega(i)\in E_i \text{ for all } i\in G\}$$

is a cylinder set based on G. Define a set function μ on the cylinder sets by

$$\mu(C\{E_i: i\in G\})=\prod_{i\in G}\lambda(E_i)$$

and extend in the same way as was done in Example 1.3 to obtain a unique measure μ on the completion $\mathfrak{F}$ of the σ-field generated by the field of all finite disjoint unions of cylinder sets.

A complete (mod 0) basis for the space $(\Sigma(I),\mathfrak{F},\mu)$ is given by the collection of all cylinder sets of the form

$$\{\omega\in\Sigma(I):\omega(i)\in B_n\}$$

where $i\in Z$ and $B_n=\bigcup_j[j/2^n,(j+1)/2^n]$, for $j=1,3,\ldots,2^n-1$.

Example 1.5 (*A general sequence space over S*, $(\Sigma(S),\mathfrak{F},m)$). This example is similar to Example 1.3 except the measure m is not a product measure.

Let $\Sigma(S)$ be as defined in Example 1.3, and let $\mathfrak{F}'$ denote the σ-field generated by the field of finite disjoint unions of cylinder sets. Assume that for each finite subset G of Z we have a probability measure P_G on the σ-field $\mathfrak{F}_G$ generated by the collection of all cylinder sets based on G. Further assume that the collection of measures $\{P_G: G \text{ a finite subset of } Z\}$ has the property that if $G_1 \subset G_2$, then P_{G_2} restricted to $\mathfrak{F}_{G_1}$ is P_{G_1}. This condition is called the Kolmogoroff consistency condition, and such a family of measures is called a consistent family of measures.

The σ-field $\mathfrak{F}'$ is generated by $\cup \mathfrak{F}_G$, where the union is taken over all finite subsets G of Z. For any set $E \in \cup \mathfrak{F}_G$, define $m(E) = P_{G'}(E)$, where G' is any finite set such that $E \in \mathfrak{F}_{G'}$. The collection $\cup \mathfrak{F}_G$ is a field, and m is countably additive on this field. It follows that m extends to a unique measure on $\mathfrak{F}'$, also denoted by m, called the Kolmogoroff extension of the consistent family P_G. The space $(\Sigma(S), \mathfrak{F}, m)$ is a Lebesgue space, where $\mathfrak{F}$ is the m completion of $\mathfrak{F}'$.

A complete (mod 0) basis for $(\Sigma(S), \mathfrak{F}, m)$ is given by the family of all sets of the form

$$\{\omega \in \Sigma(S) : \omega(i) = s\}$$

for $i \in Z$ and $s \in S$.

Example 1.6 (*A general sequence space over I*, $(\Sigma(I), \mathfrak{F}, m)$). This space is constructed in exactly the same way as Example 1.5 from a consistent family $\{P_G: G \text{ a finite subset of } Z\}$ of probability measures on $\mathfrak{F}_G$, where $\mathfrak{F}_G$ is again the σ-field generated by all cylinder sets contained in $\Sigma(I)$ based on G. This space is also a Lebesgue space with the same basis as that given for Example 1.4.

Example 1.7 (*Sequence space over a general Lebesgue space* Ω, $(\Sigma(\Omega), \mathfrak{F}, m)$). If a construction such as that used in Example 1.3, 1.4, 1.5, or 1.6 is used, but $(S, \mathbb{S}, \mu_f)$ or $(I, \mathfrak{L}, \lambda)$ is replaced by an arbitrary Lebesgue space, the resulting space is also a Lebesgue space.

on

Example 1.8 (*A space which is not a Lebesgue space*, $(Y^R, \mathfrak{F}, P)$). Let $(Y, \mathfrak{Y})$ be either $(I, \mathfrak{L})$, $(S, \mathbb{S})$, or the reals with the Borel sets. Let R denote the set of real numbers, and Y^R the set of all functions on R to Y. For each finite subset $G \subset R$, let $\mathfrak{F}_G$ denote the σ-field generated by the cylinder sets based on G. Assume that $\{P_G: G \text{ a finite subset of } R\}$ is a consistent family of probability measures on $\mathfrak{F}_G$. The Kolmogoroff extension theorem implies there exists a unique probability measure P on the σ-field $\mathfrak{F}'$ generated by the sets in the union of the σ-fields $\mathfrak{F}_G$. If $\mathfrak{F}$ denotes the completion of $\mathfrak{F}'$ with respect to P, it is not the case that $(Y^R, \mathfrak{F}, P)$ is a Lebesgue space. Intuitively, it is not a Lebesgue space because there are too many points in Y^R for the number of separating sets available in $\mathfrak{F}$. Certain restrictions on the functions allowed in Y^R, or a hypothesis of regularity such as those given by Doob [36], will make it into a Lebesgue space.

Example 1.9 (*Polish spaces*, $(X, \mathfrak{B}^-, \mu)$). Let X be a metrizable topological space complete with respect to some metric (such spaces are called Polish spaces), and let μ be a regular Borel measure on X with $\mu(X)=1$. If $\mathfrak{B}^-$ denotes the μ-completion of the Borel sets, then $(X, \mathfrak{B}^-, \mu)$ is a Lebesgue space. (See Section 2 of Chapter V in Parathasarathy [109].)

1.3 The Lattice of Measurable Partitions

In the last section we introduced the notion of a measurable partition of a probability space and the factor space associated with a partition. Rohlin proved that if ξ is a measurable partition of a Lebesgue space $(\Omega, \mathfrak{F}, P)$, then the factor space $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$ is also a Lebesgue space. This factor space can be a model for a specific random experiment, and if $(\Omega_\eta, \mathfrak{F}_\eta, P_\eta)$ is the factor space associated with another partition η, it may be the model of another random experiment. The probabilistic relationship between the two experiments can be given by knowing how the partitions ξ and η are related in the Lebesgue space $(\Omega, \mathfrak{F}, P)$.

Let $\mathfrak{Z}$ denote the collection of all measurable partitions of a fixed Lebesgue space $(\Omega, \mathfrak{F}, P)$. If ξ and η are in $\mathfrak{Z}$, we say ξ is refined by η and write $\xi \leqslant \eta$ if and only if each atom of ξ is an η-set, i.e., a union of atoms from η. It is easy to check that this defines a partial order on $\mathfrak{Z}$. A more useful partial order is obtained if we neglect sets of measure zero. We say ξ is refined by η (mod 0), and write $\xi \leqslant \eta$ (mod 0), provided there exists a set Z of measure zero such that $\xi' \leqslant \eta'$ on $\Omega - Z$, where ξ' and η' denote the partitions ξ and η with the points of Z removed from their atoms. In the sequel $\xi \leqslant \eta$ will always denote that ξ is refined by η (mod 0).

There is a very close relationship between measurable partitions of a Lebesgue space $(\Omega, \mathfrak{F}, P)$ and sub-σ-fields of $\mathfrak{F}$. For $\xi \in \mathfrak{Z}$ let $\xi^\wedge$ denote the collection of all $\mathfrak{F}$-measurable ξ-sets. It is easy to check that $\xi^\wedge$ is a σ-field, and clearly $\xi^\wedge \subset \mathfrak{F}$. Conversely, suppose $\mathfrak{F}'$ is a sub-σ-field of $\mathfrak{F}$. Since $(\Omega, \mathfrak{F}, P)$ is a Lebesgue space, there exists a basis $\{B_n\}$ for Ω which is complete (mod 0). Since $\mathfrak{F}$ is generated by the countable collection $\{B_n\}$ (i.e., it is countably generated) and since $\mathfrak{F}' \subset \mathfrak{F}$, $\mathfrak{F}'$ must be countably generated. Let $\{B_n'\}$ be such a countable generating family, and ξ the partition of Ω whose atoms are of the form $\bigcap \hat{B}_n'$ where $\hat{B}_n'$ is either B_n' or $\Omega - B_n'$. Then ξ is a measurable partition of $(\Omega, \mathfrak{F}, P)$ and $\xi^\wedge = \mathfrak{F}'$.

This technique defines a map of $\mathfrak{Z}$ to the collection of all sub-σ-fields of $\mathfrak{F}$ which will be one to one if we identify partitions which are equal (mod 0), i.e. such that $\xi \leqslant \eta$ (mod 0) and $\eta \leqslant \xi$ (mod 0). This will be denoted by $\xi = \eta$ (mod 0). It is easy to see that in this case the two partitions contain the same atoms if a set of measure zero is removed from Ω.