

第0章 概率论与测度论的关系

0.1 概率空间与测度空间

为了说明概率论与测度论的关系，我们首先复习一下随机事件，概率和概率空间的概念。

0.1.1 随机事件的概念

在自然界以及人类的工农业生产和经济社会活动中存在两种现象，一是在一定条件下必然发生或不发生的现象，如在一个标准大气压下，水加热到 100°C 就一定沸腾，不沸腾是不可能的，这样的现象称为确定性现象；还有一种现象，在一定的条件下，某个结果可能发生也可能不发生，或有多种可能的结果。比如，让硬币以自然落体的方式从某个高处下落。硬币落到地面后，可能正面朝上，也可能反面朝上，而且是我们无法事先预测的，但一旦硬币落到地上，结果就是明确的了。又如我们用同一把卡尺测量一个机械零件的尺寸，即使在同样条件下，由于受各种无法人为控制的偶然因素的影响，每次测量的结果不都完全一样，存在测量误差。我们称这类带有不确定性的现象为随机现象。

研究随机现象的第一步就是研究随机试验，这是最简单的随机现象。一个试验，如果满足以下三点：

(1) 可以在同样条件下重复进行；

(2) 试验的结果多于一个；

(3) 在试验前，其结果是不可知的，一般只知道是几个结果中的一个或在某个范围内，或只知道有某种可能性，而试验进行之后，结果是明确的。

我们就称这种试验为随机试验。如前面举的丢硬币和测量机械零件的尺寸的试验都是随机试验。

随机试验的结果称为样本点，常用 ω 表示，称所有可能的结果的集合为样本空间，常用 Ω 表示。如在丢硬币的试验中，样本点是“正面”和“反面”，样本空间是集合{正面,反面}。

若记， $\omega_1 = \text{“正面”}$ ， $\omega_2 = \text{“反面”}$ ，则 $\Omega = \{\omega_1, \omega_2\}$ 。

再考察复杂一些的随机试验。假设连续丢三次硬币，观察每次出现正面还是反面，这显然是个随机试验，因为试验结果在试验前是未知的，试验进行之后，结果是确定的。这个试验共有 8 个结果，即 8 个样本点：

“正正正”，“正反正”，“正正反”，“正反反”，

“反正正”，“反反正”，“反正反”，“反反反”

记

$\omega_1 = \text{“正正正”}$, $\omega_2 = \text{“正反正”}$, $\omega_3 = \text{“正正反”}$, $\omega_4 = \text{“正反反”}$

$\omega_5 = \text{“反正正”}$, $\omega_6 = \text{“反反正”}$, $\omega_7 = \text{“反正反”}$, $\omega_8 = \text{“反反反”}$

则样本空间为

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$$

在随机试验中, 如果我们所关心的结果可以表示为样本点的集合, 这个结果就被称为随机事件, 简称为事件。事件常用大写字母 A, B, C 等表示。例如在连续丢三次硬币试验中, 如果我们关心的是“恰好出现两次正面”这个事件, 则满足这样条件的样本点是

$\omega_2 = \text{“正反正”}$, $\omega_3 = \text{“正正反”}$, $\omega_5 = \text{“反正正”}$

显然 $\{\omega_2, \omega_3, \omega_5\}$ 是样本空间 Ω 的一个子集。因此一个事件 A 是样本空间 Ω 的一个子集。如果某次试验结果为 $\omega \in \{\omega_2, \omega_3, \omega_5\}$, 则事件“恰好出现两次正面”发生。反之, 若出现“恰好出现两次正面”的事件, 则必有样本点 $\omega \in \{\omega_2, \omega_3, \omega_5\}$ 。因此事件 A 发生, 当且仅当试验结果, 即样本点 $\omega \in A$ 。但需要说明的是, 样本空间的子集未必都能看作是一个事件, 这将在后面的讨论中看到。样本点本身也可以看成是事件, 这时可以把样本点看作是单点集 (单点集即为只有一个元素的集合), 称为基本事件, 而包含两个或两个以上的样本点的事件称为复合事件。另外, 不管随机试验的结果是什么, 都有 $\omega \in \Omega$, 所以样本空间可以看成是一个特殊的事件, 称为必然事件。又因为对任意 ω 都有 $\omega \notin \phi$ 成立, 这样, 空集也被看作是事件, 这也是一个特殊的事件, 称为不可能事件。

在上面的随机试验中, 样本空间的样本点个数是有限的, 现在我们考虑样本点的数目是无限的情形。考察这样的随机试验, 测量落在地面某个区域上的放射性粒子的数目, 这是一个随机试验。因为粒子的数目为整数, 所以样本空间为 $\Omega = \{0, 1, 2, \dots\}$, 这样的样本空间是无穷可数的, 或称为是可列的。又如前面测量机械零件尺寸的试验, 零件的尺寸取正实数, 所以样本点是 $(0, +\infty)$ 上的点, 样本空间为 $\Omega = (0, +\infty)$, 这个样本空间是无穷不可数的。

0.1.2 随机事件的运算

我们从随机现象出发提出了随机试验的概念, 进而引进了样本点和样本空间的概念, 又由样本点的集合出发定义了随机事件。现在我们考察随机事件的运算和它们之间的关系。

1 事件的包含

如果事件 A 的样本点都是事件 B 的样本点，则称事件 B 包含事件 A ，记作 $A \subset B$ 。若 $\omega \in A$ ，则 $\omega \in B$ ，所以 $A \subset B$ 的意义是： A 发生必然导致 B 发生。特别地， $A \subset B$ ，又有 $A \supset B$ ，则称事件 A 等于事件 B ，记为 $A = B$ 。

2 事件的对立

设 A 为一个事件，由样本空间 Ω 中的所有不包含在 A 中的样本点组成的事件称为 A 的逆事件或对立事件，记作 \bar{A} 或 A^c ，即 $\bar{A} = A^c = \{\omega: \omega \notin A\}$ 。这时，若 $\omega \in A$ ，必有 $\omega \notin \bar{A}$ ，若 $\omega \in \bar{A}$ ，必有 $\omega \notin A$ 。

3 事件的交

由同时属于事件 A 和事件 B 的样本点构成的事件称为 A 和 B 的交，记为 $A \cap B$ ，即 $A \cap B = \{\omega: \omega \in A \text{ 且 } \omega \in B\}$ 。若 $\omega \in A \cap B$ ，则 $\omega \in A$ 且 $\omega \in B$ ，故若 $A \cap B$ 发生，则 A 和 B 都发生。为简洁起见， $A \cap B$ 也记为 AB 。

4 事件的并

所有属于事件 A 和事件 B 的样本点的全体构成的事件称为 A 和 B 的并，记为 $A \cup B$ ，即 $A \cup B = \{\omega: \omega \in A \text{ 或 } \omega \in B\}$ 。若 $\omega \in A \cup B$ ，则 $\omega \in A$ 或 $\omega \in B$ 。换言之，事件 $A \cup B$ 发生表示 A 和 B 中至少有一个事件发生。

5 事件的差

由包含在事件 A 中而不包含在事件 B 中的样本点构成的事件称为 A 和 B 的差，记为 $A \setminus B$ ，即 $A \setminus B = \{\omega: \omega \in A, \omega \notin B\}$ 。若 $\omega \in A \setminus B$ ，则 $\omega \in A$ 且 $\omega \notin B$ ，故 $A \setminus B$ 发生表示事件 A 发生而事件 B 不发生。

6 事件的不相容

如果 $A \cap B = \phi$ ，则称事件 A 和事件 B 不相容。因为事件 A 与 B 的交为空集，所以若 $\omega \in A$ ，则 $\omega \notin B$ ；若 $\omega \in B$ ，则 $\omega \notin A$ ，故 $A \cap B = \phi$ 表示事件 A 与 B 不会同时发生。

显然有 $A \cap \bar{A} = \phi$, $A \cup \bar{A} = \Omega$ ，这是因为事件 A 和 \bar{A} 不会同时发生，所以 $A \cap \bar{A}$ 为不可能事件。而无论试验的结果是什么，都有 $\omega \in A$ 或 $\omega \in \bar{A}$ 成立，即 $A \cup \bar{A}$ 为必然事件。从集合论的角度看，这两的等式是自明的。

上述的两个随机事件的并和交可以推广到 n 个事件的情形。设有 n 个事件

A_1, A_2, \dots, A_n ，它们的并表示这 n 个事件中至少有一个发生，记为

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i$$

它们的交表示这 n 个事件同时发生，记为

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i$$

0.1.3 频率与概率

虽然随机事件在一次随机试验中可能出现，也可能不出现，呈现出很大的偶然性，但在大量试验中却呈现出一定的规律性。一个随机事件 A ，若在 n 次随机试验中出现 n_A 次，那么称 $f_A = n_A/n$ 为事件 A 出现的频率。当 n 充分大时，频率 f_A 就非常接近某个常数，而且 n 越大，接近程度越高，这种现象称为频率稳定性，这就是随机事件的规律性所在。

从表 0.1.1 可以看出，在抛硬币的试验中，出现正面的频率在 0.5 左右。随着次数的增加，

表 0.1.1 不同次数下抛硬币的实验结果

实验者	抛硬币次数	出现正面的次数	频率
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

频率越来越接近 0.5。一个随机事件出现的频率反应了该随机事件在大量重复的随机试验中出现的可能性的的大小，而频率稳定在某个固定的常数周围，这就启发我们用这个常数来度量随机事件发生可能性的的大小，并称它为概率。事件 A 发生的概率记作 $P(A)$ 。因为对每个事件 A ，都有一个概率 $P(A)$ 与之对应，所以概率是事件的函数。那么，概率作为随机事件的函数，它应该具备哪些性质？这就需要考察随机事件的频率具有哪些特征。

0.1.4 概率的基本性质

首先由频率的定义知

$$0 \leq f_A \leq 1, \quad f_\Omega = 1$$

这样我们就要求对任意随机事件 A 和样本空间 Ω , 有

性质 0.1.1 $0 \leq P(A) \leq 1$

性质 0.1.2 $P(\Omega) = 1$

再设 A, B 为两个不相容的随机事件, 即 $A \cap B = \Phi$, 考察 $A \cup B$ 的频率。设在 n 次随机试验中, A 出现的次数为 n_A , B 出现的次数为 n_B 。因为 A 和 B 不相容, 所以 A 和 B 不会同时出现, 由此知事件 $A \cup B$ 出现的次数为 $n_A + n_B$, $A \cup B$ 的频率为

$$f_{A \cup B} = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = f_A + f_B$$

由此我们也要求概率具有相同的性质, 即对于任意不相容的事件 A, B 有

性质 0.1.3 $P(A \cup B) = P(A) + P(B)$

一般地, 设 A_1, A_2, \dots, A_n 为两两不相容的随机事件, 即 $A_i \cap A_j = \Phi, i \neq j$, 由性质 0.1.3 和数学归纳法可知

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1) + P(A_2) + \dots + P(A_n)$$

这个性质被称为概率的有限可加性, 以上三个性质就是概率必须具备的性质。至此, 我们得到了有限可加概率的定义。

定义 0.1.1 设 P 是一个随机事件的函数, 如果它满足以下三个条件

- (1) 设 A 为一个随机事件, 则 $0 \leq P(A) \leq 1$;
- (2) $P(\Omega) = 1$;
- (3) 设 A_1, A_2, \dots, A_n 为两两不相容的随机事件, 即 $A_i \cap A_j = \Phi, i \neq j$, 则

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1) + P(A_2) + \dots + P(A_n)$$

那么就称 P 为有限可加概率。

因为事件对应于 Ω 上的集合, 所以概率也是集合的函数, 简称为集函数。

0.1.5 古典概型

古典概型是最简单的随机模型，这里所说的随机模型是指随机现象的数学模型。如果一个随机模型满足以下条件：

- (1) 样本空间中样本点的个数是有限的；
- (2) 每个样本点出现的概率是相同的。

那么就称其为古典概型。

现在考虑古典概型中概率的计算方法。假设样本空间为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ，记

$$P(\omega_1) = P(\omega_2) = \dots = P(\omega_n) = p$$

将 ω_i 看成单点集合 $\{\omega_i\}$ ，即只有一个元素的集合，则 $\{\omega_i\} \cap \{\omega_j\} = \Phi, i \neq j$ ，且

$$\{\omega_1\} \cup \{\omega_2\} \cup \dots \cup \{\omega_n\} = \Omega$$

由概率的有限可加性知

$$P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) = P(\Omega) = 1$$

$$np = 1, \quad p = 1/n$$

设事件 $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_m}\}$ ，则

$$A = \{\omega_{i_1}\} \cup \{\omega_{i_2}\} \cup \dots \cup \{\omega_{i_m}\}$$

$$P(A) = P(\omega_{i_1}) + P(\omega_{i_2}) + \dots + P(\omega_{i_m}) = mp = m/n \quad (0.1.1)$$

于是我们得到了有关古典概型的如下结论：

- (1) 每个样本点出现的概率为 $1/n$ ；
- (2) 随机事件 A 发生的概率为 A 中样本点的数目 m 与样本空间中样本点的数目 n 之比，即 $P(A) = m/n$ 。

需要说明的是，在古典概型中，概率的计算公式是基于概率的有限可加性得到的。

0.1.6 几何概型

许多实际问题不像古典概型那样，样本空间是有限的，考察看以下两个例子。

【例 0.1.1】 在一个 50 万平方公里的海域里，有表面积达 40 平方公里的大陆架贮藏着石油，假如在这海域中随意选定一点钻探，问钻到石油的概率有多少？

【例 0.1.2】 在一个瓶中有 400 毫升自来水，里面有一个大肠杆菌，从瓶中随机地取 2 毫升

水，放在显微镜下观察，问发现大肠杆菌的概率有多大？

在例 0.1.1 中，因为海域的每一点都可能被选中，所以样本空间 Ω 是 50 万平方公里的海域，经抽象化以后是一个平面。在例 0.1.2 中，因为瓶中每一滴水都可能被抽取，所以样本空间 Ω 是 400 毫升的自来水，经抽象化以后是一个体积。现在要重新考虑怎么定义概率。先看例 0.1.3，因为选点是随机的，海域的每一点被选中的可能性都相同，因而贮藏石油的面积越大，钻到石油的可能性越大。所以我们可以假设概率和储存石油海域的面积 $m(A)$ 成正比，即所求概率为

$$p_A = k \cdot m(A)$$

显然，如果整个海域都贮藏着石油，则钻到石油的概率为 1，即

$$p_\Omega = k \cdot m(\Omega) = 1$$

其中 $m(\Omega)$ 为样本空间 Ω ，即整个海域的面积。由此得

$$k = 1/m(\Omega)$$

所以

$$p_A = m(A)/m(\Omega) = 40/50000 = 1/1250$$

再来看例 0.1.2，瓶中水的每一滴水都可能被取到，而且取得水越多，观察到大肠杆菌的可能性越大，所以我们可以假设所求概率与所取水的体积成正比，即

$$p_g = k \cdot m(g)$$

这里 $m(g)$ 是所取自来水 g 的体积。如果我们把水全部取出来，那么发现大肠杆菌的概率为 1，即

$$p_\Omega = k \cdot m(\Omega) = 1$$

其中 $m(\Omega)$ 为样本空间 Ω ，即所有自来水的体积。于是得 $k = 1/m(\Omega)$ ，从而

$$p_g = m(g)/m(\Omega) = 2/400 = 1/200$$

需要说明的是，在这两个例子中，我们假设

$$p_A = k \cdot m(A), \quad p_g = k \cdot m(g)$$

这是为了数学上处理方便，并且在下面可以看到，这样定义的概率满足有限可加性和可列可加性，可以和古典概型统一处理。如果我们令

$$p_A = k \cdot m(A)^2, \quad p_g = k \cdot m(g)^2$$

就不满足可加性了。

从这两个例子出发，我们引出了一个新的定义概率的方法，其一般形式如下：

定义 0.1.2 记事件 A_g 为“在区域 Ω 中随机取一点，该点在 Ω 的一个小区域 g 中”，则定义概率

$$P(A_g) = \frac{m(g)}{m(\Omega)} \quad (0.1.4)$$

其中 $m(g)$ 和 $m(\Omega)$ 分别为 g 和 Ω 的度量，这样定义的概率称为几何概率。

几何概率有一个特点，样本空间 Ω 是一个区域，它可以是一维的、二维的、三维的，也可以是 n 维的。而点所在的区域 g 的概率只和 g 的度量（长度、面积、体积等）成正比，与 g 的位置和形状无关。

容易验证，几何概率满足概率的性质

$$0 \leq P(A_g) \leq 1, \quad P(A_\Omega) = 1$$

这里 A_Ω 表示事件“点落在整个样本空间 Ω 内”。

现在说明几何概率也满足有限可加性。假设把点所落在的区域 g 分割成 n 份

g_1, g_2, \dots, g_n ，则它们是两两不相交的，所以 $A_{g_i} \cap A_{g_j} = \phi$, $A_g = \bigcup_{i=1}^n A_{g_i}$ 。因为

$$P(A_{g_i}) = \frac{m(g_i)}{m(\Omega)}, \quad m(g) = \sum_{i=1}^n m(g_i)$$

所以

$$P(A_g) = \frac{m(g)}{m(\Omega)} = \frac{\sum_{i=1}^n m(g_i)}{m(\Omega)} = \sum_{i=1}^n \frac{m(g_i)}{m(\Omega)} = \sum_{i=1}^n P(A_{g_i}) \quad (0.1.5)$$

下面的例子说明几何概率还满足“可列可加性”。

【例 0.1.3】 考察一个纯理论性的例子。在区间 $[0,1]$ 上投一点，记 $A =$ “点落在 $[0,1/2]$ 中”，

$A_n =$ “点落在 $[1/2^{n+1}, 1/2^n]$ 中”， $n = 1, 2, \dots$ ，则

$$A = \bigcup_{n=1}^{\infty} A_n, \quad A_i \cap A_j = \phi, \quad i \neq j$$

我们用区间的长度作为区间的度量，则

$$P(A) = \frac{1}{2}, \quad P(A_n) = 1/2^{n+1}, \quad \sum_{n=1}^{\infty} P(A_n) = \frac{1}{2} = P(A)$$

若 $A = \bigcup_{n=1}^{\infty} A_n$, $A_i \cap A_j = \Phi, i \neq j$, 则 $P(A) = \sum_{n=1}^{\infty} P(A_n)$, 这时称概率满足可列可加性。

0.1.7 概率的公理化

所谓数学的公理化就是把数学的某分支中最基本的假设作为公理，其它结论都由这些公理经过演绎推导而出。概率论的公理化包括两个方面，一个是构造事件域 \mathcal{F} ，一个是定义概率测度 P 。

先考虑事件域。在前面的讨论中我们已经看到，事件是样本空间 Ω 的一个子集，在古典概型中， Ω 的每个子集也都可以看成是一个事件，且它的概率为子集中样本点的个数和样本空间中样本点的个数之比。所以在古典概型中，事件和样本空间的子集是一一对应的。但在几何概型的场合，样本空间的子集未必都能看做一个事件。比如，如果一个子集是不可度量的，即不可测的。（不可测集的例子见：郑维行，王声望，实变函数和泛函分析概要，第一册，第 47 页的定理 4.1），那么把它看做一个事件，就无法定义它的概率，所以事件域 \mathcal{F} 不能包含太多的子集，但子集也不能太少。那么 \mathcal{F} 应该满足什么条件才符合要求？我们从事件的关系和运算出发考虑这个问题。

首先， Ω 作为必然事件，它理所应当属于事件域 \mathcal{F} ，同时空集作为不可能事件也应该在 \mathcal{F} 中，即 \mathcal{F} 必须满足

(1) $\Omega \in \mathcal{F}$;

(2) $\phi \in \mathcal{F}$;

设 A 是事件域 \mathcal{F} 中的一个事件，即 $A \in \mathcal{F}$ ，因为 \bar{A} 也是一个事件， \bar{A} 也应该在 \mathcal{F} 中，即必须满足

(3) 若 $A \in \mathcal{F}$ ，则 $\bar{A} \in \mathcal{F}$;

又设 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$ ，因为 $\bigcup_{i=1}^n A_i$ 和 $\bigcap_{i=1}^n A_i$ 也是事件，所以它们也应该在 \mathcal{F} 中，

即必须满足

(4) 若 $A_i \in \mathcal{F}, i=1,2,\dots,n$, 则 $\bigcup_{i=1}^n A_i \in \mathcal{F}$;

(5) 若 $A_i \in \mathcal{F}, i=1,2,\dots,n$, 则 $\bigcap_{i=1}^n A_i \in \mathcal{F}$;

假设 $A \in \mathcal{F}, B \in \mathcal{F}$, 因为 $A \setminus B$ 也是事件, 所以要求 $A \setminus B \in \mathcal{F}$, 即

(6) 若 $A \in \mathcal{F}, B \in \mathcal{F}$, 则 $A \setminus B \in \mathcal{F}$;

由例 0.1.3 可知, 在几何概型中, 可列个事件的并也是事件, 所以我们要求:

(7) 若 $A_i \in \mathcal{F}, i=1,2,\dots$, 则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;

又因为可列个事件的并的逆事件也是事件, 且为可列个事件的逆事件的交, 所以可列个事件的交也应该是事件, 从而我们要求:

(8) 若 $A_i \in \mathcal{F}, i=1,2,\dots$, 则 $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ 。

由此可见, 作为一个事件域, 它必须满足以上 8 个性质。

然而, 由事件的运算性质可知, 只要上述条件 (1)、(3)、(7) 被满足, 那么其它条件自然也被满足。事实上由 (1) 和 (3) 知, $\phi = \bar{\Omega} \in \mathcal{F}$, 所以 (2) 成立。因为

$$A_1 \cup A_2 \cup \dots \cup A_n = A_1 \cup A_2 \cup \dots \cup A_n \cup \phi \cup \phi \cup \dots$$

所以, 若 $A_i \in \mathcal{F}, i=1,2,\dots,n$, 由上式和 (7) 知, $\bigcup_{i=1}^n A_i \in \mathcal{F}$, 故 (4) 成立。再设

$A_i \in \mathcal{F}, i=1,2,\dots,n$, 由 (3) 知, $\bar{A}_i \in \mathcal{F}, i=1,2,\dots,n$, 由德莫根定律知

$$\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i$$

又因为 $\bar{A}_i \in \mathcal{F}, i=1,2,\dots,n$, 所以由 (4) 知, $\bigcup_{i=1}^n \bar{A}_i \in \mathcal{F}$, 再由 (3) 知, $\bigcap_{i=1}^n A_i$ 作为 $\bigcap_{i=1}^n A_i$ 的

逆事件, 有 $\bigcap_{i=1}^n A_i \in \mathcal{F}$, 故 (5) 成立。因为 $A \setminus B = A \cap \bar{B}$, 由 (3) 知 $\bar{B} \in \mathcal{F}$, 由 (5) 知

$A \cap \bar{B} \in \mathcal{F}$, 从而 $A \setminus B \in \mathcal{F}$, 故 (6) 成立。若 $A_n \in \mathcal{F}, n=1,2,\dots$, 则 $\bar{A}_n \in \mathcal{F}, n=1,2,\dots$,

于是由德莫根定律和 (7) 知, $\overline{\bigcap_{n=1}^{\infty} A_n} = \bigcup_{n=1}^{\infty} \overline{A_n} \in \mathcal{F}$, 所以 $\bigcap_{n=1}^{\infty} A_n$ 作为 $\overline{\bigcap_{n=1}^{\infty} A_n}$ 的逆事件, 有

$\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$, 即 (8) 成立。于是我们就得到了事件域的公理化定义。

定义 0.1.3 设 Ω 为样本空间, 那么称 \mathcal{F} 是一个事件域, 如果它满足以下条件

- (1) $\Omega \in \mathcal{F}$;
- (2) 若 $A \in \mathcal{F}$, 则 $\overline{A} \in \mathcal{F}$;
- (3) 若 $A_n \in \mathcal{F}, n = 1, 2, \dots$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

而 \mathcal{F} 中的集合称为事件。

再考虑概率的公理化定义。在古典概型中, 概率必须满足有限可加性, 但在几何概型中, 概率要满足可列可加性。所以概率的公理化定义为:

定义 0.1.4 设 P 为事件域 \mathcal{F} 上的一个集函数, 如果它满足如下三个条件

- (1) 对任意 $A \in \mathcal{F}$, 满足 $0 \leq P(A) \leq 1$;
- (2) $P(\Omega) = 1$;
- (3) $A_n \in \mathcal{F}, n = 1, 2, \dots$, 且 $A_i \cap A_j = \Phi, i \neq j$, 则

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

那么就称 P 为一个概率测度, 而称三元组 (Ω, \mathcal{F}, P) 为概率空间。

现在我们可以考察概率论与测度论的关系了。在概率论中, 我们是在样本空间 Ω 上考察问题的, 而在测度论中, 我们是在一个基本集合 X 上考察问题的。由 X 的子集构成的集合称为 X 上的集族。设 \mathcal{F} 是 X 上的集族, 它满足条件

- (1) $X \in \mathcal{F}$;
- (2) 若 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$;
- (3) 若 $A_n \in \mathcal{F}, n = 1, 2, \dots$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

这样的集族在测度论中称为 σ -代数。显然这三个条件和概率论中事件域的条件是一致的, 所以当 $X = \Omega$ 时, \mathcal{F} 就是一个事件域。由此知, 事件域就是测度论中的一个 σ -代数。

再考虑概率测度。在测度论中， σ -代数 \mathcal{F} 上的集函数 $\mu(\cdot)$ ，即以集合为自变量的函数，如果满足条件

(1) 对任意 $A \in \mathcal{F}$ ，满足 $\mu(A) \geq 0$ ，且 $\mu(\emptyset) = 0$ ；

(2) $A_n \in \mathcal{F}, n = 1, 2, \dots$ ，且 $A_i \cap A_j = \emptyset, i \neq j$ ，则

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

那么就称为 \mathcal{F} 上的 σ -可加测度。而事件域就是一个 σ -代数，所以概率测度是满足 $P(\Omega) = 1$ 的 σ -可加测度。

在测度论中，称基本集合 X ， X 上的 σ -代数 \mathcal{F} 和 \mathcal{F} 上的 σ -可加测度 $\mu(\cdot)$ 构成的三元组 (X, \mathcal{F}, μ) 为测度空间。可见概率空间 (Ω, \mathcal{F}, P) 是测度空间的一个特例，它满足 $P(\Omega) = 1$ 。因此我们可以用测度论来研究概率论，测度论是研究概率论的有效工具。

0.2 随机变量与可测函数

0.2.1 随机变量的定义

先复习一下随机变量的概念。在许多时候，随机试验的结果都是和一个变量联系在一起的，变量的取值因随机试验的结果不同而不同，这样的变量称为随机变量。所以随机变量的取值在随机试验进行前是不知道的，并有多种取值的可能，但在随机试验完成后是确定的。

例如抛硬币，若出现正面则赢一元，出现反面则输一元。用 X 表示输赢的钱， $X = 1$ 表示赢一元， $X = -1$ 表示输一元， X 的取值为 $\{-1, 1\}$ 。在抛硬币前 X 取什么值是不知道的，只知道可能取1或-1，但抛了硬币后 X 的取值就是确定的，所以 X 是一个随机变量。

测量一个机械零件的直径 D （单位：cm），在测量前，直径 D 的值自然是不知道，只知道 D 的取值范围，而测量后， D 的值就已知了，这也是自然的，并不能说明 D 是随机的。但由于存在许多不能人为控制的因素的影响，每次测量得到的值是不同的，即有测量误差，也就是说每次试验的结果是不同的，测量机械零件的直径是一个随机试验，所以 D 是一个随机变量，它的取值范围为 $(0, +\infty)$ 上的一个区间。

因为随机变量 X 在一次随机试验中的取值取决于试验结果 ω ，所以随机变量可以看成

是样本点 ω 的函数，记为 $X(\omega)$ 。如在抛硬币的试验中

$$X(\omega) = \begin{cases} 1, & \omega = \text{“正面”} \\ -1, & \omega = \text{“反面”} \end{cases}$$

又如在测量机械零件直径的试验中

$$D(\omega) = \omega$$

若 $\omega = 5.1\text{cm}$ ，则 $D = 5.1\text{cm}$ ；若 $\omega = 4.95\text{cm}$ ，则 $D = 4.95\text{cm}$ 。这时样本点就是一个实数，而随机变量的值就是样本点本身。

一个随机变量，如果它只取有限个或可列个值，那么就被称为离散随机变量；如果它的取值为实数空间 R^1 或 R^1 的一个区域，那么就被称为连续随机变量。

就像对待随机事件那样，我们关心的是随机变量取值的概率。对离散随机变量，关心的是它取某个值的概率，对连续随机变量，关心的是它在实数空间 R^1 的某个区域 B 中取值的概率。但必须说明的是，这里所说的区域 B 不是 R^1 上的任意区域，而是指开区间 (a, b) ，闭区间 $[a, b]$ ，半开半闭区间 $(a, b]$ ， $[a, b)$ 和单点集 $\{a\}$ 以及由它们的有限并、可列并、有限交、可列交、差运算、补运算而得到的 R^1 上的子集。

又像随机事件那样，随机事件可以表示为样本空间 Ω 上的一个子集，但不是所有的子集都可以看成是一个随机事件，只有事件域 \mathcal{F} 中的集合才是随机事件。一个随机变量可以表示成样本点的函数，但不是所有样本点的函数都可以看成一个随机变量。我们所关心的是随机变量取某个值 a 或落在上述区域 B 上的概率，即 $\{\omega; X(\omega) = a\}$ 或 $\{\omega; X(\omega) \in B\}$ 的概率，所以要求 $\{\omega; X(\omega) = a\} \in \mathcal{F}$ 和 $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ ，也就是说， $\{\omega; X(\omega) \in B\}$ 和 $\{\omega; X(\omega) = a\}$ 必须为事件，这样才可以计算它们的概率。为此必需对样本空间 Ω 上的实函数加上条件，这个条件就是对任意常数 c ，有 $\{\omega; X(\omega) < c\} \in \mathcal{F}$ ，所以可以给出随机变量的如下正式定义。

定义 0.2.1 记 (Ω, \mathcal{F}, P) 为一个概率空间， $X(\omega)$ 为样本空间 Ω 上的实函数，若对任意常数 c ，都有 $\{\omega; X(\omega) < c\} \in \mathcal{F}$ ，则称 $X(\omega)$ 为随机变量。

现在说明定义 0.2.1 的合理性。考察单点集 $\{a\}$ ，和区间 (a, b) , $[a, b]$, $(a, b]$, $[a, b)$ 。先考

考虑集合 $\{\omega; X(\omega) \in [a, b]\}$ 。显然

$$\{\omega; X(\omega) \in [a, b]\} = \{\omega; X(\omega) < b\} \setminus \{\omega; X(\omega) < a\}$$

因为 $\{\omega; X(\omega) < b\} \in \mathcal{F}$ ， $\{\omega; X(\omega) < a\} \in \mathcal{F}$ ，所以由 σ -代数的性质知

$$\{\omega; X(\omega) \in [a, b]\} = \{\omega; X(\omega) < b\} \setminus \{\omega; X(\omega) < a\} \in \mathcal{F}$$

再考虑单点集 $\{\omega; X(\omega) = a\}$ 。我们先证明

$$\{\omega; X(\omega) = a\} = \bigcap_{n=1}^{\infty} \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}, n = 1, 2, \dots, \infty。$$

事实上，因为 $\{\omega; X(\omega) = a\} \subset \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}, n = 1, 2, \dots$ ，所以

$$\{\omega; X(\omega) = a\} \subset \bigcap_{n=1}^{\infty} \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}$$

又因为对任意 $a' \neq a$ ，不失一般性，假设 $a' > a$ ，记 $\varepsilon = a' - a > 0$ ，当 $n > \left\lceil \frac{1}{\varepsilon} \right\rceil + 1$ 时

$$\{\omega; X(\omega) = a'\} \not\subset \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}$$

$$\{\omega; X(\omega) = a'\} \not\subset \bigcap_{n=1}^{\infty} \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}$$

所以

$$\{\omega; X(\omega) = a\} = \bigcap_{n=1}^{\infty} \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\}。$$

因为 $\left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\} \in \mathcal{F}$ ，所以由 σ -代数的性质知

$$\{\omega; X(\omega) = a\} = \bigcap_{n=1}^{\infty} \left\{ \omega; X(\omega) \in \left[a, a + \frac{1}{n} \right) \right\} \in \mathcal{F}$$

于是

$$\{\omega; X(\omega) \in [a, b]\} = \{\omega; X(\omega) \in [a, b)\} \cup \{\omega; X(\omega) = b\} \in \mathcal{F}$$

$$\{\omega; X(\omega) \in (a, b)\} = \{\omega; X(\omega) \in [a, b)\} \setminus \{\omega; X(\omega) = a\} \in \mathcal{F}$$

$$\{\omega; X(\omega) \in (a, b]\} = \{\omega; X(\omega) \in (a, b)\} \cup \{\omega; X(\omega) = b\} \in \mathcal{F}$$

再考察以上单点集和区间的可列并（有限并可以用同样的方法处理）。首先证明

$$\left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\} = \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}$$

事实上，如果设 $\omega \in \left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\}$ ，则根据并集的定义，存在 i_0 ，使 $X(\omega) \in B_{i_0}$ ，所以

$$\omega \in \{\omega; X(\omega) \in B_{i_0}\} \subset \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}, \quad \left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\} \subset \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}$$

再设 $\omega \in \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}$ ，则同样由并集的定义知，存在 i_0 ，使 $\omega \in \{\omega; X(\omega) \in B_{i_0}\}$ ，故

$$\omega \in \left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\}, \quad \left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\} \supset \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}$$

所以

$$\left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\} = \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\}$$

记 $B_i (i=1, 2, \dots)$ 为以上区间或单点集，已经证明 $\{\omega; X(\omega) \in B_i\} \in \mathcal{F}$ ，于是由 σ -代数的定义知

$$\left\{\omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i\right\} = \bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\} \in \mathcal{F}$$

下面再证明

$$\{\omega; X(\omega) \in B^c\} = \{\omega; X(\omega) \in B\}^c$$

事实上，对任意 $\omega \in \{\omega; X(\omega) \in B^c\}$ ，根据补集的定义可知， $X(\omega) \notin B$ ，即

$\omega \notin \{\omega; X(\omega) \in B\}$ ，所以 $\omega \in \{\omega; X(\omega) \in B\}^c$ ，从而 $\{\omega; X(\omega) \in B^c\} \subset \{\omega; X(\omega) \in B\}^c$ ，

反之，设 $\omega \in \{\omega; X(\omega) \in B\}^c$ ，则由补集的定义知， $X(\omega) \notin B$ ，即 $X(\omega) \in B^c$ ，所以

$\omega \in \{\omega; X(\omega) \in B^c\}$ ，因此 $\{\omega; X(\omega) \in B^c\} \supset \{\omega; X(\omega) \in B\}^c$ ，从而知

$$\{\omega; X(\omega) \in B^c\} = \{\omega; X(\omega) \in B\}^c$$

记 B 为以上区间或单点集，则 $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ ，由 σ -代数的定义知

$\{\omega; X(\omega) \in B\}^c \in \mathcal{F}$ ，于是 $\{\omega; X(\omega) \in B^c\} = \{\omega; X(\omega) \in B\}^c \in \mathcal{F}$ 。

再考察可列交（有限交可用同样的方法处理）。利用德摩根定律有

$$\begin{aligned} \left\{ \omega; X(\omega) \in \bigcap_{i=1}^{\infty} B_i \right\} &= \left(\left\{ \omega; X(\omega) \in \bigcap_{i=1}^{\infty} B_i \right\}^c \right)^c \\ &= \left(\left\{ \omega; X(\omega) \in \bigcup_{i=1}^{\infty} B_i^c \right\} \right)^c = \left(\bigcup_{i=1}^{\infty} \{\omega; X(\omega) \in B_i^c\} \right)^c \\ &= \bigcap_{i=1}^{\infty} \{\omega; X(\omega) \in B_i^c\}^c = \bigcap_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\} \end{aligned}$$

所以当 $B_i, i=1,2,\dots$ 为上述的区间或单点集时，因为 $\{\omega; X(\omega) \in B_i\} \in \mathcal{F}$ ，由 σ -代数的性质知

$$\left\{ \omega; X(\omega) \in \bigcap_{i=1}^{\infty} B_i \right\} = \bigcap_{i=1}^{\infty} \{\omega; X(\omega) \in B_i\} \in \mathcal{F}$$

因为 $A \setminus B = A \cap B^c$ ，所以

$$\begin{aligned} \{\omega; X(\omega) \in A \setminus B\} &= \{\omega; X(\omega) \in A \cap B^c\} \\ &= \{\omega; X(\omega) \in A\} \cap \{\omega; X(\omega) \in B^c\} \end{aligned}$$

设 A, B 为上述区间或单点集，则 $\{\omega; X(\omega) \in A\} \in \mathcal{F}$ ， $\{\omega; X(\omega) \in B^c\} \in \mathcal{F}$ ，于是由 σ -代数的性质知

$$\{\omega; X(\omega) \in A \setminus B\} = \{\omega; X(\omega) \in A\} \cap \{\omega; X(\omega) \in B^c\} \in \mathcal{F}$$

记 F 为开区间、闭区间、半开半闭区间和单点集和它们的有限并、可列并、有限交、可列交、补运算以及差运算所得到的集合，这样我们就证明了每一种情况都满足 $\{\omega; X(\omega) \in F\} \in \mathcal{F}$ 在事件域中。

在测度论中，称对任意常数 c ，都满足 $\{\omega; X(\omega) < c\} \in \mathcal{F}$ 的集函数为可测函数，所以随机变量就是概率空间上的可测函数。这就是为什么在近代概率论中把随机变量定义为概率空间上的可测函数的原因。

0.2.2 一维波雷尔集与随机变量

现在我们解说随机变量的另一个等价定义。在上一节，我们定义了样本空间 Ω 上的事件域，那是因为不能把 Ω 的任意子集都看作事件。现在我们面临相似的问题，那就是不能把实数空间 R^1 上的任意子集用来定义随机变量，于是就考虑 R^1 上的 σ -代数。根据 σ -代数的定义， R^1 上的 σ -代数 \mathfrak{B} 是满足以下条件的集族，即

$$(1) R^1 \in \mathfrak{B};$$

$$(2) \text{ 若 } A \in \mathfrak{B}, \text{ 则 } A^c \in \mathfrak{B};$$

$$(3) \text{ 若 } A_n \in \mathfrak{B}, n=1,2,\dots \text{ 则 } \bigcup_{n=1}^{\infty} A_n \in \mathfrak{B}.$$

进一步地，我们考虑包含 R^1 上所有开区间的 σ -代数，即对任意 $(a,b) \in R^1$ ，都有 $(a,b) \in \mathfrak{B}$ 。因为所有 R^1 上的子集所构成的集族满足这个条件，所以这样的 σ -代数是存在的。记 $\mathfrak{B}_\alpha, \alpha \in I$ 为所有这样的 σ -代数。令 $\mathfrak{B}^1 = \bigcap_{\alpha \in I} \mathfrak{B}_\alpha$ ，我们证明 \mathfrak{B}^1 也包含了所有 R^1 上的开区间。事实上，对任意 $(a,b) \subset R^1$ ，都有 $(a,b) \in \mathfrak{B}_\alpha, \alpha \in I$ ，所以根据交集的定义， $(a,b) \in \bigcap_{\alpha \in I} \mathfrak{B}_\alpha = \mathfrak{B}^1$ ，因而 \mathfrak{B}^1 包含所有开区间。我们再证明 \mathfrak{B}^1 为一个 σ -代数。实际上， $R^1 \in \mathfrak{B}_\alpha, \alpha \in I$ ，根据交集的定义， $R^1 \in \bigcap_{\alpha \in I} \mathfrak{B}_\alpha = \mathfrak{B}^1$ 。又设 $A \in \mathfrak{B}^1 = \bigcap_{\alpha \in I} \mathfrak{B}_\alpha$ ，则 $A \in \mathfrak{B}_\alpha, \alpha \in I, A^c \in \mathfrak{B}_\alpha$ ，故 $A^c \in \bigcap_{\alpha \in I} \mathfrak{B}_\alpha = \mathfrak{B}^1$ 。再设 $A_n \in \bigcap_{\alpha \in I} \mathfrak{B}_\alpha = \mathfrak{B}^1, n=1,2,\dots$ ，则 $A_n \in \mathfrak{B}_\alpha, n=1,2,\dots, \alpha \in I, \bigcup_{n=1}^{\infty} A_n \in \mathfrak{B}_\alpha, \alpha \in I$ ，因而 $\bigcup_{n=1}^{\infty} A_n \in \bigcap_{\alpha \in I} \mathfrak{B}_\alpha = \mathfrak{B}^1$ 。我们称 \mathfrak{B}^1 为一维波雷尔 (Borel) σ -代数，称 \mathfrak{B}^1 中的子集为波雷尔 (Borel) 集。

设 $a \in R^1$ ，则 $\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$ ，这是因为对任意 $a' \neq a$ ，记 $\varepsilon = |a' - a|$ ，则

当 $n > \left[\frac{1}{\varepsilon}\right] + 1$ 时， $a' \notin \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$ ，所以只有 $\bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n}\right) = \{a\}$ 。因为

$\left(a - \frac{1}{n}, a + \frac{1}{n}\right) \in \mathfrak{B}^1$ ，所以 $\{a\} \in \mathfrak{B}^1$ ，于是

$$[a, b] = (a, b) \cup \{a\} \cup \{b\} \in \mathbb{B}^1, [a, b) = (a, b) \cup \{a\} \in \mathbb{B}^1, (a, b] = (a, b) \cup \{b\} \in \mathbb{B}^1$$

由此知，一维波雷尔 σ -代数包含了所有 R^1 上的开区间、闭区间、半开半闭区间和单点集，以及由它们的有限并、可列并、有限交、可列交、补运算和差运算所得到的子集。这样我们就可以把 \mathbb{B}^1 上的波雷尔集 B 作为我们要考虑的区域，要求对任意 $B \in \mathbb{B}^1$ ， $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ ，这样又得到了随机变量的另一个定义。

定义 0.2.2 设 (Ω, \mathcal{F}, P) 为概率空间， \mathbb{B}^1 为一维波雷尔 σ -代数， $X(\omega)$ 为 Ω 上的实函数，如果对任意 $B \in \mathbb{B}^1$ ，有 $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ ，那么称 X 为 (Ω, \mathcal{F}, P) 上的随机变量。

现在我们来证明以上两个随机变量的定义是等价的。事实上，设 X 是定义 0.2.1 所定义的随机变量， B 为开区间、闭区间、半开半闭区间以及由它们的有限并、可列并、有限交、可列交、差运算和补运算所得到的子集，则由前面的讨论知对 $B \in \mathbb{B}^1$ ， $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ ，符合定义 0.2.2 的条件。反之若 X 为定义 2.2 所定义的随机变量，则

$(-\infty, c) = \bigcup_{n=1}^{\infty} (-n, c) \in \mathbb{B}^1$ ，所以， $\{\omega: X(\omega) < c\} = \bigcup_{n=1}^{\infty} \{\omega: X(\omega) \in (-n, c)\} \in \mathcal{F}$ ，满足定义 2.1 的要求。

所以在测度论中，常把可测函数定义为满足对任意 $B \in \mathbb{B}^1$ ，有 $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ 的集函数，因而在近代概率论中，随机变量 X 也常被定义为满足 $\{\omega; X(\omega) \in B\} \in \mathcal{F}$ 的集函数。

0.2.3 一维波雷尔可测函数与随机变量的函数

直观地说，设 X 是一个随机变量， $y = g(x)$ 是一个函数。因为 X 可以表示为样本点 ω 的函数 $X(\omega)$ ，所以将 $X(\omega)$ 代替 $g(x)$ 中的自变量 x ，就得到函数 $Y(\omega) = g(X(\omega))$ ，于是 Y 也是一个样本点 ω 的函数，因而也是随机变量，它是随机变量 X 的函数。但从理论上讲，并不是任意函数都可以用来构造一个随机变量的函数。设概率空间为 (Ω, \mathcal{F}, P) ，根据随机变量的定义，如要使 $Y(\omega) = g(X(\omega))$ 也是一个随机变量，这就要求对任意实数 c 使得

$$\{\omega; Y(\omega) = g(X(\omega)) < c\} \in \mathcal{F}$$

即 $\{\omega; Y(\omega) = g(X(\omega)) < c\}$ 是一个事件，有确定的概率。或等价地，要求对任意一维波

雷尔集 $B \in \mathcal{B}^1$ ，有

$$\{\omega: Y(\omega) = g(X(\omega)) \in B\} \in \mathcal{F}$$

现在我们对此作进一步的讨论。首先我们证明

$$\{\omega; Y(\omega) = g(X(\omega)) \in B\} = \{\omega; X(\omega) \in B_g\}$$

这里 $B_g = \{x; g(x) \in B\}$ ，即 B_g 是使 $g(x) \in B$ 的所有 x 构成的集合。取任意

$\omega \in \{\omega; Y(\omega) = g(X(\omega)) \in B\}$ ，记 $x = X(\omega)$ ， $y = g(x)$ ，显然， $y = g(x) \in B$ ，所以

$x \in \{x; g(x) \in B\} = B_g$ ， $\omega \in \{\omega; X(\omega) \in B_g\}$ 。从而

$$\{\omega; Y(\omega) = g(X(\omega)) \in B\} \subset \{\omega; X(\omega) \in B_g\}$$

再任取 $\omega \in \{\omega; X(\omega) \in B_g\}$ ，记 $x = X(\omega)$ ，则 $y = g(x) \in B$ ，所以

$\omega \in \{\omega; Y(\omega) = g(X(\omega)) \in B\}$ ，从而

$$\{\omega; Y(\omega) = g(X(\omega)) \in B\} \supset \{\omega; X(\omega) \in B_g\}$$

故

$$\{\omega; Y(\omega) = g(X(\omega)) \in B\} = \{\omega; X(\omega) \in B_g\}$$

如果 B_g 也是波雷尔集，则因为 X 为随机变量，由随机变量的定义知，

$\{\omega: X(\omega) \in B_g\} \in \mathcal{F}$ ，这样 $\{\omega: Y(\omega) = g(X(\omega)) \in B\} \in \mathcal{F}$ ，故 $Y(\omega)$ 是 $X(\omega)$ 的函数且为

概率空间 (Ω, \mathcal{F}, P) 上的可测函数，也是一个随机变量。于是我们就得到了如下结论：

设 X 是一个概率空间 (Ω, \mathcal{F}, P) 上的随机变量， $g(x)$ 是一个实函数，它满足对任一波雷尔集 $B \in \mathcal{B}^1$ ， $B_g = \{x; g(x) \in B\}$ 也是波雷尔集，那么 $Y(\omega) = g(X(\omega))$ 为随机变量 X 的函数。

在测度论中，如果一个一维实函数 $g(x)$ ，它满足对任一波雷尔集 $B \in \mathcal{B}^1$ ，

$B_g = \{x; g(x) \in B\}$ 也是波雷尔集，则被称为一维波雷尔可测函数。由上面的讨论知，随机

变量关于一维波雷尔可测函数的复合函数也是随机变量。这就是近代概率论中只考虑可测函数与随机变量的复合函数的情形的原因。