【测度论】概率论与测度论之间联系的通俗解释(一)

张敬信
哈尔滨工业大学 基础数学博士

前言

这是我以前回答知乎问题的内容：

> **概率是什么？Sigma algebra, Borel field 是什么意思，意义何在？**
> 800 赞同 · 57 评论 回答

单独作为文章发布出来。另外，这个讨论并不完全，还有概率论中的各种分布又是怎么回事？数学期望如何是勒贝格积分？有时间再做一次讨论，所以这算是第I 篇吧。

正文：

测度论是概率论的理论基础，所以概率中的一些概念抽象化就是对应的测度论中的概念。

概率是要度量"事件发生的可能性"的大小，事件的抽象化描述就是集合，需要考察"事件的全体"，对应到测度论就是"集合系"。"事件发生的可能性"是对事件的一种度量，对应到测度论就是"集合的测度"。

不是每个事件都可以定义其概率（发生的可能性的大小）的，对应的就是不是每个集合都可以定义测度，可以定义测度集合就是可测集。同时，事件必然要涉及到事件的组合运算（复杂事件是可由基本事件表示出来），对应的就是集合的交、并、差、余、极限的运算到复杂集合，所以又需要保证做可列次这些运算不能超出全体范围（即可测集的范围要足够大，以保证集合的可列次交、并、差、余、极限的运算，之后还在里面）

那么什么样的集合系，才能保证其中的集合是可测集（可以定义测度，又对那些运算封闭）呢？测度论中讲了，只要集合系是σ-代数（也叫σ-域）就可以了。σ-代数的基本定义是：

1. 全集在里面；

2. 里面每个集合的余集在里面；

3. 里面任意可列个集合的并集在里面。

有了这三条基本定义，就可以推出：空集、可列次交、并、差、上限集、下限集运算之后都能在里面。就满足需要了。

所以，集合X + 该集合上的一个σ-代数$\mathcal{F}$，$(X,\mathcal{F})$就是一个可测空间了，即可以定义测度的空间（$\mathcal{F}$中任一集合都可以定义其测度（某种度量））。进一步再定义了测度μ，那么$(X,\mathcal{F},\mu)$就是测度空间。

对应到概率论中，样本空间Ω，事件域$\mathcal{F}$（是个σ-代数），概率测度P，放一起$(\Omega,\mathcal{F},P)$就是概率测度空间。概率测度P是满足特殊要求的一种测度：$P(\Omega)=1$.

Borel Feild就是Borel σ-代数，表示实数轴$\mathbb{R}$上的σ-代数，可由实轴上的所有开集生成（的σ-代数），也可由实数轴上所有的$(-\infty,a]$这样的区间生成（的σ-代数），是相等的。按σ-代数前面说的，实数轴上开集、闭集的至多可列次交、并、差(余)、上限集、下限集、极限集的运算，都超不出该Borel σ-代数的范围。

Borel σ-代数（用$\mathcal{B}_{\mathbb{R}}$表示）有什么用？其实概率论中的随机变量，对应测度论中的可测函数，而可测函数就是从可测空间$(X,\mathcal{F})$到$(\mathbb{R},\mathcal{B}_{\mathbb{R}})$的可测映射：即$\mathcal{B}_{\mathbb{R}}$中的任一集合在该映射下的原像都属于$\mathcal{F}$（即都是$X$上的可测集）。

怎么统一用一种简单的
）到实数轴的映射（一

一对应）就可以了，这种映射就是随机变量。有了它，基本事件映射到实数轴上就是的基本区间，基本事件经过运算生成的复杂事件，映射到实数轴上就是实数轴上Borel σ-代数中的集合。

因为有了这个对应关系，要度量“事件发生的可能性的大小”（即概率测度），只要度量“实数轴上Borel σ-代数中的集合”就可以了（前面说了$\mathcal{B}_{\mathbb{R}}$因为是σ-代数是可以定义测度的，给$\mathcal{B}_{\mathbb{R}}$中的集合定义概率测度就行了）。

所以，随机变量的测度论语言定义是这样的：设$(\Omega, \mathcal{F}, P)$为概率测度空间，若对实数轴上Borel σ-代数中的任一集合（称为Borel集）B，都有$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$, 则称$X(\omega)$为随机变量，也简记为X。

总之，随机变量就是建立了“随机事件”到“实数轴上Borel σ-代数”的一种对应，并且保证了建立了这种对应的随机事件都是可以定义概率测度的。

既然随机事件$\{\omega \in \Omega : X(\omega) \in B\}$属于$\mathcal{F}$，那么可以有概率（测度），即 $P\{\omega \in \Omega : X(\omega) \in B\}$ 是有意义的，为了简单，概率中就记 $P\{\omega \in \Omega : X(\omega) \in B\} = P\{X \in B\}$。

特别地，若取$B = (-\infty, x]$, 则事件$\{X \in B\}$的概率

$$P\{X \in B\} = P\{X \leq x\} := F(x)$$

就定义成随机变量X的分布函数。因为对任意的区间$(a, b]$, 都可表示成

$$P\{X \in (a, b]\} = P\{a \leq X \leq b\} = P\{X \leq b\} - P\{X \leq a\} = F(b) - F(a)$$

进而，由这样的区间经过至多可列次交、并、差运算的复杂的实数轴上的Borel集都可以用$F(x)$给出其概率。

当然，随机变量也可以定义为从样本空间到平面$\mathbb{R}^2$上的映射，就是二元随机变量。

作者：张敬信

编辑于 2022-08-31 23:06

概率论 测度论