随机变量的组合学

$$\sum_{i=-N}^{dN} i \times A(i) = 0.$$

$$N \times A(-N) + \cdots + 1 \times A(-1)$$

王彬◎著

中国出版集团
现代出版社

随机变量的组合学

王彬◎著

中国出版集团
现代出版社

图书在版编目（CIP）数据

随机变量的组合学 / 王彬著. -- 北京 ：现代出版社，2016.8

ISBN 978-7-5143-5252-8

Ⅰ. ①随… Ⅱ. ①王… Ⅲ. ①随机变量一组合数学 Ⅳ. ①O211.5②O157

中国版本图书馆CIP数据核字(2016)第180147号

著　　者 王　彬

责任编辑 杨学庆

出版发行 现代出版社

通讯地址 北京市安定门外安华里504号

邮政编码 100011

电　　话 010-64267325 64245264（传真）

网　　址 www.1980xd.com

电子邮箱 xiandai@cnpitc.com.cn

印　　刷 三河市金泰源印务有限公司

开　　本 710*1000　1/16

印　　张 8.5

版次印次 2016年8月第1版 2016年8月第1次印刷

标准书号 ISBN 978-7-5143-5252-8

定　　价 28.00元

著者的话

在保险业和金融业快速发展的今天，风险管理对一个公司的发展具有重要意义。在很多场景中，一个公司持有某个投资组合，或者同时面临多个风险，我们会关心总和的风险可能是什么。因此聚合风险的研究是精算学研究领域的一个重要方向。聚合风险的典型场景是考虑投资组合的风险。在一些投资组合模型 $S = X_1 + X_2 + \cdots + X_n$ 中，人们对单个的风险 X_k 拥有比较多的信息，比较容易统计和建模，可以准确地估算出 X_k 的边缘分布，但是对于这些风险之间的相依结构，往往由于数据不同步等一些原因很难估计。从某种意义上看，联合分布或者说 copula 的复杂度要远高于边缘分布的复杂度,(类似于 $[0,1]^n$ 上的函数和 $\mathbb{R}$ 上的函数的复杂度的区别)，并且维数越大复杂度越高。实际中对于相依结构方面的准确建模确实是比较困难的，尤其是很多时候不同 X_k 的数据是不同步的，例如 X_1 与 X_2 的数据来自不同时间或不同的市场，每次只拿到单个数据，无法同时拿到一组数据，就不容易对风险或者随机变量的联合分布或者相依结构建模。

因此研究者们关注这样一个问题：在已知单个风险的边缘分布 $X_k \sim F_k$，但相依结构不确定的情况下，聚合风险 $S = X_1 + X_2 + \cdots + X_n$ 可能是什么？可能有什么样的分布？在一些风险测度下的取值范围是什么？

本书就关心这一类问题，即在相依结构不确定下考虑随机变量的加和可能是什么样的分布的问题。尤其关心什么时候随机变量的加和可以是一个常数，我们称之为“混合”。我们也把这类问题叫作随机变量的组合学。这些随机变量的组合学的问题表面看

很好理解，似乎解决起来比较容易，但实际上难度很大，因为对于多元分布来说相依结构确实有很高的信息量或者说很高的复杂度，远远高于边缘分布的信息量，因而相依结构有很高的自由度，所以要确定出随机变量的加和是否可能满足目标分布这个问题，如果要一般性的解决是非常困难的。

本书主要解决了一组密度单调的边缘分布能否混合的问题，给出了充分的必要条件。并以此为出发点给出了很多类分布能否混合的条件。随机变量的组合学的问题不仅仅是风险理论或者概率论方向的问题，它们本身也是很有意思的组合数学问题。由于解决相依结构问题时，主要的想法都是在设计随机变量的取值之间的最优匹配，也就是如何把它们按某种方式分好组做好搭配，使得每组的数据加和有一些好的性质。因此把相依结构问题称之为随机变量的组合学也是名副其实的。

本书的基础要求不高，基本上是比较初等的，读者只需要最基础的概率论知识就可以看懂。书中考虑的很多问题作为组合数学或者最优化这些方向来看也是很有意思的。如果想学习组合数学的技巧，或者数学归纳法的使用技巧，本书也提供了一些优秀的例子。

本书的写作与出版得到中国国家自然科学基金的青年项目（项目号：11501017）的支持，在此表示感谢。

王　彬

2016 年 6 月　于北京工商大学

本书中主要定义和相关记号

我们用 $(\Omega, \mathcal{F}, \mathbb{P})$ 表示一个概率空间。一个随机变量 X 是指 $\Omega \to \mathbb{R}$ 的映射，随机变量 X 的分布函数 F 是 $R \to \mathbb{R}$ 的函数，具体定义是：$F(a) = \mathbb{P}(X \leqslant a), a \in \mathbb{R}$。

一个随机变量 X 的期望 EX 被定义为 $EX = \int_{-\infty}^{+\infty} x \mathrm{d}F(x)$，期望 EX 经常被称为均值，在本书中有时也被形象的称为重心。

一个分布 F 或者一个随机变量 X 的支集是指密度函数大于 0 的点的集合。可以理解为：

$$Support(X) = \{a \in \mathbb{R} : \mathbb{P}\{a - \varepsilon < X < a + \varepsilon\} > 0 \ \forall \varepsilon > 0\}$$

$$Support(F) = \{a \in \mathbb{R} : F(a + \varepsilon) - F(a - \varepsilon) > 0 \ \forall \varepsilon > 0\}$$

我们对一个分布函数 F，用 $\mu(F)$ 来表示服从 F 的随机变量的期望，用 $\sigma^2(F)$ 来表示服从 F 的随机变量的方差，即若随机变量 $X \sim F$，则

$$\mathbb{E}(X) = \mu(F), \mathrm{Var}(X) = \sigma^2(F)$$

目　录

第一章

序言

§1.1 问题的提出

在随机变量的相依结构领域，有一个经典问题：

$$\min_{X_i\sim F_i}\operatorname{Var}(X_1+\cdots+X_n)$$

这个问题就是如何使得给定边缘分布的随机变量的加和的方差最小。如果要求方差的最大值是比较容易的，只需这些随机变量 $X_1,\cdots,X_n$ 全部正相关（即当一个随机变量增大时，其他的随机变量也增大，这也叫做同单调）即可。如果是 $n=2$ 只有两个随机变量的情况也比较简单，只需让 X_1,X_2 反相关（即当一个随机变量增大时，另一个随机变量一定减小，这也叫做反单调）即可。但是对于三个或三个以上的随机变量，不可能让每两个变量都是反相关的，（类似于负负得正，反相关的反相关就是正相关了）。变量和的方差的大小，实际上就是这一组随机变量的反相关性的一种度量，什么状态下加和能有最小的方差，即某种意义上的最好的反相关性，这个在一般情况下是一个很困难的问题。

事实上，如果想让和的方差最小，最理想的情况是 $X_1,\cdots,X_n$ 的加和是一个常数，这时方差就是 0，达到了理论上的最小值。本书中把这种和为常数的状态叫做混合，如果边缘分布相同则叫作完全混合，如果边缘分布不同则叫作联合混合。这就很自然地引出了一个问题：什么样的一组边缘分布可以存在 n 个随机变量具有这种最好的反相关性？使得加和为常数从而能够混合。在已知可以混合时，如何设计出这种相关结构？这些都是很有意思的问题。

在很多情形下，加和方差等于 0 这个理论上的最小值可能是达不到的，就是说不存在 $X_1,\cdots,X_n$ 之间的一个相关结构使得加和为常数，那么在这种情况下如何使得方差最小也是个很有意思的问题。

随机变量加和方差的最小值问题，实际上是考虑变量和的平方之期望如何最小，这里如果把平方函数换成别的凸函数，问题的性质也基本是一

样的，由凸函数的琴生不等式可知加和的凸函数的期望仍然是在随机变量混合时取到最值。

§1.2 问题的研究历程和现状

在 20 世纪六七十年代，Fishman[1]，Hammersley 和 Handscomb[2] 就开始研究和方差最小问题。20 世纪八十年代，Gaffke 和 Rüschendorf[3] 解决了均匀分布的情形。Rüschendorf 和 Uckelmann[4] 在 2002 年提出了让变量和为常数的说法，并证明了单峰对称分布对于不小于 2 的阶数都是可混合的。

大约在 2010 年初，我的合作者王若度向我介绍了 $\min E(XYZ)$ 的问题，就是考虑三个 $[0,1]$ 上的均匀分布之间的相关性，使得乘积的期望最小。在做了一些初步探索之后，这个问题归结为中间的某一段分布能不能达成混合结构。幸运的是问题的答案是肯定的，我们最终解决了 $\min E(XYZ)$ 的最小值问题。

接下来，我们研究了更一般的分布能不能有混合结构的问题，彻底解决了在密度单调分布情形下的完全混合问题，即只要满足均值条件则一定能够混合。我们以此为出发点，之后的很多混合问题都会变得比较容易。

后来，我又研究了密度函数是凸函数的混合问题，证明了它们对 $n \geqslant 3$ 的阶数可以完全混合。此外还研究了类似于傅里叶展开式的径向对称函数，以及比较均匀的连续分布函数，证明了它们也都具有较好的混合性质。

2014 年，我和合作者王若度又解决了密度单调分布类的联合混合问题。

然而完全混合与联合混合领域的问题仍有很多，这个领域还有广阔的空间需要探索。

§1.3 问题的意义

完全混合和联合混合是一个很新的领域，这方面的问题也是很有意思的问题，它融合了很多方面的意义。

首先这是一个概率问题，是随机变量相依结构的问题。

如果从联合分布函数或者联合密度函数的角度来看，这实际上是 n 元分布函数在一些条件下，如何使一个泛函达到最大值或最小值的问题。(这里我们考虑的随机变量加和的方差是联合分布函数的泛函的很简单的例子)。从这个角度讲，这是一个泛函问题。

在考虑如何完全混合时，我们的问题在本质上是要设计出随机变量之间合适的相关结构，并且限制比较单纯，就是只需满足固定的边缘分布即可。这实际上是在考虑随机变量的组合结构，是一个组合数学问题。在各种看问题的角度中，组合数学的角度是我最喜欢的。

另外从设计的角度看，这也是一个规划问题，即对随机变量进行合理的匹配，这也是最优化问题。

最后，完全混合与联合混合理论在风险理论中具有巨大的应用价值，它可以解决很多风险测度方面的极值问题。因为我们在实际中，很可能对每种风险的分布都了如指掌，但对各风险之间的相关性方面的信息却知之甚少。而混合问题是研究如何使得各个风险能达到最好的反相关性，这能解决很多风险测度的最大值或最小值问题。

§1.4 本书的主要结构

本书介绍了完全混合与联合混合这些概念的来龙去脉，以及相依结构这一领域的一些研究现状。

在第一章，引入问题，提出概念，介绍了问题的研究状况和意义。

在第二章，分析了完全混合方面的各种性质，例如分布加性、指数加性、变量加性、收敛性以及一些必要条件。用这些性质分析了一些常见分布的可混合性。以期待读者对完全混合有一些理解或感性认识。

在第三章，提出了完全混合基本定理，即单调密度情形的混合问题。采取从离散到连续的逼近方法，利用归纳法构造，得到了离散单调分布的组合匹配方法。再逼近连续分布，完全解决了随机变量的密度函数单调时的混合问题，得到了这类分布能否完全混合的充要条件。

在第四章，解决了一些其他情形下的完全混合问题。首先证明了密度函数是凸函数的分布一定可以完全混合，其次解决了径向对称分布情形的混合问题。还有较均匀分布，即只要密度函数有下界就一定保证对于较大的 N 可以 N 阶混合。

在第五章，利用完全混合解决了随机变量的和的凸函数期望的最值问题，进一步的也解决了一些风险分析中的一些最值问题。

在第六章，引入了联合混合的概念，介绍了一些例子和基本性质，以及给出了一组均匀分布能否联合混合的充要条件。

在第七章，主要介绍了一组密度单调分布能否联合混合的问题，给出了充要条件。并给出了一个很长的证明。

在第八章，我们考虑当不能混合时，即一组随机变量加和不能是常数时，如何使得加和的波动很小，即能不能做到近似混合的问题。

此外，本书在最后一章还提出了混合方面的一些未解决问题，比如：混合中心是否一定唯一？多维情形的混合问题，多元随机变量问题，以及由存在性问题到更一般的构造性问题。

第二章

完全混合的性质

§2.1 完全混合的概念

对于 N 个随机变量有相同的分布 F，如果它们可以通过合适的相依结构，使得加和是一个常数，那么我们称这个分布 F 是可以 N 阶完全混合的，具体的定义如下：

定义 2.1 n 是一个正整数，对于一个 $\mathbb{R}$ 上的概率分布 F，如果存在 n 个随机变量 $X_1, \cdots, X_n \sim F$ 使得 $X_1 + \cdots + X_n$ 是一个常数，则称分布 F 是 n 阶可完全混合的（或简称为 n 阶可混合）。其中 $(X_1, \cdots, X_n)$ 的联合分布称为一个 n 阶完全混合结构，*(n-**complete mix**)*。

其实完全混合这个概念以前有人提出过，Rüschendorf 和 Uckelmann 在[4] 中提出了具有常数和的随机变量的概念（random variables with constant sums)，我们这里把它称为“完全混合”，我觉得比较形象，这个意思理解起来就是对一些随机变量设计好的相关结构，使它们充分混合，加和为常数。

§2.2 完全混合的基本性质

这一节我们陈述和证明一些关于完全混合的基本性质。

性质 2.2 *(仿射变换下的不变性)*
假设随机变量 X 的分布是 n 阶可混合的，那么对于任意常数 a 和 b，随机变量 $aX + b$ 的分布也是 n 阶可混合的。

因为完全混合考虑的是一些随机变量加和的问题，是比较线性化的。放射变换就是线性变换，所以分布的完全混合性质在仿射变换下是保持的。

性质 2.3 *(完全混合的中心)*

假设分布 F 是 n 阶可混合的，$X_i \sim F$，$i = 1, \cdots, n$ 和 $\mu = \frac{1}{n}(X_1 + \cdots + X_n)$ 是一个常数，我们称 μ 是 F 的一个中心。如果 F 满足弱大数律，那么中心 μ 是唯一的。进一步的，如果对 $\mu(F)$ 存在，那么 $\mu = \mu(F)$。

证明 2.4 假设 $E(X_1) = \cdots = E(X_n) = \mu(F)$ 存在，对等式 $\mu = \frac{1}{n}(X_1 + \cdots + X_n)$ 的两边取期望，我们将得到 $\mu = \mu(F)$。下一步我们假设分布 F 满足弱大数律，我们取随机向量 $(X_1, \cdots, X_n)$ 的一族互相独立的复制品，记为 $\{(X_{1,i}, \cdots, X_{n,i})\}_{i=1}^{\infty}$，因为满足弱大数律，我们计算它们的平均：

$$
\begin{aligned}
n\mu &= \frac{1}{k}\sum_{i=1}^{k}(X_{1,k} + \cdots + X_{n,k}) \\
&= \frac{1}{k}\sum_{i=1}^{k} X_{1,k} + \cdots + \frac{1}{k}\sum_{i=1}^{k} X_{n,k} \\
&= nE(X_1 \mathbf{1}_{\{|X_1| \leqslant k\}}) + o_p(1)
\end{aligned}
$$

当 k 趋于无穷时，我们得到 $E(X_1 \mathbf{1}_{\{|X_1| \leqslant k\}}) \to \mu$，并且 μ 是唯一的。

不过一般情况下，完全混合的中心不一定是唯一的，至少我现在还无法证明中心的唯一性。我们猜想中心唯一，但是有时候感觉可能有一些奇怪的分布可能使得中心不唯一。目前我倾向于猜想可能不成立。

性质 2.5 *(第一加性，分布角度的加性)*

假设分布 F 和 G 都是 n 阶可混合的，并且拥有一致的中心 μ，那么对于任意的 $\lambda \in [0, 1]$，$\lambda F + (1 - \lambda)G$ 是 n 阶可混合的，并且拥有相同的中心 μ。

证明 2.6 假设 $X_1 + \cdots + X_n = n\mu$，$X_i \sim F$ 并且 $Y_1 + \cdots + Y_n = n\mu$，$Y_i \sim G$，$i = 1, \cdots, n$。我们取 Z 是一个独立于 $\{X_i\}_{i=1}^n$ 和 $\{Y_i\}_{i=1}^n$