

田传俊 著



频率测度论



科学出版社
www.sciencep.com

频率测度论

田传俊 著

科学出版社

北京

内 容 简 介

本书是系统研究数列伪随机性的一门基础理论,它与概率论既有密切联系,又有明显区别。全书共分为14章,所有内容可分为两大部分:一部分是与概率论相平行的内容,包括伪随机事件,全频率公式和频率 Bayes 公式,频率分布,频率密度,二项分布和正态分布,边际分布和独立性,期望和方差,协方差、相关系数、条件期望、线性回归和矩,频率熵和互信息,频率大数定律和中心极限定理等;另一部分是与概率论不平行的内容,包括频率收敛性和频率振动性,数列的积分,分布混沌性,自相关数列和互相关数列,随机模拟等。

本书内容精练,语言朴实,主要内容参照了概率论的研究内容和方法,是数列伪随机性应用领域的基础理论。本书适合具有基本的微积分、概率论和差分方程知识的本科生、研究生、理工科教师和各类科研工作者阅读。

图书在版编目(CIP)数据

频率测度论/田传俊著. —北京:科学出版社,2010

ISBN 978-7-03-028791-5

I. 频… II. 田… III. 频率-数学方法-研究 IV. O45

中国版本图书馆 CIP 数据核字(2010)第 166762 号

责任编辑:余 丁 / 责任校对:桂伟利

责任印制:赵 博 / 封面设计:耕 者

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2010年9月第 一 版 开本: B5 (720×1000)

2010年9月第一次印刷 印张: 14 3/4

印数: 1—2 500 字数: 282 000

定价: 60.00 元

(如有印装质量问题,我社负责调换)

序 言

我们学习概率论通常都是从“频率”这个概念开始的。频率指的是在相同条件下重复若干次实验,某事件发生的次数与实验的总次数之比值。频率在相当程度上反映了一个随机事件发生的可能性的。大小。但是,频率本身是随机的,在实验之前无法确定,人们不能用它来严格准确地刻画事件发生的可能性的。大小,因此只能在相同条件下大量地重复同样的实验,希望能得出一个非随机的常数,用以刻画该事件发生的可能性的。大小。这样,从理论上来说,如果这个在相同条件下不断重复的实验给出一系列收敛到同一个常数的频率序列的话,这个极限常数值就可以用来作为这个事件出现的“概率”。在应用科学领域中,频率和概率现在都是用来作为研究随机事件发生的可能性大小的常用的特征量。在历史上,概率论作为一门严格的数学学科是 1933 年开始由著名的俄罗斯数学家 A. N. Kolmogorov 采用频率的概念通过 Lebesgue 积分而建立在现代数学测度论的基础之上的。

不妨设想一下,在上面提到的相同条件下的重复实验之中,假定实验的总次数可以有无穷多,然后每次实验观察到的次数可以用一个自然数来表示,那么我们就得到一个由一些自然数构成的集合在整个自然数集中所占的比例问题。

由深圳大学田传俊教授编著的这本《频率测度论》所依赖的一个基本概念是“频率测度”,它就是为研究由一些自然数构成的集合在整个自然数集中所占的比例问题而引入的一个数学概念。事实上,这个概念在数学文献中曾被称为“渐近密度”或“自然密度”,后来更普遍地被称为“频率测度”。注意到这种频率测度具有有限可加性,而概率测度具有无限可加性,它们之间自然就有着很多本质上的区别,因此频率测度就需要专门地加以讨论和研究,从而该书就应运而生了。

据我所知,该书是现今国内也是世界上第一本关于频率测度的专著,是为相关领域的科技工作者、特别是年轻科研人员和研究生们编写的一本很有价值的参考书,其中包括了不少作者本人在这个领域的科研成果。

该书首先利用频率测度和概率测度之间的一些共同特性,建立了与概率论相平行的一些基本概念和方法,包括伪随机事件的概念与性质、全频率公式和频率 Bayes 公式、数列的频率分布、离散数列和连续数列、二项分布和正态分布、数列的期望和方差、数列间的边际分布和独立性、数列间的协方差和相关系数以及条件期望和最小二乘法、数列的频率熵和互信息、数列的特征函数和母函数、频率大数定律和中心极限定理,等等。

同时,根据频率测度自身的特点,该书还建立了一些独特的内容,包括数列的频率收敛性和频率振动性的概念和性质以及判定准则、数列的积分、随机分布的模拟、离散系统的分布混沌性、数列的自相关和互相关、数列的单向独立性、数列的相关性与独立性检验,等等。

该书仅仅要求读者具有基本的微积分、概率论和差分方程等数学知识背景,因而适合于理工科的研究生和高年级本科生、理工科教师和相关领域的科研工作者阅读和参考。从现在数理学科的演化和应用科技的发展情况来看,频率测度论作为一门新的理论将有可能在需要数列分析方法的一些应用科技领域中逐步发挥重要的作用,特别是其中的离散和连续数列的相关知识将有可能成为频率测度论应用于实际之中的一个重要理论基础和分析方法。

特别值得提及的是,该书作者引入了与频率测度相关的离散系统的分布混沌性的内容。这是一个相当新的研究方向,可以帮助读者进一步拓宽知识领域和研究视野,对分布离散类的随机性理论有更广泛和更深入的思考,并且有可能引导读者把频率测度知识应用到混沌类随机性研究之中,为将来的一些实际应用(例如信息隐藏和数据加密)铺设一条全新的可行途径。

总而言之,该书的面世实属难能可贵,可喜可贺。希望能够看到作者本人和其他同行在这个起点之后的不久将来,有更多、更广、更深的后续研究工作,在理论和应用两方面都有更丰硕的成果,在学术研究过程中更上一层楼。

陈英荣

香港城市大学

2010年7月

前 言

数学是现代科学的基础之一,在各门学科中都有着重要应用。自然数是数学的最基本概念之一,对它的研究有着悠久的历史,直到今天仍然充满着活力。本书所利用的一个基本概念是频率测度,它是为研究由一些自然数构成的集合在整个自然数集中所占的比例问题时引入的一个概念。根据现有的一些文献,大约在 19 世纪末至 20 世纪初,它曾被应用于研究随机分布的模拟问题之中,在 20 世纪 30 年代的文献中,正式被称为渐近密度、自然密度和密度等,并在 20 世纪末的一些文献中又被称为频率测度。由于频率测度所具有的有限可加性与概率测度具有的无限可加性有着本质的区别,因此,利用频率测度建立起来的基础理论与用概率测度建立起来的理论就有一些本质区别。同时,频率测度和概率测度又具有许多共同性质,因此,自然地可参照由概率测度建立起来的概率论等来研究频率测度理论的相关内容。

数列是定义在自然数集上的函数,可分为确定数列和随机数列。确定数列常常可用离散系统来产生,随机数列通常可由随机抽样产生。因此,频率测度应是研究数列或离散系统的一个基础工具,数列或离散系统可作为频率测度理论的一个主要研究对象。在建立频率测度论的过程中,本书主要参照了概率论的研究方法和内容。特别指出:在本书建立的频率测度论和概率论之间存在着两个明显的对应关系:频率(测度)与概率(测度)相对应,数列与随机变量相对应。这导致了許多其他概念之间的对应关系,如数列的分布函数与随机变量的分布函数相对应,等等。

全书共分为 14 章,所有内容可大致分成如下两大部分:一是能与概率论相对应的内容;二是难以与概率论相对应的内容。下面分别概述如下。

利用频率测度和概率测度之间的共同特性,本书建立了与概率论相平行的一些概念和方法,包括伪随机事件的概念与性质,全频率公式和频率 Bayes 公式,数列的频率分布,离散数列和连续数列,二项分布和正态分布,数列的期望和方差,数列间的边际分布和独立性,数列间的协方差、相关系数、条件期望、回归与最小二乘法和矩,数列的频率熵和互信息,数列的特征函数和母函数,频率大数定律和中心极限定理,等等。为了方便,本书将这些能与概率论相对应和相关的性质简称为“伪随机性”。

同时,根据频率测度自身的特点,本书还建立了一些独特内容,包括数列的频

率收敛性和频率振动性的概念、性质和判定准则,数列的积分,随机分布的模拟和离散系统的分布混沌性,数列的自相关数列和互相关数列,数列的单向独立性,数列的相关性与独立性检验,等等。

还要指出:在频率测度论和概率论中,有些对应的概念和性质之间也有一些差异。例如,数列分布函数的定义比随机变量的分布函数定义所要求的条件更多;概率论中依概率 1 收敛性具有重要作用,但频率测度论中依频率 1 收敛的意义不大;数列的期望或均值共有三种不同的定义方式,而随机变量的期望只有一种定义形式,等等。另外,本书还包含小部分不可测数列伪随机性的相关内容。例如,一列数列依上频率收敛的概念和数列间的单向频率独立性,等等。

众所周知,概率论是一门研究完全随机性的基础理论,在多门学科中都有重要应用。与此相应,频率测度论是一门研究数列伪随机性的基础理论,并且利用本书的内容和方法,将有可能统一现有许多数列的常见分析方法。因此,频率测度论将有可能在数列分析方法的各种应用领域中发挥重要作用,特别是离散和连续数列的相关知识将可能成为频率测度论应用于实际之中的重要理论基础。

另外指出,在计算机科学技术已向各个领域渗透的今天,计算机计算作为科学研究的辅助手段得到了广泛的应用。特别是在随机信号的理论和应用基础研究中,科学工作者为了直观检验理论分析的效果,往往会利用计算机来模拟随机信号分析的效果。参照本书的内容和方法,在随机信号的大量应用研究中,可能只需要针对“伪随机信号”来设计“随机信号系统”就行了,因为根据大数定律和随机模拟理论,由伪随机信号设计好的确定系统基本上将以接近于 1 的概率适用于随机系统。

本书有不少内容是十几年来作者与多位合作者的共同研究成果,包括频率测度的概念和性质,频率振动性和频率收敛性,以及离散混沌等方面的研究成果。这些成果本身就构成了本书的部分重要内容,也无疑为本书其他内容的研究奠定了坚实基础。同时,在与他们的长期合作交流过程中,作者学到了许多新知识和新方法,得到过许多有益的启示,研究能力得以培养和加强,知识不断地丰富,还明白了不少做学问的道理。在此,作者要对他们表示衷心感谢!这些合作者包括谢胜利、张炳根、郑穗生和陈关荣教授。他们对本书内容的贡献和在写作过程中提供的帮助是不可磨灭的。

首先,感谢谢胜利教授在十多年前将作者引入到了差分方程的研究领域,指导作者学习了差分方程振动性的基础知识,从而帮助作者开辟了差分方程研究的新方向。当时,在他的倡导下,初步形成了部分数学教师对差分方程振动性的讨论小组。在学习和讨论过程中,逐渐激发出了作者对差分方程振动性的研究兴趣和潜力,对最初利用“均匀密度”概念来研究差分方程“均匀振动性”的思想给予了充分肯定和积极讨论。另外,他与作者等随后合作的多篇频率振动性论文中的不少成

果也是本书的重要内容之一。同时,还要感谢他在作者学习过程中的几个重要关口所给予的帮助和引导,才使作者进入到了新的研究领域,学到了更多的新知识,增强了科研能力。在此也对以前曾和作者一起讨论过差分方程论文的同事一并致谢。

第二,感谢张炳根教授曾给予作者的帮助和指导。在“频率振动性”的起草写作之初,他曾给作者提供过一篇重要参考文献,这对顺利完成“振动数列的测度”一文的初稿起到了重要作用。正是他在差分方程振动性方面的开创性研究成果,才使作者对差分方程振动性研究重要性的认识有所加强。同时,在差分方程的振动和稳定等研究方面,他与作者也曾取得了一些合作研究成果,这些拓宽了作者对差分方程的认识和研究范围,也构成了本书的部分重要内容。

第三,感谢郑穗生教授曾给予的帮助和指导。我们所发表的频率测度应用上的第一篇合作论文即由他定稿完成。该文及相关的几篇后继合作文构成了本书的部分重要内容。而且,在合作过程中,他丰富的论文组织经验、深刻的数学洞察力、规范的英语写作功力、熟练的文献检索能力等对提高作者的学术修养起到了潜移默化的促进作用。

最后,感谢陈关荣教授曾给予的帮助和指导。感谢他将作者引入到了离散系统的混沌研究领域,指导作者学习了离散混沌的相关知识,使作者顺利走上了离散混沌的研究道路,极大地扩展了作者的研究视野,也使作者对离散类随机性有了更广泛和更深入的思考,并逐渐发现了频率测度知识在混沌类随机性研究之中可能的应用前景,为随之而来的进行频率测度在数列伪随机性中的系统理论研究奠定了基础。而且,在他给作者提供的研究条件中,作者意外地获得了随机模拟理论的一本重要参考文献。该书对作者产生统一混沌类随机性和随机模拟方法的研究思路有很大的启发作用,也使得本书中的随机模拟内容顺利地完成了。同时,他所提供的分布混沌方面的文献也激发了作者利用频率测度知识研究分布混沌的兴趣。

总之,上述四位教授对作者完成本书稿都有过直接的重要帮助和贡献,并非是寥寥数语所能表达清楚的。在与他们亦师亦友的交流 and 交往过程中,作者在冥冥之中如有神助,超乎寻常地发挥出了自己的潜能,使自己都难以置信地完成了近乎不可能完成的工作。在此,再一次对他们表示由衷的谢意!同时,也对国家自然科学基金项目(No. 60374017 和 61070252)和广东省自然科学基金项目(No. 5010496)给予的相关资助一并致谢。

最后指出,由于作者的知识水平和研究能力有限,本书内容存在不足之处在所难免。这并非是作者的谦逊之词,而是作者在频率测度方面的研究过程中总结出的结论。1998年作者及合作者最初提出了频率测度的概念,并将它应用于差分方程的振动性研究之中。但是,当时在与谢胜利和郑穗生合作发表的第1篇“振动数列的测度”论文中,将频率测度误当成了一个新概念来使用,因此论文中就没有引

用早已有的概念。后来经过努力查询,在多年以后与郑穗生合作的另一文献中才指出频率测度其实是 70 年以前就有的渐近密度或自然密度概念。这件事情给作者留下了深刻印象,令作者感到在浩如烟海的数学知识面前突显出了自己的孤陋寡闻、知识面非常狭窄。更何况在过去的一百余年中,已发表过基于渐近密度的海量研究文献,但作者并未全部学习,且限于篇幅,也不能一一尽列于书中。因此,本书的内容有可能存在许多不当之处,敬请广大读者批评指正。

目 录

序言

前言

第 1 章 频率测度与数列	1
1.1 频率测度的定义	1
1.1.1 一维频率测度	1
1.1.2 二维频率测度	5
1.2 频率测度与勒贝格测度	7
1.3 数列与离散系统	8
1.4 模 1 均匀分布数列.....	10
第 2 章 数列的频率收敛性	12
2.1 几个定义和结果.....	12
2.2 频率收敛性的定义与性质.....	13
2.3 频率柯西收敛准则.....	16
第 3 章 差分方程的频率振动性	21
3.1 差分方程的频率振动性.....	21
3.2 偏差分方程的频率振动性.....	25
第 4 章 伪随机事件及其独立性	30
4.1 频率可测集合族.....	30
4.1.1 伪随机事件及性质	30
4.1.2 有限 σ -代数和有限 Borel 集	32
4.2 条件频率测度和独立性.....	35
4.2.1 条件频率测度	35
4.2.2 独立性	36
4.2.3 直接定义法	38
4.3 全频率测度公式和频率贝叶斯公式.....	42
第 5 章 数列的频率分布	44
5.1 可测数列及其频率分布.....	44
5.1.1 弱频率分布	44
5.1.2 频率分布	45

5.1.3	含例外点的频率可测数列	45
5.1.4	正则可测数列	47
5.1.5	函数数列的可测性	48
5.2	离散数列与连续数列	54
5.2.1	离散数列	54
5.2.2	连续数列	57
5.3	几种分布数列的构造	60
5.3.1	均匀分布数列及其应用	60
5.3.2	几类常见分布的构造	63
第 6 章	数列的积分	68
6.1	非负数列的积分	68
6.2	可测数列的积分	73
第 7 章	数列的数字特征和自相关数列	81
7.1	数列的均值和期望	81
7.1.1	均值和期望的定义	81
7.1.2	均值和期望的性质	84
7.2	函数数列的期望和均值	94
7.3	数列的频率方差	95
7.4	数列的自相关数列和自协方差数列	101
第 8 章	向量数列的联合分布	106
8.1	数列间的联合分布	106
8.2	离散和连续向量数列	110
8.2.1	离散向量数列	110
8.2.2	连续向量数列	115
8.3	向量数列的边际分布	121
8.3.1	边际分布和边际分布律的定义	121
8.3.2	联合分布律与边际分布律的关系	122
8.3.3	边际密度	124
8.4	向量数列的条件分布	125
8.4.1	条件分布律	125
8.4.2	条件密度	126
8.4.3	频率条件分布和频率条件分布律	127
8.5	数列间的独立性	129
8.5.1	数列间独立性定义	129
8.5.2	向量数列之间的独立性	130

8.5.3	完全自独立数列和完全互独立数列	132
8.5.4	数列间的单向独立性	133
8.6	函数数列的分布	134
8.6.1	卷积公式	134
8.6.2	最大与最小函数数列	135
8.6.3	向量函数数列	135
8.7	函数数列间的独立性	137
第9章	多维数列的数字特征和互相关数列	140
9.1	两个数列的协方差和相关系数	140
9.1.1	协方差与协方差矩阵	140
9.1.2	多元函数数列的期望公式	141
9.1.3	相关系数	142
9.1.4	两个结果	144
9.1.5	不相关	145
9.2	两个数列的互相关数列	146
9.2.1	互相关数列和互协方差数列	146
9.2.2	完全不相关性	148
9.3	数列间的条件期望	149
9.3.1	条件期望的概念	149
9.3.2	回归和最小二乘法	151
9.4	数列的矩	155
第10章	数列的频率熵和信息	157
10.1	离散数列的频率熵和互信息	157
10.1.1	离散数列的频率熵	157
10.1.2	频率熵的性质	160
10.1.3	条件熵和交互信息	163
10.2	连续数列的频率熵和互信息	165
第11章	数列的特征函数和母函数	168
11.1	母函数	168
11.2	特征函数	172
11.2.1	一元特征函数及性质	172
11.2.2	多元特征函数及性质	175
11.3	多维正态数列及其性质	176
第12章	频率大数定律和中心极限定理	179
12.1	频率大数定律	179

12.2	局部极限定理和积分极限定理	183
12.3	一系列数列的收敛性及中心极限定理	187
第 13 章	离散系统的伪随机性	194
13.1	随机分布的模拟	195
13.1.1	一维均匀数列	195
13.1.2	均匀分布的时空数列	199
13.1.3	区间上几乎处处均匀分布的数列	199
13.1.4	多维均匀分布数列	201
13.1.5	数列的渐近分布函数	202
13.1.6	一个结果	202
13.2	离散系统的分布混沌性	203
第 14 章	数列的伪随机性能与大数定律	210
14.1	确定数列的伪随机性能	210
14.1.1	几个新概念	211
14.1.2	m 数列的不相关性分析	212
14.1.3	离散数列的独立性能	215
14.2	随机样本数列的统计分析	217
参考文献		223

第 1 章 频率测度与数列

频率测度是为了研究自然数集(或整数集或有限维格点集等)子集的元素“个数”在自然数集(或整数集或有限维格点集等)中所占的比例而引入的一个概念,可分为一维或多维频率测度^[1]。从早期的一些数学文献^[2~4]来看,一维频率测度也被称为“渐近密度”、“自然密度”等,它可用于研究数列或离散系统的性质,并可应用于随机变量模拟等问题之中^[5]。

设自然数集为 $\mathbf{N}=\{1,2,\dots\}$, 整数集为 $\mathbf{Z}=\{\dots,-1,0,1,\dots\}$, 实数集为 $\mathbf{R}=(-\infty,\infty)$ 。对任意整数 $m,n \in \mathbf{Z}, m < n$, 记 $Z[m,n]=\{m,m+1,\dots,n\}$, $Z[m,\infty)=\{m,m+1,\dots\}$, $Z(\infty,n]=\{\dots,n-1,n\}$ 。此外,为方便起见,记 $Z[m,\infty]=\{m,m+1,\dots\} \cup \{\infty\}$ 。

对任意两个集合 A 和 B , 将 A 与 B 的并、交、差集分别记为 $A+B$ (或 $A \cup B$)、 $A \cap B$ (或 $A \cdot B$ 或 AB)、 $A \setminus B$ (或 $A - B$)。此外,本书将用 θ 表示空集合。

本章将介绍频率测度的基本概念及其性质、几种离散系统和模 1 均匀分布的概念,为后面各章的内容奠定基础。

1.1 频率测度的定义

本书只介绍自然数集(或整数集)和平面格点集上的频率测度的定义,可分为一维频率测度和二维频率测度。

1.1.1 一维频率测度

设 Ω 为 \mathbf{N} (或 \mathbf{Z})的一个子集, $|\Omega|$ 表示集合 Ω 中元素的个数。对任意 $n \in \mathbf{N}$ 和 $m \in \mathbf{Z}$, 记

$$\Omega^{(n)} = \{i \in \Omega \mid i \leq n\} = \Omega \cap Z(-\infty, n], \quad E^m(\Omega) = \{x + m \mid x \in \Omega\},$$

$$\sigma_m^k(\Omega) = E^m(\Omega) + E^{m+1}(\Omega) + \dots + E^k(\Omega) = \sum_{i=m}^k E^i(\Omega), \quad m \leq k,$$

其中, $\sigma_m^k(\Omega)$ 称为(从 m 到 k)由 Ω 导出的集合。容易证明: $j \in E^m(\Omega)$ 当且仅当 $j - m \in \Omega$ 。因此

$$j \in \mathbf{Z} \setminus \sigma_m^k(\Omega) \Leftrightarrow j \in \prod_{i=m}^k (\mathbf{Z} \setminus E^i(\Omega)) \Leftrightarrow j - i \in \mathbf{Z} \setminus \Omega, \quad m \leq i \leq k, \quad (1-1)$$

其中, \Leftrightarrow 表示“当且仅当”, \prod 表示集合的交集, Σ 表示集合的并集。

1. 一维频率测度的定义

在研究自然数集 N 的子集时,不难发现如下问题: N 的不同子集在 N 中所占的比例是不相同的。例如,设 $A = \{1, 3, 5, \dots\}$, $B = \{5, 10, 15, \dots\}$, $C = \{2, 2^2, 2^3, \dots, 2^n, \dots\}$, 则不难看出集合 A 的所有元素的“个数”与 N 的所有元素“个数”之比为 0.5, B 在 N 中所占的比例为 0.2, C 在 N 中所占的比例为 0。由现有的不少文献可以看出,描述 N 的子集在 N 中占有“多大”比例在许多学科分支中都是有意义的^[1~4, 6~8]。因此,有必要引入概念来描述这种比例关系。

现已有不少文献研究了这个问题,可以看出不同的文献在给出数学定义时,在定义形式和使用的名词术语上略有不同。下面的定义取自于文献[2]~[4]。

定义 1.1 对于一个正整数数列 $A = \{a_n\}_{n=1}^{\infty}$, 满足 $a_1 < a_2 < \dots$, 将极限

$$\underline{\delta}(A) = \liminf_{n \rightarrow \infty} \frac{n}{a_n}, \quad \bar{\delta}(A) = \limsup_{n \rightarrow \infty} \frac{n}{a_n}$$

分别称为该数列 A 的下渐近密度和上渐近密度。特别地,若 $\underline{\delta}(A) = \bar{\delta}(A) = \delta(A)$, 则称 $\delta(A)$ 为该正整数数列的自然密度(natural density)

注 1.1 在有些文献中,下渐近密度也被简称为渐近密度(asymptotic density)或极限密度(limit density)或密度(density)。

另外,文献[1]在研究离散系统的振动性时给出了如下另一种形式的定义。

定义 1.2 设 Ω 是自然数集 N (或 $Z[-k, \infty)$, $k \in Z$) 的一个子集, 如果

$$\mu_*(\Omega) = \liminf_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n}, \quad \mu^*(\Omega) = \limsup_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n}$$

存在, 则 $\mu_*(\Omega)$ 和 $\mu^*(\Omega)$ 分别称为集合 Ω 的下频率测度和上频率测度, 简称下频度和上频度, 或下频率和上频率, 或下测度和上测度。特别地, 若 $\mu_*(\Omega) = \mu^*(\Omega) = \mu(\Omega)$, 则称 $\mu(\Omega)$ 为 Ω 的频率测度或频度或频率或测度, 也称 Ω 是(频度或频率)可测的。若 Ω 不是可测的, 则称 Ω 不可测。

直观上看, Ω 的频率能表示集合 Ω 的元素个数在 N 中所占的“比例”的大小。

显然, 设正整数集合为 $\Omega = \{a_1, a_2, \dots\}$, 其中 $a_1 < a_2 < \dots$, 则有

$$\mu_*(\Omega) = \liminf_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n} = \liminf_{n \rightarrow \infty} \frac{n}{a_n} = \bar{\delta}(\Omega), \quad \mu^*(\Omega) = \bar{\delta}(\Omega).$$

因此, 对于无限正整数集 Ω , Ω 的上(下)渐近密度与上(下)频率测度是一样的, Ω 的自然密度和频率测度是一样的。由此可知, 在自然数集上的频率测度和自然密度是等价的。

注 1.2 在本书中, 由于需要定义很多新术语, 为了术语命名的方便和习惯, 并避免混淆, 本书将采用名词“频率测度”或“频度”或“频率”或“测度”来表示自然数子集在 N 中的“比例”。

2. 频率测度的性质

下面介绍频率测度的一些例子和性质,可参见文献[1],[9]~[11].

例 1.1 对任意 $\Omega \subseteq \mathbf{N}$, $\mu_*(\Omega)$ 和 $\mu^*(\Omega)$ 都存在,且 $0 \leq \mu_*(\Omega) \leq \mu^*(\Omega) \leq 1$. 当 Ω 是有限集合时, $\mu(\Omega) = 0$. $\mu(\mathbf{N}) = 1, \mu(\theta) = 0$, 其中, θ 表示空集合. 设 $\Omega = \{1, 3, 5, \dots\}$ 和 $\Gamma = \{2, 2^2, 2^3, \dots\}$, 则 $\mu(\Omega) = 0.5$ 和 $\mu(\Gamma) = 0$.

例 1.2 对任意 $m \in \mathbf{Z}$ 和 $\Omega \subseteq \mathbf{N}$, 都有 $\mu_*(E^m(\Omega)) = \mu_*(\Omega)$ 和 $\mu^*(E^m(\Omega)) = \mu^*(\Omega)$. 该性质称为频度的平移不变性.

例 1.3 若 $\Omega_1, \dots, \Omega_m$ 是 \mathbf{N} 的子集, $\mu(\Omega_i) = 0, 1 \leq i \leq m < \infty$, 则 $\mu(\Omega_1 + \dots + \Omega_m) = 0$.

例 1.4 下面构造一个不可测的集合 $\Omega \subseteq \mathbf{N}$, 使得 $\mu_*(\Omega) \neq \mu^*(\Omega)$. 设

$$a_m = 2 + 2^{2 \times 2} + 3^{2 \times 3} + \dots + (m-1)^{2(m-1)}, \quad m = 3, 4, 5, \dots,$$

$$A_1 = \{1\}, A_2 = \{2, 3, \dots, a_3 - 1\}, A_3 = \{a_3, a_3 + 1, \dots, a_4 - 1\},$$

$$A_n = \{a_n, a_n + 1, \dots, a_{n+1} - 1\}, \quad n = 3, 4, \dots, \quad (1-2)$$

那么,对一切整数 $m \geq 1$, 有 $|A_m| = m^{2m}$. 再设

$$\Omega = A_1 + A_3 + \dots + A_{2m-1} + \dots, \quad (1-3)$$

则 $\mu_*(\Omega) = 0$ 和 $\mu^*(\Omega) = 1$, 即 Ω 不可测. 事实上, 令

$$b_k = 1 + 2^4 + \dots + k^{2k}, \quad k = 2, 4, \dots,$$

则

$$|\Omega^{(b_k)}| = 1 + 3^6 + \dots + (k-1)^{2(k-1)}.$$

故

$$0 \leq \liminf_{n \rightarrow \infty} \frac{|\Omega^{(b_k)}|}{b_k} \leq \liminf_{n \rightarrow \infty} \frac{k(k-1)^{2(k-1)}}{k^{2k}} \leq \liminf_{n \rightarrow \infty} \frac{1}{k} = 0,$$

即 $\mu_*(\Omega) = 0$. 另外, 令

$$c_k = 1 + 2^4 + 3^6 + \dots + (k+1)^{2(k+1)}, \quad k = 1, 2, 3, \dots,$$

则

$$|\Omega^{(c_k)}| = 1 + 3^6 + 5^{10} + \dots + (k+1)^{2(k+1)}.$$

故

$$\begin{aligned} 1 &\geq \liminf_{n \rightarrow \infty} \frac{|\Omega^{(c_k)}|}{c_k} \geq \limsup_{k \rightarrow \infty} \frac{(k+1)^{2(k+1)}}{k \times k^{2k} + (k+1)^{2(k+1)}} \\ &\geq \limsup_{k \rightarrow \infty} \left(1 + \frac{1}{k} \left(\frac{k}{k+1} \right)^{2(k+1)} \right)^{-1} = 1, \end{aligned}$$

即 $\mu^*(\Omega) = 1$. 因此, Ω 不可测.

下面的频率测度性质称为单调性.

性质 1.1 对任意集合 $\Omega, \Gamma \subseteq \mathbf{N}$, 如果 $\Omega \subseteq \Gamma$, 那么 $\mu_*(\Omega) \leq \mu_*(\Gamma)$ 和 $\mu^*(\Omega) \leq$

$\mu^*(\Gamma)$ 。

证明:因为 $\Omega \subseteq \Gamma$, 所以对一切 $n \in \mathbf{N}$, 都有 $\Omega^{(n)} \subseteq \Gamma^{(n)}$ 。因此

$$\mu_*(\Omega) = \liminf_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n} \leq \liminf_{n \rightarrow \infty} \frac{|\Gamma^{(n)}|}{n} = \mu_*(\Gamma),$$

$$\mu^*(\Omega) = \limsup_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n} \leq \limsup_{n \rightarrow \infty} \frac{|\Gamma^{(n)}|}{n} = \mu^*(\Gamma)。$$

因此,频率测度具有单调性。证毕!

性质 1.2(有限次可加性) 对任意集合 $\Omega, \Gamma \subseteq \mathbf{N}$, 都有

$$\mu^*(\Omega + \Gamma) \leq \mu^*(\Omega) + \mu^*(\Gamma)。$$

此外,如果 Ω 和 Γ 互不相交,则

$$\begin{aligned} \mu_*(\Omega) + \mu_*(\Gamma) &\leq \mu_*(\Omega + \Gamma) \leq \mu_*(\Omega) + \mu^*(\Gamma) \\ &\leq \mu^*(\Omega + \Gamma) \leq \mu^*(\Omega) + \mu^*(\Gamma)。 \end{aligned}$$

证明:显然对任意 $n \in \mathbf{N}$, 都有 $|(\Omega + \Gamma)^{(n)}| \leq |\Omega^{(n)}| + |\Gamma^{(n)}|$, 故 $\mu^*(\Omega + \Gamma) \leq \mu^*(\Omega) + \mu^*(\Gamma)$ 。

如果 Ω 和 Γ 不相交,则对任意 $n \in \mathbf{N}$, 都有 $|(\Omega + \Gamma)^{(n)}| = |\Omega^{(n)}| + |\Gamma^{(n)}|$, 故

$$\begin{aligned} \mu^*(\Omega + \Gamma) &= \limsup_{n \rightarrow \infty} \frac{|\Omega^{(n)} + \Gamma^{(n)}|}{n} \geq \liminf_{n \rightarrow \infty} \frac{|\Omega^{(n)}|}{n} + \limsup_{n \rightarrow \infty} \frac{|\Gamma^{(n)}|}{n} \\ &= \mu_*(\Omega) + \mu^*(\Gamma)。 \end{aligned}$$

其他的不等式可类似证明。证毕!

性质 1.3 对任意集合 $\Omega \subseteq \mathbf{N}$, 都有 $\mu_*(\Omega) + \mu^*(\mathbf{N} \setminus \Omega) = 1$ 。

证明:因 $\mathbf{N} = \Omega + \mathbf{N} \setminus \Omega$, 故由性质 1.2, 有

$$1 = \mu_*(\mathbf{N}) \leq \mu_*(\Omega) + \mu^*(\mathbf{N} \setminus \Omega) \leq \mu^*(\mathbf{N}) = 1。$$

性质 1.4 设 Ω 和 Γ 是自然数集 \mathbf{N} 的两个子集。如果 $\Omega \subseteq \Gamma$, 则

$$\begin{aligned} \mu^*(\Gamma) - \mu^*(\Omega) &\leq \mu^*(\Gamma \setminus \Omega) \leq \mu^*(\Gamma) - \mu_*(\Omega), \\ \mu_*(\Gamma) - \mu^*(\Omega) &\leq \mu_*(\Gamma \setminus \Omega) \leq \mu_*(\Gamma) - \mu_*(\Omega)。 \end{aligned}$$

证明:因为 $\Gamma = \Omega + (\Gamma \setminus \Omega)$, 故由性质 1.2, 得

$$\mu^*(\Gamma \setminus \Omega) \geq \mu^*(\Gamma) - \mu^*(\Omega), \mu^*(\Gamma \setminus \Omega) \leq \mu^*(\Gamma) - \mu_*(\Omega)。$$

类似可证明另一个结论。证毕!

类似于上述几条性质的证明,可得到如下的性质。

性质 1.5 对于自然数集 \mathbf{N} 的任意两个子集 Ω 和 Γ , 都有

$$\begin{aligned} \mu_*(\Omega) + \mu^*(\Gamma) - \mu^*(\Omega \cdot \Gamma) &\leq \mu^*(\Omega + \Gamma) \leq \mu^*(\Omega) + \mu^*(\Gamma) - \mu_*(\Omega \cdot \Gamma), \\ \mu_*(\Omega) + \mu_*(\Gamma) - \mu^*(\Omega \cdot \Gamma) &\leq \mu_*(\Omega + \Gamma) \leq \mu_*(\Omega) + \mu^*(\Gamma) - \mu_*(\Omega \cdot \Gamma)。 \end{aligned}$$

性质 1.6 对任意集合 $\Omega \subseteq \mathbf{N}$, 如果 $\mu^*(\Omega) > 0$, 则 Ω 是无限集; 如果 $\mu_*(\Omega) < 1$, 则 $\mathbf{N} \setminus \Omega$ 是无限集。

性质 1.7 设 Ω 和 Γ 是 \mathbf{N} 的两个子集。如果 $\mu^*(\Omega) + \mu_*(\Gamma) > 1$, 则 $\Omega \cdot \Gamma$ 是无限集。