寿命表及其应用

蒋庆琅著

寿命表及其应用

〔美〕蒋庆琅 著

方积乾 译

上海翻译出版公司

内 容 提 要

寿命表是一种统计分析方法，用以研究死亡率及有关问题，广泛应用于人口统计、公共卫生、生物控制、生态学、工业管理、保险等各个领域．本书是世界卫生组织特约的书稿，阐述寿命表的意义、原理、制作方法和各种目的的寿命表的应用．全书共分十二章，可供人口统计工作者、生物统计工作者、公共卫生工作者、保险业工作人员、工业管理人员等参考．

THE LIFE TABLE and ITS APPLICATIONS

CHIN LONG CHIANG

寿命表及其应用

〔美〕蒋庆琅　著

方积乾　译

上海翻译出版公司出版

（上海瓯州路 390 号）

新华书店上海发行所发行　上海市印刷四厂印刷

开本 787×1092 1/16　印张 13.75　字数 289,000

1984 年 12 月第 1 版　1984 年 12 月第 1 次印刷

印数 1—5,000

统一书号：17311·2　　定价：3.20 元

献　给

我　的

父亲和母亲

本书英文原版曾得到联合国人口
活动基金(INT/80/P09)和
世界卫生组织的协助,
谨此敬表谢忱

中 译 本 序

生物统计学在最近三十年来经过了很大进展和演变，虽然这学科已在十九世纪末期建立，但主要的活动只限于资料的收集和整理．近年来数学、概率和数理统计逐渐应用到生物统计内各项课题的研究．其结果，生物统计已有很深的理论基础，卓然成为数理专科的一门．

虽然生物统计学的范围很广，其中主要的还是寿命表．第一，寿命表已有理论基础；第二，寿命表有广大的应用，人口统计、生存研究、医学随访等都以寿命表作为基本研究方法，寿命表已经成为统计研究方法的一门，称为"寿命表研究方法"．

1982 年夏季我在北京医学院教课时，曾用本书的英文本作教科书．在课堂中的学者是来自中国三十多个省市的大学或医学院的教师，济济一堂，足见中国教育界对寿命表的重视．他们不但对寿命表内容有深刻的认识，对它的应用亦很有经验．他们都表示有出版中译本的需要．鉴于寿命表对科学和社会研究的重要，希望本书不但对中国生物统计学有影响，而且对数理统计、人口统计等学科亦有点贡献．

最后我谢谢本书的译者方积乾先生，谢谢华东师范大学魏宗舒教授和本书的编辑、出版工作者，由于他们的大力帮忙，使本书得以顺利出版．

蒋庆琅

美国加州大学

1983 年 4 月

前　　言

七十年代中期，世界卫生组织在联合国人口活动基金的资助下，着手编写了一系列新教材，侧重于公共卫生、医学研究、统计学和人口学等广泛的专业范围．在这个系列中，第一本出版的是《死亡率分析手册》（"Manual of Mortality Analysis"），集中介绍各国的国家人口统计局常用的一些基本方法．与此同时，世界卫生组织有幸获得了应用随机方法研究死亡过程的杰出权威蒋庆琅教授的帮助，出版了一本介绍人口死亡率分析中较高级方法的专著，作为上述一书的补充． 这本书——《寿命表和死亡率分析》（"Life Tables and Mortality Analysis"）——象第一本一样，不久即告售完．鉴于该书如此深受欢迎，世界卫生组织要求蒋教授将他的书修订后出第二版．然而，在修订过程中很快就发现，仅仅修订实难以满足人们的迫切要求——希望有一本把数学的精美和清晰的讲解相结合，也适合于高等数学不怎么熟悉者阅读的书．《寿命表及其应用》（"The Life Table and Its Applications"）就是根据上述要求所作尝试的产物．它基本上是一本全新的书．书中，不但把教学实践中发现的某些缺点作了改正，还增加了这一快速发展领域内的新进展．蒋教授以其丰富的经验阐述了寿命表方法在许多社会科学和医学科学问题中广泛的应用． 这本新的教科书就是这样的一个既有牢固的理论基础又讲究"脚踏实地"的应用的很好的组合．它不但可作为教材，也适用于自学，特别适用于那些想了解最新发展并用之于日常工作的人们．

哈罗德·汉斯卢卡博士
主任统计师
日内瓦，世界卫生组织
全球流行病监测和卫生情况评估组
1982年8月5日

序

经典的寿命表最初是应用于保险科学，作寿命的随机分析，亦应用于人口学，作人口变迁的研究．相应的数学方法已先后发展起来．Gompertz (1825)和Makeham (1860) 导出了"死亡定律"，以简化养老金和保险金的计算；Euler(1760) 所引进的稳定人口的概念已成为人口规划的基础；Lotka(1922)提出了用自然增长的固有速率来度量人口的发展．本世纪初，美国政府部门开始将寿命表方法应用于生命统计和人口调查资料，以概括现时人口的死亡经历．虽然这些工作在继续进行着，但是寿命表在一个很长的时期里一直处于正规的统计学门户之外．本世纪50年代之前，人们对寿命表这一分析工具的潜力缺乏足够的认识．

本世纪50年代前期，由于卫生统计学家在医学随访方面的工作，寿命表才开始受到生物统计学家的重视．但正是由于概率和统计理论的发展才有可能以纯随机的观点来处理寿命表，并为这一主题提供了理论基础．现在，我们可以利用寿命表作生存分析，并对寿命表元素和所研究人口死亡模型的其它参数进行统计推断．"寿命表分析"已经成为严格而完整的统计学方法．事实上，寿命表分析在统计学领域里是自成一格的，给定一些个体生存经历的样本，便有一系列有关期望寿命的样本均数($\hat{e}_0$, $\hat{e}_1$, …)、一系列有关的概率估计($\hat{q}_0$, $\hat{q}_1$, …)和($\hat{p}_{01}$, $\hat{p}_{02}$, …)，这些样本值都是相应的数学期望和概率的最佳估计．

在生存分析新近的发展中考虑了许多慢性病的阶段性，病情从轻度阶段经过中度阶段发展到严重阶段，乃至死亡．死亡率随疾病的阶段而不同．疾病过程往往是不可逆的，但病人可能在任何一个阶段死去．引入阶段的概念，我们发现一种新的寿命表，其中年龄区间不是预先确定的，它具有变异性，实为随机变量．这种新的寿命表可应用于许多研究领域，只要阶段的概念有所定义，结局不一定是死亡．本书的最后一章将介绍这种新的寿命表，并通过关于人类生殖分析的生育表来说明其应用．

本书的目标是以统计学观点来阐述寿命表方法的理论和应用．有关理论分析的章节必然涉及一些数学内容，但本书的其余部分并不要求读者已经懂得统计学．为了提供若干基础知识，便于参考，本书有三章预备知识：第一章是概率初步，第二章是统计学基本概念，第三章是正态分布和统计推断．由于寿命表常常涉及大样本，对于寿命表中的生物统计函数，可以借助中心极限定理利用正态分布和卡方(χ^2)检验作统计估计和假设检验．

死亡率和率的校正与寿命表关系密切，分别见于第四、五两章．第六、七两章分别介绍完全寿命表和简略寿命表，分开叙述只是为了方便和完备，难免会有重复．第八章关于寿命表函数的统计推断主要是以第三章的统计方法为基础的．第九章有两个内容：定群寿命表和寿命表的若干应用．后者不限于定群寿命表，尤其是第5节中的胎儿寿命表和第7节中的家庭生活周期．

对于第十章寿命表的理论，一些乐于运用统计理论的数理统计工作者以及与寿命表密切

有关的保险科学和人口学工作者一定会发生兴趣，但是，一些旨在应用寿命表的读者初读时可以略去这一章．第十一章涉及医学随访研究，给出了估计存活概率的若干公式，可惜没有适当的数据来说明每个公式的应用．然而，蒙加州卫生部协助，以频数形式提供了某种肿瘤患者生存经历的数据．最后一章就是为前述新的寿命表而写的，其中给出了这种寿命表的理论基础和一个关于人类生殖分析的应用．

本书可作卫生统计学一个学期或一个学季的教材，不妨采用下列章节：

第一、二、三、四、五、七、八、九、十、十一、十二章(第6节)．

本书也可用作流行病学、人口学、生物统计学、保险科学等有关课程的参考书．

我衷心感谢许多朋友，正由于他们的鼓励和帮助，本书才得以出版．感谢联合国世界卫生组织的 H. Hansluwka，他首先提议我写出《寿命表和死亡率分析》一书；感谢 B. J. van den Berg，她和我共同讨论了胎儿和婴儿死亡率问题，并且她同意引用我们合作的生育分析作为第十二章的一个示例；感谢 C. Langhauser，她帮助完成了全部统计表和计算；感谢 J. Hughes，他读了本书的早期版本；感谢我的儿子 Robert，他编写了计算机程序；感谢我在加州大学，伯克莱的学生们，他们“校阅”了手稿．最后，我深深地感谢 B. Hutchings，她有效而熟练地处理了有关手稿的一切事务性工作．

蒋庆琅
加州大学，伯克莱
1983年3月

目　　录

第一章 概率初步

1. 引　言

透彻地理解概率的含义是对死亡资料作正确分析的基础．概率论作为有力的分析工具已越来越多地应用于生命统计和寿命表的分析．生命资料的研究不再限于一些表的描述和数值的说明；人们还可对整个人口总体的死亡和生存的模式作出统计推断．虽然概率是一个数学概念，直观上却饶有兴趣．许多自然现象可以用概率的规律来描述，日常事物的发生也可看成遵从一定的概率模式，即使象事故这样的自发事件，也能以一定程度的准确性作出事前的预测．B. Gompertz 于 1825 年提出，后由 W. M. Makeham 于 1860 年修正的死亡定律，目前已在公共卫生学、人口统计学、保险精算学等领域中用以研究人的生存和死亡规律．为此，作为本书的开端，有必要介绍概率的基本概念、有关的公式和实例．

2. 基本概念

2.1. 概率的组成部分 概率这个概念包括三个组成部分：(a)一个随机实验，(b)这个实验可能出现的种种结果，以及(c)关心的一个事件．一个随机实验有许多可能的结果，但是在实验之前不能确定哪一个结果会出现．这样，在讨论概率时，人们心目中就必须有一个要考虑的随机实验和一个所关心的事件．

2.2. 概率的定义 事件 A 发生的概率(简称"A 的概率")定义为发生事件 A 的所有结果的数目与可能结果的总数之比．

假定一个随机实验有 n 个机会均等的可能结果，而其中事件 A 发生的结果是 $n(A)$ 个．那么，事件 A 的概率定义为

$$\Pr\{A\}=\frac{n(A)}{n}. \tag{2.1}$$

这样，一个随机实验中事件 A 的概率便是事件 A 发生的可能性的一种度量．

2.3. 示例 下面的几个例子有助于阐明概率的概念．

【例 1】 把一个均匀的硬币投掷一次，正面朝上的概率是多少？这里随机实验是"把一个均匀的硬币投掷一次"；可能的结果是正面朝上和反面朝上．记事件 A 为"硬币正面朝上"．可能结果的总数是 $n=2$，出现正面结果的数目为 $n(A)=1$．因此，概率是

$$\Pr\{A\}=\frac{n(A)}{n}=\frac{1}{2}.$$

【例 2】 把一个完好的骰子投掷一次，有 6 个机会均等的可能结果．记事件 A 为"3 点朝上"．这里 $n=6$，而 $n(A)=1$；因此，

$$\Pr\{A\}=\frac{n(A)}{n}=\frac{1}{6}.$$

记事件 B 为"偶数点朝上". 因为当 2 点、4 点或 6 点朝上时, 事件 B 都发生, 所以, $n(B)=3$. B 的概率就是

$$\Pr\{B\}=\frac{n(B)}{n}=\frac{3}{6}=\frac{1}{2}.$$

【例 3】 一份名单上写有 39 个女子和 81 个男子的名字, 从这份写有 120 个人名字的单子上随机地点出一个名字. 记事件 F 为"点出一个女子的名字", 事件 M 为"点出一个男子的名字", 相应的概率就是

$$\Pr\{F\}=\frac{n(F)}{n}=\frac{39}{120}=\frac{13}{40} \quad 和 \quad \Pr\{M\}=\frac{n(M)}{n}=\frac{81}{120}=\frac{27}{40}.$$

它们的和是

$$\Pr\{F\}+\Pr\{M\}=\frac{13}{40}+\frac{27}{40}=1.$$

【例 4】 一份包括 $n=100$ 个名字的名单, 其中活人的名字 $n(s)=98$ 个, 死人的名字 $n(d)=2$ 个. 从这份名单上随机地挑一个名字, 它恰是活人的名字这一事件的概率为

$$\Pr\{s\}=\frac{n(s)}{n}=\frac{98}{100}=.98.$$

它恰是死者的名字这一事件的概率为

$$\Pr\{d\}=\frac{n(d)}{n}=\frac{2}{100}=.02.$$

显然, 两个概率之和为 1:

$$\Pr\{s\}+\Pr\{d\}=.98+.02=1.$$

2.4. 概率的值 由定义可见, 一个事件的概率是一个理想化的比值 (或相对频率). 所以, 概率只能在 0 与 1 之间取值, 即对于任何事件 A, 总有

$$0\leqslant \Pr\{A\}\leqslant 1. \tag{2.2}$$

2.5. 必然事件和不可能事件 在一个实验中必定发生的事件就是必然事件. 若 I 是一个必然事件, 则

$$\Pr\{I\}=1. \tag{2.3}$$

在一个实验中永不发生的事件是不可能事件. 若 $\emptyset$ 是一个不可能事件, 则

$$\Pr\{\emptyset\}=0. \tag{2.4}$$

2.6. 事件的补(或否定) 例 3 中当且仅当事件 M(男子的名字)不发生时, 事件 F (女子的名字)才发生, 在这个意义上, 事件 F 是事件 M 的补事件. 同样, 例 4 中的事件 s (活人)是事件 d(死人)的补事件. 一个事件的补事件还可通过一些例子来说明. 我们用记号 $\bar{A}$ 表示事件 A 的补事件.

表 1 事件及其补举例

例	A	$\bar{A}$
一个婴孩的性别	男	女
掷一个硬币	正面	反面
掷一个骰子	3 点	除了 3 点以外的任何情形
掷一个骰子	偶数点	奇数点
生存分析	生存	死亡

若事件 $\bar{A}$ 是事件 A 的补事件，则事件 A 是事件 $\bar{A}$ 的补事件．这两个事件称为互补的．

在一个随机实验中，根据 A 或 $\bar{A}$ 的出现，全部结果可以分成两组，从而

$$n=n(A)+n(\bar{A}).$$

根据定义，在一个随机实验中 $\bar{A}$ 的概率为

$$\Pr\{\bar{A}\}=\frac{n(\bar{A})}{n}.$$

所以，不论什么样的事件 A，显然有

$$\Pr\{A\}+\Pr\{\bar{A}\}=\frac{n(A)}{n}+\frac{n(\bar{A})}{n}=1, \tag{2.5}$$

或

$$\Pr\{\bar{A}\}=1-\Pr\{A\}. \tag{2.5a}$$

用文字来表述，就是 A 之补事件的概率等于 A 的概率之补．

例 3 中，事件 F 的概率和事件 M 的概率有关系式

$$\Pr\{F\}=1-\Pr\{M\} \quad 或 \quad \frac{13}{40}=1-\frac{27}{40}.$$

3．两个或多个事件——乘法定理

为了简单起见，我们从单个事件引入了概率的基本概念．现在讨论关于两个或多个事件的概率．

3.1．复合事件(A 与 B)　给定事件 A 和事件 B，若这两个事件都发生，我们就说复合事件(A 与 B)发生(或记为 $A\times B$，或 $A\cap B$，或简记为 AB)．

【例 5】　考虑 800 名新生婴儿，按性别和先天性异常来划分．记 A=异常，$\bar{A}$=正常，B=男婴，$\bar{B}$=女婴．下面的表 2 反映了这些婴儿的性别和异常性的分布：

表 2　800 个婴儿性别和先天异常性的分布

(假设的资料)

先天异常性 \ 性别	男 B	女 $\bar{B}$	边缘行和
异常 A	70 $n(AB)$	50 $n(A\bar{B})$	120 $n(A)$
正常 $\bar{A}$	330 $n(\bar{A}B)$	350 $n(\bar{A}\bar{B})$	680 $n(\bar{A})$
边缘列和	400 $n(B)$	400 $n(\bar{B})$	800 n

假定从这组婴儿中随机地选出一个，这个婴儿恰是先天性异常的男婴这一事件的概率是多少？这里，复合事件是 AB，其概率为

$$\Pr\{AB\}=\frac{n(AB)}{n}=\frac{70}{800}. \tag{3.1}$$

其它可能的复合事件还有