



生物与医学统计基础

林建忠 编

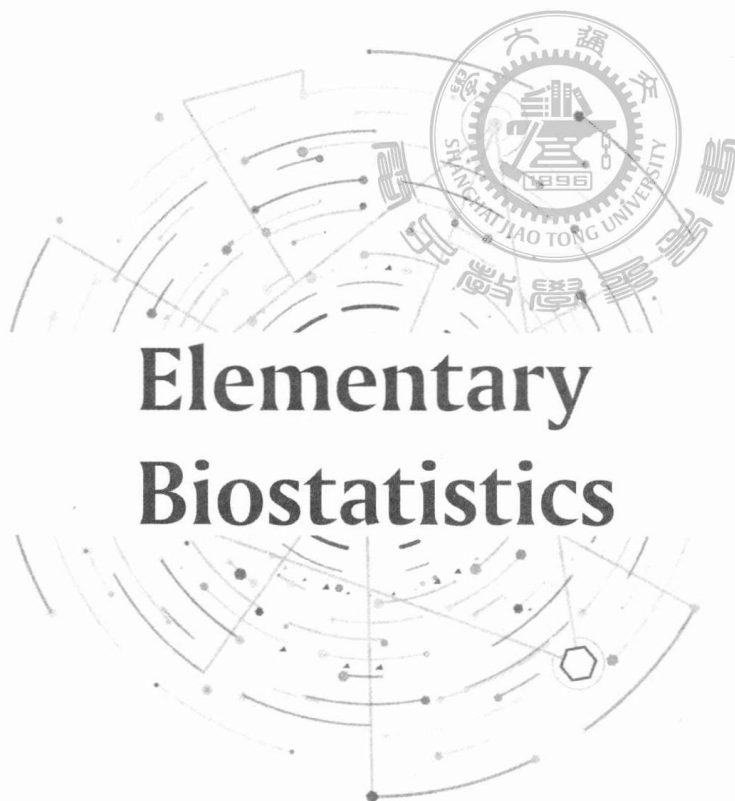
Elementary
Biostatistics



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

生物与医学统计基础

林建忠 编



Elementary Biostatistics



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书介绍了生物与医药学统计中常用的数据分析与建模方法,内容包括基础数理统计、列联表分析、回归分析、多元统计分析、生存分析和马尔柯夫链模型数据分析等统计分支学科的基础内容,每章结合生物与医药学的实际数据进行数据分析和建模方法的讲解。此外,每章还配备了一定数量的习题与上机实验题。

本书可作为普通高等院校农学、生命科学和医药学专业研究生基础课程教材,也可作为这些专业大四学生和其他非统计类学科研究生的教学参考书。

图书在版编目(CIP)数据

生物与医学统计基础/林建忠编. —上海: 上海交通大学出版社, 2019
ISBN 978-7-313-21900-8

I. ①生… II. ①林… III. ①生物医学工程—统计方法 IV. ①R318

中国版本图书馆 CIP 数据核字(2019)第 191017 号

生物与医学统计基础

编者: 林建忠

出版发行: 上海交通大学出版社

邮政编码: 200030

印制: 上海万卷印刷股份有限公司

开本: 710 mm×1000 mm 1/16

字数: 415 千字

版次: 2019 年 9 月第 1 版

书号: ISBN 978-7-313-21900-8/R

定价: 59.00 元

地址: 上海市番禺路 951 号

电话: 021-64071208

经销: 全国新华书店

印张: 23.25

印次: 2019 年 9 月第 1 次印刷

版权所有 侵权必究

告读者: 如发现本书有印装质量问题请与印刷厂质量科联系

联系电话: 021-56928178

前言

在农业、林业以及医药、卫生健康等生命科学领域的工作进程和科学实验中所产生的大量数据已不再局限于定性的描述,而是需要从大量调查和测定的数据中,应用统计学的原理和方法,分析和解释其数量的变化及其变化之间的关联性,以正确制订试验计划,科学地对试验结果进行分析,从而做出符合科学实际的推断。

本书是为生命科学类领域的研究生开设的统计学应用课程。概率论是数理统计的数学基础,本书第1章为未学习过概率论的读者提供概率论基础知识的总结性介绍。第2章至第4章介绍数据整理与展示的方法、参数估计与假设检验等基础统计学知识,第5章的四格子表分析为分析试验中两个因子彼此之间的关联关系提供了一套统计分析方法。大数据时代的数据集不仅变量繁多,而且变量性状多样化及变量间层次结构复杂化。数据分析的实践表明,拟合这类数据集中多变量复杂关系的模型往往首先考虑的是线性统计模型。第6章介绍了多自变量线性回归分析和建模的基本思想和方法。第7章的方差分析和实验设计旨在比较具有某种相同特质的多个试验对象有关指标平均值的比较,以及怎样基于方差分析进行实验设计。第8章至第10章属于多元统计分析的内容,重点介绍具有多元属性的试验对象的常用分析及分类方法。第11章介绍的分支分类方法是探索生物物种分支演化规律的基本方法。第12章介绍的生存分析原本属

于医药统计学范畴,目前已成为分析具有“生命特征”实验对象存活时间的分析工具。最后一章介绍怎样应用马尔科夫链模型分析生态系统演化以及物种世代杂交基因演化的数据变化规律。

生物与医学统计的内容十分丰富,本书的主要目的并不是把读者培养成生物医学统计分析方面的专家,而是培养学生掌握统计学的基本思想和方法,使读者对生命科学中的数据如何进行数据分析和建模有一个比较全面的基本了解,为将来能够就实践中出现的数据问题用统计学的语言与生物医学统计学的专家进行科研交流探索。正因为此,读者只要掌握了大学阶段的多元微积分、线性代数和概率论的基础知识,就可以阅读本书的全部内容。考虑到目前统计软件及介绍统计软件的书籍和网上资料繁多,因此本书对统计软件的具体操作不再做详细介绍。本书的部分内容也可作为数学专业大四学生和其他学科研究生统计课程的教学参考书以及相关业界数据分析师的参考用书。

在本书结稿之际,我要感谢所有关心和支持我写作和出版此书的人们。自2006年以来,本书的大部分内容在上海交通大学生命科学学院的研究生公共课《生物数学(I)》和《生物统计》课程中讲授过。参加过课程学习的同学对课程的成功和本书的成书做出了贡献。本书的写作和出版得到了上海交通大学2017年至2019年研究生“学科人才培养建设与创新-专业课建设”经费(经费号:WF610107101)的重点资助,以及国家自然科学基金国际(地区)合作与交流项目“开放网络下医疗资源配置和优化的模型、算法及应用研究”(项目编号:71520107003)的资助。

由于作者水平所限,书中存在缺点和错误,恳请同行和广大读者批评指正。

林建忠

上海交通大学数学科学学院

2019年6月

目 录

1	概率论基础知识	001
1.1	随机事件及其概率	001
1.2	概率的定义	003
1.3	随机变量及其分布函数	009
1.4	一元随机变量的数字特征	019
1.5	多维随机变量及其分布	025
1.6	大数定律与中心极限定理	028
1.7	习题 1	029
2	样本描述与抽样分布	031
2.1	总体和样本	031
2.2	数据类型	033
2.3	样本分布	035
2.4	样本统计量	046
2.5	一些常用的抽样分布	054
2.6	习题 2	060
3	参数估计	062
3.1	点估计方法	062
3.2	估计量的优劣标准	068
3.3	区间估计	069
3.4	习题 3	084

4	假设检验	088
4.1	基本原理与方法	088
4.2	均值的假设检验	095
4.3	总体方差的假设检验	101
4.4	单边假设检验	104
4.5	置信区间与假设检验之间的关系	108
4.6	习题 4	110
5	χ^2 检验与四格表分析	113
5.1	分布假设检验——拟合优度检验	113
5.2	成组设计两样本率比较的四格子表 χ^2 检验	115
5.3	配对设计两样本率比较的配对四格表 χ^2 检验	123
5.4	独立性的卡方检验	125
5.5	Fisher 精准检验	131
5.6	习题 5	135
6	线性回归模型	138
6.1	一元线性回归模型	138
6.2	可线性化的一元非线性回归与 Logistic 种群增长模型	147
6.3	多元线性回归模型	154
6.4	回归方程的检验	159
6.5	回归诊断与治疗	163
6.6	回归方程的选择	173
6.7	习题 6	177
7	方差分析模型与正交试验设计	180
7.1	单因素方差分析	180
7.2	两因素方差分析	188
7.3	正交试验设计与方差分析	197

7.4	习题 7	202
8	主成分分析	204
8.1	总体与样本的主成分	204
8.2	主成分的计算与实际意义	211
8.3	习题 8	222
9	判别分析	224
9.1	费希尔的判别分析法	224
9.2	多类群时的贝叶斯判别法	231
9.3	习题 9	233
10	聚类分析	235
10.1	分类的基本概念和原始数据的获得	235
10.2	数据变换和数据标准化	237
10.3	相似性概念的数量化	239
10.4	系统聚类法	244
10.5	k 均值聚类	251
10.6	习题 10	254
11	生物演化的分支分类	255
11.1	演化集合及其基本定理	255
11.2	分支性状与编码	261
11.3	演化的定量表示与简约性公理	267
11.4	性状演化的和谐性与和谐性分析方法	272
11.5	生物演化历史的重构	281
11.6	习题 11	287
12	生存分析	289
12.1	基本数据与变量类型	289
12.2	生存分析的基本函数	295
12.3	生存数据建模常用的参数模型	300

12.4	删失和截尾数据似然函数的构造	308
12.5	估计基本特征函数的寿命表法	311
12.6	右删失数据的生存函数和累积死亡力 函数的估计	313
12.7	双删失数据的生存函数估计	318
12.8	比较生存函数的非参数方法,两个生存 函数的比较	322
12.9	分层情形下的 Mantel-Haenszel 检验	327
12.10	比例危险率模型	329
12.11	习题 12	333
13	马尔柯夫链数学模型	337
13.1	马尔柯夫链的基本概念及其表示	337
13.2	正则马尔柯夫链	346
13.3	吸收马尔柯夫链	351
13.4	习题 13	360
	参考文献	363

1 概率论基础知识

概率论是研究随机现象数量规律的数学分支,是数理统计学的数学基础。本章仅介绍联系于生物与医药学有关的一些概率论基础知识,需要完整了解这部分内容的读者可以参考概率论教科书,比如文献[1]、[2]和[3]等。

1.1 随机事件及其概率

1. 试验、随机事件、样本空间

自然现象,一般可以分为确定性现象和非确定性现象。在某些确定条件下,可能发生也可能不发生的现象,称为非确定性现象,也称为随机现象。例如,小麦种子在播种后可能发芽也可能不发芽;投掷一枚匀质硬币,其结果可能是硬币正面朝上也可能是背面朝上。统计学所研究的是随机现象,而其研究过程往往基于试验的结果。这里“试验”一词是一个比较宽泛的概念。我们把为发现某事物特征所进行的观察过程统称为试验。若这个试验具有以下三个特点:①可在相同的条件下重复进行;②试验结果不止一个,但能明确所有的结果;③试验前不能预知出现哪种结果,则称这种试验为随机试验,用 E 表示。随机试验中所伴随的随机现象具有三个特点:①每次试验前不能预言出现什么结果;②每次试验后出现的结果不止一个;③在相同的条件下进行大量观察或试验时,出现的结果有一定的规律性,这种规律性称之为统计规律性。

随机试验 E 所有可能的结果所组成的集合称为样本空间,记为 Ω 。样本空间的元素,即 E 的直接结果称为样本点(或基本事件),常记为 ω ,即 $\Omega = \{\omega\}$ 。 Ω 的任意一个子集合称为随机事件,简称事件,记为 A, B, \dots 它是满足某些条件的样本点所组成的集合。 A 与 Ω 的关系一般可用文氏图表示。

表1-1给出一组随机试验及其相应样本空间的例子。

表 1-1 随机试验结果和样本空间

试 验	试 验 结 果	样 本 空 间
检查一瓶药剂	合格,不合格	$\Omega = \{\text{合格,不合格}\}$
一天内繁殖的细菌数	1, 2, ..., N	$\Omega = \{1, 2, \dots, N\}$
掷一骰子	正面号码为 1, 2, 3, 4, 5, 6	$\Omega = \{1, 2, 3, 4, 5, 6\}$
小白鼠注射药剂后存活时间	T	$\Omega = \{x \mid 0 \leq T < \infty\}$

在随机事件中,仅由一个样本点组成的子集合称为基本事件,它是随机试验的直接结果,每次试验必定发生且只可能发生一个基本事件。组成某一随机事件的样本点同时发生了,称该随机事件发生了。全体样本点组成的事件称为必然事件,即为样本空间 Ω ,它是每次试验必定发生的事件。不包含任何样本点的事件称为不可能事件,记为 \emptyset ,它是每次试验必定不发生的事件。

2. 事件的关系和运算

1) 随机事件之间的关系

(1) 包含关系: $A \subset B \Leftrightarrow x \in A \Rightarrow x \in B$ 。它意味着,事件 A 发生必导致事件 B 发生。

(2) 相等关系: $A = B \Leftrightarrow A \subset B$ 且 $B \subset A$

(3) 互斥关系: 两事件 A 和 B 称为互斥的或称互不相容,假如 $A \cap B = \emptyset$ 。

(4) 对立关系: 如果两事件 A 和 B 满足 $AB = \emptyset$ 且 $A \cup B = \Omega$,则称 A 与 B 互为对立事件。

(5) 完备事件组: 如果 n 个事件 A_1, A_2, \dots, A_n 是相互互斥的,且 $\Omega = \bigcup_{i=1}^n A_i$,那么 $\{A_1, A_2, \dots, A_n\}$ 称为完备事件组,或称为 Ω 的剖分。

2) 随机事件之间的运算

(1) 积事件: 事件 A 与 B 的积事件,记为 $A \cap B \Leftrightarrow$ 指事件 A 和事件 B 同时发生。

(2) 和事件: 事件 A 和 B 的和事件,记为 $A \cup B \Leftrightarrow$ 指事件 A 或 B 至少有一个发生。

(3) 事件 A 的补集: 由 Ω 中不属于集合 A 的样本点组成,记为 \bar{A} 或 A^c 。

(4) 差集运算: 事件 A 与事件 B 的差集,记为 $A - B$ 或 $A \setminus B$,它由属于集合 A 而不属于集合 B 的样本点组成 \Leftrightarrow 事件 A 发生,但事件 B 不发生。

3) 随机事件的运算律对应于通常的集合运算律

1.2 概率的定义

1. 概率的定义

历史上概率(probability)的定义经历了古典定义、统计定义和公理化定义三个阶段。事件的概率是指某一事件发生的可能性大小,这种“概率”的描述性定义称为古典定义。“概率”的统计定义是基于“频率”这一概念进行的。

定义 1.1 在 n 次试验中,设事件 A 发生了 m 次,则称 $f_n = \frac{m}{n}$ 为事件 A 发生的频率。

概率的统计定义是指:在相同条件下重复进行的 n 次试验中,事件 A 发生的频率稳定地在某一常数 p 附近摆动,且随 n 越大摆动幅度越小,则称 p 为事件 A 的概率,记作 $P(A)$ 。

性质 1.1 频率具有以下四个性质:

(1) 非负性: $0 \leq f_n(A) \leq 1$ 。

(2) 归一性: $f_n(\Omega) = 1$ 。

(3) 可加性:如果事件 A, B 互斥,则 $f_n(A \cup B) = f_n(A) + f_n(B)$ 。可推广到有限个两两互斥事件的和事件。

(4) 稳定性: $\lim_{n \rightarrow \infty} f_n(A) = P(A)$ 。

概率的公理化理论由苏联数学家柯尔莫哥洛夫(Kolmogorov)于 1933 年建立。

定义 1.2 设 Ω 是随机试验 E 的样本空间,若能找到一个法则,使得对于 E 的每一事件 A 赋予一个实数,记为 $P(A)$,称之为事件 A 的概率,这种赋值必须满足下面的三个条件:

(1) 非负性: $\forall A \subset \Omega, P(A) \geq 0$ 。

(2) 归一性: $P(\Omega) = 1$ 。

(3) 可列可加性: $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$,其中 A_1, A_2, \dots 为两两互斥事件。

概率本质上是以事件 A 为自变量,以 $[0, 1]$ 为因变量取值范围的函数,这个函数必须满足以上定义中的三个条件。

性质 1.2 根据概率的定义, 易见

(1) $P(\emptyset) = 0$ 。

(2) 假如 $A_1, A_2, \dots, A_n \subset \Omega$ 两两不相交, 那么 $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$ 。

(3) $P(\bar{A}) = 1 - P(A) \Rightarrow P(A) \leq 1$ 。

(4) 假如 $A \subset B \Rightarrow P(B - A) = P(B) - P(A) \Rightarrow P(A) \leq P(B)$ 。

(5) 对于 A 和 B , 由关系式 $B - A = B \cap \bar{A}$ 可得 $P(B - A) = P(B) - P(AB)$ 。

例 1.1 在有两个孩子的家庭中, 孩子性别的组成有 4 种类型, 即男男, 男女, 女男, 女女。它们是 4 个基本事件, 而且是互斥且等可能的。设 A_1 为事件“两个男孩”, 它是 4 个基本事件 (n) 中的一个 (m), 概率为 $P(A_1) = \frac{1}{4} = 0.25$ 。设 A_2 为事件“第一个是男孩”, 该事件包括男男, 男女两个基本事件。 A_2 的概率为 $P(A_2) = \frac{2}{4} = 0.5$ 。

2. 患病率与发病率

在临床医学中, 患病率及发病率常用来描述一些特殊情形的概率。

定义 1.3 患病率 (prevalence) 是指某人在考察或研究的瞬时患病的概率, 它不依赖于人们是何时得的病, 其值为现时有病的人数除以所研究的总体中人数的总数。

例 1.2 美国政府 1974 年调查了总数为 144 380 位 17 岁及以上的人群, 发现患高血压的人数为 22 626 人, 则患病率为 15.7%。

定义 1.4 某病的累加发病率 (cumulative incidence, 简称为发病率) 是指人们以前没有此病而在某个指定时间内患有该病的概率。

例 1.3 某地区 40 岁到 44 岁的妇女在 1970 年 1 月 1 日到 1970 年 12 月 31 日止的时间内乳腺癌的累加发病率是千分之一。这个值意味着, 40 岁到 44 岁的妇女 1000 人, 在 1970 年 1 月 1 日以前未曾有乳腺癌, 而到 1970 年 12 月 31 日时, 大约有 1 名妇女患有乳腺癌。

3. 条件概率

实践中经常需要研究在事件 A 已发生的条件下, 事件 B 发生的概率。这种概率称为已知事件 A 发生的条件下, 事件 B 发生的条件概率, 记为 $P(B|A)$ 。

例 1.4 施用两种不同药物杀灭螟虫, 结果如表 1-2 所示。

表 1-2 不同药物杀灭螟虫统计

	死亡(A)	存活(\bar{A})	合 计
药物甲(B)	96	24	120
药物乙(\bar{B})	64	16	80
合 计	160	40	200

从 200 只虫子中任取 1 只,这虫是死虫的概率为 $P(A) = \frac{160}{200} = 0.80$ 。从 200 只虫子中任取 1 只,该虫接受了药物甲的概率为 $P(B) = \frac{120}{200} = 0.60$ 。该虫接受药物甲且死亡的概率为 $P(AB) = \frac{96}{200} = 0.48$ 。如果已知某只虫子死亡了,试问该虫子接受药物甲的概率是多少? 这就是一个计算条件概率问题。因为已知死亡总数为 160 个,其中接受药物甲的虫子个数为 96,故

$$P(B | A) = \frac{96}{160} = \frac{96/200}{160/200} = \frac{P(AB)}{P(A)} = \frac{0.48}{0.80} = 0.60 \quad (1-1)$$

关系式(1-1)具有普遍性。在一般情况下,我们将上述关系式作为条件概率的定义。

定义 1.5 设 A, B 是两个事件,且 $P(A) > 0$, 称

$$P(B | A) = \frac{P(AB)}{P(A)} \quad (1-2)$$

为在事件 A 发生的条件下事件 B 发生的条件概率。

条件概率也是概率,它具有类定义 1.3 和性质 1.2 的概率基本运算公式。

4. 概率乘法法则与独立事件

由式(1-2)可以得到概率乘法法则为

$$P(AB) = P(B)P(A | B) \text{ 或 } P(AB) = P(A)P(B | A) \quad (1-3)$$

该法则表明: 两事件交的概率等于前一事件的概率乘以另一事件在已知前一事件已发生条件下的条件概率。

定义 1.6 若事件 A 的发生,并不影响事件 B 发生的概率,即

$$P(B | A) = P(B) \text{ 或 } P(A | B) = P(A) \quad (1-4)$$

此时称 A 和 B 是独立事件。

在例 1.4 中, $P(A)=0.80$, $P(A|B)=0.80$, 即 $P(A|B)=P(A)$, 这说明死亡与否并不受是否接受甲药物的影响。因此 A 和 B 是独立的。对于独立事件, 概率乘法公式为

$$P(AB) = P(A)P(B) \quad (1-5)$$

在例 1.4 中, 由 A 和 B 的独立性也可计算得 $P(AB) = P(A)P(B) = 0.48$ 。

5. 贝叶斯法则与筛选检验、ROC 曲线

定义 1.7 一个筛选检验的预测值阳性(predictive value positive, PV^+)是指一个人在该检验中呈阳性(A)条件下而患病(B)的概率, 即定义为

$$PV^+ = P(B|A) = P(\text{疾病} | \text{检验}^+)$$

而预测值阴性(predictive value negative, PV^-)是指一个人在该检验中呈阴性(\bar{A})条件下而未患病(\bar{B})的概率, 即定义为

$$PV^- = P(\bar{B}|\bar{A}) = P(\text{无疾病} | \text{检验}^-)$$

定义 1.8 一个症状(或一组症状, 或筛选检验)的灵敏度(sensitivity)是疾病发生后出现症状的概率, 即灵敏度 $= P(A|B)$ 。

定义 1.9 一个症状(或一组症状, 或筛选检验)的特异度(specificity)是疾病不发生时不出现症状的概率, 特异度 $= P(\bar{A}|\bar{B})$ 。

定义 1.10 某实验的结果是阴性但实际是阳性, 称为假阴性(false negative), 实验结果是阳性但实际是阴性, 称为假阳性(false positive)。

在预测疾病中称某个症状是有效的, 是指该症状的灵敏度和特异度两个指标都是高的。如何使一个(或一组)症状的灵敏度或特异度指标能被医生所估计且能用于计算预测值? 由条件概率的定义 1.5, 有

$$PV^+ = P(B|A) = \frac{P(B \cap A)}{P(A)} \quad (1-6)$$

又由 $P(A \cap B) = P(A|B) \times P(B)$, 定义 1.2 的(3)及乘法公式(1-3)可得全概率公式为

$$P(A) = P(A|B) \times P(B) + P(A|\bar{B}) \times P(\bar{B})$$

把 $P(B \cap A)$ 及 $P(A)$ 的表达式代入式(1-6)得贝叶斯公式, 有

$$PV^+ = P(B | A) = \frac{P(A | B) \times P(B)}{P(A | B) \times P(B) + P(A | \bar{B}) \times P(\bar{B})} \quad (1-7)$$

若记 $x = P(B)$ 表示有关总体中疾病的患病率,式(1-7)可以改写成

$$PV^+ = \frac{\text{灵敏度} \times x}{\text{灵敏度} \times x + (1 - \text{特异度}) \times (1 - x)} \quad (1-8)$$

也就是说, PV^+ 能表示成灵敏度、特异度及患病率的函数。类似地,有

$$PV^- = \frac{\text{特异度} \times (1 - x)}{\text{特异度} \times (1 - x) + (1 - \text{灵敏度}) \times x} \quad (1-9)$$

在某些情况下,一个实验可提供多个结果而不是简单的阳性或阴性。在另外情形中,实验结果是连续性变量。在上述任一情形下,为识别实验是阳性或阴性的切断点(cut-off-point)常是任意的。

例 1.5 下面材料由 Hanley 及 McNeil 提供。疑似有神经系统疾病的 109 名受试者接受某放射学家的 CT 成像术的检查,结果用等级(rating)形式表示。每一个受试者实际上是否有病是早已知道的。数据列于表 1-3,我们如何计算诊断的准确性?

表 1-3 某放射学家做的 CT 成像术的等级结果

疾病的 实际状态	CT 成像诊断					总 数
	肯定正常 (1)	可能正常 (2)	有问题 (3)	可能不正常 (4)	肯定不正常 (5)	
正常	33	6	6	11	2	58
不正常	3	2	2	11	33	51
总数	36	8	8	22	35	109

不同于过去例子的是,表 1-3 中未给出明确的切断点用于诊断疾病。对表 1-3 我们对切断点给出不同的值,就会有不同的结果。比如,如果我们规定“可能不正常”与“肯定不正常”作为检验的阳性(即把等级 4 或 5,或 4 及以上作为阳性),则这个检验的灵敏度为 $(11 + 33)/51 = 44/51 = 0.86$,而同时特异度是 $(33 + 6 + 6)/58 = 45/58 = 0.78$ 。表 1-4 就是按照这种不同的准则来计算灵敏度及特异度的。

表 1-4 对表 1-3 数据放射学家使用不同准则去确定阳性时的灵敏度与特异度

检验的阳性准则	灵敏度	特异度
1 ⁺	1.0	0
2 ⁺	0.94	0.57
3 ⁺	0.90	0.67
4 ⁺	0.86	0.78
5 ⁺	0.65	0.97
6 ⁺	0	1.0

定义 1.11 接受者操作特性(receiver operating characteristic, ROC)曲线是灵敏度相对于(1-特异度)的筛选检验曲线,此处曲线上的不同点对应于用不同的切断点去识别阳性。

例 1.6 图 1-1 是对表 1-4 中的数据构建的 ROC 曲线,其以“1-特异度”为横坐标,以灵敏度为纵坐标构建而成。

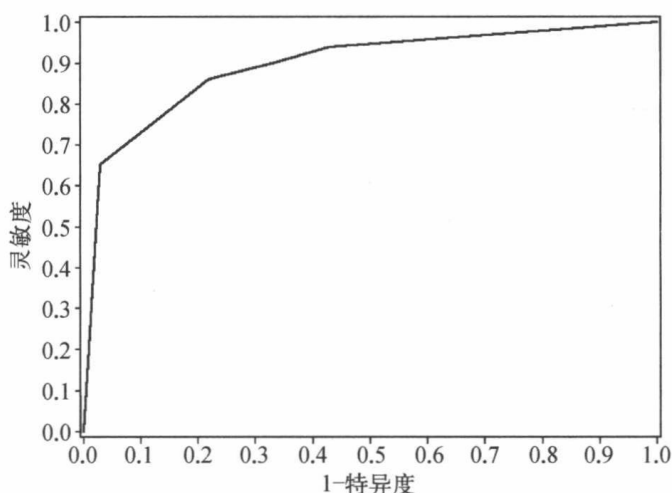


图 1-1 ROC 曲线图

ROC 曲线下的面积是检验整个诊断精度的一个合理指标。ROC 曲线作出后,对于任何一个等级(即为切断点),若接受试者的检验等级值低于上述切断点,则诊断受试者为正常,若高于切断点则认为是不正常。

使用梯形面积法,则可算出该面积为 0.89,这意味着放射学家能按 CT 等级的相对顺序从而正确地把一个正常人从不正常人中识别出来的概率是 89%。当正常受试者与不正常受试者有相同等级时,放射学家可以随意地做出决定。