

实用 统计学

11
12

福建科技出版社

实用医学统计

林求诚

福建科学技术出版社

一九八五年·福州

2626

实用医学统计

林 求 诚

*

福建科学技术出版社出版

(福州得贵巷27号)

福建省新华书店发行

福建新华印刷厂印刷

开本787×1092毫米 1/32 8.5625印张 186千字

1985年3月第1版

1985年3月第1次印刷

印数: 1—5,740

书号: 14211·79 定价: 1.51元

前 言

临床工作者每天都要接触许多数据。有数据,就有统计的问题。老医师之所以有许多宝贵的经验,就是因为他在临床实践中,对各种客观现象进行了长期的观察和积累,从而在脑子里留下了这些客观现象统计分布的印象,并且自觉地或不自觉地运用了统计方法的结果。所以,学习和掌握医学统计的知识和方法,不但能有效地推动科学实验的进行,还有助于医疗水平的提高。

过去国内出版的有关医学统计的书籍,内容多偏重于预防医学或基础医学。有感于此,作者乃决心编写一本偏重于临床医学的较为实用的医学统计小册子。这本小册子主要是根据作者在多年的临床医疗和科研中所应用的统计方法,总结编写而成的,着重介绍各种医学统计方法的应用,并尽可能用直观性的逻辑说明代替繁琐的理论推导,多举一些实际例子,以使读者能在较短的时间内学会并掌握临床医学中常用的统计方法。

本书大部分内容承蒙第四军医大学郭祖超教授审阅,在此表示衷心的感谢。

由于作者的水平和实践经验有限,书中不足和错误之处在所难免,恳请读者批评指正。

林 求 诚

一九八四年一月

目 录

第一章 概论	(1)
一、几个基本概念与知识	(1)
二、正确运用统计方法分析试验结果	(7)
第二章 统计工作的步骤	(9)
一、统计资料的收集	(9)
二、原始资料的检查	(12)
三、统计资料的整理	(14)
四、统计资料的分析	(19)
第三章 计量资料的计算与分析	(21)
一、平均数	(21)
二、变异指标	(25)
三、正常值范围	(29)
四、t检验	(42)
五、t'检验	(49)
六、F检验	(51)
第四章 计数资料的计算与分析	(57)
一、“率”与“比”的计算	(57)
二、相对数的正确应用	(59)
三、“率”的标准误和可信限	(60)
四、“率”的差别显著性检验	(62)
五、“率”的标准化	(64)

六、卡方检验	(71)
七、零反应检验	(76)
八、四格表精确检验	(78)
第五章 非参数统计	(81)
一、符号检验	(82)
二、秩和检验	(83)
三、Ridit检验	(89)
第六章 简单相关与回归	(98)
一、相关系数	(98)
二、直线回归	(103)
三、等级相关	(110)
第七章 多变量分析	(113)
一、多指标的差异显著性检验	(113)
二、逐步回归分析	(128)
三、判别分析	(140)
第八章 诊疗统计	(145)
一、正常值或异常值的确定	(145)
二、二项分布的临床应用	(147)
三、普哇松分布的临床应用	(153)
四、关于血细胞分类数据的处理	(156)
五、动态观察	(158)
第九章 疾病统计和医疗质量统计	(167)
一、常用的疾病统计指标	(167)
二、医疗质量统计	(177)
三、医疗质量管理	(184)
第十章 医学科研设计、实验设计与实验统计方法	(189)

一、科研设计与实验设计	(189)
二、实验误差与避免误差的方法	(206)
三、实验室误差与误差的控制	(210)
四、有效数字及其运算法则	(215)
五、统计表	(220)
六、统计图	(224)
附表	(232)
附表一 正态性D检验界值表	(232)
附表二 t值表	(234)
附表三 F值表 (方差齐性检验用)	(240)
附表四 F值表	(242)
附表五 F值表 ($\alpha = 0.3$)	(250)
附表六 q值表	(252)
附表七 χ^2 值表	(253)
附表八 秩和检验表	(261)
附表九 相关系数之5%界与1%界	(262)
附表十 等级相关系数的显著性界值表	(263)
附表十一 游程总数检验表	(263)
附表十二 游程检验 I 值的显著界限	(265)

第一章 概 论

医学统计学就是把数理统计应用于医学科学研究和疾病防治实际工作的一门科学。医学统计方法可以帮助医务人员合理地安排实验设计，准确地搜集整理资料，科学地分析资料，以便通过众多的、偶然的、次要的因素，揭示出事物内在客观规律，把感性的认识提高到理性认识。

一、几个基本概念与知识

（一）总体与样本

根据研究目的所确定的研究对象的全体称为总体。样本是总体中被选为代表的一部分，而样本的每个个体称为样本单位。根据样本所包含的样本单位数，可分为大样本和小样本，习惯上把样本单位数超过30的叫做大样本，等于或者小于30的样本叫做小样本。样本的作用在于估计总体。总体所包含的样本数可以是无穷多，而且往往只是设想的或者是抽象的，但也可以是有限的、实际的。在实际工作中，手头上所掌握的资料经常只是总体的一小部分，只是一个样本。科学研究工作，可以根据一个样本所得到的结果，而期望知道总体的情况。如何正确地从样本来推测总体，这就是统计学所要解决的根本问题。

（二）抽样误差

从有变异性的总体中随机抽取样本所计算出的统计指标和总体所计算出来的同样统计指标并不完全相等，也就是说它们之间存在差异。这种差异称为抽样误差。它主要是个体差异和抽样时的机遇所造成的，因此也是客观存在的，是不可避免的。举一个最简单的实验来说明：有一个大口袋，内盛有一万个圆滑的同样大小的玻璃球，红球两千个，白球八千个，即红球的比例占20%，白球占80%。这是总体的比例数。把口袋的玻璃球充分地混匀后，若从中随机抽取100个球，可能红球占21%，白球占79%；然后把100个放回口袋中再充分混匀后，再从中随机抽取了100个球，结果红球占24%，白球占76%；这样反复抽取多次，红球占的比例数可能是18%、19%、16%、25%、23%……也可能是20%。这种样本比例数与总体比例数之间的差异，就是抽样误差。只要有抽样，就会有抽样误差，但是这些样本比例数总是围绕总体比例数呈一定规律的分布。同样，样本的其它统计值（如均数）与总体相应数值之间的数学关系，也有一定的规律。影响抽样误差的因素，是事物变异的程度和样本的大小。一般说来，样本越大，则抽样误差越小，统计值和总体的情况越接近，正确性也越高，换言之，越能阐明事物的规律性。

（三）随机化

随机化是抽样研究和抽样分配中十分重要的原则。从研究对象的总体中抽取样本进行研究时，为了使样本对总体有较好的代表性，并使其抽样误差的大小可以用统计方法来估计（因为一般数理统计所提供的方法是在随机抽样的基础上推演而来的），就必须遵守随机化原则。即在抽取样本以前，要使总体中的每个单位都有同等的被抽取的机会。同样的，在临床分组对照实验中，也必须用随机的方法，使各个

病例有同等的机会被分配到各组去，以免人为地造成各组间的不齐同。

(四) 概率 (机率、或然率)

它是反映某一事件发生的可能性大小的量，统计学上常用符号 P 来表示。必然发生的事件概率为1，不可能发生的事的概率为0，一般的概率范围是在0与1之间。如 $P=0.05$ ，即表示可能发生 (或不可能发生) 的机会为5%。

(五) 自由度 (符号df)

自由度是数学上的名词。用通俗的比喻，穿在一根线上的玻璃珠只能沿着线来回移动，所以玻璃珠运动的自由 (活动) 度只有一个；滑冰运动员在滑冰场上滑行，就有前后、左右两个自由度；鸟或飞机在空中飞行，就有上下、左右、前后三个自由度。数学上把 x_1 、 x_2 、 x_3 …… x_n 作为 n 度空间里的一个点，如不加任何条件限制，这个点子有 n 个自由度。现在如果加上以下一个约束条件： $x_1 + x_2 + x_3 + \dots + x_n = \Sigma x$ (这里 Σx 为 n 个值的总和)，则自由度就减少了一个，即这个点子的运动只有 $n-1$ 个自由度了。这就是自由度等于样本数减1的初步原理。例如，设 $x_1 + x_2 + x_3 + x_4 = 20$ ，如果 $x_1 = 1$ ， $x_2 = 2$ ， $x_3 = 3$ ，则 x_4 必然为14；如 $x_1 = 4$ ， $x_2 = 5$ ， $x_3 = 6$ ，则 x_4 必然为5。即在总和固定为20的情况下，任意取值的可能性只有三个，剩下一个是不能任意取值的，此时自由度为 $4-1=3$ 。

(六) 显著性检验

在医学科学研究中，进行比较、鉴别是经常应用的手段。有些调查研究的结果，可以通过直接比较得出结论；而有些实验研究的结果，则不能只凭表面数字直接比较得出结论，需经一定的统计方法处理。例如，用甲、乙两种疗法各治疗

观察50例同类慢性支气管炎患者，甲疗法的有效率为80%，乙疗法的有效率为70%，两者相差10%。这可能包括两种情况：一是甲疗法真正优于乙疗法，二是甲疗法并不优于乙疗法，这里所高出之10%，纯属偶然因素所造成的。这就必须用“差异显著性检验法”估算出相差的10%属于后一种情况的可能性有多大。如果属于后一种情况的可能性只有5%以下，属于前一种情况的可能性便有95%以上了，这时便记为 $P < 5\%$ ，说明甲疗法确实优于乙。

由此可见，显著性检验的目的在于求出 P 值，然后根据 P 值的大小作出结论。统计学给 P 值规定两个界限，表明三种情况，即：

- (1) $P > 5\%$ ，差异不显著；
- (2) $1\% < P \leq 5\%$ ，差异显著；
- (3) $P \leq 1\%$ ，差异非常显著。

应该指出，统计的结论永远是推断性的。因此，即使是 $P < 1\%$ 或 $P < 0.1\%$ ，也不是说就绝对有差异了，只是说在这种情况下做出有差异的结论时，搞错的可能性很小，小到平均每做100次或1000次这样推断时，可能只有1次搞错。反之， $P > 5\%$ 或10%，也并不是说就绝对没有差异了，只是说在这种情况下做出有差异的结论时，搞错了的可能性太大了一些，大到平均每20次或10次就有1次搞错了，这样研究结果的再现性较小。

当两组的差别已经没有实际意义时，可以不必进行显著性检验。例如对降压药的研究，如果某降压药在人体只能降低血压2毫米汞柱时，即使肯定是药物的作用，也无实际用途，因而去推敲它究竟是否显著便没有意义了。

显著或不显著，是相对的，而不是绝对的。实验者可以

根据具体情况，适当选择显著性概率水平，不一定拘泥于 $P = 0.05$ 的界限，尽管这是最常用的公认水平。例如，肿瘤药物的筛选，有时可采用 $P = 0.10$ ，甚至退至 $P = 0.20$ 。

样本之间的可比性，是作显著性检验的必要前提。可比性是指两样本除要比较的因素外，其他影响观察结果的主要因素、次要因素都要控制“齐同”。否则，进行显著性检验是徒劳的。例如，在比较两种药物的疗效时，先要看两组治疗的时间是否相同？地点是否相同？中西医病型是否相同？病期是否相同？病情是否相同？治疗条件是否相同？必要时先要进行“齐同性”的显著性检验，在基本一致的情况下再进行疗效的差异显著性检验。

(七) 正态分布曲线

数学上的正态分布曲线，是钟形对称的曲线。平均数所在处最高，正态曲线下面积表示频数或机率。横轴上的距离，从平均数所在处为原点，大于平均数的一边为正值，小于平均数的一边为负值。正态分布具有集中性、对称性和均匀变动性三大特点。由于正态曲线左右对称，故平均数左右的面积各占50%。曲线下面积有一定的分布规律，图中 μ 为总体均数， σ 为总体标准差。我们可以应用正态曲线的理论，对类似正态分布的变量频数分布作出概括的估计，并用样本均数 $\bar{x}$ 作为 μ 的估计值，用样本标准差 S 作为 σ 的估计值。曲线下的面积分布百分比，常用以下几种：

(1) 均数减一个标准差到均数加一个标准差（即 $\mu \pm \sigma$ ）的范围内，包括全部面积的68.27%；超出此范围的面积，占31.73%。

(2) 均数减1.96个标准差到均数加1.96个标准差（即 $\mu \pm 1.96\sigma$ ）的范围内，包括全部面积的95%；超出此范围

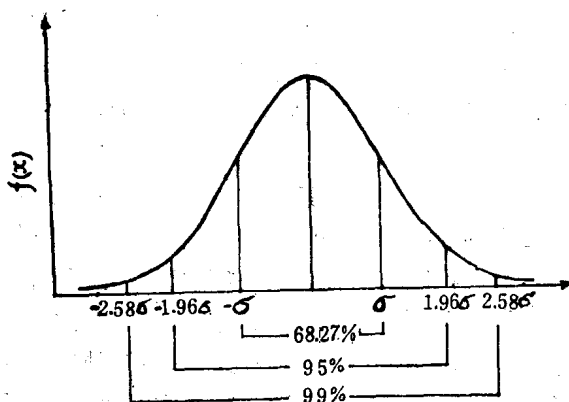


图1-1 正态曲线下的概率分布图

的面积，占5%。

(3) 均数减2.58个标准差到均数加2.58个标准差（即 $\mu \pm 2.58\sigma$ ）的范围内，包括全部面积的99%；超出此范围的面积，占1%。

也就是说，在一次从总体随机抽样中，样本数值位于 $\mu \pm \sigma$ 区间内的概率为68.27%，位于 $\mu \pm 1.96\sigma$ 区间内的概率为95%，位于 $\mu \pm 2.58\sigma$ 区间内的概率为99%。医学上常根据这个原理应用样本的均数和标准差来估计正常值范围。例如，1973年某地120名12岁健康男孩的身高，其平均数为143.10厘米，标准差为5.67厘米，其95%正常值范围为 $143.10 \pm 1.96 \times 5.67$ 即132.0~154.2厘米，99%正常值范围为 $143.10 \pm 2.58 \times 5.67$ 即128.5~157.7厘米。

(八) 单侧检验与双侧检验

例如用甲、乙两种药物对比观察治疗阑尾炎的疗效。如果在实验设计时没有足够专业方面的理论根据判断两种药物中何者为优，这时在作统计推断时，就要考虑在疗效方面

可能甲药比乙药好，也可能乙药比甲药好，在处理实验数据时就应作双侧检验；如果有足够专业方面的理论根据，认为只可能甲药比乙药好，而不可能乙药比甲药好，也就是在否定了无效假设后作统计推断时，只有一种可能的结果出现，这时，统计上称这种检验为单侧检验。如甲药为抗菌素加增效剂，乙药为单用抗菌素，目的在于确定甲药比乙药好这样一种情况，即为单侧检验。

在直接计算概率的检验方法（如零反应检验法等）中，计算之 P 值为单侧概率，并规定单侧检验的显著性水平 $P=0.05$ ，双侧检验的显著水平为 $P=0.025$ ，即单侧检验的概率是双侧检验的一倍。由于两种检验涉及到判断的显著性水平的选择，因此，要求在实验设计时就要确定该试验属何种检验，不能事后由研究者任意选定。

二、正确运用统计方法分析试验结果

有些人认为“没有统计也过得去”，仅凭数字的表面值就可作出判断。当实验资料数量多而且两组间的差别相当大的时候，凭常识判断和统计处理的结论可能是一致的。但如果实验资料少，特别是在两组差别不太大的情况下。如不考虑机遇问题，仅凭常识判断下结论，那就难免要发生错误。例如某些药品并没有什么特殊疗效或预防效果，但被错误地认为有效，并且推广应用，以致造成一定的损失。这说明科学研究需要正确的统计方法处理。

数据处理仅是一种工作方法，统计不能代替原来科研工作的科学性，如果数据本身的来源、收集、分配有严重的缺

点和错误，统计是统计不出来的，更不用说去纠正这些错误了。恰好相反，不用统计，作者不得不作许多定性的描述，从中可能发现问题，从而可以去纠正错误，科学性也因而提高。现在有那么一种倾向，似乎科研结果只要报出一个显著性检验的P值，于是万事大吉。用了统计，表面上看起来方法好象很“科学化”，数据更整齐了，有什么t值、 χ^2 值、R值等等，似乎令人信服，即使有心的读者或编辑，最多也不过重复一下计算过程，多数发现是正确无误。然而数据处理的科学化丝毫不能代替数据本身的科学性，有时甚至相反。如果数据来源有严重的非科学性，统计处理反而会起到掩盖作用。

还应该注意，统计的方法和演算技术是用来贯彻统计概念和原理的，前者是工具，后者才是主体。如果舍本逐末地去追求方法和演算技术，不认真考虑在什么具体情况下应该如何对待，不正确掌握概念和原理而机械地乱套统计公式，有时会犯错误的。

第二章 统计工作的步骤

收集资料、整理资料和分析资料是统计工作的三个基本步骤。收集资料就是根据研究目的进行调查或实验设计，然后根据设计要求收集准确与完整的原始资料，这是统计工作的前提与基础。整理资料就是把收集的资料，有目的、有计划地进行科学加工，使资料系统化、条理化，以便进行统计分析。分析资料就是把统计整理的结果，进一步计算出相应的指标，结合专业知识，运用统计方法进行分析对比，阐明事物的规律性，应用到医疗卫生实践中去。这三个步骤是紧密联系的，某一环节发生错误，都可影响整个统计工作，特别是收集资料如果有明显的错误或缺点，在以后两个步骤中将无法补救或纠正。

一、统计资料的收集

收集资料事先要有明确的目的、切实的设计，对所研究的问题要有概括的了解。收集资料常用的方法有以下几种：

(一) 统计报表

统计报表用表格形式，按照一定的时间程序，系统收集统计资料。卫生工作统计表是由国家卫生领导部门制定，并指定各个卫生机构填报，自下而上地提供给有关卫生部门的

统计资料。如卫生基本情况年报表、医院工作年报表、居民病伤死亡原因年报表、传染病年（月）报表、疫情旬（月、季）报表等。统计报表的填写要注意以下几点：

（1）必须按照报表中规定的项目和格式填写报表，未经批准，不得自行修改。

（2）严格按照规定日期和程序提交报表。

（3）保证报表资料的正确性和完整性，不得伪造统计数字。

（4）保守国家机密。

（二）日常医疗卫生工作记录和报告卡

日常医疗卫生工作记录有：门诊病历、住院病历、健康检查记录、门诊工作日志、病区工作日志、检验工作记录等。

报告卡有：传染病报告卡、肿瘤报告卡、职业病报告卡、出生报告卡、死亡报告卡等。报告卡容易出现重复、遗漏或项目不清等情况，要经常检查、核对，发现错误及时纠正。

应用医院病历资料时，要注意以下几个问题：

（1）医院病历可能具有挑选性。大医院的住院病例往往多是病情较重的或疑难的病例，若用以代表这种疾病的一般性是不全面的。

（2）时期不同，医院病历资料也往往有所不同。例如，收治某种病的病人除了治疗方法的不断改进外，医务人员的业务水平、服务态度和护理质量的不断提高等，对于治疗效果都有不同程度的影响。在分析资料时，应予以注意。

（3）一个地方医院病历资料的总和，往往不等于该地总的发病人数，用以统计发病率时要估计到这个问题。