



实用医学统计方法

长春生物制品研究所

实用医学统计方法

李 守 成 编

长春生物制品研究所

1 9 7 6

毛主席语录

胸中有“数”。这是说，对情况和问题一定要注意到它们的数量方面，要有基本的数量的分析。任何质量都表现为一定的数量，没有数量也就没有质量。我们有许多同志至今不懂得注意事物的数量方面，不懂得注意基本的统计、主要的百分比，不懂得注意决定事物质量的数量界限，一切都是胸中无“数”，结果就不能不犯错误。

伟大领袖毛主席教导我们：对情况和问题一定要注意到它们的数量方面，要有基本的数量的分析。任何质量都表现为一定的数量，没有数量也就没有质量。我们有许多同志至今不懂得注意事物的数量方面，不懂得注意基本的统计、主要的百分比，不懂得注意决定事物质量的数量界限，一切都是胸中无“数”，结果就不能不犯错误。毛主席关于数量和质量辩证关系的教导对于医学科学研究工作，如医学免疫制品的生产、研究、使用、流行病学调查和卫生防疫及临床医学的研究工作等均具有非常重要的指导意义。

当前，卫生战线和全国各条战线一样形势一片大好。在毛主席革命卫生路线的指引下，广大卫生人员遵照毛主席光辉的“六、二六”指示，认真贯彻执行“面向工农兵，预防为主，团结中西医，卫生工作与群众运动相结合”的方针，在防治人民疾病和开展医学科学研究方面都取得了可喜的成果。赤脚医生、合作医疗等社会主义新生事物欣欣向荣。农村的缺医少药状况和城乡的卫生面貌都有了显著改善。从而提高了广大工农兵群众的健康水平。这是无产阶级文化大革命在卫生战线取得的丰硕成果，也是对党内最大的不肯改悔的走资派邓小平刮起的右倾翻案风的有力回击。为了深入批判邓小平鼓吹的“三项指示为纲”这一修正主

义纲领，巩固和发展无产阶级文化大革命的胜利成果，继续搞好卫生革命，在开展各项卫生工作中，做到胸中有“数”广大卫生人员迫切希望了解一些简明实用的医学统计方法。因此，我们在“东北协作区免疫工作会议”与会代表的热情鼓励下，把会议上介绍的医学统计讲稿经补充、修改，整理成这本《实用医学统计方法》。医学统计方法近年来在理论和实践的结合上有了很大的进展，内容很丰富。但是，对广大初学者来说，当前急需了解一些通俗、易懂、便于实用的统计方法。考虑到我们的工作仅能在整个医学工作中，从事某个方面的研究，因而，我们对整个医学统计的内容也仅能涉及某些部分，这样我们可以根据工作需要选择常用的一些统计方法有针对性的学习。在学习和应用中逐步了解某些基本原理。就像显微镜一样，它是医学研究的一个工具，我们应当掌握它的操作方法，并一般地了解一些应用原理，但无须像显微镜研制人员那样必须精通这方面的工艺和光学原理。医学统计也同样，它也是我们进行医学研究的一门辅助工具。因此，无须对复杂的数学公式都进行深入的探讨，只要把什么样的材料，选用什么样的统计方法，为什么用这种方法，说明什么问题等基本知识弄清楚，一般就可以了。

我们就是从实用的角度，把几年来在实际工作中接触到的一些统计学知识，加以整理成这本《实用医学统计方法》，仅供初学者参考。其内容，一般都是通过举例，介绍一些常用的方法，有些是把我们工作中的素材引为例证，有些是直接引用了参考资料中的某些素材。力求浅显、易懂、实用，没更多的涉及医学统计的原理。特别是为了使用方便，减少繁复的数学推导，除介绍一般的统计方法外，还着重介绍了吉林医科大学王广仪同志根据医学统计的

原理研制的一些统计便查图表和一些简便方法。希望大家通过实际应用加以验证，提出意见，以臻完善。

总之，这是从实际应用着眼整理的一个一般性的医学统计知识，对于致力于钻研、提高和发展医学统计这门科学的同志，这还是很肤浅的。

本材料在整理过程中，曾得吉林医科大学王广仪老师的热情帮助与指导并在百忙中予以审稿，吉林医大张志军老师利用业余时间协助校阅样稿，本所马学严同志为本稿绘制图表，在此一并表示感谢！

这本小册子，仅是我个人学习与应用医学统计的一个粗浅理解，因此，在概念的解释，基本原理的阐述和逻辑性等方面都不像医学统计专著那样内容深广、严谨、系统，这样就难免有错误的地方。特别是由于工作性质的需要，有的地方在统计方法上做了一点摹拟补充，在计算方法上还做了一点大胆的尝试，希从事医学统计专业的老师和阅过本文的同志提出批评指正。

编 者

1976年春

于长春生物制品研究所

总 目 录

总目录

第一章 绪 论

- 一、医学统计的基本意义 3
- 二、医学统计的几个基本概念及其意义 5

第二章 计数资料的统计分析

- 一、计数资料有关基本概念 33
- 二、用样本百分率估计总体百分率 37
- 三、差数百分率的差异显著性测验 45
- 四、两个样本百分率的差异显著性测验 50
- 五、几个样本百分率的差异显著性测验 62
- 六、两列样本百分率的差异显著性测验 67
- 七、多列样本百分率的差异显著性测验 75

第三章 计量资料的统计分析

- 一、计量资料有关基本概念 85

二、用样本平均数估计总体平均数	95
三、差数平均数的差异显著性测验	97
四、两个样本平均数的差异显著性测验	102
五、几个样本平均数的差异显著性测验	106
六、两组样本平均数的差异显著性测验	120
七、多组样本平均数的差异显著性测验	124

第四章 半计量资料的统计分析

一、根据资料性质选择适宜的方法	132
二、按等级分组资料的统计分析方法的比较	135
三、多列数据的等级型资料统计分析	140

第五章 相关回归资料的统计分析

一、函数与相关——不同类型的变量关系	161
二、相关分析与回归分析	163
三、直线相关——两个标识间的相互关系	163
四、直线回归——自变与倚变的关系	173
五、求相关系数及回归直线的简易方法	180
六、半数致死量与半数有效量	183

第六章 实验设计等有关基本知识

一、实验设计的一般常识	209
二、比较简单的几种实验设计	218
三、应用统计学方法处理实验结果	245
四、利用辅助工具帮助统计分析	252

第七章 附 录

一、置信限与显著界限表 (附表 1 ~ 8)	269~301
二、 t 值、 X^2 值和 F 值表 (附表 9 ~ 12)	304~308
三、随机、对数和符号检验表 (附表 13 ~ 15)	318~334
四、计量单位、常数与希腊和拉丁字母 (附表 16 ~ 21)	335~353
五、置信限与显著界限和二项概率纸 (附图 1 ~ 5)	354~358

I 绪 论

第一章

一、医学统计的基本意义	(3)
二、医学统计的几个基本概念及其意义	(5)
(一) 总体、个体和样本	(5)
(二) 抽样误差、随机抽样和统计推断	(6)
(三) 参数估计和统计假设检验	(6)
(四) 精密性、准确性和置信度	(7)
(五) 置信度和置信限的选择与运用	(8)
(六) 概率“P”的概念与意义	(9)
(七) 概率分布	(11)
(八) 抽样误差、本质差异与无效假设	(15)
(九) 统计显著性与实际显著性	(16)
(十) 计数资料与计量资料	(17)
(十一) 单侧概率与双侧概率	(17)
(十二) X^2 测验	(17)
(十三) t 测验	(23)

(十四) f 测验 (27)

(十五) 自由度 (df) (29)

I 绪 论

一、医学统计的基本意义

医用数理统计（简称医学统计）是建立在概率论基础上的一种科学的分析方法。在医学科学研究工作中（实验研究和调查研究），恰当地运用这种方法，可用较少的人力、物力和时间，取得可靠的资料和做出比较正确的结论。但是，医学统计的主要作用是逐渐认识的。18到19世纪的医学研究人员还没有完全认识医学研究和数理统计的关系以及数理统计方法对于科学研究的重要性。像英国坏血病的研究，由于缺乏数理统计的理论设计，以至拖了一个多世纪才得到解决。然而，也必须认识到，数理统计也并非是无能的，数理统计只有结合专业的特点才能发挥和发展它的作用。

医学科学实验研究的对象是有生命的——人或动物——其特点具有变异性。像利用同质气溶胶在相同条件下，进行免疫或治疗，其效果和反应不尽相同；在同样条件下生活的人们，其身体的发育及其生理指标亦不尽相同；在同样条件下进行的动物实验，其结果总有一定的差异，这些都是有变异性的现象。然而在这些变异性的现象中却蕴藏着必然的规律性。

只是这种规律性不易从个别的现象中发现，须通过大量的观测才能揭示出来。这就是说，参差不齐的现象，总是稳定在一定的范围——百分率或平均数。对于大量观测的结果，只算出较为稳定的统计数值，用其阐述研究对象的规律性的方法，称为“描述统计”，属于大样本的统计方法。

但是，在实际医学研究中除必要的普查外，往往不易大量观测或不允许进行大量观测。如气溶胶免疫后的效果观察，总不能把所有被免疫的对象都进行观测，这里不仅是费时费力，而且观测的量过大，条件不易控制，准确性反而差。因此，理论上大样本的稳定性，实际很难做到。其次，有些情况即或是能做到，但也不允许，如医学生物制品的生产，总是从制品中抽取小量的样品做检定，如果100%的制品都做了检定，那么生产为了检定就没有意义了，这样做出来的检定结果再精确也是毫无价值的。

大量观测有困难或不允许，而小量观测的结果又总是有波动，尤其用小量观测的结果推断总体总是有误差，这就有了矛盾。现代医学统计的任务就是为解决这个矛盾而着重小样本统计方法的研究。

总之，无论大量观测或小量观测，只要不是全部观测，就都是样本，仅有大有小罢了。医学统计的基本任务在于从样本去推断总体，应用概率论的方法探讨样本与总体的关系——揭示偶然现象中隐藏着的必然规律性——从而对总体做出比较正确的结论。

但是，不能错觉，医学数理统计的正确应用并非仅在实验数据取得之后，而应在研究工作开始之前，研究课题一经提出，就应从实验设计开始，如果事前没有一个以数理统计理

论为指导的实验设计和抽样方法，最后就很难进行统计推断和提出一个比较正确的结论。

二、医学统计的几个基本概念及其意义

(一) 总体、个体和样本

我们实验的目的在于探求适用于全体的客观规律，而我们实验的对象仅能为全体中很小的一部份，对这一小部份称为样本(Sample)。而样本所代表的全体称为总体(Population)。总体中的一个单位称为个体(统计单位)。如果实际观测的范围和研究结论所包括的范围完全(或基本)一致，为全面观测，这时则不存在总体和样本问题。

这些概念是容易理解的，浏览医学统计的著述，几乎都无遗漏地首先交待这一概念。这是因为容易理解的东西，往往不加重视，不注意分辨其间的关系，结果在实际应用时，又常常混淆了概念。

【例1. 1】 以某种疫苗对某地区500名儿童进行气溶胶免疫，一定时间后于其中抽取100名采血测定血清阳转率结果为90%。这里100名儿童可称为样本，因为它 是被观测的部份。而500人中虽然抽出100人，但却不是总体，因实验的目的决非仅为这500个儿童。总体则是许多，甚至全世界同条件的儿童。

【例1. 2】 测定在某种条件下发生的某种气溶胶的粒谱，抽取200个粒子，其粒径1—5 μ 的粒子占80%。该百分率为样本百分率，200个粒子称样本，而要推论的则是本次发生的或同样条件再发生的全部气溶胶粒子为其总体。

【例1. 3】 对某地区进行某种疾病的普查,发现全地区100,000人中有5,000人患有该种疾病,其患病率为5%。因实际调查和结论所要说明的范围完全一致,故不存在总体和样本的问题。前两例属于抽样观测,本例则属于全面观测。

从上述三例可知,我们研究样本的目的是要推断总体,那么样本就必须具有代表性,即样本能够成为总体的一个“缩影”——反映出总体的本质特点及生物差异等规律性。否则,不能依据样本去推断总体。

(二) 抽样误差、随机抽样和统计推断

实验证明,从同一总体中抽取几个样本,而各样本的结果却并不一样,像从同一气雾箱里做数次粒谱测定仅有接近的趋势,但都不一样,这些样本间的差异,反映了样本与总体间的差异,因为如果样本间毫无差异,则样本与总体间也就无差异了(当然也就无须进行统计研究了)。从总体中抽取样本带来的误差,即样本与总体间的误差。统计上称为“抽样误差”。

既然存在抽样误差,用样本推断总体也有推断误差。统计处理的核心就是解决“抽样误差”的问题。这种抽样误差实际是“随机抽样误差”(注意:随机不能随意),否则不能用统计方法处理,当然也谈不上推断总体了。

(三) 参数估计和统计假设检验

“参数估计”是用样本的统计数值(统计量)去估计总体的统计数值(参数)。如某气溶胶样本有效粒子占80%,那么气溶胶总体的有效粒子占百分之多少?需用80%进行估计。

统计假设检验(差异显著性测验)是根据样本之间的差异推断其总体之间是否有差异。如

甲、乙两批活菌苗的样本活菌率分别为35%和40%，相差5%，试问乙批活菌率是否一定高于甲批？要根据样本活菌率之间的差异推断其总体（甲、乙两批）之间是否有差异。

（四）精密性、准确性和置信度

参数估计的方法有两种：点估计法和区间估计法。可用下面两个例子说明：

【例1.4】 做某人年令的估计——

“他大概有33岁！”（点估计法）

“他大概有三、四十岁！”（区间估计法）

【例1.5】 用样本活菌率估计总体活菌率——

“抽检活菌率38.5%，总体活菌率亦可能为38.5%”

“抽检活菌率38.5%，总体活菌率可能在31.5—45.5之间，做这样估计有95%的把握。同样，前者属于点估计法，后者属于区间估计法。看来点估计法比较具体、精密，但究竟有多大的可靠性为38.5%呢？说不出个程度来。然而区间估计法虽不如点估计那么具体，却能说出估计的可靠程度来，因此具有实用意义。

通过上述例证，可对两种估计法的精密性与准确性加以比较如表(1.1)：

可见，就一个样本自身来说，精密性与准确性成反比关系；

准确性与区间大小成正比关系；

精密性与区间大小成反比关系。

因此，缩小估计区间，可提高精密性。

表 (1. 1)

两种估计法其精密性与准确性的比较

估 计 方 法	精 密 性	准 确 性
点 估 计 法	高	低
区 间 估 计 法	低	高
(区 间 越 大)	(越 低)	(越 高)
(区 间 越 小)	(越 高)	(越 低)

但是，不能顾此失彼。正如估计某人年龄时——

如说“就是33岁!”精密性为100%而准确性较小；

如说“在1—100岁之间!”精密性近乎0%，而准确性几乎可达100%。这种极端的估计方法都是毫无意义的。判定一个事物，总要保持一定的准确性，这就需要人为地加以规定。近代统计学多数都把估计的准确性定为不得低于95%，通用的有两种：

(1) 定估计的准确性为95%的估计区间，亦称95%的置信区间；

(2) 定估计的准确性为99%的估计区间，亦称99%的置信区间。

95%和99%亦称置信度，指每做这样估计100次平均有95次或99次是正确的，尚有5次或1次估计错的可能性。总之，参数估计是从正面指出推断的可靠性有95%或99%的置信度。而差异显著性测验是从反面指出推断错的可能性为5%或1%，实为一个意思。

(五) 置信度和置信限的选择与运用

在医学研究中，对置信度与置信限的选择与合理运用是很重要的。比如，我们估计某种