

医学统计基本方法

(自学读本)

广东医药学院 卢鑑焜 编

广东科技出版社

内 容 提 要

本书深入浅出、生动有趣地介绍了常用的医学统计基本方法。作者力求从初等数学入门，用一些简单的数学运算，结合运用统计学的基本原理，以解决一般认为难度大的医学统计方法问题，从而帮助和启发读者在医学各科的科学实验中，正确地分析运用各种医学数据资料，作出正确的判断和结论。

电子计算机的问世，更加速了医学科学发展的步伐。为了适应这一形势，书中还特意介绍了电子计算机在医学统计上的应用，便于读者在医学各科中快速地进行运算。

为了帮助读者及时检查自学的接受情况，作者除在论述中举例以示启发外，还特意在每章节后附加练习题，书末附有练习题标准答案，使读者能有机会自己做练习，边学习边提高，达到自学成材的目的。

医学统计基本方法

(自学读本)

广东医药学院 卢镛焜 编

*

广东科技出版社出版

广东省新华书店发行

广东新华印刷厂印刷

787×1092毫米32开本 6,375印张 130,000字

1982年5月第1版 1982年5月第1次印刷

印数1—13,000册

统一书号14182·53 定价0.60元

前 言

医学各分科，无论总结经验或是科学实验，都经常用到医学统计调查和统计分析。阅读中外医学文献，也常常会碰到统计分析的讨论。医学统计方法是一门比较新兴的科学，目前各级医药卫生工作者，为了提高医学各专业技术水平，为了我国医学现代化，都迫切要求学习一些常用的基本医学统计方法，希望有这一科的自学书籍问世，以适应边工作，边学习的需要。本书正是为适应这一要求而编写的。

编写本书，力求做到如下几点：

1. 深入浅出，生动有趣，切合实际，把最基本最常用的医学统计方法介绍出来。对疑难较大的地方加以通俗说明，提出一些简易的处理方法。

2. 适合未学过医学统计方法的各级医药卫生工作人员自修学习；也适合已学过一些医学统计方法，但日久不用已遗忘的人。

3. 从普通的到高深尖的医学科学研究，不一定都要用上高深尖的统计方法，多数是常用的最基本的统计方法。本书力求做到用简单、基本的统计方法，去解决一些较高深尖的医学科学研究问题。若需用到本书论述之外的统计方法，则本书可作为学习其他统计方法的基础。

4. 学习医学统计方法，不能只靠阅读，必须有实例做练习，多实践才能真正学懂。所以，本书除论述中举例以示启发外，每章节还附有练习题，书末附有标准答案。便于读者

自己做练习，自己修改，逐步熟练，达到正确运用的目的。

5. 本书除了从正面举例说明每种统计方法适用范围之外，还从反面举例说明了每种统计方法的不适用性。帮助读者避免误用不适合的统计方法，作出错误的判断。

6. 书末附有“常用对数简介”、“对数表”、“反对数表”、“平方表”和“平方根表”，便于计算应用。

7. 小型电子计算机在统计分析上有特殊用途，在某些例子中列出，以加快计算速度和提高准确性。

卢 镒 焯

一九八一年十二月

目 录

正确判断与错误判断(相对数指标)	1
一、构成比(构成指标)	2
二、率(频率指标)	3
三、相对比	5
四、相对数的正确使用	6
请你做公正人(率的标准化)	10
一、另选一个作共同标准	11
二、以各组资料之和作共同标准	12
三、任选一组作共同标准	13
正确判断防治效果、诊断检验和科研成果的数据	16
一、卡方(χ^2)检验	16
二、“行列表”的 χ^2 检验	22
三、对同一批对象做两种处理的 χ^2 检验	30
四格表直接计算机率法	34
一、分子分母对消法	37
二、阶乘对数法	38
三、比较法	41
四、简捷法	42
五、直接计算机率法的特殊情况	45
再谈常见医学数据的处理	49
一、资料例数较多的平均数计算	50
二、知道了平均数,还应该知道标准差	60
三、常见的医学数据分布形状	69
四、其它平均数	77
五、医学数据中均数差别的显著性	84

偏态分布的数据处理	104
一、优势检验法(或称符号检验法)	106
二、两个样本的符号顺序检验(或称秩和检验)	109
三、非参数统计法的优缺点	116
医学数据的收集和整理	120
一、统计资料的收集	120
二、统计资料的整理	124
医学常用的统计表和统计图	127
一、统计表	128
二、统计图	130
练习题答案	143
附录:	167
一、常用对数简介	167
二、平方表	170
三、平方根表	176
四、常用对数表	186
五、反对数表	192

正确判断与错误判断

(相对数指标)

医学工作者经常会接触到很多数据资料。医学越进步，应用的数据就越多越充分。怎样去分析利用数据资料，是每个医学工作者必须解决的重要问题，同时，也是关系到医学现代化的大问题。

就以常见的百分比与百分率来说，若不学一点统计分析方法，往往会得出错误的结论。例如：

表 1—1 某年某县恶性肿瘤死亡统计

年龄组(岁)	人 口 数	死 亡 数	死亡百分比	死亡专率/10万
0—14	356,980	11	1.52	3.08
15—29	232,505	22	3.03	9.46
30—49	205,032	142	19.59	69.26
50—69	121,882	443	61.10	363.47
70岁以上	20,047	107	14.76	533.75
合 计	936,446	725	100.00	77.42

若根据上表死亡百分比，50岁到69岁为61.10%，70岁以上为14.76%，就下判断说：50岁到69岁的人死于恶性肿瘤的情况比70岁以上的人更严重，那末，这个结论是错误的。

但若根据恶性肿瘤死亡专率来看，则可见年龄越高，死于恶性肿瘤的情况越严重。即70岁以上的恶性肿瘤死亡专率（533.75/10万）为最严重。那末，这个判断才是正确的。

这是什么道理呢？从基本的数理统计来看，我们对任何事物进行调查研究，常得出说明研究对象的原始实际数据，称为“绝对数”。如表1—1的人口数和恶性肿瘤死亡数。为了科学地说明问题，就要把有关系的绝对数建立一定的比例关系，这个比例关系称为“相对数”。如表1—1的死亡百分比和死亡专率。

常用的相对数可分为“构成比”、“率”和“相对比”三种。

一、构成比（构成指标）

说明事物内部某一组成部分的比重。常用百分比（%）来表示。计算公式为：

$$\text{构成比} = \frac{\text{事物内部某一组成部分的个体数}}{\text{事物各构成部分个体数的总和}} \times 100\%$$

（公式1—1）

若用简明示意来说：

$$\text{构成比} = \frac{\text{部分}}{\text{全体}} \times 100\%$$

例如：表1—1某年某县恶性肿瘤死亡人数为725人，其中70岁以上的恶性肿瘤死亡人数为107人。

$$\text{构成比} = \frac{107}{725} \times 100\% = 14.76\%$$

即恶性肿瘤死亡人数之中，70岁以上的人占14.76%。

构成比的特点：“分子”是部分，“分母”是全体。部分小于全体，分子必小于分母。

二、率（频率指标）

说明某种现象发生的频繁程度（频率或强度）。常以百分率（%）、千分率（‰）、万分率（1/万）、十万分率（1/10万）等表示。计算公式为：

$$\text{率} = \frac{\text{某种现象发生的例数}}{\text{可能发生该现象的主体个数}} \times 100\% \text{ (或 } 1,000\% \dots\dots) \text{ (公式 1-2)}$$

若用简明示意来说：

$$\text{率} = \frac{\text{发生例数}}{\text{主体个数}} \times 100\% \text{ (或 } 1,000\% \dots\dots)$$

例如：表 1-1 某县某年 70 岁以上的人口数为 20,047 人，因患恶性肿瘤死去的有 107 人，计算其死亡专率为：

$$\begin{aligned} \text{70 岁以上恶性肿瘤死亡专率} &= \frac{107}{20,047} \times 100,000/10 \text{ 万} \\ &= 533.75 \text{ (1/10 万)} \end{aligned}$$

医学上常用的频率指标及其计算公式如下：

$$\begin{aligned} 1. \text{发病率} &= \frac{\text{某年(期)内所发生的新病例数}}{\text{同年(期)平均人口数}} \\ &\times 1,000\% \text{ (或 } 100,000/10 \text{ 万等)} \end{aligned}$$

$$\begin{aligned} 2. \text{患病率} &= \frac{\text{受检者中患某种疾病人数}}{\text{受检人数}} \\ &\times 1,000\% \text{ (或 } 100,000/10 \text{ 万等)} \end{aligned}$$

患病率常用于慢性病。

$$3. \text{感染率} = \frac{\text{受检者中感染某种病原体人数}}{\text{受检人数}}$$

$$\times 100\% \text{ (或 } 1,000\% \text{ 等)}$$

$$4. \text{治愈率} = \frac{\text{治愈病人数}}{\text{受治病人数}} \times 100\%$$

$$5. \text{有效率} = \frac{\text{治疗有效病人数}}{\text{受治病人数}} \times 100\%$$

$$6. \text{受治病死亡率} = \frac{\text{受治病死亡人数}}{\text{受治病人数}} \times 100\%$$

$$7. \text{某病病死率} = \frac{\text{因某病死亡人数}}{\text{同期某病病人数}} \times 100\% \text{ (或 } 1,000\%)$$

$$8. \text{某病死亡率} = \frac{\text{某年内因某病死亡人数}}{\text{同年平均人口数}} \\ \times 1,000\% \text{ (或 } 100,000/10 \text{ 万)}$$

$$9. \text{因病(伤)缺勤率} = \frac{\text{某年(月)因病(伤)缺勤人日数}}{\text{同期内应出勤人日数}} \\ \times 100\%$$

率(频率指标)的特点:“分子”是“分母”所发生的现象。分子可以小于、等于或大于分母。例如一年的发病总率,以人口数为分母,以新发病例为分子。因为在一年之内,一个人可同时患几个不同的新病例,或患几次同样的病(如感冒)。所以有可能分子大于分母。以之明确区别“构成比”与“率”的不同。因为未学过统计学的人,往往会把“构成比”(构成指标)误作“率”(频率指标)使用。上例提到恶性肿瘤死亡百分比,50岁到69岁为61.10%,70岁以上为14.76%,而判断50岁到69岁的人死于恶性肿瘤的情况比70岁以上的人更严重,从而得出错误的结论,就是把“构成指标”误作“频

率指标”使用。若要判断问题的严重程度，就应该用频率指标。即用年龄组的人口数除以年龄组的死亡数，得各年龄组的“死亡专率”来判断。这样得30岁到49岁的恶性肿瘤死亡专率为69.26/10万；50到69岁为363.47/10万；70岁以上为533.75/10万。进而得出年龄越高，恶性肿瘤死亡情况越严重的正确判断。这里“构成比”和“率”为什么不一致呢？因为年龄组70岁以上的人口较少，虽然死亡专率最高，但死亡人数相对地比30岁到49岁和50岁到69岁两个年龄组少，所以死亡人数百分比反而较低。

三、相 对 比

相对比就是两个有关的、独立的、互不包括的事物的数量比例。若是两个同类事物的绝对数或相对数之比，称为“对比指标”；两个有关的不同类的事物的数量之比，称为“关系指标”；实际数与计划数之比，称为“计划完成指标”。

$$\text{相对比} = \frac{\text{甲事物的数量}}{\text{乙事物的数量}} (\text{或} \times 100\%) \quad (\text{公式1-3})$$

例：某地流行乙型肝炎病死率1962年为23.78%，1972年为12.39%，可以说某地“乙脑”病死率1972年下到1962年的 $\frac{12.39}{23.78} \times 100\% = 52.1\%$ ，或者说某地“乙脑”病死率1962年为1972年的 $\frac{23.78}{12.39} = 1.9$ 倍。

把同一事物在不同时期数量的绝对值或相对值作一系列的对比，称为“动态数列”。动态数列能说明事物数量的动态变化。

动态数列可分为定基式和环比式两种。以某一年的数值作为基数，然后将各年数值与之相比，称为“定基式动态数列”。如果将各年数值都与其前一年的数值相比，称为“环比式动态数列”。构成动态数列的各个指标，称为“动态指标”。例如：

表 1—2 某地婴儿死亡率动态分析

年 份 (1)	婴儿死亡率 (%) (2)	动 态 数 列	
		定 基 式 (以1951年为100) (3)	环 比 式 (以上一年为100) (4)
1951	81.8	100.0	—
1952	57.8	70.7	70.7
1953	52.4	64.1	90.7
1954	49.4	60.4	94.3
1955	59.7	73.0	120.9
1956	45.0	55.0	75.4
1957	35.9	43.9	79.8
1958	35.2	43.0	98.1
1959	44.3	54.2	125.9
1960	35.6	43.5	80.4

四、相对数的正确使用

1. 各种相对数指标，不能互相代替，否则会得出错误的结论。常见的错误是把构成指标（构成比）代替频率指标（率）。如表 1—1 资料，有人会误认为50岁到69岁的恶性肿瘤死亡情况比70岁以上的更严重。

2. 各构成比一般用百分比(%)，其总和必等于100。若某一部分构成比减少时，其他部分构成比必然相应增多。例如：

表 1—3 某工厂两年传染病分类百分比

病 类	1977年		1978年	
	病例数	%	病例数	%
肠道传染病	700	70	200	40
呼吸道传染病	130	13	130	26
其他传染病	170	17	170	34
合 计	1,000	100	500	100

已知该工厂人数两年中没有什么变动，由于1978年改善了食堂卫生，肠道传染病由1977年的700例减少到1978年200例；而呼吸道传染病和其他传染病的例数虽无变化，但所占比重（构成比）却明显增加。如果从“构成比”来比较，就会得出呼吸道传染病及其他传染病增加的 错误结论。因为每100个传染病例中肠道传染病例由70%降到40%，而构成比的总和要等于100%，就影响到呼吸道传染病和其他传染病的比重相应增加，并不是后两者的发病数增加。

3. 计算相对数的分母（例数）不宜过少。例数过少时，用绝对数表示较好。例如：用某种方法治疗某病6人，其中5人痊愈，1人恶化等。一般地说，相对数的基数最好小于例数，如例数超过100时才用百分数。但在医学实验中，往往由于周密设计，精选对象，严格控制各种条件。因此，实验时，即使是每组选用10只、20只小动物，也同样可以用百分数来表示反应率和死亡率。

4. 各分组的“构成比”可以相加,而且合计应等于100;但各分组的“率”是不能相加的。例如表1—1合计的死亡专率,并不等于各年龄组死亡专率相加,而是死亡总数与人口总数之比。即:

$$\frac{725}{936,446} \times 100,000/10万 = 77.42/10万$$

5. 要在相同的条件下进行“率”的比较。决定率的因素往往是多方面的,除了被研究的因素之外,其他有影响的因素都应相同或接近,才能进行“率”的比较。例如:用某种病的病死率来分析两个单位医疗技术水平的高低时,就要考虑治疗的及时程度和病情的严重程度等条件是否相同。比较不同时期的资料时,应注意客观条件是否有变化,要考虑疾病登记报告制度的改善和资料的完整性提高,可以使某病发病率“升高”。

研究方法、诊断标准不同,观察时间不相等时,所得的“率”就不能比较。例如:两地疟疾治愈率,甲地以临床症状消失为治愈,乙地以疟原虫阴性为治愈,这样,就不能作比较。又如死因统计中,甲地诊断条件较乙地差,将较多的癌死者列入“死因不明”或“老衰”,则两地的癌死率也不能比较。

6. 内部构成的两个不同“率”,要标准化后才能比较(本书以下章节论述)。

7. “率”的比较,不能只凭表面数值的大小来作结论,而应进行显著性检验,此问题将在本书以下章节详述。

练 习 题

1. <锑剂短程疗法治疗血吸虫病51例死亡病例的临床分析>一文

中，根据下表资料认为：“其中10岁到19岁组死亡率最高，其次为20岁到29岁组”。你是否同意？为什么？

死亡病例分类

性 别	年 龄						合 计
	<10岁	10—	20—	30—	40—	50—	
女	3	11	4	5	1	5	29
男	3	7	6	3	2	1	22
合 计	6	18	10	8	3	6	51

2. 有人根据下表资料说10年到14年工龄的慢性支气管炎的情况最严重，以后反而有所下降，你同意吗？如不同意，理由是什么？

某化工厂慢性支气管炎与工龄的关系

工龄(年)	检 查 人 数	患 者 数	百 分 比
1—	340	17	11.5
5—	254	30	20.4
10—	432	73	49.7
15—	136	27	18.4
合 计	1,162	147	100.0

请你做公正人

(率的标准化)

下面表2—1是甲乙两医院疾病治愈率的比较，假设两个医院病人病情轻重情况相同，治疗及时情况相同，算出“治愈率”如下：

表2—1 甲乙两医院疾病治愈率

科别	甲 医 院			乙 医 院		
	出院人数	治愈人数	治愈率%	出院人数	治愈人数	治愈率%
内 科	600	450	75	100	70	70
外 科	300	255	85	300	240	80
五官科	100	95	95	600	540	90
合计	1,000	800	80	1,000	850	85

这个资料出现的矛盾在于：从“每科治愈率”来看，认为甲医院治疗效果高于乙医院；而从“合计的治愈率”来看，则认为乙医院治疗效果高于甲医院。如果你做公正人，作出正确的判断，使双方都心服口服，你怎么办？

要解决这个矛盾，必须用统计学的“标准化法”。当两组资料内部各小组的“率”明显不同，而且各小组观察例数的“构成比”也明显不同时，则不能比较两组的“总率”。要采用统一

的“标准构成”对两组资料进行校正后，再作比较，这种方法称为“标准化法”。标准化法可用于比较两个地区年龄组人口构成不同的死亡率、发病率、患病率，或具有不同科别病床比例的两个医院的病死率和治愈率的评价等。

选择共同的内部构成比例标准的方法有三种：

一、另选一个作共同标准

另选一个通用的或便于比较的标准为两者的共同标准。如表 2—1 中，可另选一个标准的出院人数构成：内科出院人数为 400，外科出院人数为 400，五官科出院人数为 200，作为两院共同的标准。

表 2—2 标准化率算法（一）

科别 (1)	标准 出院 人数 (2)	甲 医 院		乙 医 院	
		原治愈率% (3)	预期治愈人数 (4) = (2) × (3)	原治愈率% (5)	预期治愈人数 (6) = (2) × (5)
内 科	400	75	300	70	280
外 科	400	85	340	80	320
五官科	200	95	190	90	180
合计	1,000	—	830	—	780

表 2—2 甲医院内科预期治愈人数 = $400 \times 75\% = 300$ 人，外科为 $400 \times 85\% = 340$ 人，五官科为 $200 \times 95\% = 190$ 人，合计预期治愈人数为 830 人。同样算得乙医院合计预期治愈人数为 780 人。分别计算两院预期总治愈率（即标准化总率）为：