

计算机时代《医学统计学》教学的几点思考

华中科技大学同济医学院公共卫生学院流行病与卫生统计学系 (710032)

宇传华

随着计算机性价比的不断提高, 电脑已成为科研办公教学、家庭娱乐生活等不可缺少的必备工具。与其它医学学科相比, 《医学统计学》与计算机的关系显得更为密切。如何与时俱进, 利用现代化手段引导学员准确理解《医学统计学》的基本概念, 正确领悟统计分析软件所输出的结果, 是计算机时代《医学统计学》教学所面临的机遇与挑战。下面从《医学统计学》的教学形式与教学内容两个方面, 提出几点个人在教学中的一些体会。

一、关于教学形式的思考

1. 电脑实验

许多医学课程都有实验课, 《医学统计学》需不需要实验呢? 早于 1997 年, 我国著名生物统计学家方积乾教授主编的《医学统计学与电脑实验》以书名的形式强调了“医学统计学”与“电脑实验”之间的密切联系^[1], 并在教材中利用 SAS 软件实现了许多统计学实验, 将抽象的、难以理解的统计学概念变成了形象生动的图形或容易理解的数据结果。

除了 SAS 软件以外, 许多其他现有软件也可简单地用来制作统计学电脑实验课件, 特别是微软办公软件 Excel。

因为用 Excel 绘制的“图形”或编制的“公式”具有随所引用单元格数据改变而改变的特点, 所以可通过改变有关单元格数据(如统计学参数: 总体均数、标准差等), 让学员观察“图形”或“公式”所产生结果数据的变化, 理解统计学参数的实际意义。参考文献 2 中制作了抽样的中心极限定理实验, 正态分布、二项分布、Poisson 分布、 t 分布、 F 分布、卡方分布实验, 直线回归实验等等。读者也可仿此制作更多更好的电脑实验。

2. 教学幻灯片

PowerPoint 幻灯片是现代化教学的重要辅助工具。通过幻灯片, 将主要教学内容以不同字体、不同颜色、不同线条、不同图形等形式传递给学员, 可节约大量写黑板时间, 对于提高教学的效率具有重要意义。尤其对于《医学统计学》的教学, 制作幻灯片的意义更大。《医学统计学》涉及大量实际数据、数学公式和统计图表等, 如果花费大量时间在黑板上陈列这些数据、书写这些公式、绘制这些图形显然不切实际。Internet 国际互联网上有很多免费的统计学幻灯片, 具体可参见“诊断试验评价与数据挖掘 (<http://statdtedm.6to23.com>)”的“网上免费统计资源”栏目, 该网页有整套“医学统计学”PowerPoint 幻灯片供免费下载。

3. 网络教学

互联网是近年来兴起的、集办公经商与娱乐资讯等为一体的多元化媒体, 由于信息的海量、反馈的及时、入网时间的自由等特点, 互联网已日益成为现代人日常生活不可缺少的重要组成部分。互联网络对《医学统计学》知识的普及与传播作用日渐突出, 特别是统计学论坛, 通过发送帖子, 可以相互请教、各抒己见, 达到取长补短、共同提高的目的; 学员可以发送自己在学习或使用统计学时遇到的问题, 教师可以通过这些问题发现自己教学中的不足, 并有针对性地加强实际问题的教学。目前国内有关统计学教学的网页很多, 如“医学统计之星 (<http://www.MedStatStar.com>)”等网站已在国内医学界产生了一定的影

响。欲了解其他实用的统计学相关网址者, 还可参见<http://vip.6to23.com/statdtedm/friend.htm> 网页。

二、关于教学内容的思考

1. 实验设计

统计学的主要目的可归纳为二: ①创建“最好”的研究设计产生新的数据, ②通过数据分析从噪声或变异中摘取真实信息。然而在教学中往往容易忽略第一点, 结果是: 收集的数据不满足统计学要求, 成为一堆“垃圾”数据; 有了数据不知道究竟应该采用何种统计学分析方法; 只要数据相似, 便张冠李戴, 错误选用所学方法。随着统计学软件的普及, 各种统计学指标的计算已不再困难, 教学时不应将大量的时间浪费在公式的推导或数据的运算上。

2. 数据类型

数据类型、特别是结果变量的数据类型, 是数据分析时选择统计学方法的重要依据。目前教材绪论中常将数据分类成“计量资料、计数资料、等级资料”, 而在正文章节中又将数据分类成“连续型变量与离散型变量”。那么这两种分类之间究竟有何联系? 如何理解“计数资料”这一术语?

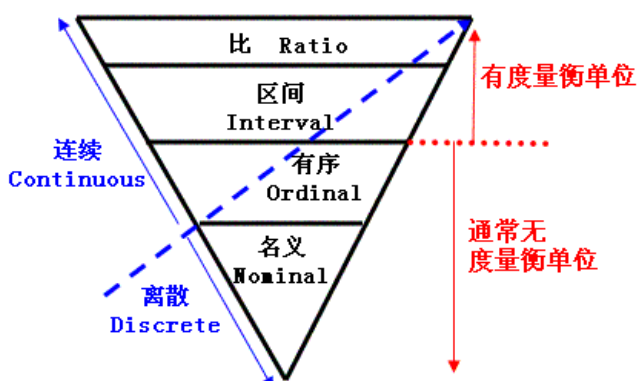


图1 数据分类

“计量、计数、等级”的数据分类方法是数据整理后的分类形式, 对于日常收集的原数据(如病案首页数据)难以分类。Stevens (1946) 提出的数据分类^[3](见图 1), 也许对我们具有借鉴意义。他按变量测量的精确程度由低到高, 将数据分类为: 名义变量(如性别)、有序变量(如疗效, 类别间差别大小难以度量)、区间变量(如摄氏体温, 类别间差别有实际意义)、比变量

(如身高, 除具有区间变量的特征外, 还具有真实意义的零点。摄氏温度的零点为水结冰时温度, 并非绝对意义的零点, 所以它不属于比变量)。根据统计分析的实际需要, 通常将区间变量与比变量合并为 1 类, 只将数据分类为: 名义、有序和区间变量三类^[3,4]。

更精确类别的数据可以转换为较粗类别的数据, 如年龄一般为区间变量, 分析时出于实际需要, 也可将它当作有序或名义变量处理; 反过来不成立, 如血型为名义变量, 就不能当作其他类型变量处理。

根据观察数据之间有无缝隙 (gap), 常将数据分类为离散型变量(有缝隙)与连续型变量(无缝隙)两大类, 名义变量一定是离散型变量; 连续型变量只能是比、区间和有序变量, 但比、区间和有序变量也可以是离散型变量(见图 1)。

3. 数据的标准格式

过去计算机很少, 主要靠手工计算, 加上文献篇幅的限制, 教材所列数据一般都是经整理加工后的数据, 很多人学完统计, 遇到教材上的数据类型尚能应付, 一旦自己收集科研数据时, 就不知道自己的数据应该如何记录, 更不用说该采用哪一种统计学方法进行数据分析。标准格式数据的第 1 行通常为变量(字段)名, 第 1 列通常为观察单位(记录)编号; 每一列代表一个变量, 除第 1 行外的每一行代表一个观察记录(见表 1)。

表1 原始数据的标准记录格式

住院号	年龄	身高	体重	住院天数	职业	文化程度	分娩方式	妊娠结局
8925655	27	165	71.5	5	无	中学	顺产	足月
8925653	22	160	74.0	5	无	小学	助产	足月
8925830	25	158	68.0	6	管理员	大学	顺产	足月
8922466	25	159	62.0	11	商业	中学	剖宫产	足月
8924535	27	157	68.0	2	无	小学	顺产	早产
8925834	20	158	66.0	4	无	中学	助产	早产
8919464	24	158	70.5	3	无	中学	助产	足月
...

其实,大多数标准统计软件(如SAS、SPSS)对标准格式数据进行统计分析时,均无需进一步整理加工便可直接进行分析;此外,这些标准格式数据也可利用标准统计软件整理成教材中呈现的数据形式。例如,为了反映表1中“文化程度”与“妊娠结局”间的关系,可利用表1的第7、9两列将数据整理成表2的形式。

表2 由表1数据整理的行×列表

文化程度	妊娠结局				合计
	过期产	死产	早产	足月	
大学	0	2	23	183	208
文盲	3	2	12	24	41
小学	9	11	57	147	224
中学	9	6	120	794	929
合计	21	21	212	1148	1402

4. 关于具体P值的问题

P值是以零假设(H_0)成立为条件,概率密度分布曲线下等于检验统计量(如t值、F值、卡方值等)以及比检验统计量更极端(或绝对值更大)的概率(或面积)。计算机软件均可以给出具体P值,与事先设定的I类错误概率 α (即检验水准,一般规定为0.05)比较,便可下差异有无统计学意义的结论。过去由于统计计算软件的缺乏,难以得到具体P值,便以计算获得的统计量(如t值、F值、卡方值等)与事先规定的 α 值对应的临界值(查阅教材末尾附表)比较,统计量 \geq 临界值,便有 $P \leq \alpha$,拒绝 H_0 ,下差异有统计学意义的结论。

目前计算机统计软件如此普及,获得具体P值不再困难。即使无原始数据,只要有统计量及有关参数,也能较容易得到具体P值。如方差分析得到F值=8,分子、分母自由度分别为2、20,便可利用Excel函数“=FDIST(8,2,20)”得到具体P值为0.0028(见参考文献2所配光盘“000统计用表与Excel计算公式”文件)。因此,推广使用具体P值是可行的,当然也是必要的。具体P值的教学可减少将P与 α 混为一团的现象。

5. Levene 检验

Levene检验的计算原理并不难,标准软件采用的方法是:对每组原数据作离均差平方 $(X_{ij} - \bar{X}_i)^2$ 变换(如SAS软件,ANOVA或GLM过程步中使用“means group/hovtest;”语句;其中 X_{ij} 为个体观察值, \bar{X}_i 为第i组均数)或离均差绝对值 $|X_{ij} - \bar{X}_i|$ 变换^[5][如SPSS软件以及上述SAS语句hovtest后加上“(type=abs)”]变换后,对变换数据执行完全随机设计方差分析,然后采用常规F统计量来检验组间方差是否齐同。Levene检验与Bartlett检

验都是方差齐性检验的方法,但各有其特点,当已知数据服从正态分布或近似正态分布时,Bartlett检验(统计量服从卡方分布)更合适;但通常情况下数据分布是未知的,由于Levene检验并不要求数据服从正态分布,对偏离正态分布的假定不敏感,所以SAS与SPSS软件作多组均数比较的方差齐性检验时,以Levene检验作为默认方法。SPSS软件还用Levene检验比较两组均数的方差是否齐同。

6. 秩相关与秩回归

Spearman秩相关与Monotonic秩回归实际上是对 X 与 Y 变量原数据分别作秩变换(从小到大按1、2、3、…排顺序,相同秩次取平均秩次),然后按简单直线相关(即Pearson相关)与直线回归方法进行的分析。这样得到的Spearman秩相关系数 r_s ,与教材中相同

秩次校正公式计算结果相同,此外还可对得到的秩相关系数 r_s 按 $t = r_s / \sqrt{(1-r_s^2)/(n-2)}$ 公式作假设检验(t 值对应的自由度为 $n-2$)。

对秩变换后的 X 与 Y 变量按一般直线回归方法获得回归方程后,再按内插法对给定 X 值进行 Y 变量的预测,这就是教材中所叙述的秩回归。如果在课堂上推导或讲授大量公式,只会让学员越学越模糊。

7. McNemar 检验

McNemar检验英文全称为McNemar's test for correlated proportions,主要用于配对资料率的检验。教材中提到四格表配对资料的假设检验可采用McNemar检验,然而采用SPSS等软件计算结果与教材有所不同。原来SPSS软件计算时采用的是一种精确检验(详细内容见<http://vip.6to23.com/statdtedm/common/mcnemar.htm>)。令 $m = b + c$ (b 、 c 为配对四格表中不同对子的频数), $k = \min(b, c)$ (\min 为取最小函数)。对于双侧检验,可计算具体 P 值为:

$$P = 2 \times \sum_{j=0}^k \binom{m}{j} \left(\frac{1}{2}\right)^m$$

其中

$$\binom{m}{j} = \frac{m!}{j!(m-j)!}, \quad m! = m \times (m-1) \times \cdots \times 1, \quad 0! = 1$$

如果在传统配对卡方检验后,补充上述计算具体 P 值的内容,也许能让学员更多理解这类数据的假设检验方法。

8. z 检验还是 u 检验

标准正态分布的均数为0(zero),方差为1(unit),所以对应于标准正态分布的检验有 z 检验或 u 检验之称。国外英文统计学书籍大多采用 z 检验,而我国大多数统计学教材大多沿用 u 检验这一名词。笔者觉得还是采用“ z 检验”更确切。原因有:①谈到 u 检验,英文文献一般认为是Mann-whitney U检验;②SAS、SPSS、Excel等软件输出的 z 值就是我们通常意义的 u 值;③ u 与总体均数 μ 的书写很相似,容易混淆;④“正(zheng)态”的汉语拼音首写字母为“ z ”,也好记忆。为了与国际接轨,并基于以上原因,建议使用“ z 检验”这一名词。

《医学统计学》面对的学员主要是(未来的)医生,如何正确引导他们在医疗实践和科学研究中寻找证据,提高医生的业务素质和服务质量,是计算机时代《医学统计学》教学的主要任务。因此,在基础医学统计学知识教学的基础上可适当增加循证医学有关的统

计学知识(如 Meta 分析、诊断试验评价等)的教学,特别是对硕士研究生。在讲课中尽量强调数据的实验设计形式和统计分析结果的解释,让学员学会选用统计学方法,并能理解分析的结果,解决实际医疗科研中所遇到的统计学问题。

为了解决以上所述的教学内容中的有关问题,未来《医学统计学》教材内容的革新显得更为重要,但愿此文能对此有抛砖引玉的作用。

参考文献

1. 方积乾, 徐勇勇, 余松林等. 医学统计学与电脑实验. 上海:上海科学技术出版社, 1997.
2. 宇传华, 颜杰 编著. Excel 与数据分析(配光盘). 北京: 电子工业出版社, 2002.
3. Kleinbaum DK, Kupper LL, Muller KE, et al. Applied Regression Analysis and Other Multivariable Methods(3rd Edition), Brooks/Cole, 1998.
4. Lattin JM, Carroll JD, Green PE. Analyzing Multivariate Data. Brooks/Cole, 2003.
5. 孙振球主编. 医学统计学(供研究生用). 北京: 人民卫生出版社, 2002, 68-69.