

实用学校卫生统计方法
汇 编
(第二辑)

《学校卫生》编辑部编印

1984年5月

编 辑 说 明

本汇编是帮助学校卫生工作者运用初步掌握的医学统计方法，解决学校卫生专业实际问题的参考资料。

本辑共收集44篇文章及有关附录，内容有：学校卫生统计资料的初步分析指标；生长发育及青春发育的资料收集、整理、分析与评价方法；儿少营养状况评价方法；疾病防治资料统计方法，学校保健组织与管理的统计研究方法以及进一步的统计研究方法等。最后附录有常用统计方法名词解释，常用医学统计方法流程图及医学生物统计方法公式用法指南。分两辑出版。

本书以普及对学校卫生专业实用的统计方法为目的，故未照搬一般卫生统计教科书内容。各篇文章虽独立，但编排由浅入深，尽量保持前后衔接与一定的系统性。每篇文章均简要介绍该方法的基本概念，并列出示范性实例演算步骤供读者参考。文字叙述力求通俗，以便读者通过自修方式学会运用。为便利广大读者，有些方法可能重复，但各作者阐述角度各异。未及汇编文章，有待今后陆续出版。

本书统计学符号主要参照上海第一医学院卫生统计学教研组《医学统计方法》和医科院卫生研究所等主编《卫生统计学》。本汇编曾邀请部分专家审阅，最后由阎承锐副教授初步校阅统一编排，原则上保持原著内容。但由于缺乏经验，水平有限，缺点错误难免，尚祈读者指正。

《学校卫生》编辑部《实用学校卫生统计方法汇编》编辑小组

1984年5月

目 录

一、疾病防治统计方法

1. 观察学生近视眼矫治效果的研究设计……………周达生 戴梅竞(1)
2. 学生青春期视力变化规律数学模型的研究……………戴梅竞 周达生(10)
3. 医学科研数据置换与处理——近视眼矫治效果的观察
与数据分析……………周达生 戴梅竞(18)
4. 龋齿、牙周疾病统计方法……………毕连华 张永治等(24)

二、进一步统计研究及其他方法

1. 逐步回归分析简介…………… 何大卫 赵融(33)
2. 课桌椅高度与读写姿势关系的数学模型……………唐锡麟 关诚(47)
3. 最大似然判别法和贝叶斯公式判别法……………黄正南 曾庆光(52)
4. 正交试验法在医学科研中的应用浅介……………周达生(59)
5. 怎样用函数型计算器简化计算直线相关与回归……………解宝鑫(67)

三、学校保健组织与管理

1. 学校卫生室基本资料的积累与统计……………葛存洁(73)
2. 学生健康状况的统计方法……………张耀云(84)
3. 评价学校卫生工作效果指标的探讨……………阎承锐(94)

四、附 录

1. 医学生物统计方法公式用法指南……………周达生 陈祥生译(102)
2. 常用统计名词解释……………李无为 黄桂兰(136)
3. 常用医学统计方法流程图……………陈祥生 周达生(附页)

观察学生近视眼矫治效果的研究设计

周达生

戴梅竞

防治学生近视眼是社会普遍关心的问题,也是国内外儿少卫生工作者及广大校医热心致力研究的课题^(1,2,3)。近几年来在《学校卫生》、《青少年视力保护》等专业杂志上介绍了大量“近防”的方法,不少学者对观察方法作了有效的探讨。如钟润先等对眼保健操的科学性研究⁽⁴⁾,胡诞宁等对环境与遗传因素在近视发病中所起的作用——前瞻性研究⁽⁵⁾等都强调了“近防”研究的科学性问题。此外,不少学者对防治近视眼的疗效评价问题提出了很好的意见。^(6,7,8)

郭秉宽氏⁽⁶⁾指出,近几年来,无论在近视眼发病的研究或防治方法的探索上,都已进行了许多工作,并提出了各种各样的观点和措施,形势是喜人的。但就其科学性和效果而言,还有待于进一步提高。国内“近防”研究的现况正是如此。尽管许多论文报道其防治近视眼的有效率达90%~100%,但是究竟能否经得起实践检验呢?这确是值得重视和研究的。其中最突出的是防治效果观察的实验设计问题。

本文针对学生近视眼矫治效果观察的研究设计问题,提出几点意见,供参考。

一、影响“防近”效果的几个因素

1. 学校近视的性质:

近视的性质是影响矫治效果的一个基本因素,是分析评价矫治效果的可比性前提。为实际工作的方便,近视性质按胡诞宁氏提出的“近视分型”(青少年近视眼防治(1):28,1980)分为:

(1)变性近视(病理性、进行性近视):需作眼底检查,常由遗传或疾病所致,与调节无关。

(2)一般近视(单纯性近视)

①假性近视:主要由于调节紧张所致。

②真性近视:主要由于眼轴延长所致。

③中间性近视。

近视眼疗效观察主要以一般近视,特别是假性近视为观察研究对象。

2. 性别、年龄、年级:

在观察疗效时,要考虑性别、年龄、年级的因素,因为不同性别、年龄、年级的基础近视状况可严重影响视力疗效的判断。根据我们研究结果⁽⁹⁾,以小学四年级至初三的学生为观察对象较为相宜。一般应视研究目的而定。

3. 近视的诊断方法:

常用近视诊断方法有远、近视力表法、排镜法、检影法等。各种诊断方法也可影响

疗效。

4. 检测条件与时机:

检测视力表应保持有足够的照度, 治疗前后应保持一致。此外, 读书期间与假期, 一年四季、一日之内上午或下午均可影响视力检测结果。

二、观察指标与疗效评价

观察指标分计数、半定量、定量三类。

计数指标: 以百分比(率)相对数表达, 如有效率、恢复正常视力百分率等。

半定量指标: 以等级表达, 如治愈(复常)、显效、有效、无效、恶化等。

定量指标: 以平均数(\bar{x} 、G、M)表达, 如平均增进(减退)行数、对数视力增进(减退)平均值等。

在疗效考核时, 应将定性与定量指标结合起来(如有效率与平均增进 ≥ 2 行数或对数视力平均值)综合评定。拟定指标时, 应根据研究目的, 选择有实际意义的、客观的、灵敏的、无害的, 且具有特异性、重现性、无偏性、易推广以及技术上可能的指标。在此须指出, 作为近视眼疗效考核, 仅以远、近视力测定是不够的, 应至少在治前及疗程结束时均用近视确诊方法(见三、视力检测与近视确诊方法)加以判定, 否则难以作出确切结论。

根据统计研究, 在大样本($n_1 = n_2$, $n > 100$ 时)的场合, 两组有效率相差 12% 即有统计学显著意义。经周密设计而进行有控制的对照比较研究是疗效研究中获得科学结论最经济的办法。

一项疗效观察研究完成后, 评价所得结果须进行统计学的显著性检验, 以判断治疗组与对照组疗效之间的差别是否由于偶然机会所造成。同时, 还要考虑第 I 类错误(假阳性, 以 α 表示)和第 II 类错误(假阴性, 以 β 表示)的概率问题, 这可用图 1、2 来描述。

由图 1、2 可知, 显著性检验固然重要, 但尚需结合专业和实际条件(临床意义与推广价值)谨慎地下结论, 因统计学检验仅提供概率的结论。(参见表 5)

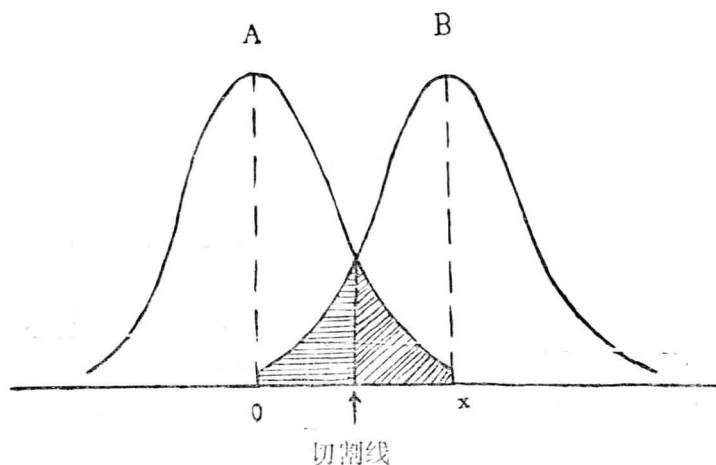


图 1 显著性检验的第 I 类和第 II 类错误示意图

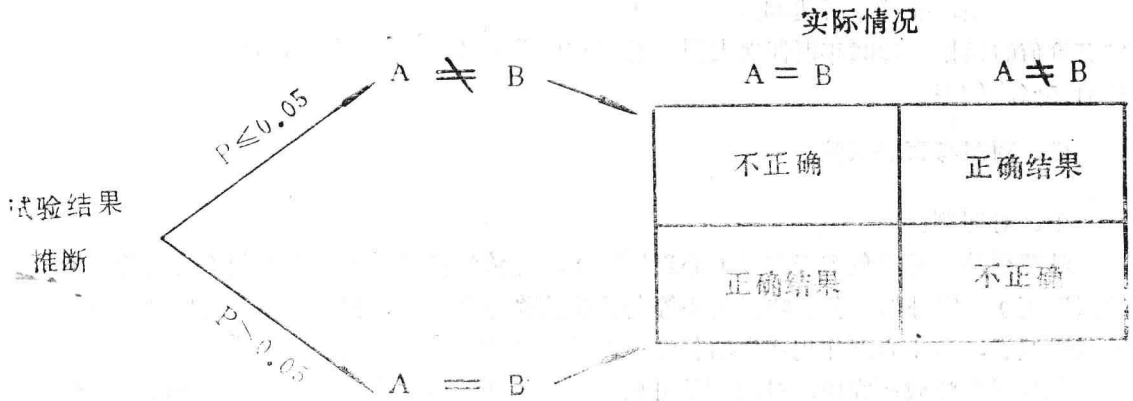


图2 显著性检验的可能结果

三、视力检测与近视确诊方法

1. 视力表:

可采用对数(E字)远、近视力表综合判断是否近视。判断标准如下:

表1 利用对数视力表初步判断近视标准

屈光性质	远视力	近视力
正 视	≥ 5.0	≥ 5.0
远 视	≥ 5.0	< 5.0
近 视	< 5.0	≥ 5.0
散光及其他	< 5.0	< 5.0

采用该法初步判定是否近视, 简便易行。据文献报道, 该法可将90%的近视者检出, 因而基层均可采用作初步筛检。

近视程度分为轻度: 4.9~4.7; 中度: 4.6~4.4; 重度 ≤ 4.3 。

由于对数视力表的五分记录值具有几何级数的性质, 其均值可直接按 $\bar{x} = \frac{\sum x}{n}$ 计算; 其差值能直接以 $x_i - x_j$ 来表示视力的增减幅度。

2. 排镜(插片)检查:

用以诊断屈光不正的类型(如近视、远视等), 可初步定量(屈光度数), 大致区分出假性近视($< 3D$)和真性近视($> 6D$), 介于两者之间为中间型(混合性)近视($3 \sim 6D$), 但准确性甚差。

3. 检影与云雾试验⁽¹¹⁾:

采用小瞳检影法、扩瞳检影法以及云雾试验法进一步区分假性、真性与中间型近视的存在。其中以云雾法较为简便, 兼有诊断与治疗的作用。各地可根据科研目的与要求选择使用。必要时作眼底检查等, 以排除其他眼疾。

4. 1980年吴奕灿教授提出“真性和假性近视眼的新鉴别法”⁽¹³⁾,不必麻痹睫状肌,简单方便,准确率高达93%,适用于在中小學生中进行大规模普查,为矫治措施提供有价值的根据。1982年浙江医大附院眼科在中學生506只眼中验证,诊断准确率达91.5%。⁽¹⁴⁾

四、对照与盲法试验

1. 对照剂:

对照剂是一种避免主观因素(心理影响),无治疗及药理作用的药剂(有人称其为“假”药);它外形、色、味、大小等与试验药物完全一样,仅在心理上安慰对照组病人或受试者,防止心理作用产生的偏性;当然切忌被受试者知道。

在近视眼疗效研究中,对照剂是可取的一种方法。如考核“夏天无”眼药水的疗效,可取一种与“夏天无”眼药水同样颜色的灭菌生理盐水或黄连素眼药水,而在瓶上以甲、乙号眼药水代之,受试者不知道他滴的是何种眼药水。这种生理盐水或黄连素液就可称为对照剂。

2. 盲法试验:

如果在一项防治效果观察中,须用对照剂且不让病人知道自己接受的是什么治疗方法,执行者也不知道何药是防治药物或对照剂。这种试验称为盲法试验。

单盲法——只是不让受试者知道所试验的是何药的观察方法。

双盲法——受试者和执行者都不让知道试的是何种药物的观察方法。

三盲法——受试者、执行者以及统计处理数据者均不清楚何为实验组,何为对照组的盲法试验。(当然实验设计者、负责人知道)。

这样做法的目的是尽可能避免下结论时有任何主观偏见,使效果评价更为客观、可靠,但要以不贻误治疗及伤害病人为原则。

近年来,还有人使用改良双盲法:即将实验分成两期:第一期为公开期,让试验有效者留下,无效者弃去;第二期为双盲试验期,将“有效”者按随机方法分成实验组与对照组作双盲试验。若疗效差异有显著性,则表示该效果系所试药物的效果。

对照剂与盲法试验是国内外公认的科学试验方法,其研究结果较一般方法准确、公正和减少偏性。

五、对照与均衡性测定

1. 对照

对照比较研究是近视眼防治效果观察的一条重要原则。只有设立对照观察才有利于效果的鉴定⁽¹²⁾。在对照中可采用治疗自身前后对照、配对比较对照、组间比较对照、双盲对照等。

(1)自身对照设计:对每个观察者进行治疗前后视力差数的比较。可采用处理前后差数的t检验。

(2)配对比较设计:将观察对象同性别、同年龄、同年级、同视力者配成对子,这样两组的受试者条件比较均衡。可采用配对比较的显著性检验一般可用1:1或1:2配对。

(3) 组间比较设计：用多种方法同时比较，也可多种疗法与对照组比较。宜用 F 检验。

(4) 综合疗法设计：若观察综合疗法(如甲 + 乙)的效果，应该设置以下几个组：①甲 + 乙法；②甲法；③乙法；④丙法(不用甲，也不用乙，而用丙作对照)，可用 F 检验。

(5) 双盲试验设计：实验组与对照组所用药物采用双盲方法进行效果观察。可用 t 检验。

2. 均衡性测定：

实验组与对照组应尽可能条件一致，除处理(治疗方法)因素外，其它各项因素应保持一致。均衡性测定可采用目测法、卡方检验、t 检验、F 检验、中位数检验及 Ridit 分析法等。

六、随机分组方法

在实际近视眼疗效观察研究中，要使实验组与对照组完全均衡往往不易办到。此时可采用随机化方法(即每个对象都有同等机会被抽取)。下面介绍几种简便易行的随机化分配方法：

1. 随机分配卡法：

可根据学生学号(末两位数字)个位数与十位数查“随机分配卡”以确定观察分在何组：T——实验，C——对照组，这样可不必去查很大篇幅的“随机数字表”即可作到随机分配。

表 2 病 例 随 机 分 配 卡

个 位 数	十 位 数									
	1	2	3	4	5	6	7	8	9	10
0	T	C	C	C	T	C	T	T	C	T
1	C	C	T	T	C	T	T	C	T	C
2	T	C	T	C	C	T	C	T	T	C
3	C	T	T	T	T	C	C	C	C	T
4	T	T	T	C	T	C	T	C	C	C
5	T	C	C	T	T	C	T	T	C	C
6	T	T	C	T	C	T	C	C	T	C
7	C	T	T	C	T	C	T	C	C	T
8	C	T	C	C	C	T	C	T	T	T
9	C	C	C	T	C	T	C	T	T	T

例：若有两个近视眼，甲生学号为 0041，乙生学号为 0047，查表 2，得甲分在 T 组，乙分在 C 组，余类推。

2. 随机数字表法：

如果设计将一批近视眼患者随机分配在三组或四组，或更多组内，则可将患者先编号(1、2、3……n)，然后从随机数字表上抄录相应数目的随机数字(如 27、38、49、

13……)，将随机数字用k去除（k为组数，如分3组，则k为3），余数为1分为第一组，余数为2分为第二组，除尽为0分为第三组，余类推。

例如设将60例近视眼患者分三组，随机分配方法见表3。

表3 多组随机分配表(60例)

患者编号	1	2	3	4 ……………	60
随机数字	27	38	49	13……………	53
除3的余数	0	2	1	1……………	2
分 组	丙	乙	甲	甲……………	乙

随机数字表请参阅各类统计专著。

3. 配对随机分配法:

将研究对象条件相同者配成对子，并标出甲、乙，再按随机化原则分配。如随机数字为单数，将对子中的A编入实验组，同时B编入对照组；如随机数字为双数，则将对子中的B编入对照组，同时A编入实验组。见表4所示。

表4 配对随机分配表(20例)

对子编号	1	2	3	4 ……………	20
随机数字	27	38	49	13……………	16
对子A	甲	乙	甲	甲……………	乙
对子B	乙	甲	乙	乙……………	甲

4. 其他设计

根据实验设计要求尚可采取随机区组设计正交设计、序贯设计计算，见有关统计方法专著。

七、样本大小——观察例数的估算

每次实验观察例数太少，则数据不可靠；例数太多，易产生混采杂，而且费时费事，因此确定合理的样本含量是重要的。所谓合理的样本含量系指可得到科学结论的最小样本含量。

样本含量大小可以通过某些公式计算，也可自统计方法书所附表中查得。举例简介如下。

1. 相比较的各组例数愈接近，误差愈小，显著性水平愈高；反之亦然。这是实验设计时须注意的。从t值公式可见，

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

t 与 $\left(\frac{1}{n_1} + \frac{1}{n_2} \right)$ 成反比，即 $\left(\frac{1}{n_1} + \frac{1}{n_2} \right)$ 愈小，t值愈大，显著水平就愈高。

当两组例数相等时，比如 $n_1 = n_2 = 25$ ， $n = 50$ ，

则 $\frac{1}{n_1} + \frac{1}{n_2} = \frac{1}{25} + \frac{1}{25} = 0.08$ 可达到最小,

当 $n_1 = n_2$ 时, 比如 $n_1 = 40, n_2 = 10, n = 50$,

则 $\frac{1}{n_1} + \frac{1}{n_2} = 0.25$ (比0.08大)。

由此可见当 $n_1 = n_2$ 时, 显著性水平最高, 故比较的各组, 当以例数相等时, 统计效率最高。

2. 两个率比较时, 样本含量的估计:

$$n = \frac{8PQ}{(P_1 - P_2)^2}$$

式中 P_1, P_2 为两个样本率, $P = \frac{P_1 + P_2}{2}$, $P + Q = 1(100\%)$

例1 预初试验比较甲、乙两种疗法治疗假性近视眼患者各10例的结果如下: 甲法有效率为80% 乙法有效率为60% 要比较两法疗效差异有无显著性时, 各组应治疗多少例?

已知: $P_1 = 0.8$ $P_2 = 0.6$ $P = 0.70$ $Q = 0.30$

代入公式得:

$$n = \frac{8 \times 0.7 \times 0.3}{(0.8 - 0.6)^2} = 42$$

每组应治疗42例。

3. 治疗前后均数比较时, 样本含量的估计:

$$n = \frac{4s^2}{d^2}$$

式中 s 为标准差, d 为允许误差。

例2 预初试验20例青少年近视眼的疗效, 假设某法治疗前后平均视力差数的均数为1.23, 标准差为0.41, 允许误差为0.1, 问正式试验需观察多少例?

已知, $s = 0.41, d = 0.1$

代入公式得:

$$n = \frac{4 \times (0.41)^2}{(0.1)^2} = 67$$

正式试验需各观察67例。

4. 两个均数比较时, 样本含量的估计:

$$n = \frac{8S_c^2}{(\bar{x}_1 - \bar{x}_2)^2}$$

式中 S_c^2 为合并方差, \bar{X}_1, \bar{X}_2 为两样本均数。

例3 预初试验甲、乙两法治疗青少年近视眼患者均用对数视力表检测视力。治疗后视力均值甲法为4.8, 乙法为4.4, $S_c^2 = 0.88$, 问正式试验需治多少例?

已知 $\bar{X}_1 = 4.8, \bar{X}_2 = 4.4, S_c^2 = 0.88$,

$$n = \frac{8 \times 0.88}{(4.8 - 4.4)^2} = 44$$

每组需观察44例。

5. 确定样本含量, 还需根据研究者要求, 确定第 I、II 类错误的概率 α 和 β 。把握度 $(1-\beta)$, 即指若两组确有差别, 则在 100 次实验中平均能发现出差别来的概率; 显著性水平 α 通常订为 0.05。归纳 α 和 β 的含义见表 5。

表 5 α 和 β 的 意 义

编号	α	β
1	误诊率	漏诊率
2	假阳性率	假阴性率
3	第 I 型误差	第 II 型误差
4	显著性水平	把握度 $(1-\beta)$
5	“以真作假”	“以假当真”

关于近视眼矫治效果观察的数据处理, 我们将另作专题讨论。

(附) 矫治学生近视眼记录表。

参 考 文 献

1. 钟润先, 等: 20年来防治青少年近视眼的探讨。青少年近视眼防治(1): 3, 1980
2. 李子良: 美国近视研究动向。青少年视力保护(3): 28, 1982
3. Ismail H, et al: Visual acuity of school entrants. Gbr Child Care. Health Dev (3): 127, 1981
4. 钟润先, 等: 眼保健操的科学性研究探讨。青少年近视眼防治(1): 61, 1981
5. 胡诞宁, 等: 环境与遗传因素在近视发病中所起的作用—前瞻性研究, 青少年视力保护(4): 1, 1982
6. 郭秉宽: 我的一点希望。青少年近视眼防治(2): 1, 1980
7. 王怡然, 等: 夏天无疗效综述。青少年视力保护(1): 44, 1982
8. 金梓良: 试论评价矫治近视眼疗效的若干问题。青少年视力保护(4): 62, 1982
9. 戴梅竞, 等: 学生青春期视力变化规律的数学模型的研究。青少年视力保护(3): 19, 1982
10. Corbett MF: The role of clinical trials in dermatology. Clin and Exp Dermatol 6(4): 411, 1981
11. 钟润先, 等: 雾视法对青少年近视的诊断和治疗作用。青少年近视眼防治(2): 19, 1980
12. 戴梅竞: 课中眼保健功对视力保护效果的观察。中华卫生杂志(6): 355, 1964
13. 吴奕灿: 真性和假性近视的新鉴别法, 中华眼科杂志16: 1, 1980
14. 赵瑾等: 真性和假性近视新鉴别法在中学生中的验证, 眼屈光专辑, 1: 117, 1983

(本文蒙阎承锐副教授审阅, 谨谢。原载青少年视力保护杂志1983年第3期)

学生视力变化规律的数学模型研究

戴梅竞 周达生

近年来, 数学模型(Mathematical Models)在研究外界因素对儿童生长发育影响以及多元分析——聚类分析方法用于青春发育分期与多指标的综合评价已有报道。的^(1,2)在青春期视力变化规律的研究中, 尚未见有数学模型运用的报道。

本资料来源于本系儿少卫生教研组等, 于1980年5—6月对南京市铁路系统大、中、小学生3260人(男1600人, 女1660人)青春发育与视力检测的数据⁽³⁾。研究的数学模型限于一元或二元曲线回归方程经验公式。目的是试图为阐明学生青春期视力变化规律的一些理论问题提供生物数学依据和可供参考的数学模型。

研究方法

一、对数视力及其数据特点

本研究统一采用对数视力表(E字表, 分远距视力表与近距视力表)测定, 每个对象均先测远视力表, 后测近视力表。视力检查按全国青少年近视眼科研课题工作会议所规定的要求测定, 一律采用人工照明视力箱, 使视标获得固定照度。

为适应大规模现场调查研究的需要, 我们采用近、远视力结合的方法作粗略判断:

视力状况	远视力	近视力	视力不良 = 远视 + 近视 + 其他视力低下(系指远视力 < 5.0者)程度:
正 视	≥ 5.0	≥ 5.0	轻 度 4.9—4.7
远 视	≥ 5.0	< 5.0	中 度 4.6—4.4
近 视	< 5.0	≥ 5.0	重 度 4.3以下
其 他	< 5.0	< 5.0	

对数视力表的五分记录值具有几何级数性质, 其均值可直接按 $\bar{X} = \frac{\sum X}{n}$ 算术平均数计算法计算; 其差值能直接以 $X_i - X_j$ 来表达视力的增减幅度; 对视力低下者均可作统计分析。

二、经验公式数学模型的建立

(一) 设计:

根据专业知识, 将不同年龄(每岁一组)从6—20岁共14个组的学生正视率、近视率、远视率、视力不良率等相互之间关系的点图的分布趋势描出来, 选择一个适合变量间关系的曲线类型; 再由曲线直线化或其他数学方法, 据实测数据求出经验公式的曲线回归方程。

在医学科研中, 往往需要对反映变量间联系的实测数据统计处理分析, 初步总结出经验性的公式, 这就是经验公式。青春期生长发育与视觉器官机能发育(简称视力发

育)的关系及其发育过程是很复杂的,在完全弄清其中各部分相互关系原理之前,我们可以提出一些与实测资料出入不大的近似公式,用以指导儿少卫生和视力保护工作实践,同时,这些数学模型亦是在理论研究中实践与理论之间必要的纽带^(4·5·6)。

(二) 曲线类型与经验公式:

1. 幂函数——抛物线曲线配合
抛物线的回归方程是 $y = a + bx + cx^2$ 可先变成二元线性方程的形式来处理,令 $x_1 = x$, $x_2 = x^2$ 。则 $Y = a + bx_1 + cx_2$ (1)

2. 指数函数——指数曲线配合

采用曲线直线化的方法(即先用数字转换,使曲线变为直线后,再按直线回归处理),对下式两边取对数:

$$Y = e^{a+bx} \quad \ln Y = a + bx$$

$$\text{令 } y = \ln Y, \text{ 则 } y = a + bx \quad (2)$$

3. 对数函数——对数曲线配合

对数曲线配合也可先用对数转换法,使对数曲线直线化后,再按直线回归进行处理。

对数曲线回归方程为:

$$\text{在 } Y = a + b \lg X \text{ 中, 令 } x = \lg X,$$

$$\text{则得 } Y = a + bx \quad (3)$$

(三) 配合适度检验——相关指数⁽⁴⁾

实际观察数据的频数分布与由曲线配合该数据的理论频数分布的符合情况,需作配合适度检验。

配合适度检验可用相关指数来检查:

$$R^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2} \quad (4)$$

式中 R^2 为相关指数, Y 为实测值, \hat{Y} 为理论值, \bar{Y} 为因变量的均值。 R^2 愈接近 1, 拟合度越好。

数学模型实例计算

根据我们调查研究资料分析,大、中、小学生视力变化不能用简单的“随年龄增长或下降”的结论来解释,青少年的正视率、近视率、远视率和视力低下程度百分比等与年龄(或学习阶段)之间变化规律反映在变量间呈各种曲线回归关系。用曲线回归分析方法(即根据实测资料找到能反映变量间关系的曲线回归方程),我们可以发现不同年龄正视率呈抛物线回归关系;不同年龄远视率呈指数曲线关系;不同年龄近视率,视力低下重度百分比呈对数曲线关系等。下面以实例介绍具体计算步骤与配合方法。为使计算程序简化,全部运用薛仲三教授倡导的计算公式进行运算⁽⁵⁾。

(一) 不同年龄正视率的抛物线配合
抛物线配合,可采用由曲线方程转变为二元线性方程的形式来处理。

表1

学生年龄与正视率的关系

编号	X: 年龄	Y: 正视率	x: X-12	x ²	x ³	x ⁴	\hat{Y}
1	6	0.0303	-6	36	-216	1296	0.06514
2	7	0.0708	-5	25	-125	625	0.22386
3	8	0.4380	-4	16	-64	256	0.35905
4	9	0.6812	-3	9	-27	81	0.47074
5	10	0.6050	-2	4	-8	16	0.55890
6	11	0.6189	-1	1	-1	1	0.62354
7	12	0.5333	0	0	0	0	0.66466
8	13	0.6989	1	1	1	1	0.68226
9	14	0.6132	2	4	8	16	0.67634
10	15	0.5550	3	9	27	81	0.64690
11	16	0.5553	4	16	64	256	0.59394
12	17	0.5669	5	25	125	625	0.51746
13	18	0.4748	6	36	216	1296	0.41746
合计	156	6.4916	0	182	0	4550	6.50098

试用适当曲线配合之。

本例 $n=13$ $\sum x=0$ $\sum x^2=182$ $\sum x^3=0$ $\sum x^4=4550$ $\sum Y=6.4916$

$\sum xY=5.3345$ $\sum x^2Y=67.3353$ 。($\sum xY$ 和 $\sum x^2Y$ 由电子计算器连续运算而得。)

$$D = \begin{vmatrix} n & 0 & \sum x^2 \\ 0 & \sum x^2 & 0 \\ \sum x^2 & 0 & \sum x^4 \end{vmatrix}$$

$$D = [(\sum nx^4 - (\sum x^2)^2)] \cdot \sum x^2 = (13 \times 4550 - 182^2) \times 182 = 4736732$$

$$a = \begin{vmatrix} \sum Y & 0 & \sum x^2 \\ \sum xY & \sum x^2 & 0 \\ \sum x^2Y & 0 & \sum x^4 \end{vmatrix} \div D$$

$$a = \{[(6.4916 \times 4550 - 182 \times 67.3353) \times 182] \div 4736732 = 0.66402$$

$$b = \begin{vmatrix} n & \sum Y & \sum x^2 \\ 0 & \sum xY & 0 \\ \sum x^2 & \sum x^2Y & \sum x^4 \end{vmatrix} \div D$$

$$b = \{[(n \cdot \sum x^4) - (\sum x^2)^2] \sum xY \div D = [(13 \times 4550 - (182)^2) \times 5.3345] \div 4736732 = 0.02931$$

$$C = \frac{\begin{vmatrix} n & 0 & \sum Y \\ 0 & \sum x^2 & \sum xY \\ \sum x^2 & 0 & \sum x^2 Y \end{vmatrix}}{D}$$

$$C = \{[13 \times 67.3353 - 182 \times 6.4916] \times 182\} \div 4736732 = -0.01176$$

(1) 计算

数据来源 (1)	D (2)	a (3)	b (4)	c (5)
本例	4736732	0.66402	0.02931	-0.01176

(2) 结论

$$\hat{Y} = a + bX + CX^2 = 0.66402 + 0.02931(X - 12) + (-0.01176)(X^2 - 24X + 144) = -1.3811 + 0.3116X - 0.01176X^2$$

(3) 作图

各点 (X, Y) 的散布图和配合的曲线见图 1

结果表明, 不同年龄正视率并非随年龄增长而增长(或降低)的直线关系, 而是呈一种曲线回归关系, 6—14岁正视率随年龄增长而升高, 15岁以后正视率有所下降, 10岁以前正视率较低, 主要由于生理性远视而形成的暂时现象, 6—12岁正视率迅速提高, 随着眼球发育, 由生理性远视自然转为正视。因此, 在作近视眼疗效观察时, 应考虑到14岁以前学生正视率自然增高的趋势。还须指出, 本例抛物线只代表6—18岁年龄阶段的正视率变化规律, 不宜任意外推。

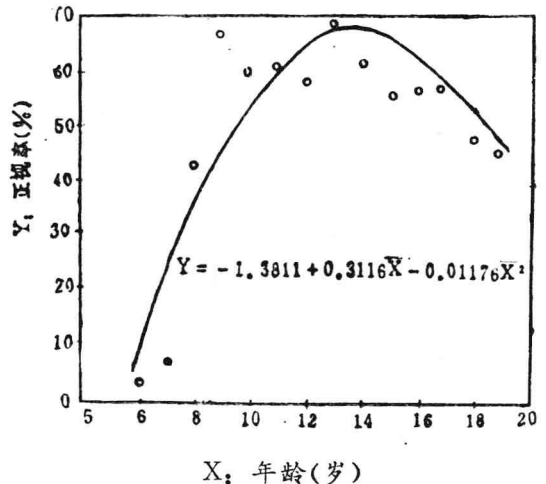


图 1 不同年龄正视率的实际值和配合的曲线

(二) 不同年龄远视率的指数曲线配合

配合指数曲线时, 可先用对数转换法, 使指数曲线直线化, 再按直线回归方法进行处理。

试用适当曲线配合之。

本例: $n = 13$ $\sum X = 156$

$$\sum X^2 = 2054 \quad \sum Y = 11.0679$$

$$\sum Y^2 = 13.9127 \quad \sum XY = 107.9102$$

$$\bar{X} = 12 \quad \bar{Y} = 0.8514$$

(1) 计算:

数据来源 (1)	D (2)	D' (3)	b (4)	b' (5)	r (6)
本 例	2366	58.3667	-0.1368	-5.5470	0.8711

表2 学生年龄与远视率的关系

编 号	X: 年龄	Y: 远视率%	y: lgY	\hat{y}	\hat{Y}
1	6	63.64	1.8037	1.6722	47.0110
2	7	61.06	1.7858	1.5354	34.3084
3	8	38.84	1.5893	1.3986	25.0380
4	9	13.75	1.1383	1.2618	18.2726
5	10	14.50	1.1614	1.1250	13.3352
6	11	8.07	0.9069	0.9882	9.7320
7	12	4.41	0.6444	0.8514	7.1023
8	13	2.27	0.3560	0.6912	4.9113
9	14	2.09	0.3201	0.5526	3.5694
10	15	1.02	0.0086	0.4140	2.5942
11	16	2.51	0.3997	0.2754	1.8854
12	17	3.12	0.4942	0.1368	1.3703
13	18	2.98	0.4594	0.0018	0.9959
合 计	156	218.10	11.0679

$$D = \left| \begin{array}{cc} n & \sum X \\ \sum X & \sum X^2 \end{array} \right| = n \sum X^2 - (\sum X)^2$$

$$= 13 \times 2054 - (156)^2 = 2366$$

$$D' = \left| \begin{array}{cc} n & \sum Y \\ \sum Y & \sum Y^2 \end{array} \right| = n \sum Y^2 - (\sum Y)^2$$

$$= 13 \times 13.9127 - (11.0676)^2$$

$$= 58.3667$$

$$b = \left| \begin{array}{cc} n & \sum X \\ \sum Y & \sum XY \end{array} \right| \div D = [13$$

$$\times 107.9102 - 156 \times 11.0679]$$

$$\div 2366 = -0.1368$$

$$b' = \left| \begin{array}{cc} n & \sum X \\ \sum Y & \sum XY \end{array} \right| \div D' = [13$$

$$\times 107.9102 - 156 \times 11.0679]$$

$$\div 58.3667 = -5.5470$$

$$a = \bar{Y} - b\bar{X} = 0.8514 - (-0.1368)$$

$$\times 12 = 2.493$$

$$r = \sqrt{b \cdot b'}$$

$$= \sqrt{(-0.1368) \times (-5.5470)}$$

$$= 0.8711$$

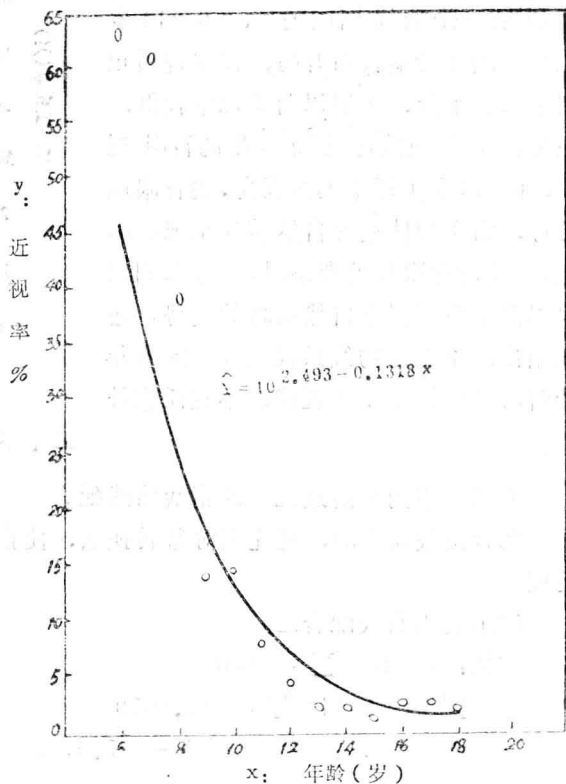


图2 男女生不同年龄远视率的实际值和配合的曲线