



Xiandai Shiyong Weisheng
Tongjixue

现代实用卫生 统计学

高 歌 郭秀花 黄水平 主编

目 录

第一章 绪论

第一节 卫生统计学的作用和地位	1
第二节 卫生统计学的定义和内容	2
第三节 统计学的几个基本概念	3
第四节 学习卫生统计学应注意的问题	5

第二章 统计表和统计图

第一节 统计表	7
第二节 统计图	8

第三章 定量资料的统计描述

第一节 频数分布表和频数分布图	15
第二节 集中趋势的描述	18
第三节 离散趋势的描述	22
第四节 正态分布及其应用	25

第四章 分类变量资料的统计描述

第一节 常用相对数	33
第二节 应用相对数的注意事项	34
第三节 标准化法	35
第四节 动态数列及其分析指标	39

第五章 参数估计基础

第一节 均数的抽样误差与标准误	44
第二节 t 分布	46
第三节 总体均数的估计	47
第四节 二项分布和 Poisson 分布	49
第五节 总体率的估计	53

第六章 假设检验的基本概念

第一节 假设检验的基本思想及步骤	57
第二节 I 型错误与 II 型错误	59
第三节 单侧检验与双侧检验	60
第四节 假设检验需注意的问题	62
第五节 假设检验与区间估计的联系	64

第七章 t 检验——两组均数的比较

第一节 t 检验的基本概念	67
第二节 样本均数与总体均数的比较	67
第三节 配对设计均数的比较	68

一、点估计

用样本统计量直接作为总体参数的估计值就是点估计(point estimation),如直接用随机样本的样本均数 $\bar{x}$ 作为总体均数 μ 的点估计值。点估计方法简单,但未考虑抽样误差的影响,估计的正确程度很难评价,即认为 48.65kg 为该地 16 岁男孩体质量的总体均数。

二、区间估计

区间估计(interval estimation)是按事先给定的 $(1-\alpha)$,估计包含未知总体参数的一个区间范围,该范围称为参数的置信区间(confidence interval, CI)。 $(1-\alpha)$ 称为置信度(confidence level),也可表示为 $100(1-\alpha)\%$,常取 95%(也可取 90%、99%)。置信区间通常由两个数值即两个置信限(confidence limit, CL)表示,较小者称为置信下限(lower limit, L),较大者称为置信上限(upper limit, U)。

总体均数的 95% 置信区间的实际涵义是:如果从同一总体中重复抽取 100 份样本含量相同的独立样本,每份样本分别计算 1 个置信区间,在 100 个置信区间中,将大约有 95 个置信区间包含了总体均数,大约有 5 个置信区间不包含总体均数。对于某一次估计的置信区间,置信度为 95%,那么总体均数有 95% 的可能包含在我们求得的区间内。

如果总体标准差 σ 已知或 σ 未知但样本含量足够大,可按标准正态分布估计总体均数 μ 的置信区间;如果总体标准差 σ 未知,采用样本标准差 s 取代总体标准差 σ ,则可按 t 分布估计总体均数 μ 的置信区间。

1. σ 已知

σ 已知,在一般正态分布曲线下, $(1-\alpha)$ 置信区间的范围可表示为图 5-3 的中间部分。

总体均数 μ 的 $(1-\alpha)$ 置信区间计算公式为:

$$(\bar{x} - z_{\alpha/2} \sigma_{\bar{x}}, \bar{x} + z_{\alpha/2} \sigma_{\bar{x}}) \quad (5-5)$$

或简写为:

$$\bar{x} \pm z_{\alpha/2} \sigma_{\bar{x}}$$

令 $\alpha = 0.05$, $z_{0.05/2} = 1.96$, 总体均数 μ 的 95% 置信区间为:

$$(\bar{x} - 1.96 \sigma_{\bar{x}}, \bar{x} + 1.96 \sigma_{\bar{x}})$$

σ 未知但 n 足够大(如 $n > 50$) 时,公式(5-5)中的 $\sigma_{\bar{x}}$ 可近似用 $s_{\bar{x}}$ 代替。

2. σ 未知

当 σ 未知,采用样本标准差 s 替代 σ 时,由于样本均数的分布不再服从 z 分布,而是服从 t 分布,总体均数 μ 的 $1-\alpha$ 置信区间的计算公式应为:

$$(\bar{x} - t_{\alpha/2, \nu} s_{\bar{x}}, \bar{x} + t_{\alpha/2, \nu} s_{\bar{x}}) \quad (5-6)$$

或简写为:

$$\bar{x} \pm t_{\alpha/2, \nu} s_{\bar{x}}$$

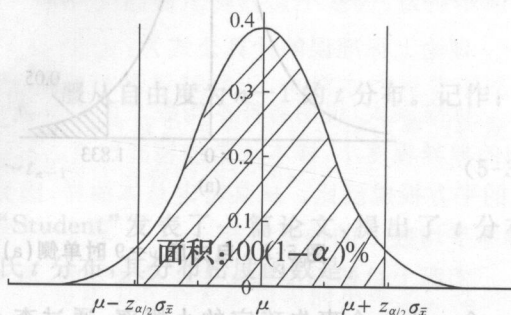


图 5-3 总体均数 μ 的双侧 $(1-\alpha)$ 置信区间

表 8-5 例 8-3 的 SNK 检验计算表

对比组 A 与 B	$ \bar{x}_A - \bar{x}_B $	q 值	组数 a	q 界值		P 值
				0.05	0.01	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 与 2	0.425	2.617	2	2.95	4.02	>0.05
1 与 3	0.875	5.388	3	3.58	4.64	<0.01
2 与 3	0.450	2.771	2	2.95	4.02	>0.05

3. 确定 P 值, 作出统计推断

q 界值不但考虑自由度, 而且考虑组数 a , 即任意两对比组包含的组数, 见表 8-5 中第(4)栏。以组数 a 和 $\nu_{\text{误差}} = 21$ (本例取 20) 查附表 7 即 q 界值表, 当 $a=2$ 时, $q_{0.05(20,2)} = 2.95$, $q_{0.01(20,2)} = 4.02$; 当 $a=3$ 时, $q_{0.05(20,3)} = 3.58$, $q_{0.01(20,3)} = 4.64$, 见表 8-5 第(5)栏和第(6)栏。将第(3)栏计算的 q 值与相应的 q 界值进行比较得各组的 P 值, 见表 8-5 第(7)栏。由上表可以看出, 按 $\alpha=0.05$ 水准, 甲组和丙组的样本均数间差异有统计学意义, 而甲组和乙组、乙组和丙组的样本均数间差异均无统计学意义。

二、LSD- t 检验

LSD- t 检验即最小显著性差异 (least significant difference) t 检验, 适用于某一对或某几对在专业上有特殊意义的均数间的比较, 如多个处理组与对照组的比较, 或某几个处理组间的比较, 一般在设计阶段确定哪些均数需进行多重比较。检验统计量 t 值的计算公式为:

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{S_{\bar{x}_A - \bar{x}_B}} = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{MS_{\text{误差}} \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}}, \nu = \nu_{\text{误差}} \quad (8-12)$$

同样对例 8-3, 若事先计划比较煤炭仓库与地面办公楼、矿井下与地面办公楼染尘大鼠肺湿质量, 介绍 LSD- t 检验的步骤如下:

1. 建立检验假设, 确定检验水准

$H_0: \mu_A = \mu_B$, 即所研究的两个对比组的总体均数相等

$H_1: \mu_A \neq \mu_B$, 即所研究的两个对比组的总体均数不等

$\alpha = 0.05$

2. 计算检验统计量

已知: $n_A = n_B = 8$, $MS_{\text{误差}} = 0.211$

(1) 计算差值的标准误: 本例各组例数相等, 故任意两组均数差值的标准误相等, 即

$$S_{\bar{x}_A - \bar{x}_B} = \sqrt{0.211 \left(\frac{1}{8} + \frac{1}{8} \right)} = 0.2297$$

(2) 计算统计量 LSD- t 值: 如表 8-6 所示。

表 8-10 血液放置不同时间的血糖浓度 (mmol/L)

受试者	放置时间 (min)			
	0	45	90	135
1	5.27	5.27	4.94	4.61
2	5.31	5.26	4.93	4.70
3	5.88	5.83	5.38	5.01
4	5.49	5.43	5.32	5.04
5	5.66	5.44	5.38	4.88
6	6.22	5.72	5.61	5.22

2. 两因素重复测量设计资料

例 8-6 为比较两种同类药物 A、B 在体内的代谢速度, 将 10 名受试对象随机分为两组, 一组给予 A 药, 另一组给予 B 药。分别在服药后 1、2、4、6 h 测定血中的药物浓度, 结果见表 8-11。此研究对同一受试者在服药后不同时间点上重复测量了 4 次血中的药物浓度, 是为了分析药物及时间两个因素对药物浓度的影响, 故称之为两因素重复测量设计资料。

表 8-11 两种同类药物在血中的浓度 ($\mu\text{g/mL}$)

药物	受试对象	服药后时间 (h)			
		1	2	4	6
A	1	9.12	53.16	54.89	45.88
A	2	5.10	48.78	77.39	61.29
A	3	3.37	20.65	71.22	75.08
A	4	6.51	30.07	74.26	75.69
A	5	1.99	5.84	80.26	69.82
B	6	13.55	27.10	46.88	53.21
B	7	1.02	25.12	54.03	44.15
B	8	2.15	14.23	68.74	64.65
B	9	6.18	25.83	45.70	54.25
B	10	1.68	40.29	32.34	71.00

(二) 重复测量设计与随机区组设计的区别

(1) 重复测量设计中, 各观察对象各时间点是固定的, 同一观察对象依次接受各时间点的测量。两因素重复测量设计中, 处理 (如例 8-6 中的药物因素) 是在观察对象间随机分配, 不能在各时间点随机分配。在随机区组设计中, 各区组内有多个条件一致的观察对象, 每个观察对象随机分配给一种处理。

(2) 随机区组设计中, 各区组内各处理水平相互独立, 各区组内各观察对象所得数据相互独立。重复测量设计中各行内的数据重复测自同一观察对象, 它们之间往往有较高的相关性。

本资料分析如下:

H_0 : 两种血型的分布无关联

H_1 : 两种血型的分布有关联

$\alpha=0.05$

$$\chi^2 = 6.094 \left(\frac{431^2}{1823 \times 1451} + \frac{490^2}{1823 \times 1666} + \dots + \frac{325^2}{641 \times 2977} - 1 \right) = 8.5952$$

$$\nu = (4-1)(3-1) = 6$$

$$r = \sqrt{\frac{8.5952}{6.094 + 8.5952}} = 0.0375$$

从上述结果中可以看出,列联系数比较小,且 $\chi^2 < \chi_{0.05,6}^2 = 12.59$, $P > 0.05$, 所以两种血型之间无关,即某人按 ABO 血型系统测定为 A、B、O、AB 四种血型之一时,再按 MN 血型系统来划分,他(她)属于 M、N、MN 血型的可能性是相等的,没有一定的倾向性。

2. 双向有序且属性不同的列联表

例 9-9 某医生观察依沙酰胺治疗皮肤真菌感染的临床试验,结果见表 9-12。试分析该病的疗效是否与病程有关。

表 9-12 依沙酰胺治疗皮肤真菌感染疗效

病程	痊愈	好转	无效	合计
<1 月	79	24	8	111
1 月~	30	13	1	44
3 月~5 年	102	83	30	215
>5 年	29	26	10	65
合计	240	146	49	435

对于此类资料可以采用 Spearman 秩相关、CMH 中非零相关等方法进行检验行变量和列变量是否存在相关。若采用 Spearman 秩相关分析(统计方法详见第十一章),等级相关系数 $r_s = 0.21479$, $P < 0.0001$, 相关有统计学意义。所以依沙酰胺治疗皮肤真菌感染疗效与该病的病程有关,即病程越长,疗效越差。

3. 双向有序且属性相同的列联表

例 9-10 某医生回顾收集了 2000~2005 年间该院手术治疗的 206 例子宫内膜癌患者的临床病理资料,其临床分期和手术病理分期的情况见表 9-13。试问:两种方法的分期结果是否一致?

表 9-13 临床分期和手术病理分期的情况

手术前临床分期	手术后病理分期			合计
	I	II	III	
I	110	20	0	130
II	8	41	9	58
III	0	7	11	18
合计	118	68	20	206

为 50%，甲、乙两法一致的检出率为 35%。试问：两种方法何者为优？两种方法的检出结果是否相关？

3. 在一项关于吸烟与肺癌关系的研究中，用 1：1 配对的病例—对照设计，得到下列资料：

表 9-16 病例与对照组的吸烟情况

病例	对照		合计
	吸烟	不吸烟	
吸烟	120	25	145
不吸烟	7	20	27
合计	127	45	172

试分析吸烟是否与肺癌有关。

4. 某研究者想了解某地区城区居民就业状态是否会影响慢性病就诊治疗情况，随机抽查了 867 人，结果见表 9-17：

表 9-17 企事业单位职工家庭不同就业状态慢性病患者就诊情况

就业状况	患病例数	就诊人数	就诊率(%)
在岗	544	377	69.3
下岗/失业	31	22	71.0
离退休	251	209	83.3
学生	11	11	100.0
其他	30	24	80.0
合计	867	643	74.2

试问：不同就业状况的慢性病患者就诊率是否相同？

5. 某卫生防疫站对屠宰场及肉食零售点的猪肉检查其表层沙门菌带菌情况，结果见表 9-18：

表 9-18 屠宰场及肉食零售点的猪肉沙门菌带菌情况

采样地点	检查例数	阳性例数	带菌率(%)
屠宰场	26	2	7.69
零售点	13	5	38.46
合计	39	7	17.95

试问：两者带菌率有无差别？

6. 某研究者搜集了 164 名工人的工龄和铅中毒情况，结果见表 9-19：

第一节 配对设计 Wilcoxon 符号秩和检验

当配对设计数值变量资料差值 d 不服从正态分布时,假设检验应选择非参数检验,称为符号秩和检验,也称为 Wilcoxon 符号秩和检验(Wilcoxon signed-rank test)。用于配对设计的样本差值的中位数和 0 的比较,还可以用于单样本中位数和已知总体中位数的比较。

一、配对设计的样本差值的中位数和 0 的比较

(一) 方法步骤

例 10-1 将 24 只家兔按体重质量相近和性别相同配成 12 对,按随机化原则将每对中的两只家兔分到甲、乙两组,用某种放射线的 0.5Gy 和 1.0Gy 两种剂量分别对甲、乙两组家兔进行局部照射,观察放射性急性皮肤损伤,损伤程度用评分指标反映,结果如表 10-1 中的第(2)、(3)栏。请问:该放射线的这两种剂量对家兔局部照射的急性皮肤损伤程度有无差别?

表 10-1 24 只家兔受两种剂量放射线局部照射的皮肤损伤程度的评分

编号 (1)	0.5Gy (2)	1.0Gy (3)	差值 (4)=(2)-(3)	秩次 (5)
1	39	55	-16	-10
2	42	47	-5	-9
3	51	53	-2	-3
4	43	41	2	3
5	55	54	1	1
6	45	63	-18	-11
7	22	42	-20	-12
8	48	46	2	3
9	40	43	-3	-6
10	45	49	-4	-8
11	40	37	3	6
12	49	52	-3	-6

$$T_+ = 13, T_- = 65$$

1. 建立假设

H_0 : 差值总体中位数 $M_d = 0$

H_1 : 差值总体中位数 $M_d \neq 0$

$\alpha = 0.05$ (双侧)

2. 求差值

见表 10-1 第(4)栏。

$$\chi^2 = \frac{|R_i - R_j|}{c[N(N+1)/12][1/n_i + 1/n_j]} \tag{10-20}$$

相同秩次校正数 $c = 1 - \sum (t_j^3 - t_j) / (N^3 - N)$, t_j 为第 j 个相同秩次的个数; N 为各处理组的总例数。

3. q 检验法

用多样本均数比较的方差分析后的两两比较的 q 检验的方法对任两样本的秩进行分析。

$$q = \frac{R_i - R_j}{s_{R_i - R_j}} \tag{10-21}$$

式中, $s_{R_i - R_j}$ 为秩和差异标准误, 计算公式为:

$$s_{R_i - R_j} = \sqrt{\frac{n(na)(na+1)}{12}} \tag{10-22}$$

n 为样本含量, 即处理的重复数; a 为比较的两秩和差数范围内所包含的处理数。这里的 q 检验法只适用于重复数相等的试验资料。

计算 q 值后, 以 $\nu = \infty$ 和 a 查附表 7, 得临界值 $q_{\alpha(\infty, a)}$, 作出统计推断。

例 10-5 中进行两两比较的结果详见表 10-9。

表 10-9 例 10-5 三组样本两两比较的秩和检验

对比组	$ R_i - R_j $	$s_{R_i - R_j}$	q_{ij}	a	q 界值		P
					0.05	0.01	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A 和 B	65	16.416	3.959	3	3.31	4.12	<0.05
A 和 C	61	11.068	5.511	2	2.77	3.64	<0.01
B 和 C	4	11.068	0.361	2	2.77	3.64	>0.05

H_0 : 任两组来自的总体分布位置相同

H_1 : 任两组来自的总体分布位置不同

$\alpha = 0.05$

任两组的 $|R_i - R_j|$ 、 $s_{R_i - R_j}$ 和 q_{ij} 的计算结果见表 10-9。 a 为比较的任两组包含的组数。

按 $\alpha = 0.05$ 水准, A 组和 B 组、A 组和 C 组的差别有统计学意义, 而 B 组和 C 组的差别无统计学意义。

二、随机区组设计多样本两两比较的秩和检验

当经过多个相关样本比较的 Fridman M 检验拒绝 H_0 , 接受 H_1 , 认为多个总体分布位置不全相同(或全不相同)时, 若要进一步推断是哪两两总体分布位置不同, 可用 q 检验。

下面对例 10-7 资料(三门医学基础课程的教学效果)进行两两比较。

H_0 : 任两门课程的教学效果分布的位置相同

H_1 : 任两门课程的教学效果分布的位置不同

$\alpha=0.05$

设 g 个相关样本, 当区组个数 n 较多时, 按下式求第 i 个样本和第 j 个样本比较的 q 值。

$$q = \frac{R_i - R_j}{\sqrt{n \cdot MS_{\text{误差}}}} \quad (10-23)$$

$$\text{其中, } MS_{\text{误差}} = \frac{\frac{ng(g+1)(2g+1)}{6} - \frac{1}{n} \sum R_i^2 - \frac{1}{12} \sum (t_j^3 - t_j)}{(n-1)(g-1)} \quad (10-24)$$

q 的自由度 $\nu = (n-1)(g-1)$, 样本间跨度 a 指把 g 个样本秩和从小到大排列后 R_i 和 R_j 之间涵盖的秩和个数(包括 R_i 和 R_j 自身在内)

本例根据表 10-7 有: $n=10, g=3, \sum R_i^2 = 15^2 + 18^2 + 27^2 = 1278, \sum (t_j^3 - t_j) = [(2^3 - 2) + (2^3 - 2) + (2^3 - 2) + \dots + (2^3 - 2)] = 36$

$$MS_{\text{误差}} = \frac{\frac{10 \times 3 \times (3+1) \times (2 \times 3 + 1)}{6} - \frac{1}{10} \times 1278 - \frac{36}{12}}{(10-1) \times (3-1)} = 0.5111$$

$$\frac{1}{\sqrt{n \cdot MS_{\text{误差}}}} = \frac{1}{\sqrt{10 \times 0.5111}} = 0.4423$$

$$\nu = (n-1)(g-1) = 18$$

三门医学基础课程的教学效果两两比较结果详见表 10-10。

由于三门医学基础课程的教学评分的秩和从小到大排列顺序为: 解剖学 < 生理学 < 组织胚胎学, 按 $\alpha=0.05$ 水准, 解剖学和生理学的教学效果无统计学差别, 解剖学和组织胚胎学的教学效果有高度统计学差异, 生理学和组织胚胎学教学效果有统计学差异。

表 10-10 例 10-7 三门医学基础课程的教学效果两两比较的秩和检验

对比组	$ R_i - R_j $	q_{ij}	a	q 界值		P
				0.05	0.01	
(1)	(2)	(4)	(5)	(6)	(7)	(8)
解剖学和生理学	3	1.327	2	2.97	4.07	>0.05
解剖学和组织胚胎学	12	5.308	3	3.61	4.70	<0.01
生理学和组织胚胎学	9	3.981	2	2.97	4.07	<0.05

练习题

一、最佳选择题

- 配对设计计量资料差值的 Wilcoxon 符号秩和检验, 确定 P 值的方法为()。
 - T 越大, P 越大
 - T 越大, P 越小
 - T 值在界值范围内, P 小于相应的 α
 - T 值在界值范围内, P 大于相应的 α
- 以下检验方法, 不属于非参数检验的方法是()。
 - t 检验
 - T 检验
 - H 检验
 - M 检验
- 完全随机设计两独立小样本计量资料比较的假设检验, 首先应考虑()。
 - 用 t 检验

$$l_{yy} = \sum (y - \bar{y})^2 = \sum y^2 - \frac{(\sum y)^2}{n} = 3.294\ 834 - \frac{5.726\ 6^2}{10} = 0.015\ 4$$

$$l_{xy} = \sum (x - \bar{x})(y - \bar{y}) = \sum xy - \frac{(\sum x)(\sum y)}{n} = 77.559\ 5 - \frac{(134.4)(5.726\ 6)}{10} = 0.594\ 0$$

3. 按式(11-1)计算相关系数

$$r = \frac{l_{xy}}{\sqrt{l_{xx}} \sqrt{l_{yy}}} = \frac{0.594\ 0}{\sqrt{24.904} \sqrt{0.015\ 4}} = 0.959\ 2$$

4. 相关系数的假设检验

上述相关系数 r 是样本相关系数,它是总体相关系数 ρ 的估计值。要判断 x 、 y 间是否具有相关关系,就要检验 r 是否来自总体相关系数 ρ 为零的总体(零相关)。因为即使从 $\rho=0$ 的总体作随机抽样,由于抽样误差的影响,所得 r 值也不一定等于零;反之,一个不为零的样本相关系数是有可能来自零相关的总体的。因此,若要判断该样本 r 是否有统计学意义,必须作 $\rho=0$ 的假设检验。

相关系数的假设检验,可用 t 检验,公式如下:

$$t = \frac{r-0}{s_r} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \quad (11-5)$$

式中 s_r 为相关系数的标准误。求得 t 值后,查 $\nu=n-2$ 的 t 界值表(见附表 2)得 P 值,按所取检验水准推断结论。

$H_0: \rho=0$, 两变量间无直线相关关系

$H_1: \rho \neq 0$, 两变量间有直线相关关系

$\alpha=0.05$ (双侧)

本例 $n=10$, $r=0.959\ 2$, 按式(11-5)计算 t 值:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.959\ 2}{\sqrt{\frac{1-0.959\ 2^2}{10-2}}} = 9.595\ 9$$

$\nu=n-2=10-2=8$, 查附表 2, 得 $P<0.01$ 。按 $\alpha=0.05$ 水准, 拒绝 H_0 , 接受 H_1 , 可认为 3 岁儿童体表面积与体质量间存在直线相关关系。

相关系数的假设检验亦可按 $\nu=n-2$ 直接查 r 界值表(见附表 15), 得出相同的结论。

四、相关分析中应注意的问题

(1) 进行相关分析的资料应有实际意义。医学实践中存在将两个毫无关联的变量来进行相关分析, 这是滥用统计方法的表现, 应该坚决杜绝此类现象的发生。

(2) 相关系数的计算适用双变量正态分布资料, 如果资料不符合该条件, 可首先通过变量转换使之正态化, 再根据变换值计算相关系数。在实际工作中, 完全符合上述要求的资料并非经常遇到, 但只要与正态分布相差不远, 在实际应用中仍有相当好的效果。如果变量(一个或两个)明显呈偏态分布, 就须经过适当的变量变换, 使资料满足相关分析的条件或者直接采用等级相关法。

表 11-3 15 名健康人凝血酶浓度(x)与凝血时间(y)的测量值

编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x	1.1	1.2	1.0	0.9	1.2	1.1	0.9	0.6	1.0	0.9	1.1	0.9	1.1	1.0	0.7
y	14	13	15	15	13	14	16	17	14	16	15	16	14	15	17

(1) 试进行相关与回归分析。

(2) 当凝血酶浓度 x 取 1.0 时,估计 μ_y 的 95% 可信区间。

2. 就表 11-4 资料分析血小板数与出血症状的关系。

表 11-4 12 例患者的血小板浓度($10^9/L$)和出血症状的关系

病例号	1	2	3	4	5	6	7	8	9	10	11	12
血小板数	120	130	160	310	420	540	740	1 060	1 260	1 230	1 440	2 000
出血症状	++	+++	±	-	+	+	-	-	-	-	++	-

3. 为研究某药物浓度与肾上腺素释放量的关系,选取 10 个药物浓度水平,每种药物剂量水平上重复 5 次试验,观测结果见表 11-5:

表 11-5 药物剂量对肾上腺素释放量影响的试验结果

药物剂量(mg)	肾上腺素释放量(pg/mL)				
	1	2	3	4	5
15	19.26	14.29	17.60	18.36	16.53
20	21.20	21.78	20.77	20.65	23.38
25	21.77	22.61	22.70	21.17	21.65
30	23.47	23.22	21.74	24.02	24.05
35	23.88	25.32	22.90	24.84	23.70
40	25.27	24.69	24.67	24.48	25.24
45	24.20	24.94	25.52	25.02	27.43
50	27.98	25.88	26.67	26.31	25.94
55	27.42	24.91	26.42	28.24	25.49
60	28.41	27.09	29.04	28.85	27.89

试用合适的回归模型描述该药品剂量与反应的规律。

(李君荣)

关于临床试验的分期、方案设计、对象选择和多家研究单位的抽样方式等具体问题,读者可参阅药政管理的有关文件,如《中华人民共和国药品管理法》、《新药审批办法》、《药品临床试验管理规范》和《药品临床试验的管理规定》等。

二、盲法和安慰剂

临床试验中产生偏倚是一个既重要又普遍的问题,这是由有意识和无意识种种原因造成的,在施加处理因素、测量反应指标、分析、评估研究结果等任一阶段都会发生。解决办法大多靠随机化措施。但是当所测的反应指标是由受试者主观感觉(如患者的症状改善与否)和研究者主观判断(如医生认为“新疗法总比旧疗法优越”这样一个先入为主的印象后),随机化措施就无能为力了。解决这一类问题,只能使受试者不知道自己接受的是何种处理(具体用的是哪种药),而观察、测量者是知道的,称为“单盲”。这与前面讲过的“透明度”是两回事,透明度指的是让受试者明白这次临床试验的目的、意义,可能会出现哪些预期结果等。有时候还要使直接进行观察、测量的工作人员也不知道受试者接受的是什么处理。这就被称为“双盲”(double blind)。

安慰剂(placebo)是一种虚拟药物,其剂型、大小、重量、气味和口味等特征与试验药物尽可能保持一致,但不含试验药物的成分。它是临床试验中众多对照方法中的一种,类似于“空白”对照,但是空白对照不给受试者任何处理,本质上属于“非盲”的,往往影响到试验结果的正确评价。用安慰剂(属盲法)对照就能克服受试者、研究者和参与评价疗效及安全性的工作人员因心理因素造成的偏倚,最大限度地减少受试者和研究者的主观期望效应,避免对照组患者产生与受试组患者不同的心理作用和精神异常。同时,通过直接比较试验药物组和安慰剂组,起到消除疾病自然进程(自愈趋势等)的影响,证实试验药物究竟有无疗效(包括副作用)。盲法是使用安慰剂的基础,但是与空白对照一样,安慰剂对照也须注意伦理道德,决不能延误患者的治疗时机。

三、诊断试验的评价

(一) 概述

临床试验的受试对象必须有明确可靠的诊断,不能有可疑病例,更不允许混入误诊病例。这就需要最佳的诊断标准,这里的“最佳”是讲某个诊断指标既灵敏又特异。诊断的正确与否不仅影响临床试验结果,更关系到日常临床治疗中患者的康复和预后。

随着科学技术的发展,医学上曾不断出现过一些新的技术与方法,在其刚开始应用于临床时,由于缺乏科学的设计与评价,往往不恰当地夸大其价值。随着时间的推移和经验的积累,渐渐对它们有了正确的认识,甚至发现这些诊断试验并不理想。给诊断试验一个科学的评价是临床医师选择和采用该诊断试验的依据。

(二) 诊断试验的研究设计

1. 确定金标准(gold standard)

临床试验的金标准是指当前医学界所公认的诊断某病最为正确可靠的方法,如病理学诊断、外科手术的发现、特殊影像学诊断或长期临床随访以后获得的肯定诊断等。金标准选择不当会造成分类(有病、无病)错误,从而影响诊断试验的正确评价。

- (1) 统一编号: 对 15 000 名学生统一编号: 1, 2, 3, ..., 15 000;
- (2) 抽一数字: 本例抽样比例为十分之一, 故在 1~10(号)内随机抽得一数字为 6;
- (3) 确定样本: 则对应编号为 6, 16, 26, ..., 14 986, 14 996 的学生被抽中。

(三) 特点

系统抽样的优点是易于理解、操作简便; 被抽到的观察单位在总体中分布均匀, 抽样误差一般比单纯随机抽样小。其缺点是在某些特殊情形下会出现偏性或周期性变化; 无计算抽样误差的专用公式, 一般用单纯随机抽样的公式来代替。

三、分层抽样

(一) 概念

所谓分层抽样(stratified sampling), 是指先按对观察指标影响较大的某项或某几项特征, 将总体分成若干层(strata), 该特征的测定值在层内变异较小、层间变异较大; 然后分别从每一层内随机抽取一定数量的观察单位结合起来组成样本。

(二) 操作

例 13-2 的项目中, 将该大学新校区的学生划分为本科生、研究生两层, 每层均采用单纯随机抽样抽取二十分之一的学生调查其身体健康状况, 具体操作如下:

(1) 分层: 因本科生与研究生在年龄、学习、体育锻炼、生活等方面差别较大, 这些因素都与健康状况有关, 故将学历层次作为分层的特征;

(2) 抽本科生: 在 15 000 名本科生中采用单纯随机抽样抽取学生 750 人, 具体操作同单纯随机抽样;

(3) 抽研究生: 在 5 200 名研究生中采用单纯随机抽样抽取学生 260 人, 具体操作同单纯随机抽样。

(三) 特点

分层抽样的优点是抽样误差小; 各层可以独立进行统计分析; 尤其适合大规模的调查。其缺点是事先要对总体进行分层, 操作麻烦。如果只调查一个指标, 最好是用该指标作为分层的特征。但实际调查中同时须调查多个指标, 且考虑调查的行政组织需要, 往往用经济状况(与健康水平有关)、地理位置(与经济状况有关)或行政区划(与调查的组织管理有关)作为分层的特征。

四、整群抽样

(一) 概念

所谓整群抽样(cluster sampling), 是指将总体划分为群(初级观察单位, 各群由次级观察单位组成), 然后随机抽取一部分群, 调查抽中群的全部次级观察单位。

(二) 操作

例 13-2 的项目中, 将该大学新校区的学生以班为群, 采用整群抽样从中随机抽取 38 个班, 对这 38 个班的全体学生进行考试作弊情况的调查。具体操作如下:

(1) 编号: 该大学新校区共有 448 个班, 进行统一编号: 1, 2, ..., 448;

(2) 随机抽号: 采用随机数字表法, 从 448 个编号中随机抽取 38 个编号, 操作同单纯随机抽样;

机内,调查员可通过边询问、边输入的方式进行调查。当整个调查完成后,资料库也就建立起来了。目前,国外常用这种方式进行社会调查。但电话调查中因调查对象看不见书面调查表,对调查项目的理解有时会发生偏差。

4. 信访

信访也称函调,即将问卷邮寄或通过电子邮箱发送给调查对象,请调查对象填好后寄、发回。信访节省人力、财力,但会发生调查对象理解偏差、失访、选择偏倚等问题。

5. 集中调查法

将调查对象集中起来,当场发放调查表,各被调查者独立完成后交回调查表。该方法的优点是快速、费用低,但一般只适用于有组织、集中程度高的人群,如学校学生。它对于诸如卫生知识之类的调查测量尤其有效。

八、制订资料的整理与分析计划

对资料进行整理分析是将漫无头绪的粗糙资料去粗取精、去伪存真的精细加工过程,从而揭示事物内部的规律性和本质。

(一) 问卷的接受与核查

问卷接受是整理工作的第一步。其工作要点是认真地管理好收回的问卷,并做专门的登记,记录下完成日期和回收日期,掌握每天回收的问卷数。

问卷核查分为完整性核查和逻辑核查。所谓完整性核查是对调查表全部项目进行检查,核对填写是否完整无缺。如有漏项,应立即补填。完整性检查应在调查现场进行,否则难以弥补。所谓逻辑核查,主要是检查逻辑上的矛盾,例如某人的“死亡年月”不应早于其“出生年月”。

(二) 数据的计算机录入

可使用先进的数据处理软件 Epidata、Access 等录入调查表数据。为了保证资料的录入质量,可采用同一份资料由两人分别输入,然后通过软件核对输入的数据;也可根据调查项目间的逻辑关系编写程序进行逻辑查错,如女性不可能出现前列腺癌等。对数值变量资料,可作简单的统计描述,编制频数表发现异常值,或作散点图发现异常点等。总之,对调查资料在输入计算机阶段需要反复检查、核对,以确保资料的完整、准确和可靠。

(三) 资料的分组

分组的原则是把同质的事物(观察单位)结合在一起,把不同质的事物分离出去,以便显示组内的共性和组间的差异,最后揭露出事物内部的规律性。分组前先区别资料的类型,定量资料按量的大小分组,定性资料按性质或类别分组。如例 13-3 项目中,年龄按 18 岁~、30 岁~、40 岁~、50 岁~、60 岁~69 岁分组;吸烟按现在吸、曾吸过现不吸、从未吸过分组。

(四) 统计学分析

根据调查目的、资料类型、适用条件选用合适的统计分析方法,应用先进的统计软件进行分析。在调查设计阶段就应明确采用哪些卫生统计学分析方法,符合参数统计条件时采用哪些方法,不符合参数统计条件时又采用哪些相应的方法。例 13-3 项目在进行大量统计描述以后,以是否患心血管疾病为二分类反应变量;以年龄,腌熏食品、水果类食品摄入频率,情绪紧张有压力、烦躁不安,嗜好吸烟,酗酒,超重等为自变量,采用 SAS 软件进行 Logistic 回归分析,发现该省农民患心血管疾病的主要影响因素为年龄、水果类食物摄入量、

酗酒、超重。

(五) 医学专业分析

医学调查研究的主要目的是了解人群健康状况,探寻影响人群健康的因素,提出卫生策略与措施,医学调查研究结果应从这三个方面进行医学专业分析。例 13-3 项目的研究结果表明:该省农民心血管疾病的患病率为 24.1%;每天对水产品、禽类、猪牛羊肉类、豆制品、乳制品类摄入者占总人数的 10%以下,而每天摄入腌熏食品者达总人数的 22.0%;人群中基本没有体育运动者占 92.3%;超重率高达 37.8%;年龄、水果类食物摄入量、酗酒、超重等因素是该省农民心血管疾病的主要相关因素。因此,应加强农村地区人群的健康教育,改变人群的不良生活方式和饮食习惯,提倡农民每日坚持体育运动。

九、调查的组织措施

调查的组织措施贯穿于调查的全过程。调查研究是一项社会性很强的研究工作,调查的组织措施十分重要,它是调查研究得以顺利实施的重要保证。需要制订好组织领导、宣传发动、时间进度、人员培训、分工协调、经费预算、调查表格的印制、试点调查、现场安排、调查质量控制、调查表的回收、资料的检查核对、数据的上机输入、统计计算分析、结题报告的撰写等方面的组织措施计划,以书面或文件形式发到有关人员手中。主要方面如下:

(1) 组织领导:对大的调查研究项目,需成立调查项目领导小组、技术指导小组,明确各项工作的负责人。为取得当地行政领导的理解和支持,调查项目领导小组成员中可考虑有当地行政领导人员参加。对大型调查研究,有几家单位协作参加,则要选出牵头单位,明确分工,做好经费分配方案,协调好财务管理办法,明确调查成果、知识产权的归属与署名。

(2) 宣传发动:在调查项目组全体成员、调查员、调查地区的协作单位人员、调查对象中,进行广泛的动员宣传,充分调动全体成员的积极性,取得调查对象的信任和配合。

(3) 时间进度:对整个调查研究项目的全过程,制定切实可行的统一时间进度表,严格按照时间进度表开展各项工作。当某一环节出现拖延情况,应及时分析原因、设法补救和赶上进度。

(4) 培训管理:对调查员的挑选及人数确定,对调查员培训的教师、教材、方式、时间、地点,对培训中试调查现场的安排,对调查员培训的住宿、交通、场地、用餐、教材等经费,均须做出合理的计划,以提高培训的质量。

(5) 分工协调:一项调查研究尤其是一项大的调查研究,涉及单位、人员众多。对牵头单位及人员、协作单位及人员、调查现场有关单位及人员、调查项目领导小组及人员、技术指导小组及人员、各项工作的负责人、一般研究成员、调查督导员、调查员等,经广泛征求意见和认真讨论,研究制定出合理明确、具体细致的分工协调方案,各司其职,奖惩有据。

(6) 经费预算:调查研究项目的费用包括仪器设备添置费、调查表格的印刷费、调查员培训费及劳务费、调查交通费、协作单位及调查现场有关单位的协作费、差旅费、会议费、专家咨询费、通信费、办公用品费等,一定要有明细的预算。统筹兼顾,按财务制度合理使用经费。在多阶段抽样调查中,可在调查研究总费用有限的条件下使抽样误差达到最小时,确定出各阶段的调查样本大小。

(7) 现场安排:调查现场的安排落实,是调查研究人员尤其是高校、科研单位的研究者普遍感到棘手的问题。特别受市场经济、人员流动的影响,调查对象的配合较困难。调查项