

# 预防医学统计学

主 编 徐宝华 张一鸣

副主编 金传元 景学安 等

主 审 束怀符



山东大学出版社

# 预防医学统计学

徐宝华 张一鸣 主编

山东大学出版社出版发行

(济南市山大南路27号)

山东省泰安市第三印刷厂印刷

787×1092毫米 16开本 31印张 714千字

1991年8月第1版 1991年8月第1次印刷

印数1—6000

ISBN7—5607—0566—9/O·36

---

定价：11.50元

## 编 著 者

(按姓氏笔划排列)

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| 于景琮 | 马晓莉 | 王竹民 | 王 秀 | 王坤林 |
| 付文焕 | 孙宝春 | 刘光俞 | 刘爱国 | 乔建博 |
| 杜心芙 | 宋丹华 | 辛宝家 | 张一鸣 | 张志香 |
| 金传元 | 庞华明 | 赵跃进 | 徐宝华 | 景学安 |
| 傅瑞轩 | 楚贻康 |     |     |     |

## 前 言

本书是根据现代预防医学的发展,为了预防医学科研和实际工作的需要而编写的。全书共25章分上、下两篇:上篇19章为预防医学研究中常用的统计方法,除基本统计计算外,还介绍了拉丁方设计、析因设计、圆形分布、寿命表及减寿年数分析、直线和曲线拟合、二项分布和Poisson分布、Ridit分析、正常值制定、实验设计、调查设计、半数效量、相对危险度分析以及Logistic回归等多因素分析资料的统计方法。下篇6章为预防医学业务统计,包括卫生防疫统计、环境卫生统计、营养卫生统计、儿童少年卫生统计、劳动卫生与职业病统计和实验质量控制等。

本书内容丰富,理论联系实际,侧重阐述统计方法在预防医学实践中的应用,统计方法中着重阐述基本概念及应用条件,公式推导从略。本书对市、地、县、乡镇、厂矿从事预防医学工作人员是一本很有实用价值的参考书,也可供各级医疗卫生和妇幼保健人员学习参考,还可作为高、中等医学院校卫生专业学生的补充教材。

本书的作者们以其多年的教学和卫生防疫工作的实践经验,用自己的辛勤劳动,编写出了具有创新、实用特色的《预防医学统计学》一书,为建立预防医学统计体系迈出了可喜的一步。相信本书的出版,对预防医学工作将起到一定的推动作用。希望读者对本书的不足之处多加指教,再版时进行修改、充实与提高,以便更好地为各地市预防医学工作者进行科学研究和推动工作服务。

山东医科大学教授

束怀符

1991年1月5日

# 目 录

|                 |       |
|-----------------|-------|
| 绪 言             | ( 1 ) |
| 一、预防医学统计学的任务和内容 | ( 1 ) |
| 二、统计工作的步骤       | ( 1 ) |
| 三、统计资料的类型       | ( 2 ) |
| 四、统计中的几个基本概念    | ( 2 ) |
| 五、统计工作中应注意的问题   | ( 5 ) |

## 上篇 基本统计方法

|                    |        |
|--------------------|--------|
| 第一章 平均数与变异指标       | ( 7 )  |
| 1.1 频数表及其编制        | ( 7 )  |
| 1.2 平均数            | ( 8 )  |
| 一、算术均数             | ( 9 )  |
| 二、几何均数             | ( 11 ) |
| 三、中位数及百分位数         | ( 12 ) |
| 四、众数               | ( 14 ) |
| 1.3 变异指标           | ( 15 ) |
| 一、极差和四分位数间距        | ( 15 ) |
| 二、方差和标准差           | ( 16 ) |
| 三、变异系数             | ( 19 ) |
| 1.4 正态分布及其应用       | ( 19 ) |
| 一、正态分布的概念和特征       | ( 19 ) |
| 二、正态曲线下面积的分布规律及其应用 | ( 20 ) |
| 第二章 均数的抽样误差与假设检验   | ( 23 ) |
| 2.1 均数的抽样误差        | ( 23 ) |
| 一、抽样误差的概念          | ( 23 ) |
| 二、均数的标准误           | ( 23 ) |
| 2.2 t分布            | ( 24 ) |
| 2.3 总体均数的区间估计      | ( 25 ) |
| 2.4 假设检验的意义和一般步骤   | ( 27 ) |
| 2.5 均数的u检验         | ( 29 ) |

|                              |        |
|------------------------------|--------|
| 一、样本均数与总体均数的比较·····          | ( 29 ) |
| 二、两个大样本均数的比较·····            | ( 29 ) |
| 三、两个中位数的比较·····              | ( 30 ) |
| 2·6 均数的t检验·····              | ( 32 ) |
| 一、样本均数与总体均数的比较·····          | ( 32 ) |
| 二、配对计量资料的比较·····             | ( 32 ) |
| 三、两样本均数的比较·····              | ( 34 ) |
| 四、两样本几何均数的比较·····            | ( 36 ) |
| 2·7 两个方差的齐性检验与 $t'$ 检测·····  | ( 37 ) |
| 一、两个方差的齐性检验·····             | ( 37 ) |
| 二、 $t'$ 检验·····              | ( 38 ) |
| 2·8 两类错误·····                | ( 39 ) |
| 2·9 应用假设检验的注意问题·····         | ( 40 ) |
| <b>第三章 方差分析</b> ·····        | ( 41 ) |
| 3·1 方差分析的基本概念·····           | ( 41 ) |
| 3·2 单因素多个样本均数比较的方差分析·····    | ( 42 ) |
| 3·3 双因素多个样本均数比较的方差分析·····    | ( 44 ) |
| 3·4 各组均数间的两两比较·····          | ( 46 ) |
| 3·5 拉丁方设计资料的方差分析·····        | ( 48 ) |
| 3·6 析因设计资料的方差分析·····         | ( 51 ) |
| 3·7 多个方差的齐性检验·····           | ( 54 ) |
| 3·8 变量变换·····                | ( 56 ) |
| 一、平方根变换·····                 | ( 56 ) |
| 二、对数变换·····                  | ( 57 ) |
| 三、反正弦变换·····                 | ( 58 ) |
| <b>第四章 相对数</b> ·····         | ( 60 ) |
| 4·1 相对数的概念·····              | ( 60 ) |
| 4·2 常用相对数·····               | ( 60 ) |
| 一、率·····                     | ( 60 ) |
| 二、构成比·····                   | ( 60 ) |
| 三、相对比·····                   | ( 61 ) |
| 4·3 动态数列·····                | ( 61 ) |
| 4·4 应用相对数应注意的问题·····         | ( 63 ) |
| 4·5 标准化法·····                | ( 65 ) |
| 一、标准化的意义·····                | ( 65 ) |
| 二、标准化率的计算·····               | ( 66 ) |
| <b>第五章 率的抽样误差与假设检验</b> ····· | ( 70 ) |
| 5·1 率的抽样误差·····              | ( 70 ) |

|                                |         |
|--------------------------------|---------|
| 一、率的标准误                        | ( 70 )  |
| 二、总体率的区间估计                     | ( 70 )  |
| 5.2 率的u检验                      | ( 72 )  |
| 一、样本率与总体率的比较                   | ( 72 )  |
| 二、两个样本率的比较                     | ( 72 )  |
| 5.3 $\chi^2$ 检验                | ( 73 )  |
| 一、四格表资料的 $\chi^2$ 检验           | ( 73 )  |
| 二、行 $\times$ 列表资料的 $\chi^2$ 检验 | ( 77 )  |
| 三、配对计数资料的 $\chi^2$ 检验          | ( 78 )  |
| 5.4 四格表的确切概率法                  | ( 81 )  |
| 5.5 行 $\times$ 列表的 $\chi^2$ 分割 | ( 83 )  |
| 5.6 同组中构成比的比较                  | ( 85 )  |
| 一、同组中两个构成比的比较                  | ( 86 )  |
| 二、同组中多个构成比的比较                  | ( 86 )  |
| 三、同组中多个构成比的两两比较                | ( 87 )  |
| <b>第六章 二项分布与Poisson分布</b>      | ( 88 )  |
| 6.1 二项分布                       | ( 88 )  |
| 一、二项分布的概念及应用条件                 | ( 88 )  |
| 二、二项分布的图形                      | ( 90 )  |
| 三、二项分布的均数与标准差                  | ( 90 )  |
| 四、二项分布的应用                      | ( 91 )  |
| 五、疾病的聚集性研究                     | ( 91 )  |
| 6.2 Poisson分布                  | ( 93 )  |
| 一、Poisson分布的概念及应用条件            | ( 93 )  |
| 二、Poisson分布的概率函数及图形            | ( 93 )  |
| 三、Poisson分布的性质                 | ( 94 )  |
| 四、Poisson分布的应用                 | ( 94 )  |
| <b>第七章 秩和检验</b>                | ( 99 )  |
| 7.1 非参数统计的概念                   | ( 99 )  |
| 7.2 配对资料比较的秩和检验                | ( 99 )  |
| 7.3 两样本比较的秩和检验                 | ( 101 ) |
| 7.4 单因素多个样本比较的秩和检验             | ( 105 ) |
| 7.5 单因素多个样本间两两比较的秩和检验          | ( 107 ) |
| 一、各样本含量相等的两两比较                 | ( 107 ) |
| 二、各样本含量不等时的两两比较                | ( 108 ) |
| 7.6 双因素多个样本比较的秩和检验             | ( 110 ) |
| <b>第八章 Ridit分析</b>             | ( 114 ) |
| 8.1 Ridit值的计算方法                | ( 114 ) |

|             |                 |       |
|-------------|-----------------|-------|
| 8·2         | R值的假设检验         | (115) |
| 8·3         | 两组平均Ridit比较     | (117) |
| 8·4         | 多组平均Ridit的比较    | (118) |
| <b>第九章</b>  | <b>直线相关与回归</b>  | (121) |
| 9·1         | 直线相关            | (121) |
|             | 一、相关系数的意义       | (121) |
|             | 二、相关系数的计算       | (122) |
|             | 三、相关系数的假设检验     | (123) |
|             | 四、两直线相关系数的比较与合并 | (124) |
| 9·2         | 直线回归            | (125) |
|             | 一、求直线回归方程       | (125) |
|             | 二、绘制回归方程        | (126) |
|             | 三、回归系数的假设检验     | (126) |
|             | 四、两直线回归方程的比较    | (127) |
| 9·3         | 进行相关和回归分析应注意的问题 | (129) |
| 9·4         | 等级相关            | (130) |
| <b>第十章</b>  | <b>曲线回归</b>     | (132) |
| 10·1        | 曲线回归的用途与拟合步骤    | (132) |
| 10·2        | 指数与对数曲线的拟合      | (133) |
|             | 一、指数曲线的拟合       | (133) |
|             | 二、对数曲线的拟合       | (136) |
| 10·3        | 双曲线的拟合          | (138) |
| 10·4        | 抛物线的拟合          | (140) |
|             | 一、二次抛物线的拟合      | (140) |
|             | 二、三次抛物线的拟合      | (143) |
| 10·5        | 曲线方程的假设检验和拟合优度  | (145) |
| <b>第十一章</b> | <b>正常值范围的制定</b> | (148) |
| 11·1        | 制定正常值的原则        | (148) |
| 11·2        | 制定正常值范围的方法      | (150) |
|             | 一、正态分布法         | (150) |
|             | 二、百分位数法         | (153) |
| 11·3        | 如何选取正常值范围的界限    | (157) |
| <b>第十二章</b> | <b>调查设计</b>     | (161) |
| 12·1        | 医学科研中的统计设计      | (161) |
| 12·2        | 调查计划            | (161) |
| 12·3        | 调查方法            | (165) |
| 12·4        | 整理分析计划          | (167) |
| 12·5        | 抽样设计            | (168) |

|             |                   |       |
|-------------|-------------------|-------|
| 12·6        | 样本含量估计            | (172) |
| 12·7        | 非抽样误差的控制          | (174) |
| <b>第十三章</b> | <b>实验设计</b>       | (176) |
| 13·1        | 实验设计的基本要素         | (176) |
| 一、          | 处理因素              | (176) |
| 二、          | 受试对象              | (177) |
| 三、          | 实验效应              | (178) |
| 13·2        | 实验设计的基本原则         | (179) |
| 一、          | 随机化原则             | (179) |
| 二、          | 对照的原则             | (180) |
| 三、          | 重复的原则             | (181) |
| 四、          | 均衡的原则             | (183) |
| 13·3        | 常用的几种实验设计方法       | (184) |
| 一、          | 交叉设计              | (184) |
| 二、          | 完全随时设计            | (185) |
| 三、          | 配对设计              | (185) |
| 四、          | 配伍组设计             | (186) |
| 五、          | 拉丁方设计             | (186) |
| 六、          | 正交设计              | (187) |
| 七、          | 盲法设计              | (188) |
| <b>第十四章</b> | <b>统计表与统计图</b>    | (189) |
| 14·1        | 统计表               | (189) |
| 一、          | 制表的基本要求           | (189) |
| 二、          | 统计表的种类            | (190) |
| 14·2        | 统计图               | (190) |
| 一、          | 制图的基本要求           | (190) |
| 二、          | 常用统计图及其绘制方法       | (191) |
| <b>第十五章</b> | <b>寿命表及减寿年数分析</b> | (201) |
| 15·1        | 寿命表的概念            | (201) |
| 15·2        | 寿命表指标及其计算         | (202) |
| 一、          | 寿命表指标、符号及含义       | (202) |
| 二、          | 寿命表各项指标的计算        | (202) |
| 15·3        | 简略寿命表的编制方法        | (204) |
| 15·4        | 寿命表的分析应用          | (205) |
| 一、          | 评价居民健康状况          | (205) |
| 二、          | 研究某种死因对居民预期寿命的影响  | (207) |
| 三、          | 分析人口再生育情况         | (211) |
| 四、          | 研究疾病的发展过程及预后      | (211) |

|             |                         |              |
|-------------|-------------------------|--------------|
| 15·5        | 减寿年数及其在死因分析中的应用·····    | (211)        |
| 一、          | 减寿年数的计算及性质·····         | (211)        |
| 二、          | 减寿年数指标与其他死亡指标的比较·····   | (212)        |
| <b>第十六章</b> | <b>圆形分布资料的统计分析·····</b> | <b>(214)</b> |
| 16·1        | 圆形分布的一般概念·····          | (214)        |
| 16·2        | 圆形分布资料的图示·····          | (215)        |
| 16·3        | 角的均数和标准差·····           | (216)        |
| 16·4        | 平均角的假设检验·····           | (218)        |
| 16·5        | 两个或多个样本角均数间的比较·····     | (220)        |
| 一、          | 两个样本角均数的比较·····         | (220)        |
| 二、          | 多个样本角均数的比较·····         | (223)        |
| <b>第十七章</b> | <b>相对危险度分析·····</b>     | <b>(224)</b> |
| 17·1        | 概述·····                 | (224)        |
| 17·2        | 队列研究中累积发病率资料的分析·····    | (225)        |
| 一、          | 不分层资料的分析·····           | (225)        |
| 二、          | 分层资料的分析·····            | (227)        |
| 17·3        | 队列研究中发病密度资料的分析·····     | (229)        |
| 一、          | 不分层资料的分析·····           | (229)        |
| 二、          | 分层资料的分析·····            | (231)        |
| 17·4        | 成组病例对照研究资料的分析·····      | (234)        |
| 一、          | 不分层资料的分析·····           | (234)        |
| 二、          | 分层资料的分析·····            | (235)        |
| 17·5        | 配对病例对照研究资料的分析·····      | (237)        |
| 一、          | 1:1配对资料的分析·····         | (237)        |
| 二、          | 1:R配对资料的分布·····         | (239)        |
| <b>第十八章</b> | <b>半数效量·····</b>        | <b>(241)</b> |
| 18·1        | 概述·····                 | (241)        |
| 18·2        | 概率单位法·····              | (242)        |
| 一、          | 目测法(图解法)·····           | (242)        |
| 二、          | 最小二乘法(直线回归法)·····       | (245)        |
| 18·3        | 加权概率单位法·····            | (247)        |
| 18·4        | 寇氏法·····                | (249)        |
| 18·5        | 序贯法·····                | (251)        |
| <b>第十九章</b> | <b>多元分析简介·····</b>      | <b>(253)</b> |
| 19·1        | 多元分析的特点、内容与方法·····      | (253)        |
| 一、          | 多元分析的特点·····            | (253)        |
| 二、          | 多元分析的主要内容·····          | (253)        |
| 三、          | 几种常用的分析方法·····          | (254)        |

|                     |         |
|---------------------|---------|
| 19·2 多元线性回归         | ( 254 ) |
| 一、多元回归方程的用途         | ( 254 ) |
| 二、方法和步骤             | ( 255 ) |
| 三、例 题               | ( 256 ) |
| 19·3 多元线性逐步回归       | ( 257 ) |
| 一、用 途               | ( 258 ) |
| 二、方 法               | ( 258 ) |
| 三、方法和步骤             | ( 258 ) |
| 四、例 题               | ( 259 ) |
| 19·4 Logistic回归模型   | ( 259 ) |
| 一、Logit变换           | ( 260 ) |
| 二、比数比OR             | ( 260 ) |
| 三、Logistic回归模型的种类   | ( 261 ) |
| 四、Logistic回归模型的参数估计 | ( 261 ) |
| 五、Logistic回归模型的建立   | ( 262 ) |
| 六、Logistic回归模型的主要指标 | ( 263 ) |
| 七、例 题               | ( 263 ) |

## 下篇 预防医学业务统计

|                   |         |
|-------------------|---------|
| <b>第一章 防疫工作统计</b> | ( 265 ) |
| 1·1 疫情统计          | ( 265 ) |
| 一、概 述             | ( 265 ) |
| 二、疫情报告管理质量指标      | ( 266 ) |
| 三、疫情报告质量的综合评价     | ( 269 ) |
| 四、疫情统计常用指标        | ( 272 ) |
| 五、疫情资料的分析         | ( 273 ) |
| 六、疫情预测预报          | ( 277 ) |
| 1·2 预防接种统计        | ( 280 ) |
| 一、预防接种资料的收集       | ( 280 ) |
| 二、制定预防接种的指标       | ( 281 ) |
| 三、预防接种完成情况指标      | ( 281 ) |
| 四、接种反应观察指标        | ( 283 ) |
| 五、效果观察指标          | ( 284 ) |
| 六、综合评分法           | ( 288 ) |
| 1·3 疾病防治统计        | ( 291 ) |
| 一、痢疾防治统计          | ( 291 ) |
| 二、疟疾防治统计          | ( 297 ) |

|                                |         |
|--------------------------------|---------|
| 三、病毒性肝炎防治统计·····               | ( 302 ) |
| 四、流行性脑脊髓膜炎防治统计·····            | ( 304 ) |
| 五、流行性出血热防治统计·····              | ( 310 ) |
| 六、钩虫病防治统计·····                 | ( 316 ) |
| 七、丝虫病防治统计·····                 | ( 320 ) |
| 八、地方性甲状腺肿与地方性克汀病防治统计·····      | ( 323 ) |
| 九、克山病防治统计·····                 | ( 331 ) |
| 十、心血管疾病防治统计·····               | ( 334 ) |
| 十一、恶性肿瘤防治统计·····               | ( 338 ) |
| 十二、精神病防治统计·····                | ( 356 ) |
| 十三、意外死亡统计·····                 | ( 358 ) |
| 十四、结核病防治统计·····                | ( 362 ) |
| <b>第二章 环境卫生监测统计</b> ·····      | ( 367 ) |
| 2·1 环境卫生监测统计的意义和任务·····        | ( 367 ) |
| 2·2 资料来源·····                  | ( 367 ) |
| 一、环境质量资料·····                  | ( 367 ) |
| 二、环境污染源情况资料·····               | ( 368 ) |
| 三、人群健康状况资料·····                | ( 368 ) |
| 2·3 环境监测常用指标·····              | ( 369 ) |
| 一、常用的率·····                    | ( 369 ) |
| 二、移动平均数·····                   | ( 369 ) |
| 2·4 环境质量统计的评价·····             | ( 371 ) |
| 一、环境质量指数的统计分析评价·····           | ( 371 ) |
| 二、大气污染指标·····                  | ( 372 ) |
| 三、水质质量指数·····                  | ( 375 ) |
| 四、土壤质量指数·····                  | ( 375 ) |
| 2·5 环境质量综合评价指标·····            | ( 376 ) |
| 2·6 环境质量预测预报·····              | ( 377 ) |
| 一、利用动态数列进行预测预报·····            | ( 377 ) |
| 二、回归方程预测法·····                 | ( 377 ) |
| <b>第三章 营养卫生及食品卫生监测统计</b> ····· | ( 379 ) |
| 3·1 营养调查·····                  | ( 379 ) |
| 一、膳食调查·····                    | ( 379 ) |
| 二、体格营养状况检查结果评价·····            | ( 382 ) |
| 三、实验室检查结果评价·····               | ( 387 ) |
| 四、营养缺乏症检查结果的综合评价·····          | ( 387 ) |
| 3·2 食品卫生监测工作统计·····            | ( 389 ) |
| 一、资料来源·····                    | ( 389 ) |

|                              |                |
|------------------------------|----------------|
| 二、常用统计指标.....                | ( 390 )        |
| <b>第四章 儿童少年生长发育统计.....</b>   | <b>( 391 )</b> |
| 4·1 资料的来源.....               | ( 391 )        |
| 一、生长发育统计指标的建立.....           | ( 391 )        |
| 二、调查设计.....                  | ( 391 )        |
| 4·2 生长发育的评价.....             | ( 392 )        |
| 一、指数法.....                   | ( 392 )        |
| 二、离差法.....                   | ( 393 )        |
| 三、百分位数法.....                 | ( 396 )        |
| 四、相关回归分析法.....               | ( 396 )        |
| 五、体型图法.....                  | ( 399 )        |
| 六、发育年龄评价法.....               | ( 399 )        |
| 七、生长速度曲线.....                | ( 402 )        |
| 八、其它单项评价方法.....              | ( 402 )        |
| 九、综合评价法.....                 | ( 404 )        |
| <b>第五章 劳动卫生及职业病防治统计.....</b> | <b>( 408 )</b> |
| 5·1 劳动卫生及职业病防治统计内容.....      | ( 408 )        |
| 5·2 资料的收集、整理与分析.....         | ( 408 )        |
| 一、基础资料.....                  | ( 408 )        |
| 二、生产环境和生物材料监测资料.....         | ( 409 )        |
| 三、监测及资料整理时的几个问题.....         | ( 409 )        |
| 四、监测结果分析的常用指标.....           | ( 410 )        |
| 五、职业病体格检查及发病资料的统计分析.....     | ( 411 )        |
| 六、专题调查分析.....                | ( 412 )        |
| 5·3 矽肺防治统计.....              | ( 422 )        |
| 一、生产性粉尘资料的收集和评价.....         | ( 422 )        |
| 二、矽肺健康检查及防尘效果的评价.....        | ( 425 )        |
| 三、矽肺的长期动态观察.....             | ( 428 )        |
| <b>第六章 卫生检验质量控制.....</b>     | <b>( 431 )</b> |
| 6·1 概述.....                  | ( 431 )        |
| 6·2 内容及方法.....               | ( 431 )        |
| 一、内 容.....                   | ( 431 )        |
| 二、基本条件的控制.....               | ( 432 )        |
| 三、工作人员素质和自我控制.....           | ( 432 )        |
| 四、室内质控.....                  | ( 432 )        |
| 五、室内质控的评价方法.....             | ( 434 )        |
| 六、室间质量控制及其评价.....            | ( 434 )        |
| 6·3 质量控制常用统计指标.....          | ( 437 )        |

|                                        |         |
|----------------------------------------|---------|
| 一、室内精密度与室间精密度.....                     | ( 437 ) |
| 二、变异系数 ( CV ) .....                    | ( 440 ) |
| 三、变异指数.....                            | ( 440 ) |
| 6·4 质量控制常用的统计方法与质控图.....               | ( 441 ) |
| 一、质量控制常用统计方法.....                      | ( 441 ) |
| 二、质控图.....                             | ( 442 ) |
| <b>附 统计用表</b> .....                    | ( 453 ) |
| 附表1 标准正态分布曲线下的面积.....                  | ( 453 ) |
| 附表2 t界值表.....                          | ( 455 ) |
| 附表3 F界值表 ( 方差齐性检验用 ) .....             | ( 456 ) |
| 附表4 F界值表 ( 方差分析用 ) .....               | ( 457 ) |
| 附表5 q界值表 ( Newman—Keuls 检验用 ) .....    | ( 461 ) |
| 附表6 百分率的可信限.....                       | ( 462 ) |
| 附表7 $\chi^2$ 界值表.....                  | ( 466 ) |
| 附表8 阶乘的对数表.....                        | ( 467 ) |
| 附表9 poisson分布 $\lambda$ 的可信区间.....     | ( 468 ) |
| 附表10 T界值表 ( 配对比较的符号秩和检验用 ) .....       | ( 469 ) |
| 附表11 T界值表 ( 两样本比较的秩和检验用 ) .....        | ( 470 ) |
| 附表12 H界值表 ( 三样本比较的秩和检验用 ) .....        | ( 472 ) |
| 附表13 M界值表 ( 配伍组比较的秩和检验用 ) .....        | ( 473 ) |
| 附表14 r界值表.....                         | ( 474 ) |
| 附表15 r界值表.....                         | ( 476 ) |
| 附表16 百分数与概率单位对照表.....                  | ( 477 ) |
| 附表17 正态性D检验界值表.....                    | ( 479 ) |
| 附表18 随机数字表.....                        | ( 480 ) |
| 附表19 随机排列表 ( $n=20$ ) .....            | ( 481 ) |
| 附表20 配对比较 ( t检验 ) 时所需样本含量.....         | ( 482 ) |
| 附表21 两样本均数比较 ( t检验 ) 时所需样本含量.....      | ( 483 ) |
| 附表22·1 两样本率比较时所需样本含量 ( 单侧 ) .....      | ( 484 ) |
| 附表22·2 两样本率比较时所需样本含量 ( 双侧 ) .....      | ( 485 ) |
| 附表23 圆形分布 r 的界值.....                   | ( 486 ) |
| 附表24 Watson与Williams检验用的校正因子, K值表..... | ( 487 ) |
| 附表25 加权系数.....                         | ( 488 ) |

# 绪 言

## 一、预防医学统计学的任务和内容

统计学是认识世界的一种重要工具。它是研究数据的收集、整理、分析与推断的科学。卫生统计学是把统计理论、方法应用于居民健康状况研究、医疗卫生实践和医学科研的一门应用科学。预防医学统计学是卫生统计学在预防医学领域中的应用。

预防医学研究的核心问题，是环境与人群健康的关系。为了研究不同环境条件下人群健康状况、效应或反应（包括疾病、死亡、寿命、生育、儿童生长发育及生理病理反应等）的分布和变化规律以及影响这些规律的因素，作为制订预防保健策略和措施的基础，必须应用统计学的方法，才能完成上述任务。也只有应用预防医学统计方法才能够对人群健康和疾病给予定量评估。

预防医学统计学主要包括二方面的内容：

（一）基本统计方法 如收集、整理统计资料的方法，计量资料的统计方法，计数资料的统计方法和等级资料的统计方法以及预防医学科研的调查设计和实验设计等。

（二）业务统计 如防疫工作统计、环境卫生监测统计、营养卫生工作统计、儿童少年健康状况统计、劳动卫生与职业病防治工作统计及卫生检验质量控制等。

## 二、统计工作的步骤

预防医学统计工作和一般统计工作类似，可分为四个步骤：即先要有一个全过程的设计，然后按照设计的要求去收集资料、整理资料和分析资料。这四个步骤是互相联系，不可分割的，任何步骤的缺陷，都会影响统计分析的结果。

（一）设计（design） 首先要明确研究目的，要对被研究的事物有一定的了解，可根据已往工作的经验和参考文献，或通过试查和预备试验，掌握较多的信息。对工作的全过程要有一个全面的设想。例如根据研究目的需要收集哪些资料？人力、财力和客观条件是否可能办到？用什么方式和方法来取得原始资料？怎样对取得的原始资料作进一步的整理汇总？怎样对汇总后资料进行有关统计指标的计算？预期会得到什么结果？凡此种种，都要经过周密地考虑，明确地回答，结合实际情况，作出科学的、细致的安排，才能用较少的人力、财力取得较大的效果，特别是要用电子计算机来整理、分析资料，这个全过程的设计，显得更为重要。

（二）收集资料（Collection of data） 它的任务是按研究设计的要求，及时取得准确、完整的原始数据。当然只有原始数据可靠，才能得出可靠的结论。因此，这一步骤具有极重要的基础意义。

预防医学统计资料的来源主要有三个方面：①统计报表。如疫情报表、卫生工作报

表等。②卫生工作记录和报告卡片。如各种卫生学检查记录及传染病报告卡等。③专题调查或实验研究等。

(三)整理资料 (Sorting data) 它的任务是按研究设计的要求,使原始数据系统化、条理化,便于进一步的计算指标和分析。为此,必须认真地核校原始资料,细心地分组和归纳,才能消除或减少整理中引入的误差。

(四)分析资料 (analysis of data) 它的任务是按研究设计的要求,计算有关指标,阐明事物的内在联系和规律。为此,必须对整理出的基础数据,作进一步的计算分析和统计处理,并用适当的统计图表陈示资料,最后结合专业作出恰如其分的结论。

### 三、统计资料的类型

(一)计量资料 (measurement data) 对每个观察单位用定量方法测定某项指标量的大小,所得的资料称为计量资料,一般有度量衡等单位。如调查某地10岁女童的身体发育状况,以人为观察单位,每个人的身高(cm)、体重(kg)和血压(kPa)等;又如以每个采样点为观察单位,测得不同采样点的二氧化碳浓度(mg/L)。

(二)计数资料 (enumeration data) 将观察单位按某种属性或类别分组,所得各组的观察单位数,称为计数资料。例如对某小学全体学生进行蛔虫卵粪检,每个学生是一个观察单位,将每个学生按粪检结果阳性与阴性分组,得每组人数;又如调查某人群的血型分布,按A、B、AB、O四型分组,得各血型组的人数。

(三)等级资料 (ranked data) 将观察单位按某种属性的不同程度分组,所得各组的观察单位数,称为等级资料。例如用某药治疗若干痢疾病人,其中治愈、好转、无效人数;测定某项血清反应的一、±、+、++人数。这类资料与计数资料不同的是:属性的分组有程度的差别,各组按大小顺序排列;与计量资料不同的是:每个观察单位未确切定量,因而称为半计量资料。

(四)圆形资料 (circle graph data) 用角度表示的资料和可以换算成角度的资料,称为圆形资料。如心电图量的方位角,脑血流图的上升角、主峰角,另外有的数据以年、月、日或时间的时、分表示。如疾病的发病时间、死亡时间、新生儿出生时间、微丝蚴在血液中出现的时间等,这类资料具有周期性,呈圆形分布,故称为圆形分布资料。

根据分析的需要,计量资料、计数资料和等级资料可以互相转化。例如每个人的血红蛋白,原属计量资料;若按血红蛋白正常与异常分为两组,得各组人数,则是计数资料;若将血红蛋白按量(g/dl)的多少分为五个等级:<6(重度贫血)、6~(中度贫血)、9~(轻度贫血)、12~16(血红蛋白正常)、>16(血红蛋白增高),得各等级人数,就是等级资料。再如多变量的线性回归分析中,要求将定性指标数量化,如男、女按性别分组,可得计数资料;若将女用0表示,男用1表示,这时就转化为计量资料了。

### 四、统计中的几个基本概念

(一)同质与变异 医学科学实验研究的对象是有生命的人或动物,其特点是有变

异性,即使是性质相同的事物,就同一观察指标来看,各观察单位(亦称个体)之间,也各有差异,称为变异(variation)。例如,研究儿童的身体发育,同性别、同年龄儿童(统计上称为“同质”观察单位)的身高,有高有低,各不相同,称为身高的变异。日常生活中,有时也可以看到,同性别中,有的低年龄儿童的身高高于高年龄儿童,但总的说来,儿童的身高总是随年龄而增加,这是客观规律。同理,同种属、同性别、年龄相近的小白鼠,喂以同种饲料,其所增体重(g)亦各不相同,亦称为变异,但客观规律总是饲料所含营养成分愈多,所增体重愈多。上例中的个体变异表现为定量,如身高的高低,所增体重的多少;有时亦可表现为定性的,如用某药治疗某种病人后的痊愈、好转、恶化等。

同质观察单位之间的个体变异,是生物的重要特征,是偶然性的表现,是由于生物体内外环境中,多种因素的综合影响造成的,其中许多因素是未知的,也是难于控制的。然而在这些变异性的现象中却蕴藏着必然的规律性。统计的任务就是在同质分组的基础上,通过对个体变异的研究,透过偶然现象,反映同质事物的本质特征和规律。

(二) 总体(population)与样本(sample) 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分,对这一小部分称为样本。而样本所代表的全体称为总体。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,浏览卫生统计的书刊,几乎都无遗漏地首先介绍这一概念。但容易理解的东西,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。例如,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。又如,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 $1\sim 5\mu\text{m}$ 的粒子占80%。该百分率为样本百分率,200个粒子称样本,而要推论的则是与本次同样条件下发生的全部气溶胶粒子,即其总体。再如,对某地区进行某种疾病的普查,全地区共100 000人中发现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3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差异等规律性。否则,不能依据样本推断总体。

(三) 抽样误差(sampling error) 由于总体中各观察单位间存在个体变异,抽样研究中抽取的样本,只包含总体的一部分观察单位,因而样本指标不一定恰好等于相应的总体指标。例如从某市1982年7岁男童的总体中随机抽取110名儿童,调查得此样本的身高均数为119.95cm,这个数值不一定恰好等于该市7岁男童的总体均数。又如从某地随机抽取500人,查出乙型肝炎病毒携带率为10%,这个数字不一定恰好等于该地人群乙型肝炎病毒携带率。这种样本指标与总体指标的相差称为抽样误差。显然,抽样误