

主 编 李平凡 张功员

卫生统计学学习指导

河南大学出版社

(豫)新登字 09 号

卫生统计学学习指导

主 编 李平凡

张功员

责任编辑 张玉琦

河南大学出版社出版发行

(开封市明伦街 85 号)

黄委会设计院印刷厂印刷

开本:850×1168 毫米 1/32 印张:7.75 字数:194 千字

1993 年 8 月第 1 版 1993 年 8 月第 1 次印刷

印数:1—2500 定价:5.50 元

ISBN 7-81041-029-6/R·19

主 审 李叔民

主 编 李平凡 张功员

副主编 张延安 杨维奇 徐荣战 霍天帮

编 委 (以姓氏笔划为序)

王永生 王启俊 仝兴涛 邢华燕

宋延堂 张艳玲 张效梅 胡书孝

赵胜利 曾国保 程黎明

编写说明

《卫生统计学》是把统计学原理与方法广泛地应用于到医疗、卫生领域的一门实用学科。为帮助中等卫生学校师生更好地学习和应用这门科学,作者根据“周士楷主编《卫生统计学》”和“卫生统计学教学大纲”,编写了这本《卫生统计学学习指导》。

全书共分三部分:第一部分为教学参考,包括教学重点,教学提示和复习思考题解答;第二部分为实习作业解答,对教材末实习二至实习十全部作业题给出解答;第三部分为试卷,是根据大纲要求,编写了五份卫生统计学试卷和参考答案,供教师参考,供同学们自测。为方便读者,全书的章节顺序、习题、统计符号、缩写等均与原教材一致。由于本书运算繁多,许多章节都是由多位作者共同合作完成的。最后,第一部分和第三部分由李平凡主编统稿,第二部分由张功员主编统稿。

在本书的编写过程中,得到了河南医科大学、西安医科大学、焦作市博爱卫生职业中专、焦作卫校、郑州铁路卫校、博爱县卫生防疫站、焦作市第二人民医院、博爱县人民医院等单位的热情支持和大力帮助;李炽民教授在百忙中抽出时间对本书进行了细致的审查;河南大学出版社的领导和编辑对本书的出版给予了大力支持。在此向上述单位及同志表示衷心的感谢。

由于成书时间仓促,编者水平有限,错误之处在所难免。希广大读者不吝批评指教,编者深表谢忱,并将在修订时努力改正。

编者

1993年9月

目 录

教学参考部分

第一章	绪论	(2)
第二章	计量资料频数分布的特征与规律	(7)
第三章	总体均数的估计和假设检验	(13)
第四章	方差分析	(18)
第五章	相对数	(24)
第六章	二项分布及其应用	(29)
第七章	泊松分布及其应用	(32)
第八章	χ^2 (卡方)检验	(39)
第九章	非参数统计	(43)
第十章	直线相关与回归	(47)
第十一章	多元线性回归简介	(56)
第十二章	半数致死量	(70)
第十三章	统计表与统计图	(79)
第十四章	调查设计	(85)
第十五章	实验设计	(88)
第十六章	居民健康统计	(92)

实习题解部分

实习一	函数型电子计算器的使用方法	(99)
-----	---------------------	------

实习二	计量资料频数分布、集中趋势指标和离散程度指标	(100)
实习三	总体均数的估计和假设检验	(108)
实习四	方差分析	(131)
实习五	相对数	(148)
实习六	二项分布及其应用	(153)
实习七	χ^2 检验	(157)
实习八	秩和检验	(166)
实习九	直线相关与回归	(178)
实习十	统计表与统计图	(193)

试卷部分

试卷(一)	(207)
试卷(二)	(210)
试卷(三)	(213)
试卷(四)	(216)
试卷(五)	(220)
试卷参考答案	(224)

教学参考部分

第一章 绪 论

(讲授 3 学时)

教学重点(一)

基 本 概 念	变异与变量:变异就是个体差异。对变异观测的数值称变量。
	总体与样本:根据研究目的确定的研究对象的全体称总体;样本是从总体内随机抽取的一部分。
	概率与频率:概率是事件发生的可能性大小的度量,是对总体而言的;频率也是事件发生的可能性大小的度量,是对样本而言的。
	偶然误差与系统误差:偶然误差包括抽样误差和随机测量误差;系统误差是由某种确定原因引起的测得值与真值的偏差。

教学重点(二)

统计工作	步 骤	1. 全过程设计; 2. 收集资料; 3. 整理资料; 4. 分析资料。
	资料类型	1. 计数资料; 2. 计量资料; 3. 等级资料。

教学提示

1. 本章是《卫生统计学》学习的开始,要用大量时间使学生认识本学科在医疗、卫生科研当中所发挥的巨大作用,启发学生们的学习兴趣,才能达到应有的教学目的。要求多举例。
2. 必须弄清统计学的实质;透过现象(数据)看本质(病因、规律、关系等)。
3. 对所有的基本概念都要在充分理解的基础上重点记忆。
4. 要当一名合格的卫生医士(师),要在医学、卫生领域内有所造诣,必须学好《卫生统计学》这门工具学科。

复习思考题解答

1. 卫生统计学的任务是什么? 卫生医士专业为什么要学习卫生统计学?

卫生统计学的任务是借助于统计学的原理和方法,从有限的观察中,从表现为偶然的数据中,把所研究的事物现象的本质特征、整体情况和相互关系一一揭示出来。

卫生医士专业学生将来要从事卫生防疫工作,工作中为了解居民健康状况和评价医疗卫生措施的效果;为了判断社会、环境和生物因素对居民健康的影响;为了改善卫生条件,提高居民健康水平,经常要做流行病学、卫生学调查和实验研究。卫生统计学是完成这些任务的重要手段。所以作为一名合格的卫生医士,必须学习,而且要学好卫生统计学。

2. 卫生统计工作的步骤,各步骤的意义与要求如何?

卫生统计工作的步骤分为:

①统计全过程设计:

意义:用尽可能少的人力、物力和时间,获得准确、可靠的结论。

要求:科学、周密、简明

②收集资料

意义:为整理资料之基础。

要求:完整、准确、及时。

③整理资料

意义:为分析资料之基础。

要求:系统化、条理化。

④分析资料

意义:揭示事物的规律。

要求:方法准确、计算精确、判断合理。

3. 统计资料的类型分哪几种? 各举实例说明。

统计资料分为三种,即:

①计数资料:如粪检人体寄生虫卵时,所得阳性、阴性的人数;又如调查化验人群血型,所得 A、B、O、AB 各型的人数。

②计量资料:如体重测量时所得的公斤数;测量身高所得的厘米数。

③等级资料:如尿糖测定时,分为一、+、++、+++、++++,然后清点各组个数所得资料,又叫半定量资料。

4. 观察值变异是由什么原因引起的? 试举例说明。

观察值的变异原因是个体与外界环境关系极为复杂,影响个体特征的内、外因素很多,不同因素对于个体的影响偶然地结合在一起,形成了差异。使得生物体在相同条件下,出现了互不一致的结果。如:同一年龄组的青少年女学生血红蛋白含量各不相同。

5. 举例说明抽样研究的目的。

抽样研究的目的是通过样本了解总体情况。如:抽样测量某地 40 名慢性四乙基铅中毒患者的脉搏,就可以应用统计学的原理与

方法,推断出所有该地四乙铅中毒患者的脉搏。

6. 系统误差与偶然误差的性质如何? 各举例说明。

系统误差与偶然误差的性质比较如下:

表 1.1 系统误差与偶然误差比较表

比较项目	偶然误差	系统误差
大小	一般较小	一般较大
方向	双向	单向
大小和方向重现性	不一定重现	可重现
产生原因	多种原因	少数确定的原因
可否消除	不可避免但可控制	消除原因,即可避免
统计规律性	无	有

例如:用水银柱式血压计测量血压时偶然误差与系统误差比较,如表 1.2。

表 1.2 血压计测量血压的两种误差比较

比较项目	偶然误差	系统误差
大小	很小	较大
方向	可高可低	偏高(或偏低)
大小和方向重现性	不一定重现	再三出现
产生原因	情绪波动、运动等	水银流失等
可否消除	不可	校准后避免
统计规律性	无	有

7. 为什么说偶然发生的误差不一定是偶然误差？试举例说明

因为偶然发生的误差有时具有系统误差的性质，且通过加大观察例数基本上可以消除，而偶然误差却无法消除。如在测量身高的过程中，偶然读错了一个人的数据，使得某一个值偏高了；但随着测量者的增多，总体均数却无多大变化，即偶然发生的误差消失了，而每个人的身高值与均数之差，尽管很小，却是无法消除的。

8. 研究误差的性质与统计学有何关系？

研究误差的性质有助于弄清误差产生的原因，进而进行消除或控制，使事物内部的规律性能更准确地被揭示出来，这正是统计学的要求。

第二章 计量资料频数分布的特征与规律

(讲授 7 学时, 实习 4 学时)

教学重点(一)

计量资料的频数分布	频数表的编制 频数分布图 频数分布类型
-----------	---------------------------

教学重点(二) 正态分布

概 念	变量服从于均数所在处峰值最高, 以均数为中心, 两侧逐渐下降并完全对称, 两端永远不与横轴相交的光滑的钟形曲线分布。
规 律	$\mu \pm 1.0\sigma$ 区间面积占正态曲线下面积的 68.27% $\mu \pm 1.96\sigma$ 区间面积占正态曲线下面积的 95.00% $\mu \pm 2.58\sigma$ 区间面积占正态曲线下面积的 99.00%
应 用	1. 估计频数分布情况; 2. 作正常值范围估计; 3. 作质量控制图。

教学重点(三): 集中趋势指标

意义	平均数: 平均水平、代表值。 1. 算术均数: 观察值的平均水平; 2. 几何均数: n 个观察值的连乘积开 n 次方之根; 3. 中位数(M): 观察值按大小顺序排列, 位居中间数。	
计	算术均数	$\bar{x} = \frac{\sum X}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$ (小样本, X 为观察值) $\bar{x} = \frac{\sum fX}{\sum f} = \frac{f_1X_1 + f_2X_2 + \dots + f_nX_n}{f_1 + f_2 + \dots + f_n}$ (大样本, X 为组中值) $\bar{x} = X_0 + \frac{\sum fx}{\sum f} \cdot i$ (简捷法, x 为缩减值)
算	几何均数	$G = \lg^{-1}\left(\frac{\sum \lg X}{n}\right)$ (小样本, X 为观察值) $G = \lg^{-1}\left(\frac{\sum f \lg X}{\sum f}\right)$ (大样本, X 为组中值) $G = \lg^{-1}\left(\lg G_0 + \frac{\sum fx}{\sum f} \cdot i\right)$ (简捷法, x 为缩减值)
	中位数	$M = L + \frac{i}{f_m} \left(\frac{n}{2} - \sum f_L \right)$ $[\text{百分位数: } P_x = L + \frac{i}{f_x} \left(\frac{nX}{100} - \sum f_L \right)]$

应用条件	算术均数:对称分布、正态分布。 几何均数:对数正态分布、等比数列资料。 中位数:①正负偏态资料;②开口资料;③频数分布性质不明资料。
------	---

教学重点(四) 离散程度指标

意义	标准差:各观察值与均数之差平方和的平均值的方根。 变异系数:离散系数、组间比较。
计算	标准差(s): $s = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}} \quad (\text{小样本, } X \text{ 为观察值})$ $s = i \sqrt{\frac{\sum fX^2 - \frac{(\sum fX)^2}{\sum f}}{\sum f - 1}} \quad (\text{大样本, 简捷法, } X \text{ 为组中值})$ 变异系数(CV): $CV = \frac{s}{\bar{x}} \times 100\%$
应用条件	标准差:对称(正态)分布; 变异系数:均数相差较大或计量单位不同,组(种类)间比较。

教学提示

1. 本章要在充分理解的基础上,一定要让学生亲自动手,动笔进行运算,否则眼高手低收不到预期效果,本章是统计学的基础之基础。
2. 必须学会频数分布表的编制;善于判断资料类型。
3. 必须弄清集中趋势和离散程度之内涵。
4. 掌握每个公式的适用条件,并弄清每个公式中各个符号所代表的意义。
5. 正态性检验为选学内容。

复习思考题解答

1. 比较几个常用集中趋势指标的意义与应用条件。

常用的集中趋势指标有算术均数(均数)、几何均数、中位数等,它们的意义和应用条件比较如下:

①意义:

算术均数:观察值的平均水平;

几何均数: n 个观察值的连乘积开 n 次方之根;

中位数:观察值按大小顺序排列,位居中间的那个数。若观察值为奇数,中间只有一个数即中位数;若观察值为偶数,中间有两个数,取此两个数的算术平均数作为中位数。

②应用条件:

算术均数:对称分布、正态或近似正态分布。

几何均数:对数正态分布、等比数列资料。

中位数:①正、负偏态资料;②开口资料;③频数分布性质不明的资料。

2. 比较几个常用离散程度指标的意义与应用条件。

常用离散程度指标有极差、方差、标准差和变异系数等。它们的意义和应用条件比较如下：

①意义：

极差：全距。最大值与最小值之差，资料信息利用不充分。

方差：离均差平方和的平均值。与总体有偏差。

标准差：各观察值与均数之差值平方和的平均值的平方根。

变异系数：离散系数，组间或不同计量单位比较。

②应用条件：

极差：只编制频数分布表。

方差：只用做方差分析。

标准差：对称分布、正态或近似正态分布；

变异系数：均数相差较大或计量单位不同组间比较。

3. 什么是正态分布？什么是标准正态分布？什么是对数正态分布？

正态分布：变量服从于均数所在处峰度最高，以均数为中心，两侧逐渐下降并完全对称，两端永远不与横轴相交的光滑的钟形正态曲线的分布。

标准正态分布：将变量值 X 与其均数的离差用其标准差为单位进行度量，转换，所得到的 u 值分布，即“标准正态分布”

对数正态分布：某些偏态分布资料，其观察值 X 经过对数变换后，可以转变为正态分布，称此资料服从对数正态分布。

4. 怎样应用正态曲线下面积分布的规律？

正态分布规律的应用：

①查阅标准正态曲线下的面积表，估计频数分布情况；

②估计正常值范围时，必须注意：a. 资料呈正态或近似正态分布（包括可变换为正态分布）；b. 样本含量足够大；c. 总体必须有明确的定义和范围；d. 观察仪器和方法必须统一。