

全国普通高等教育医学类系列教材

申 杰 主编

中医统计学

Statistics for Traditional Chinese Medicine



科学出版社

www.sciencep.com

全国普通高等教育医学类系列教材

中医统计学

申 杰 主 编

科 学 出 版 社

北 京

内 容 提 要

本书是由科学出版社牵头,并组织全国14所高等院校在统计学教学中具有丰富经验的一线教师编写的高等中医院校全日制本科生的统计学规划教材。

全书内容包括绪论、统计资料的搜集与整理、统计描述、参数估计、假设检验概述、 t 检验、 F 检验、 χ^2 检验、秩和检验、直线相关与回归、研究设计基础、统计表和统计图等共12章及统计用表、多元统计分析简介、统计学报告准则简介和SPSS应用简介4个附录,并且在正文前给出了希腊字母表和常用统计符号。

本书重点讲述概念和方法,力求通过例题和SPSS电脑实验,引导学生学习和运用统计学的基本理论、基本知识和基本技能,培养学生从事科学研究的能力,提高科研素质,为培养能够胜任中医科研工作及原创性研究的高素质人才奠定基础。

图书在版编目(CIP)数据

中医统计学/申杰主编. —北京:科学出版社,2009
全国普通高等教育医学类系列教材
ISBN 978-7-03-024358-4

I. 中… II. 申… III. 中国医药学—医学统计—中医院—教材 IV. R2-32

中国版本图书馆CIP数据核字(2009)第052028号

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

上海敬民实业有限公司长阳印刷厂印刷

科学出版社发行 各地新华书店经销

*

2009年5月第 一 版 开本:889×1194 1/16

2009年5月第一次印刷 印张:9½

印数:1—5 300 字数:317 000

定价:20.00元

全国普通高等教育医学类系列教材

《中医统计学》编委会名单

主 编 申 杰(河南中医学院)

副主编 石 晶(河北医科大学)
黄品贤(上海中医药大学)
张跃林(陕西中医学院)

编 委 (以姓氏笔画为序)
王晓波(辽宁中医药大学)
王淑媛(长春中医药大学)
史周华(山东中医药大学)
邢建民(北京中医药大学)
孙春阳(河南中医学院)
李国春(南京中医药大学)
余增丽(郑州大学公共卫生学院)
张雪飞(湖北中医学院)
陈晓凡(江西中医学院)
郑国华(福建中医学院)
莫传伟(广州中医药大学)
黄 杏(湖北中医学院)

前 言

我国著名教育家马寅初指出：“学者不能离开统计而研学，政治家不能离开统计而施政，事业家不能离开统计而执业，军事家不能离开统计而谋略。”由于诸多原因，既往有很多中医药重大科研课题因为实验设计的不严谨和统计方法的错误，严重影响了现代中医药事业的发展，使得广大中医科研工作者及在校学生越来越认识到，现代中医药事业的发展必须注重统计艺术和科学证据。因此，中医统计学由中医药科学研究的需要而产生，在中医药研究中的广泛使用而发展，并且必将推动中医药事业的科学发展。

全书以认识社会和自然界中随机现象的数量特征及其客观规律为主线，贯彻教材编写的思想性、科学性、先进性、启发性和适用性，强调教材的整体优化及编写的标准化、规范化，注重强化统计学的基础理论、基本知识和基本技能，突出继承性，紧扣教学需要、紧跟时代发展。在保持统计学知识体系的相对完整性的同时，联系中医药院校的实际，充实了例题和 SPSS 电脑实验的内容，以便于学生理论联系实践，为日后统计学的应用奠定基础。

本书的编写得到了河南中医学院、湖北中医学院、山东中医药大学、河北医科大学、上海中医药大学、陕西中医学院、广州中医药大学、南京中医药大学、北京中医药大学、长春中医药大学、辽宁中医药大学、郑州大学、福建中医学院、江西中医学院等院校领导的大力支持。科学出版社、湖北中医学院和上海中医药大学对教材编委会、定稿会给予了热情接待、周到安排。在本书编写过程中，借鉴和吸收了国内外有关文献和科研资料。在审稿期间，石晶教授、黄品贤教授付出了极其艰辛的劳动。在此，我谨代表编委会一并对他们的关心、支持、帮助和贡献表示最诚挚的敬意和谢意！

鉴于既往没有《中医统计学》规划教材,《中医统计学》的出版发行将成为全国高等中医药院校教材建设的一个里程碑。因此,特别感谢科学出版社对本书出版发行的鼎力支持!

限于我们的水平、能力和经验,书中难免存在许多不妥或错漏,恳请同行及广大师生、中医药工作者不吝赐教。

申 杰

2009年1月于郑州

shenjie007@126.com

希腊字母表

序 号	希 腊 字 母		英 文 注 音	国际音标注音	音 标
	大 写	小 写			
1	A	α	alpha	a:lf	[ˈælfə]
2	B	β	beta	bet	[bɪtə, ˈbeɪtə]
3	Γ	γ	gamma	gɑ:m	[ˈgæmə]
4	Δ	δ	delta	delt	[ˈdeltə]
5	E	ε	epsilon	epsilon	[epˈsaɪlən, ˈepsɪlən]
6	Z	ζ	zeta	zat	[ˈzɪtə]
7	H	η	eta	eita	[ˈɪtə, ˈeɪtə]
8	Θ	θ	theta	θit	[θɪtə]
9	I	ι	iota	aiot	[aiˈoutə]
10	K	κ	kappa	kap	[kæpə]
11	Λ	λ	lambda	lambd	[læmdə]
12	M	μ	mu	mju	[mjuː]
13	N	ν	nu	nju	[njuː]
14	Ξ	ξ	xi	ksi	[gzai, ksai, zai]
15	O	ο	omicron	omikron	[ouˈmaɪkrən]
16	Π	π	pi	pai	[pai]
17	P	ρ	rho	rou	[rou]
18	Σ	σ	sigma	sigma	[ˈsɪgmə]
19	T	τ	tau	tau	[təː]
20	Υ	υ	upsilon	jupsilon	[juːpˈsaɪlən, ˈjuːpsɪlən]
21	Φ	φ	phi	fai	[fai]
22	X	χ	chi	phai	[kai]
23	Ψ	ψ	psi	psai	[psai]
24	Ω	ω	omega	omiga	[ˈoumɪgə]

常用统计符号

参数(希腊字母)

统计符号	含 义
α	检验水准(显著性水准),第Ⅰ类错误,假阳性错误, 回归方程常数项
α'	校正检验水准
$1-\alpha$	可信度
β	总体回归系数,第Ⅱ类错误,假阴性率
$1-\beta$	检验效能(把握度)
δ	允许误差,差值
μ	总体均数
μ_d	差值的均数
μ_i	各样本均数所代表的总体均数
μ_0	某已知总体均数
μ_p	频率的总体均数
ν	自由度
ζ	随机变量
π	总体率
π_0	某已知的总体率
ρ	总体相关系数
Σ	求和
Σf	总例数
σ	总体标准差
σ^2	总体方差
σ_d	差值的总体标准差
σ_p	频率的标准误
$\sigma_{\bar{x}}$	均数的标准误
φ	标准正态分布曲线函数
χ^2	卡方检验统计量

统计量(拉丁字母)

统计符号	含 义
A	实际频数
a	组数,截距
b	回归系数,区组数
C	列数

CI	可信区间
CV	变异系数
C_X^2	二项系数
c	校正系数
D	Kolmogorov-Smirnov 检验统计量
d	差值
df	自由度
$\bar{d}$	差值的均数
e	自然对数的底
F	方差分析或方差齐性检验统计量
f	频数
f_x	第 x 位百分位数所在组段的频率
$fn(A)$	事件 A 发生的频率
G	几何均数
g	处理的不同水平
H	$K-W$ 检验统计量
H_c	校正的 H 统计量
H_0	无效假设
H_1	备择假设
i	组距,各种下标
k	组段数,比例基数,处理组数,样本率(构成比)的个数
L	下限,可信下限
L_n	正交表符号
l_{xx}	x 的离均差平方和
l_{xy}	x 与 y 的离均差积和
l_{yy}	y 的离均差平方和
M	中位数, Friedman 检验统计量
MS	均方
MS_e	误差的均方
$MS_{\text{回归}}$	回归均方
$MS_{\text{剩余}}$	剩余均方
$MS_{\text{组间}}$	组间均方
$MS_{\text{组内}}$	组内均方
n	样本量,配对设计资料的对子数
n_C	第 C 列的合计数
n_R	第 R 行的合计数
n_i	各组的例数
P	概率
p	样本频率,阳性率,合并的率
P_x	百分位数
Q_L	下四分位数间距
Q_U	上四分位数间距
Q_R	四分位数间距
q	阴性率, SNK 检验统计量
R	全距(极差),行数
R_k	样本秩和
$\bar{R}_A$	A 组的平均秩和

r	相关系数
r^2	确定系数
SE	标准误
SEM	样本均数的标准误
$SS_{\text{回归}}$	回归平方和
$SS_{\text{剩余}}$	剩余平方和(残差平方和)
$SS_{\text{总}}$	总差异
$SS_{\text{组间}}$	组间离均差平方和
$SS_{\text{组内}}$	组内离均差平方和
$SS_{\text{处理}}$	处理组间离均差平方和
$SS_{\text{区组}}$	区组间离均差平方和
$SS_{\text{误差}}$	误差的离均差平方和
s	标准差
s_b	样本回归系数的标准误
s_d	差值 d 的标准差
$s_{\bar{d}}$	差值 d 的标准误
s_i	各样本标准差
s_p	样本频率的标准误
s^2	样本方差
s_c^2	两样本联合方差
$s_{\bar{x}}$	均数的标准误
$s_{\bar{x}_1 - \bar{x}_2}$	两样本之差的联合标准误
s_{yx}	剩余标准差
T	理论频数, 秩和, 处理因素
T_{RC}	第 R 行、 C 列格子的理论数
t	t 检验统计量
t'	校正 t 检验统计量
t_r	相关系数 r 检验用 t 统计量
$t_{\alpha, \nu}$	单侧 t 界值
$t_{\alpha/2, \nu}$	双侧 t 界值
t_j	第 j 个相同秩次的个数
U	上限, 可信上限
u	标准正态变量, u 检验统计量
u_c	校正的 u 值
u_{α}	单侧 u 界值
$u_{\alpha/2}$	双侧 u 界值
W	Shapiror-Wilk 检验的统计量
$\bar{x}$	算术均数, 样本均数
$\bar{x}_i$	各处理组均数
x_{ij}	第 i 处理组的第 j 个测量值
x'	x 变量的转换值
$x_{\max}$	最大值
$x_{\min}$	最小值
$\hat{y}$	y 的估计值
Z	标准正态分布的检验统计量

目 录

前言
希腊字母表
常用统计符号

第一章 绪论

1

第一节 什么是统计学 /1

- 一、统计的涵义、统计学的发展简史及分类 /1
- 二、中医统计学的定义与研究要素 /1
- 三、统计学的基本思想、方法和特点 /2
- 四、统计学的主要内容 /2
- 五、统计学的基本概念 /3

第二节 为什么学习中医统计学 /4

- 一、顺应中医药学的发展趋势 /4
- 二、认识和处理医学现象的不确定性 /4
- 三、培养统计思维 /4
- 四、了解中医统计学的作用 /5
- 五、熟悉统计工作的基本步骤和研究过程 /5
- 六、辨明统计学与数学的区别与联系 /5
- 七、明确学习目标和学习要求 /6

第二章 中医统计资料的搜集与整理

7

第一节 中医统计资料的搜集 /7

- 一、中医统计资料的来源 /7

- 二、中医统计资料的类型 /8
- 三、搜集资料的基本要求 /8
- 四、搜集资料的注意事项 /9

第二节 中医统计资料的整理 /9

- 一、整理方案设计 /9
- 二、资料审核 /10
- 三、设计分组 /10

第三章 统计描述

13

第一节 频数分布 /13

- 一、频数分布的特征与类型 /13
- 二、频数表和频数图 /13

第二节 数值变量的统计描述 /15

- 一、集中趋势描述 /15
- 二、离散趋势描述 /15

第三节 正态分布及其应用 /16

- 一、正态分布的图形及其概率密度函数 /16
- 二、正态分布的特征和正态曲线下的面积分布规律 /17
- 三、正态分布的应用 /17

第四节 分类变量的统计描述 /19

- 一、常用的相对数指标 /19
- 二、率的标准化 /20

第五节 例题和 SPSS 电脑实验 /21

第四章 参数估计

23

第一节 抽样分布与抽样误差 /23

- 一、样本均数的抽样分布与标准误 /23
- 二、 t 分布及其应用 /24
- 三、样本率的抽样分布与标准误 /24
- 四、二项分布及其应用 /25
- 五、泊松分布及其应用 /27

第二节 总体均数的估计 /29

- 一、总体均数的点估计 /29
- 二、总体均数的区间估计 /29

第三节 总体率的估计 /31

- 一、总体率的点估计 /31
- 二、总体率的区间估计 /31

第四节 例题和 SPSS 电脑实验 /31

第五章 假设检验概述

33

第一节 假设检验的分类、思维方法与步骤 /33

- 一、假设检验的分类 /33

- 二、假设检验的思维方法 /33
- 三、假设检验的步骤 /34
- 第二节 假设检验的两类错误和注意事项 /35
 - 一、I 型错误和 II 型错误 /35
 - 二、应用假设检验的注意事项 /36
- 第三节 正态性检验与数据转换 /38
 - 一、正态性检验 /38
 - 二、数据转换 /38
- 第四节 例题和 SPSS 电脑实验 /39

第六章 t 检验

41

- 第一节 单样本 t 检验 /41
 - 一、单样本的正态性检验 /41
 - 二、单样本 t 检验 /41
- 第二节 配对设计资料均数的 t 检验 /42
 - 一、配对设计数值变量的正态性检验 /42
 - 二、配对设计数值变量 t 检验 /43
- 第三节 独立样本的方差齐性检验与 t 检验 /43
 - 一、独立样本的方差齐性检验 /43
 - 二、独立样本 t 检验 /44
- 第四节 例题和 SPSS 电脑实验 /44

第七章 F 检验

49

- 第一节 F 检验概述 /49
 - 一、因素与水平 /49
 - 二、对多样本均数重复进行 t 检验的风险性 /49
 - 三、 F 检验的基本思想 /49
 - 四、 F 检验的类型和应用条件 /50
 - 五、多个样本方差的齐性检验 /51
- 第二节 完全随机设计和随机区组设计资料的 F 检验 /51
 - 一、完全随机设计资料的 F 检验 /51
 - 二、随机区组设计资料的 F 检验 /52
 - 三、多个样本均数的多重比较 /54
- 第三节 重复测量资料的 F 检验 /55
 - 一、离均差平方和与自由度的分解 /55
 - 二、重复测量资料方差分析的基本步骤 /56
- 第四节 例题和 SPSS 电脑实验 /56

第八章 χ^2 检验

59

- 第一节 χ^2 检验的基本思想 /59
 - 一、相关概念 /59
 - 二、 χ^2 检验的基本公式 /59

三、 χ^2 检验的基本思想	/60
第二节 2×2 表资料的 χ^2 检验	/60
一、成组设计 2×2 表资料的 χ^2 检验	/60
二、配对设计 2×2 表资料的 χ^2 检验	/61
第三节 $R \times C$ 表资料的 χ^2 检验	/62
一、多个样本率比较的 $R \times 2$ 表资料的 χ^2 检验	/62
二、多个构成比比较的 $2 \times C$ 表资料的 χ^2 检验	/63
三、双向无序分类资料关联性的 $R \times C$ 表资料的 χ^2 检验	/63
四、 $R \times C$ 表资料的 χ^2 检验的注意事项	/64
第四节 多个样本率或构成比间的多重比较	/64
一、 χ^2 分割法	/64
二、基于 χ^2 分割原理的多重比较	/66
第五节 例题和 SPSS 电脑实验	/67

第九章 秩和检验

70

第一节 配对设计资料的符号秩和检验	/70
一、基本思想	/70
二、检验步骤	/70
第二节 完全随机设计资料的两样本比较的秩和检验	/71
一、两组数值变量资料的秩和检验	/71
二、两组有序分类变量资料的秩和检验	/72
第三节 完全随机设计资料的多个样本比较的秩和检验	/73
一、多组有序分类变量资料的秩和检验	/73
二、多组数值变量资料的秩和检验	/74
第四节 随机区组设计资料的秩和检验	/74
一、基本思想	/75
二、检验步骤	/75
第五节 多个样本两两比较的秩和检验	/75
一、完全随机设计资料的多个样本两两比较	/75
二、随机区组设计资料的两两比较	/76
第六节 例题和 SPSS 电脑实验	/77

第十章 直线相关与回归

80

第一节 直线相关	/80
一、相关系数的意义与计算	/80
二、相关系数的假设检验	/81
三、注意事项	/82
第二节 直线回归	/82
一、直线回归分析的步骤	/82
二、线性回归方程的主要用途	/83
三、直线回归方程的假设检验	/83
四、直线相关与回归的区别与联系	/85
第三节 例题和 SPSS 电脑实验	/85

第十一章 研究设计基础

87

第一节 专业设计的基本要素 /87

- 一、受试对象 /87
- 二、处理因素 /87
- 三、实验效应 /88

第二节 统计设计的基本原则 /88

- 一、随机 /88
- 二、对照 /90
- 三、均衡 /91
- 四、重复 /91

第三节 样本量的估算 /91

- 一、样本量估算的主要参数 /91
- 二、样本均数与总体均数比较的样本量的估算 /91
- 三、两样本均数比较的样本量的估算 /92
- 四、配对设计和交叉设计数值变量资料的样本量的估算 /92
- 五、随机区组设计的样本量的估算 /92
- 六、样本率与总体率比较的样本量的估算 /93
- 七、两样本率比较的样本量的估算 /93
- 八、配对计数资料的样本量的估算 /93
- 九、多个样本均数比较的样本量的估算 /94
- 十、多个率样本比较的样本量的估算 /94
- 十一、估计总体均数的样本量的估算 /95
- 十二、估计总体率的样本量的估算 /95
- 十三、直线相关分析的样本量的估算 /95
- 十四、动物实验和临床试验的基本样本量 /96

第四节 常用实验设计方案 /96

- 一、完全随机设计 /96
- 二、配对设计 /96
- 三、交叉设计 /97
- 四、随机区组设计 /98
- 五、拉丁方设计 /98
- 六、析因设计 /98
- 七、正交设计 /99

第五节 调查设计概述 /101

- 一、调查设计的目的、类型与方法 /101
- 二、调查方案设计的基本原则 /102
- 三、调查方案的主要内容 /102
- 四、调查指标的设计 /102
- 五、调查方案的可行性研究 /103
- 六、调查问卷设计 /103
- 七、量表的应用 /103

第十二章 统计表和统计图

105

第一节 常用统计表 /105

一、统计表的基本格式与种类 /105	
二、统计表的制表原则与结构 /106	
第二节 常用统计图 /106	
一、统计图的种类与绘制要求 /106	
二、常用统计图的适用条件与绘制 /106	
第三节 例题和 SPSS 电脑实验 /109	
附录 A 统计用表 /112	
附表 1 标准正态分布曲线下的左侧面积 $\varphi(u)$ 值 /112	
附表 2 t 分布界值表 /113	
附表 3 百分率的 95% 可信区间 /114	
附表 4 F 界值表(方差齐性检验用, 双侧界值) /115	
附表 5 F 界值表(方差分析用) /116	
附表 6 χ^2 分布界值表 /117	
附表 7 T 界值表(配对比较的符号秩和检验用) /118	
附表 8 T 界值表(两样本比较的秩和检验用) /119	
附表 9 H 界值表(三样本比较的秩和检验用) /120	
附表 10 M 界值表(随机区组比较的秩和检验用) /121	
附表 11 q 界值表 /122	
附表 12 r 界值表 /123	
附表 13 随机数字表 /124	
附表 14 随机排列表($n=20$) /125	
附录 B 多元统计分析简介 /126	
附录 C 统计学报告准则简介 /130	
附录 D SPSS 简介 /133	
主要参考文献 /135	

导 学

1. 掌握中医统计学的主要内容和步骤。
2. 熟悉统计学的基本概念;统计的涵义、分类;统计学的基本思想、方法和特点。
3. 了解中医统计学的定义;中医统计学的作用;学习目标和学习要求。

世界上最伟大的统计学家之一 C. R. 劳先生在统计学哲理论著《统计与真理——怎样运用偶然性》中指出:“在终极的分析中,一切知识都是历史;在抽象的意义下,一切科学都是数学;在理性的基础上,所有的判断都是统计学。”因此,我们在对中医药学进行努力挖掘、加以提升的同时,应该注重统计学的学习与运用。

第一节 什么是统计学

统计学是处理数据中变异性(homogeneity)和复杂性(complicacy)的科学和艺术。作为一门科学,统计学必须如实反映现状,无论从实施统计方法或进行科学研究的角度,目的都在于得到真实的数据;作为一门艺术,统计学必须在概率论和数理统计理论的指导下,正确运用统计学思维,针对数据特点,选用适宜的统计分析方法,让科研工作变得井然有序,让科学的发展更缜密、更完美、更实用。

一、统计的涵义、统计学的发展简史及分类

1. 统计的涵义 统计(statistics)一词有三种内涵:①统计工作(statistical work),指搜集、整理、分析和研究统计数据的工作过程,是统计数据与统计理论的基础和源泉。②统计数据(statistical data),指统计工作研究的主体及成果。③统计学(statistics),是对研究对象的数据进行搜集、整理、分析和研究,以展示其总体特征和规律性的方法论科学。

2. 统计学的发展简史 人类最初利用手指、石子、贝壳、小木棍以及绳索等工具进行的计数活动中就蕴藏着统计学的种子。但是,由统计实践上升到统计学,却只有 300 多年的历史。按统计方法及特征的历史演变顺序,统计学的发展史可分为古典统计学的萌芽时期(17 世纪 70 年代至 19 世纪初期)、近代统计学的形成时期(19 世纪初至 20 世纪初)和现代统计学的发展时期(20 世纪初至今)3 个阶段。正是由于古典统计学时期的政治算术学派和国势派以及近代统计学时期的数理统计学派和社会统计学派之间的相互争论、相互渗透,使数理统计学与社会统计学最后融合成为统一的现代统计学。

统计学的发展有 4 个明显趋势:①随着数学的发展,统计学依赖和吸收的数学方法越来越多。②统计方法与计算机技术相结合,已渗透到了科学的所有学科部门,使得以统计学为基础的边缘学科不断形成。③统计学与实质性学科、统计软件、现代信息相结合,所发挥的功效日益增强。④统计学的作用与功能已从描述事物现状、反映事物规律,朝着抽样推断、预测未来变化的方向发展,已成为具有方法论性质的综合性学科。

3. 统计学的分类

(1) 理论统计学(theoretical statistics):即数理统计学(mathematical statistics),主要探讨统计学的数学原理和统计公式的来源。它把研究对象一般化、抽象化,以概率论为基础,从纯理论的角度,对统计方法加以推导论证,中心的内容是统计推断问题,实质是以归纳方法研究随机变量的一般规律。

(2) 应用统计学(applied statistics):是由于统计方法在自然科学和社会科学领域的应用而产生的学科。例如,在生物学中的应用形成了生物统计学(biostatistics);在医学中的应用形成医学统计学(medicinal statistics)、卫生统计学(hygeian statistics)和中医统计学等不同的分支学科。

二、中医统计学的定义与研究要素

1. 中医统计学的定义 中医统计学是处理中医理论与实践中的同质性和变异性的科学与艺术。它以概率论和