

全国高等医药院校实验教材

预防医学 实习和学习指导

供医学及相关专业使用

主编 张青碧 叶运莉 雷章恒

清华大学出版社

中国医药出版社

辅助医学

实践与学习指南

（第二版）

主编 王树强 副主编 王树强 王树强

中国医药出版社

• 全国高等医药院校实验教材 •

预防医学 实习和学习指导

主编 张青碧 叶运莉 雷章恒

清华大学出版社
北京

内 容 简 介

《预防医学实习和学习指导》分为实习指导和复习思考题两篇。实习篇分为5个部分,即医学统计学、流行病学、卫生学、社会医学和卫生事业管理、扩展性知识,每部分又分节编写。复习思考题篇包含大量习题。本书以内容覆盖全面、层次分明、通俗易懂、简明实用,强化校内外结合、理论与实践结合、基本知识与能力培养结合为特色。

本书可供临床、口腔、麻醉、中西医结合、中医、护理、卫生事业管理、儿科、检验等专业使用,也可作为预防医学专业教学教师参考书和学生学习复习资料。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

预防医学实习和学习指导/张青碧,叶运莉,雷章恒主编. —北京:清华大学出版社,2013.3

(全国高等医药院校实验教材)

ISBN 978-7-302-31568-1

I. ①预… II. ①张… ②叶… ③雷… III. ①预防医学—医学院校—教学参考资料 IV. ①R1

中国版本图书馆CIP数据核字(2013)第030720号

责任编辑:罗健 王华

封面设计:戴国印

责任校对:王淑云

责任印制:杨艳

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印装者:北京嘉实印刷有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:15.5 字 数:364千字

版 次:2013年3月第1版 印 次:2013年3月第1次印刷

印 数:1~4000

定 价:29.80元

产品编号:049073-01

《预防医学实习和学习指导》

编审委员会名单

主 审 贾 红

主 编 张青碧 叶运莉 雷章恒

副主编 甘仲霖 李爱玲 刘军祥 陈 润

编 委 (按姓氏拼音排序)

陈 润	泸州医学院	甘仲霖	泸州医学院
韩知峡	泸州医学院	黄 宵	泸州医学院
雷章恒	泸州医学院	李爱玲	泸州医学院
李 红	遵义医学院	李 祥	泸州医学院
梁海荣	广东医学院	刘 杰	泸州医学院
刘军祥	泸州医学院	刘林华	广东医学院
刘 娅	泸州医学院	潘池梅	川北医学院
绕朝龙	成都中医药大学	苏红卫	泸州医学院
孙 雪	泸州医学院	汤 艳	泸州医学院
唐焕文	广东医学院	汪春梅	泸州医学院
王 琼	泸州医学院	谢惠波	泸州医学院
熊 伟	泸州医学院	杨 超	泸州医学院
杨 慧	广东医学院	杨 艳	泸州医学院
叶运莉	泸州医学院	张春莲	泸州医学院
张俊辉	泸州医学院	张青碧	泸州医学院
张学峰	川北医学院	周远忠	遵义医学院

Preface 前言

预防医学是一门跨学科、综合性和社会实践性很强的课程，被国家教育部确定为医学院校 15 门主干课程之一，是培养与现代医学模式相适应的 21 世纪新型医师的重要课程，也是我校临床、口腔、麻醉、护理、中西医结合（皮肤、骨伤、耳鼻喉）、心血管、医学心理学、儿科、病理诊断等专业本科生和专科生等多个层次学生的公共基础课，属于上述专业学生的必修考试课。目前我校将该课程分为卫生学、医学统计学、流行病学、社会医学和卫生事业管理、循证医学 5 个部分，课程总学时为 144 学时，其中实验学时为 40 学时左右，目前我校学生使用的主要是分部分自编的内部教材，没有预防医学综合实验教材。预防医学是临床执业医师考试的必考内容，所占比重不轻，针对我校近几年临床学生执业医师考试预防医学部分丢分比较严重的情况，同时为了加强预防医学精品课程的建设，拟编写一本适合我校学生具体情况和培养目标的《预防医学实习和学习指导》。

本书的特色是紧扣临床医学生执业医师考试大纲和时代发展要求，将预防医学的几个部分有机结合，突出预防为主的理念、大卫生观念和可持续发展观念；通过案例分析和现场卫生学调查，强调在临床环境下的三级预防措施策略；既培养学生的基本实验技能，更着眼于综合能力和综合素质的培养，以提高学生应对突发公共卫生事件的能力。

本书是针对各类医学院校的临床医学专业、口腔专业、麻醉专业、检验专业、影像专业、中医专业、中西医结合专业等非预防医学专业本科生的预防医学教学需要而编写的，包括实验技术和实验方法，既包括该门课程教学的基本技能和基本知识，也包括创新性能力和综合能力的培养。并且针对预防医学教学大纲和临床执业医师预防医学部分考试大纲编写了各种复习思考题，有利于提高医学生的整体素质。

本书分为医学统计学、流行病学、卫生学、循证医学、社会医学和卫生事业管理五方面内容的实习和复习指导，另附有 2012 年临床执业医师预防医学考试大纲和各部分复习题的参考答案。

由于编写时间较紧、编写经验不足等，难免存在疏漏之处，敬请读者谅解。

本书的编写得到泸州医学院分管教学领导、教务处以及公共卫生学院各位领导和全体教师的大力支持，在此一并致谢！

目 录

Contents

第 1 篇 实 习

第 1 部分 医学统计学..... 3

- 实习 1 定量资料的统计描述 3
- 实习 2 定量资料的统计推断 8
- 实习 3 方差分析 12
- 实习 4 分类资料的统计分析 15
- 实习 5 秩和检验 18
- 实习 6 直线相关与回归分析 21
- 实习 7 统计图表 22

第 2 部分 流行病学..... 27

- 实习 8 疾病的分布 27
- 实习 9 现况调查 29
- 实习 10 病例对照研究 31
- 实习 11 队列研究 34
- 实习 12 筛检 36
- 实习 13 爆发调查 37

第 3 部分 卫生学..... 44

- 实习 14 大气的采样方法 44
- 实习 15 大气中二氧化硫的测定 46
- 实习 16 漂白粉中有效氯含量的测定 49
- 实习 17 饮用水的消毒及卫生评价 51
- 实习 18 环境砷污染案例分析 54
- 实习 19 全血胆碱酯酶活性的
分光光度测定方法 58

实习 20 尘肺案例分析 61

实习 21 金属类中毒案例分析 61

实习 22 窒息性气体中毒案例分析 62

实习 23 尿中汞的冷原子吸收
光谱测定方法 65

实习 24 鲜奶的卫生检验 67

实习 25 膳食调查及营养素计算
与评价 70

实习 26 食谱编制 72

实习 27 食物中毒案例讨论 78

实习 28 食品中亚硝酸盐(硝酸盐)
含量测定 80

实习 29 油脂中酸价、过氧化值
测定 82

第 4 部分 社会医学和卫生事业管理..... 84

实习 30 问卷设计与实施 84

实习 31 健康危险因素评价 85

实习 32 生命素质指数的计算与分析 89

实习 33 卫生服务需求调查 90

实习 34 社区居民常见慢性病
干预研究设计 96

第 5 部分 扩展性知识..... 97

知识 1 突发公共卫生事件的预防与
控制 97

知识 2	自然灾害后预防流行性 疾病工作优先顺序	101
知识 3	自然灾害中的疾病预防 控制问题与技术	103
知识 4	自然灾害后疫情监测与 防病基本措施	106

第 2 篇 复习思考题

第 1 部分 医学统计学	113
第 1 章 绪论	113
第 2 章 定量资料的统计描述	114
第 3 章 总体均数的估计和 假设检验	116
第 4 章 方差分析	118
第 5 章 分类资料的统计描述	120
第 6 章 分类资料的统计推断	122
第 7 章 秩和检验	123
第 8 章 直线相关与回归分析	124
第 9 章 统计图表	126
第 2 部分 流行病学	127
第 10 章 绪论	127
第 11 章 疾病的分布	130
第 12 章 现况调查	132
第 13 章 病例对照研究	134
第 14 章 队列研究	136
第 15 章 实验流行病学	137
第 16 章 偏倚控制及病因推断	138
第 17 章 筛检和诊断试验	141
第 18 章 传染性疾病的预防 与控制	143
第 19 章 慢性非传染性疾病的 预防与控制	150

第 3 部分 卫生学	161
第 20 章 绪论	161
第 21 章 环境卫生学	164
第 22 章 职业卫生与职业医学	183
第 23 章 营养卫生学	192
第 24 章 突发公共卫生事件 及其应急策略	201
第 4 部分 社会医学和卫生事业管理	209
第 25 章 循证医学	209
第 26 章 公共卫生监测	212
第 27 章 临床预防服务	213
第 28 章 健康相关行为干预	215
第 29 章 吸烟的控制	219
第 30 章 体力活动促进	220
第 31 章 人群健康与社区预防 服务	221
第 32 章 医院安全管理	223
第 33 章 卫生系统及其功能	226
第 34 章 医疗保险	228
第 35 章 卫生政策与卫生资源配置	229
第 36 章 全球卫生保健策略	231
附录 1 参考答案	233
第 1 部分 医学统计学	233
第 2 部分 流行病学	233
第 3 部分 卫生学	234
第 4 部分 社会医学和卫生事业 管理	235
附录 2 2012 年临床执业医师预防医学 考试大纲	237

○ 实 习

第1篇

实习1 定量资料的统计描述

【实习目的】

1. 掌握定量资料集中趋势、离散趋势各指标的意义、计算方法和用途。
2. 掌握正态分布的概念、特征、应用及正态曲线下的面积分布规律。
3. 标准正态分布的概念和标准化变换。
4. 熟悉频数表的编制、特征以及医学参考值范围的制定。

【知识点】

(一) 定量资料的频数分布

定量资料又称为计量资料，它是测量每个观察单位某项指标值的大小所得的资料，一般有计量单位。描述定量资料分布规律的统计方法有两种：一是统计图表，主要是频数表；二是选用适当的统计指标。

1. 频数表的编制 变量的取值范围可划分为几个区间，每个区间称作一个组段，将各组段与相应的频数列表，即为频数表。组段之间的距离称为组距，一般为等距。对于离散数据，每一个观察值即对应一个频数，如某医院某年度一天内死亡0, 1, 2, …, 20个患者的天数。如描述某学校学生性别的分布情况，男、女生的人数即为各自的频数。对于散布区间很大的离散数据和连续型数据，数据散布区间由若干组段组成，每个组段对应一个频数。制作连续型数据频数表一般步骤如下所述。

(1) 求全距：全距又称为极差，指全部观察值中最大值与最小值之差，用符号 R 表示：

$$R = X_{\max} - X_{\min}。$$

(2) 根据极差划分“组段”数（通常8~10个）：确定组段和组距。每个组段都有下限 L 和上限 U ，数据 x 归组统一定为 $L \leq x < U$ 。

(3) 统计各组段频数：统计出各组段相应的观察单位个数（频数），将各组段与相应的频数列表即得到频数表。

2. 频数分布的特征 频数分布有两个重要特征：集中趋势和离散趋势。集中趋势指一组

数据向某一个位置聚集或集中的倾向；离散趋势指一组数据的分散程度或变异度。

3. 频数分布的类型 频数分布可分为对称分布和偏态分布两种类型。对称分布指集中位置在中间，左右两侧的频数基本对称。偏态分布指频数分布不对称，集中位置偏向一侧：若偏向数值较小的一侧，称为正偏态；若偏向数值较大的一侧，则称为负偏态。定量资料的频数分布类型不同，描述其集中趋势和离散程度的指标也不同。

4. 频数表的用途 频数表可用于揭示资料的分布特征和分布类型，便于发现某些特大或特小的可疑值，也便于进一步计算指标和统计分析处理。

(二) 集中趋势的描述

统计指标可从数量上较准确地描述数据分布的集中趋势和离散程度。描述定量资料集中趋势的指标统称为平均数，常用的平均数有均数、几何均数及中位数。

1. 均数 均数是算术均数的简称，总体均数用 μ 表示，样本均数用 \bar{X} 表示。均数适用于描述单峰对称分布，特别是正态分布或近似正态分布资料的集中趋势。

2. 几何均数 几何均数适用于描述原始观察值呈偏态分布，但经过对数变换后呈正态分布或近似正态分布的资料，如血清抗体滴度等。

3. 中位数 中位数是指将一组观察值按从小到大的顺序排列后位次居中的观察值。中位数适用于描述各种分布的资料，实际工作中常用来描述偏态分布资料、一端或两端无确切值或分布不明确资料的集中趋势。

上述 3 种常用平均数的关系如下：对于正态分布资料，中位数等于均数；对于对数正态分布资料，中位数等于几何均数；对于正偏态分布资料，中位数小于均数；对于负偏态分布资料，中位数大于均数。3 种常用平均数的意义及其用途见表 1-1。

表 1-1 常用平均数的意义及其用途

平均数	意义	用途
均数	平均数量水平	应用甚广，最适用于对称分布，特别是正态分布
几何均数	平均增(减)倍数	等比资料；对数正态分布
中位数	位次居中的观察值水平	偏态分布；分布不明；分布末端无确定值

(三) 离散程度的描述

描述定量资料离散程度的常用指标有：极差、四分位数间距、方差、标准差和变异系数。

1. 极差 (range) 即最大值与最小值之差： $R = X_{\max} - X_{\min}$ ，用于资料的粗略分析，其计算简便但稳定性较差。样本含量越大，抽到最大值和最小值的可能性较大，导致极差也较大，因此当样本含量相差较大时，不宜用极差来描述其离散程度。

2. 四分位数间距

(1) 百分位数：指将观察值从小到大排列后处于第 x 百分位置上的数值，用符号表示为 P_x 。一个百分位数将全部数据分成两部分，有 $x\%$ 的数据小于 P_x ，有 $(100-x)\%$ 的数据大于 P_x 。百分位数的一个重要用途是确定医学正常参考值范围。

(2) 四分位数间距 (QR): 指第 75 百分位数与第 25 百分位数之差, 即 $P_{75}-P_{25}$, 反映了一组数据从小到大排列后中间一半观察值的变动范围。常与中位数一起使用, 描述偏态分布资料、一端或两端无确切值或分布不明确资料的离散程度, 比极差稳定。四分位数间距越大, 说明资料的离散程度越大。

3. 方差 方差是描述对称分布, 特别是正态分布或近似正态分布资料的离散程度的常用指标。其值越大, 说明观察值变异程度越大。总体方差用 σ^2 表示, 样本方差用 S^2 表示。在实际工作中总体方差往往未知, 常用样本方差来估计。

4. 标准差 方差开算术平方根即得到标准差 (standard deviation), 使用的量纲与原量纲相同, 是描述对称分布, 特别是正态分布或近似正态分布资料离散程度的最常用指标, 适用于近似正态分布的资料, 大样本、小样本均可, 最为常用。标准差越大, 说明资料的变异程度越大。总体标准差用 σ 表示, 样本标准差用 S 表示。

5. 变异系数 变异系数 (coefficient of variation), 用 CV 表示, 为标准差与均数之比, 公式为: $CV = \frac{S}{\bar{X}} \times 100\%$ 。极差、四分位数间距和标准差都有单位, 其单位与观察值的单位相同; 而变异系数为相对数, 没有单位, 更便于资料间变异程度的比较。变异系数主要用于下列两种情况:

- (1) 比较度量衡单位不同的几组资料的变异程度。
- (2) 比较均数相差悬殊的几组资料的变异程度。

平均指标和变异指标分别反映资料的不同特征, 作为资料的总结性统计量, 两类指标要求一起使用, 如常用 $\bar{X} \pm S$ 或 $M(QR)$ 。

(四) 正态分布

1. 正态分布的概念和特征

(1) 正态分布: 在医学卫生领域中, 有许多变量为连续型随机变量, 如身高、体重、血压等, 这些变量的频数分布特点是中间频数多, 两边频数少, 且左右对称, 其频数分布规律往往可用正态分布来描述。若某指标 X 服从正态分布, 记为 $X \sim N(\mu, \sigma^2)$ 。正态曲线呈钟形, 两头低, 中间高, 左右对称, 曲线与横轴间的面积总等于 1。

(2) 正态分布的特征

- 1) 正态密度函数曲线在横轴上方均数处最高。
- 2) 正态分布以均数为中心, 左右对称。
- 3) 正态分布有两个参数, 即位置参数 μ 和形态参数 σ 。不同的 μ 和 σ 对应于不同的正态分布。若固定 σ , 改变 μ 值, 曲线就会沿着 X 轴平行移动, 其形态不变。若固定 μ , σ 越小, 曲线越陡峭; 反之, σ 越大, 曲线越低平, 但中心在 X 轴的位置不变。
- 4) 正态分布曲线的面积分布有一定规律。

2. 正态分布曲线下的面积分布规律 无论 μ 、 σ 取什么值, 正态分布曲线与横轴间的面积恒等于 1 或 100%; 且其对称轴为直线 $X=\mu$, $X>\mu$ 与 $X<\mu$ 范围内曲线下的面积相等, 各占 50%; 在 $(\mu-\sigma, \mu+\sigma)$ 范围内曲线下面积为 68.27%, 在 $(\mu-1.645\sigma, \mu+1.645\sigma)$ 范围内曲线下



面积为 90%，在 $(\mu-1.96\sigma, \mu+1.96\sigma)$ 范围内曲线下面积为 95%，在 $(\mu-2.58\sigma, \mu+2.58\sigma)$ 范围内曲线下面积为 99%。

3. 标准正态分布

(1) 标准正态分布是一种特殊的正态分布，标准正态分布的均数为 0，标准差为 1，通常用 u (或 Z) 表示服从标准正态分布的变量，记为 $Z \sim N(0, 1^2)$ 。

(2) 标准化变换： $Z = \frac{X-\mu}{\sigma}$ ，此变换有特性，任何正态分布 $X \sim N(\mu, \sigma^2)$ 经过 Z 变换均服从标准正态分布，故 Z 变换又被称为标准化变换。

4. 正态分布的应用

(1) 估计频数分布。

(2) 制定参考值范围：医学参考值范围指绝大多数“正常人”的某项解剖、生理、生化指标的波动范围。这里的“绝大多数”可以是 90%、95% 或 99% 等，最常用的是 95%。所谓“正常人”不是指完全健康的人，而是指排除了对所研究指标有影响的疾病和有关因素的同质人群。

对于一个指标，随机抽取一个含量足够大的样本后，可参照表 1-2 采用正态分布法或百分位数法制定其医学参考值范围。

(1) 正态分布法：适用于服从正态（或近似正态）分布的指标及可以通过转换服从正态分布的指标。

(2) 百分位数法：适用于偏态分布的指标。

表 1-2 医学参考值范围的制定方法

参考值范围(%)	正态分布法			百分位数法		
	双侧	单侧		双侧	单侧	
		只有下限	只有上限		只有下限	只有上限
95	$\bar{X} \pm 1.96S$	$\bar{X} - 1.645S$	$\bar{X} + 1.645S$	$P_{2.5} \sim P_{97.5}$	P_5	P_{95}
99	$\bar{X} \pm 2.58S$	$\bar{X} - 2.326S$	$\bar{X} + 2.326S$	$P_{0.5} \sim P_{99.5}$	P_1	P_{99}

【习题】

1. 某研究者测定了 176 例燃煤型砷中毒患者的尿总砷含量 ($\mu\text{g/L}$)，资料如下：

0.0169	0.0262	0.3433	0.0505	0.2266	0.1690	0.0165	0.0356	0.0968	0.1628
0.0904	0.1059	0.0582	0.0211	0.0867	0.0318	0.0256	0.0267	0.1592	0.1364
0.0583	0.0275	0.2285	0.0246	0.0508	0.1076	0.0195	0.0400	0.0646	0.1109
0.0212	0.0164	0.1401	0.0646	0.0139	0.0377	0.0161	0.0121	0.0617	0.2686
0.0532	0.0724	0.1280	0.0143	0.0980	0.5678	0.0228	0.1279	0.0872	0.0675
0.0361	0.0680	0.0591	0.0821	0.1418	0.1051	0.0662	0.1033	0.1188	0.0887
0.0102	0.0154	0.1775	0.0223	0.0319	0.0986	0.1019	0.0419	0.0678	0.0347
0.0753	0.0532	0.0151	0.0219	0.1139	0.1124	0.0524	0.0290	0.0376	0.1510

0.1250	0.0339	0.0549	0.0974	0.0753	0.2902	0.0222	0.0204	0.1325	0.0462
0.3047	0.0464	0.1486	0.0271	0.3953	0.0288	0.1520	0.0559	0.1244	0.1264
0.0576	0.0112	0.0222	0.4085	0.1128	0.0463	0.1240	0.0226	0.0809	0.0371
0.0183	0.1430	0.0559	0.0353	0.1333	0.2383	0.0929	0.0209	0.2748	0.0189
0.4542	0.0782	0.0741	0.1460	0.1317	0.0456	0.0499	0.0317	0.0863	0.0505
0.2691	0.3570	0.0227	0.0392	0.0406	0.0596	0.0260	0.0906	0.1516	0.0695
0.0723	0.0389	0.0810	0.2326	0.0311	0.0174	0.0868	0.0516	0.0970	
0.0372	0.0126	0.0678	0.2133	0.5265	0.4385	0.0357	0.3706	0.0621	
0.2330	0.0947	0.1591	0.0636	0.1845	0.0445	0.0430	0.0236	0.0429	
0.0134	0.5805	0.0600	0.0387	0.0392	0.0747	0.0470	0.0425	0.2218	

(1) 绘制频数分布图，并简述分布类型和分布特征。

(2) 计算适当的集中趋势指标。

2. 抽样调查某市 45 ~ 55 岁健康男性居民的血脂水平，184 名 45 ~ 55 岁健康男性居民的血清总胆固醇 (TC) 的 $\bar{X}=4.84\text{mmol/L}$, $S=0.96\text{mmol/L}$ 。已知健康人的血清总胆固醇服从正态分布，请完成下列计算：

(1) 估计该市 45 ~ 55 岁健康男性居民的血清总胆固醇的 95% 参考值范围。

(2) 估计该市 45 ~ 55 岁健康男性居民中，血清总胆固醇在 3.25 ~ 5.25mmol/L 范围内的比例。

(3) 估计该市 45 ~ 55 岁健康男性居民中，血清总胆固醇低于 3.80mmol/L 所占的比例。

3. 测得某地 300 例正常人尿汞值，其频数表见表 1-3。请计算均数、中位数，并回答何者代表性更好？

表 1-3 300 例正常人尿汞值 (μg/L) 的频数表

尿 汞 值	例 数	尿 汞 值	例 数
0 ~	49	36 ~	4
4 ~	27	40 ~	5
8 ~	58	44 ~	—
12 ~	50	48 ~	3
16 ~	45	52 ~	—
20 ~	22	56 ~	2
24 ~	16	60 ~	—
28 ~	9	64 ~	—
32 ~	9	68 ~	1

4. 表 1-4 为 10 例垂体催乳素微腺瘤经蝶手术前后的血催乳素浓度，请分别求出术前、术后的均数、标准差及变异系数。比较手术前后数据的变异情况应该采用何指标？能否说明手术前数据的变异大？为什么？

表 1-4 手术前后患者血催乳素浓度 (mg/ml)

例号	血催乳素浓度		例号	血催乳素浓度	
	术前	术后		术前	术后
1	276	41	6	266	43
2	880	110	7	500	25
3	1600	280	8	1700	300
4	324	61	9	500	215
5	398	105	10	220	92

5. 某地微丝蚴血症患者 53 例治疗后 7 年用间接荧光抗体试验测得抗体滴度情况见表 1-5, 试求平均抗体滴度。

表 1-5 用间接荧光抗体试验检测某地微丝蚴血症患者抗体滴度

抗体滴度的倒数	10	20	40	80	160
例数	5	18	15	9	6

(张俊辉)

实习 2 定量资料的统计推断

【实习目的】

1. 掌握均数抽样误差的概念、意义和计算; t 分布的特点; 区间估计的意义与应用; I 型错误与 II 型错误、检验效能的概念。
2. 掌握均数的标准误与标准差的区别。
3. 掌握均数的置信区间与医学参考值范围的区别。
4. 熟悉假设检验的基本思想、步骤和注意事项, 检验水准 (α) 和 P 值的区别。
5. 掌握 3 种常用的 t 检验的目的、公式、适用条件并能正确应用: ① 样本均数与总体均数比较的 t 检验; ② 配对 t 检验; ③ 成组 t 检验。

【知识点】

(一) 样本均数的抽样误差

1. 样本均数的抽样误差 在抽样研究中, 由于同质总体中的个体间存在差异 (即个体变异), 即使从同一总体中随机抽取若干份样本, 样本均数也常常不等于总体均数, 且各个样本均数之间也存在差异。这种由于随机抽样造成的样本均数与总体均数的差别, 就称为样本均数的抽样误差。在抽样研究中, 抽样误差是不可避免的。

2. 样本均数的标准误 样本均数围绕总体均数 μ 的离散程度, 可以用样本均数的标准差来描述。样本均数的标准差也称为样本均数的标准误 $\sigma_{\bar{x}}$, 其计算公式: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ 。

标准误是描述样本均数抽样误差大小的指标, $\sigma_{\bar{x}}$ 越小, 抽样误差越小, 用样本均数估计总体均数的可靠性越大。由式 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ 可知, $\sigma_{\bar{x}}$ 的大小与 σ 成正比, 与 \sqrt{n} 成反比, 即通过适当增加样本含量可以减少标准误, 降低抽样误差。在抽样研究中, 总体标准差 σ 常常是未知的, 因此常用样本标准差 S 来估计抽样误差的大小, 故样本均数标准误的估计公式为 $S_{\bar{x}} = \frac{S}{\sqrt{n}}$ 。

均数的标准误与标准差的区别见表 2-1。

表 2-1 均数的标准误与标准差的区别

	均数的标准误	标准差
意义	反映样本均数 \bar{X} 的抽样误差的大小	反映一组数据的离散情况
符号	总体标准误 $\sigma_{\bar{x}}$, 样本标准误 $S_{\bar{x}}$	总体标准差 σ , 样本标准差 S
计算	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $S_{\bar{x}} = \frac{S}{\sqrt{n}}$	$\sigma = \sqrt{\frac{\sum(X-\mu)^2}{n}}$ $S = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$
控制方法	增大样本含量可减小标准误	个体差异或自然变异, 不能通过统计方法来控制

(二) t 分布的特征

t 分布与标准正态分布相比有以下特征: ① 都是单峰、对称分布。② t 分布峰值较低, 而尾部较高。③ t 分布曲线是一簇曲线, 其形态与自由度 ν 的大小有关; 随自由度增大, t 分布逐渐逼近标准正态分布, 当 $\nu \rightarrow \infty$ 时, t 分布的极限分布是标准正态分布。

(三) 总体均数的估计

总体均数的估计有点估计和区间估计两种方法。

1. 点估计 点估计指用相应样本统计量直接作为总体参数的估计值, 如用样本均数 \bar{X} 估计总体均数 μ 。

2. 区间估计的意义与应用 区间估计是指按预先给定的概率 $(1-\alpha)$ 确定的包含未知总体参数的可能范围, 该范围称为总体参数的置信区间 (confidence interval, CI)。它的确切含义是: CI 是随机的, 总体参数是固定的, 所以, CI 包含总体参数的可能性是 $1-\alpha$ 。不能理解为 CI 是固定随机的, 总体参数是随机固定的, 总体参数落在 CI 范围内可能性为 $1-\alpha$ 。当 $\alpha=0.05$ 时, 称为 95% 置信区间, 记作 95% CI。当 $\alpha=0.01$ 时, 称为 99% 置信区间, 记作 99% CI。如无特别说明, 一般取双侧 95%。95% 置信区间即按 95% 置信度估计总体均数的可能范围, 此时估计正确的概率为 95%。

置信区间的计算方法有两种, 可根据资料的具体情况选择:

(1) σ 未知: 按 t 分布的原理估计置信区间, 则总体均数的 $(1-\alpha)$ 置信区间的公式: $(\bar{X} - t_{\alpha/2, \nu} S_{\bar{X}}, \bar{X} + t_{\alpha/2, \nu} S_{\bar{X}})$ 。