

(第四版)

医学科研中的统计方法

Statistical Method in Medical Research

马斌荣 主编 罗艳侠 副主编



科学出版社

医学科研中的统计方法

(第四版)

主 编 马斌荣

副主编 罗艳侠

科 学 出 版 社

北 京

《医学科研中的统计方法》(第四版)编委名单

主 编 马斌荣

副主编 罗艳侠

编 委 (按姓氏汉语拼音排序)

陈 卉 (首都医科大学生物医学工程学院)

李 霞 (首都医科大学公共卫生与家庭医学学院)

罗艳侠 (首都医科大学公共卫生与家庭医学学院)

马斌荣 (首都医科大学生物医学工程学院)

吴立娟 (首都医科大学公共卫生与家庭医学学院)

第四版前言

21 世纪是生命科学和信息技术快速发展的新时代。医学科研中正确收集、整理和统计分析数据对基础医学、临床医学、流行病学、医院信息管理和卫生经济学的发展将起重要的促进作用。医学科研工作可以提高医疗诊断和治疗水平,促进医学发展,继而产生巨大的社会效益。

《医学科研中的统计方法》作为医学硕士研究生的必修课是非常必要的。它将全过程指导研究生的学位科研课题研究工作,包括总体设计、数据采集、数据处理和论文书写。

本书保留了第三版深入浅出地讲解医学研究生所必需的统计学中的基本概念、基本理论、基本知识和基本技术的内容。另外,根据医学科学的发展和医药研究生研究课题的多样性,增加了如下实用的内容。

(1) 在生物医学研究中,常常对同一观察对象在多个时点进行测量,这种设计称为重复测量设计,宜采用“重复测量设计方差分析”(见第六章)。

(2) 在生物医学研究中,其结局指标或效应指标常是分类变量,如“生存或死亡”,“发病或不发病”,并且与其关联的是非线性关系的影响因素(自变量)。此时,宜用分类资料分析的非条件概率“Logistic 回归”(见第十二章)。

(3) 研究生存期、生存时间及不同治疗方法对生存时间比较的统计方法常用的有“Kaplan-Meier 生存分析”、“COX 比例风险回归模型”(见第十三章)。

(4) 用于新的诊断试验方法及新的诊断指针的有效性、适用性及癌症的筛查和精神病的诊断评价的“反映诊断试验的诊断价值的 ROC 曲线分析”(见第十六章)。

(5) 在申请国家自然科学基金及其他医学科研申请书中,经常需论述“为什么该样本量是合理的”,本书提供了“样本例数的估计方法”(见第十九章)。

(6) 用于流行病学、社会医学、心理学、卫生管理学的“调查量表设计及其信度效度分析”(见第十八章)。

(7) 目前很多高水平的杂志,都先发表了统计学指南。本书为了规范医学研究生论文书写,参考近期的一些统计学指南,增加了“医学论文书写的统计学指南”(见第二十章)。

(8) 本书中的大部分例题与《SPSS(PASW) 17.0 在医学科研中的应用》一书共享。

(9) 本书中的例题,凡可用 SPSS 软件操作的,都编写了 SPSS 数据文件,需要的读者可与出版社联系,免费提供。

本书是由首都医科大学生物医学工程学院和公共卫生与家庭医学学院具有长期教学经验的教师编写而成,对医学研究生教学具有较好的实用价值。

限于我们的学识和能力,本书还可能有许多不足之处,我们愿意虚心听取广大读者的批评和建议,以利改进。

编 者

2012 年春节

第三版前言

21 世纪是信息技术和生命科学快速发展的时代。医学信息的正确收集、整理和分析对基础医学、临床医学、流行病学、医院信息管理和卫生经济的发展将起到重要的促进作用。医学科研工作能够提高医疗诊断水平和质量,促进医学发展,具有巨大的社会效益。《医学科研中的统计方法》作为医学硕士研究生的必修课是非常必要的。它将全程指导研究生的学位科研课题,包括总体设计、数据采集、数据处理和论文书写。

随着计算机科学的迅猛发展,医学统计软件包的日益完善和广泛应用,原来需要很长时间才能计算出结果的内容,现在使用统计软件包瞬间就能准确地计算出来,大大促进了医学统计学的发展。对于熟悉统计学的基本原理和基本方法的用户来说,使用统计软件包将为他们带来极大的方便;但是,那些统计概念不清、不能正确选择统计方法的用户面对统计软件包将感到困惑和不知所措。

因此,本书尽可能深入浅出地讲解医学统计学的基本概念、基本理论、基本知识和基本技术,尤其对各种统计方法的使用条件、使用方法及统计结果的理解等方面进行了正确而详尽的解释,而不拘泥于大量的繁杂的计算过程。

本书在本科生教学的基础上加强了多元统计的内容,比如,多元回归、协方差分析、聚类分析和主成分分析等。

本书第三版增加了诊断试验的评价、ROC 曲线等内容,并对样本例数估计提供了按公式计算和查表两种方法。

限于我们的学识和能力,本书可能有许多不足之处,我们愿意虚心听取广大读者的批评和建议,以利改进。

编 者

2004 年 12 月 12 日

目 录

第四版前言

第三版前言

第一章 医学统计中的基本概念	1
第一节 绪论	1
第二节 医学统计工作的内容及资料类型	2
第三节 医学统计中的基本概念	3
第二章 定量资料的统计描述	10
第一节 频数表与直方图	10
第二节 正态分布的概念和特性	13
第三节 集中趋势指标	15
第四节 离散趋势指标	19
第五节 正态分布的应用	23
第三章 定性资料的统计描述	26
第一节 相对数的意义	26
第二节 常用相对数	26
第三节 应用相对数的注意事项	30
第四节 标准化法	31
第四章 抽样误差与假设检验	35
第一节 抽样误差	35
第二节 定量资料总体均数的估计	36
第三节 定性资料总体率的估计	37
第四节 假设检验的概念	37
第五节 假设检验的两类错误	39
第五章 t 检验	41
第一节 单样本的 t 检验	41
第二节 配对样本的 t 检验	43
第三节 两独立样本均数比较的 t 检验	44
第四节 使用 t 检验的注意事项	46
第六章 方差分析	48
第一节 完全随机设计的方差分析	48
第二节 随机区组设计的方差分析	51

第三节	多个样本均数间的两两(多重)比较	53
第四节	析因设计的方差分析	55
第五节	重复测量设计方差分析	57
第七章	χ^2 检验	64
第一节	行 $\times$ 列表的 χ^2 检验	64
第二节	成组四格表资料的 χ^2 检验	67
第三节	配对四格表资料的 χ^2 检验	70
第四节	使用 χ^2 检验注意事项	72
第八章	秩和检验	75
第一节	配对比较的秩和检验	75
第二节	两独立样本比较的秩和检验	76
第三节	多个样本比较的秩和检验	79
第四节	多个样本间的两两比较的秩和检验	81
第五节	使用秩和检验注意事项	81
第九章	直线相关与回归	84
第一节	线性相关	84
第二节	线性回归	89
第三节	应用线性相关与回归的注意事项	94
第四节	等级相关	94
第十章	协方差分析	97
第一节	完全随机设计的协方差分析	97
第二节	随机配伍组设计的协方差分析	100
第三节	协方差分析的应用条件及注意事项	104
第十一章	多重线性回归	105
第一节	多重线性回归	105
第二节	多重逐步线性回归	114
第三节	使用多重线性回归的注意事项	116
第十二章	Logistic 回归	120
第一节	Logistic 回归的基本概念	120
第二节	Logistic 回归模型的参数估计与检验	122
第三节	Logistic 回归的统计分析实例	122
第四节	使用 Logistic 回归的注意事项	131
第十三章	生存分析	133
第一节	生存分析的基本概念	133
第二节	寿命表法	135

第三节 Kaplan-Meier 法	136
第四节 生存曲线的比较	139
第五节 Cox 比例风险回归模型	140
第十四章 聚类分析	146
第一节 常用的聚类统计量	146
第二节 样品聚类	148
第三节 指标聚类	151
第四节 使用聚类分析注意事项	152
第十五章 主成分分析	155
第一节 主成分分析	155
第二节 因子分析	162
第十六章 诊断试验的评价与一致性检验	164
第一节 诊断试验的评价	164
第二节 诊断试验的一致性检验	168
第三节 ROC 曲线及其曲线下面积	172
第十七章 统计表与统计图	181
第一节 统计表	181
第二节 统计图	185
第十八章 医学常用的调查设计	193
第一节 实验研究设计	193
第二节 调查研究设计	199
第三节 量表的信度效度分析	204
第十九章 样本含量估计	219
第一节 影响样本含量的因素	219
第二节 配对设计均数比较的样本含量估计	219
第三节 两独立样本设计均数比较的样本含量估计	220
第四节 配对设计率比较的样本含量估计	221
第五节 两独立样本设计率比较的样本含量估计	222
第六节 调查研究中的样本含量估计	223
第二十章 医学论文书写的统计学指南	225
第一节 样本例数和数据的质量控制	225
第二节 统计方法的选择和检验水准的表达	225
第三节 使用的统计方法及其结果的确切表达	226
第四节 科研工作中的偏性及敏感问题的处理	227
参考文献	229
附表	230

第一章 医学统计中的基本概念

第一节 绪 论

统计分析是科研工作中的一种有力工具。它与科研的总体设计、资料采集、资料整理、资料分析直到最后作出正确结论都有密切关系。掌握这个工具可以使用较少的人力、物力和时间获得比较可靠的结果。需要运用统计方法的场合而不去运用它可能造成不应有的缺陷或得出错误的结论。

北京某医院某大夫使用“乌贝散”(乌贼骨 3 钱,白芨 3 钱,川贝 3 钱,甘草 3 钱,共研细末过 120 目筛,冲服)治疗胃溃疡病出血 107 例,有效 101 例,有效率为 94.4%。那么别的医院,其他大夫使用“乌贝散”来治疗胃溃疡病出血,其有效率也一定是 94.4%吗?显然,不一定恰为 94.4%,那么是多少呢?比它高还是比它低?这就是一个求置信区间的问题。

北京某医院有位老大夫,用“冠心病灵”治疗冠心病,其对照组用单纯西药,观察结果如表 1.1 所示。

表 1.1 冠心病灵与单纯西药疗效对比

分组	显 效	有 效	无 效	合计
单纯西药	9	25	6	40
冠心病灵	19	18	5	42

请问冠心病灵是否比单纯西药有效?这就涉及统计学中的假设检验(hypothesis and test),俗称要算 P 值。

北京 15 所医院,联合对心肌梗死抢救治疗后的患者进行了随访调查,其随访观察的部分资料如表 1.2 所示。

表 1.2 心肌梗死抢救治疗后的随访结果

随访年份	随访数	死于本病	中断随访
第一年	1342	130	41
第二年	1171	48	53
第三年	1070	42	73
第四年	955	56	48
第五年	851	35	

根据这些资料,可提出如下几个问题:

出院后第一年死亡率与以后几年的死亡率的差异是否有统计学意义?

2 年生存率? 3 年生存率? ……

其中有些患者中断随访了,中断随访的患者如何进行统计分析? 怎样的随访才算有效? 这些问题的解决,将有赖于一系列统计方法。

某杂志中一篇文章报道,用某种中草药治疗玫瑰糠疹,有效率为 78%,平均疗程为 3 周,问此药治疗玫瑰糠疹的确有效吗? 由于玫瑰糠疹具有自愈的情况,如果不服药,多喝水,到 3 周左右有些患者也会自愈。所以,此药的疗效有待探讨。这就要求设立具有可比性的对照组。

再如,用某药治某病,治疗 1 例有效,能说 100%有效吗? 治疗 2 例都有效,能说 100%有效吗? 治疗 10 例,其中 9 例有效,能说 90%有效吗? 显然,观察的例数将是个重要因素,例数太少不足以说明问题;观察例数太多,将费时并且加重科研经济负担,那么观察多少例能说明问题? 这是一个样本可靠性问题。

第二节 医学统计工作的内容及资料类型

一、医学统计工作的内容

实验设计(experiment design)、收集资料(collection of data)、整理资料(sorting data)和分析资料(analysis of data)是医学统计工作的主要内容。

(1) 实验设计就是根据研究的目的,制订总的研究方案,包括研究对象的纳入标准和排除标准、样本量和样本获取方法、实验组与对照组的分组原则、确定观察指标及精度、实验过程中的质量控制、拟使用的统计方法等。

(2) 收集资料就是根据研究的目的、实验设计的要求,收集准确的、完整的、充满信息的原始资料(raw data)。

医学统计资料主要有实验数据和现场调查资料报表、医疗卫生工作记录和报告卡等。

在这些资料的收集过程中,必须进行质量控制,包括它的统一性、确切性、可重复性。对这些原始数据的精度(precision)和偏性(bias)应有明确的控制范围。

(3) 整理资料就是把收集到的原始资料,有目的地进行科学加工,使资料系统化、条理化,以便进行统计分析。一般来说,要注意如下几点:①从医学专业的角度对资料之间的逻辑一致性进行检查,如婴儿有母乳喂养的,应有相应的喂养时间,没有母乳喂养的,不应该有喂养时间;②从专业的角度对资料的合理性进行检查,如退休时的年龄不应出现小于 20 岁,出生婴儿的体重不应出现大于 12 000g,男性患者的调查表中不应出现妇科疾病等;③从医学专业的角度对资料的一致性进行检查,如诊断和疗效的评定标准是否统一,胃镜下萎缩性胃炎的严重程度的评定标准是否一致等。

④分析资料就是把经过整理的资料,做一系列统计描述和统计推断,阐明事物的规律性。应该注意,不同的资料使用的统计描述和统计推断的方法也会不一样。

学习医学统计学应着重于:

- (1) 理解医学统计学的基本原理和基本概念;
- (2) 掌握收集、整理与分析资料的基本知识与基本技能;
- (3) 重视原始资料的完整性、可靠性及处理数据时的实事求是的科学态度。

二、资料的类型

医学统计资料一般可分为定量变量(quantitative variable)(也称为数值变量)(measurement variable)和定性变量(qualitative variable)(也称为分类变量,enumeration variable)两大类。不同的统计资料应采用不同的统计分析方法。

定量变量是对每个观察对象用定量方法测定观察指标的数值大小所得的资料,一般有度量衡单位,如身高(cm)、体重(kg)、浓度(mg/L)、脉搏(次/min)、血磷(mmol/L)、血红蛋白(g/L)等。

定性变量是先将观察对象的观察指标按性质或类别进行分组,然后清点各观察对象的指标落入各组的数目。例如,调查某人群的血型分布,按照A、B、AB、O 4型分组,计数所得该人群的各血型组的人数就是定性变量(分类变量)。

在定性变量的类别之间出现顺序或等级关系,则该变量为有序变量(ordinal variable),在医学实践中,如用“生脉散”治疗脑血管病,其疗效可分为治愈、显效、有效、无效,那么其对应的人数将是定性有序变量。又如,临床化验中,将化验结果按一、+、++、+++等级分组,计算每组患者数,也是定性有序变量。

根据分析的需要,定量变量、定性变量、定性有序变量可以互相转化。例如,每个人的血红蛋白含量,原属定量变量;若按血红蛋白正常与异常分为两组,得出各组的人数,是定性变量;若按血红蛋白含量的多少分为5个等级:<60g/L(重度贫血)、60~90g/L(中度贫血)、90~125g/L(轻度贫血)、125~160g/L(血红蛋白正常)、>160g/L(血红蛋白增高),计算各等级人数,就是定性有序变量。

第三节 医学统计中的基本概念

一、变 异

医学研究的对象是有机的生命体,其机能是十分复杂的。不同的个体在相同的条件下,对外界环境因素可以发生不同的反应。例如,同种族、同年龄、同性别的健康人,在相同的条件下测其脉搏、呼吸、体温等生理指标可以有很大差异。在临床治疗

中,用同样的药物治疗病情相同的患者,疗效也不尽相同,即使在实验室里,动物与动物之间也有明显的差异,这种现象称为个体差异或称为变异。变异是由众多的、偶然的、次要的因素造成的。

由于医学统计研究的对象是有变异的事物,因此,用观察 1~2 例的结果来推论出一般规律是不恰当的。例如,我们不能用某一健康成人的红细胞数作为一般健康成人的红细胞数;也不能因为用“青木香”治疗一例高血压患者有效,而断定“青木香”治疗高血压有效率为 100%,因为观察例数太少,受个别的、偶然的因素影响太大。科学研究工作的主要任务之一就是要从表现为偶然性的大量数据中,分析出其中必然性的规律。而概率论和数理统计就是要通过这种偶然性去寻找隐藏在其内部的必然规律性。

二、总体和样本

总体是同质的个体所构成的全体。医学研究的对象,一般都是数量巨大的群体。例如,我们要研究华北地区成年人血液的红细胞数的正常值,那么,华北地区全部健康成年人的血液都是我们的研究对象,称为总体。但是,我们不可能对华北地区所有健康成年人都进行血液检查,而只能按照一定的科学方法从中抽取一部分人来进行血液检查。然后,根据检验结果来推论华北地区健康成年人血液的医学参考值。

这种从总体中抽取部分个体的过程称为抽样(sampling),所抽得的部分称为样本。在一个样本里含有的个体数可以不同,样本包含的个体数目称为样本含量(sample size)。如何正确地从样本观察到的情况来推测总体,这就是医学统计学要解决的问题之一。

三、抽 样

从总体中抽取样本,一定要遵循科学原则。一般来说,样本应具有“代表性”(representative)、“随机性”(randomization)和“可靠性”(reliability)。

(一)代 表 性

代表性就是要求样本中的每一个个体必须符合总体的规定。这就要求对总体有一个明确的规定,这种规定是根据研究目的而具体确定的。例如,要调查华北地区健康成年男子红细胞数的正常值,我们可以规定研究对象为体温正常、无急性病、无血液消耗性疾病、肝在肋缘下 0.5cm 以内、血红蛋白为 125g~160g/L、血小板在 10 万以上的华北地区 18 岁以上的男子。必须做肝功、X 射线胸透及某些血象检查……,那么所抽取的样本中的每个个体都必须做上述检查,并符合上述规定。

(二) 随 机 性

随机性就是要保证总体中的每个个体有相同的概率被抽作样本,重要的是要避免主客观的“偏性”。必须指出,随机化抽样绝不等于随意抽样。例如,将 40 只小鼠分为两组。如果闭上眼睛,随意抓取 20 只作为第一组,留下的作为第二组,表面上看来是随机的,实际上是不随机的。因为体壮、活泼性强的小鼠难于抓到,故大部分留在第二组,两组的活泼性是不同的。

为了保证抽样的随机性,可用抽签法、机械抽样法、分层抽样法及利用随机数字表抽样等方法。

(1) 抽签法。例如,要把 40 个动物分成两组,首先把每个动物编上号码(1~40),然后做 40 个签(1~40),混匀后,再随机把它分成两组。

(2) 机械抽样法。又称为等间隔抽样,即先将总的观察对象按某种顺序编号,再从这些编号中采用等间隔抽样。例如,在某工厂 1000 名健康女工中抽取 100 人进行血液检查,可先将工人编号,然后从 0~9 共 10 个数字中,用抽签法任意抽取一个数字,如抽到数字“8”,那么工人编号末尾逢“8”的即为检查对象,即 8、18、28、38、…、998 为检查对象。如果想在 1000 名工人中抽取 200 人做检查,那么可在 0~9 共 10 个数字中随机抽取两个数字,则工人编号末尾的数字与这两个数字相同者即为受检对象。机械抽样的优点是易于理解,简便易行,容易得到一个按比例分配的样本。

(3) 分层抽样(stratified sampling)法,即先按某种性质将总体分为若干组别、类别或区域(在统计上称为层),再从每一层内按比例进行随机抽样,组成样本。例如,调查北京市胃癌发病情况,可先根据北京城区和农村人口比例,确定城区和农村选取的人数,然后,在城区和农村中用机械抽样或抽签法进一步抽样。当然,在城区和农村的抽样中,还应考虑各年龄段的人口构成,按不同比例,再分别随机抽样,即分层随机抽样(stratified random sampling)。分层随机抽样可根据实际科研题目要求分成多层,但一定要注意各层的比例和按比例分配抽取样本。

(4) 利用随机数字表进行随机抽样。现在的计算机统计软件根据医学科研的要求,也能产生一系列随机数字并进行分组。如何用计算机统计软件进行随机分组,读者可参阅有关书籍,本书从略。

综上所述,随机化抽样,绝不是随便抽样,应根据科研要求,做好设计,按设计抽样。

(三) 可 靠 性

可靠性即实验的结果要具有可重复性,即由科研课题样本的结果所推测总体的

结论有较大的可信度。由于个体之间存在差异,只有观察一定数量的个体方能体现出其客观规律性。如果根据少数几例就下结论,这种结论可靠性差,体现不了规律,可能被后人所否定,这种先例是曾经发生过的。诚然,每个样本的含量越多,可靠性会越大,但是例数增加,得到所需的人力、物力就会比较困难,所以应以“足够”为准。需要例数的多少,与所观察的指标的变异程度有关。有关计算方法可在本书的“样本含量的估计”一章中找到。

(四) 可比性

如果进行两个或多个样本之间的比较,那么要求任意两个样本之间应具有可比性(comparable)。

可比性是指处理组(临床设计中称为治疗组)与对照组之间,除处理因素不同外,其他可能影响实验结果的因素基本齐同,也称为齐同对比原则。

不同医院的住院患者差别较大,相互作比较时要注意其可比性。大医院或有特色的医院,接受治疗的患者中可能重症患者、疑难杂症患者比例较大,小医院或基层医院可能一般的轻症患者比例较大,所以两个医院的住院患者病死率不具有可比性,不能直接进行比较。要进行标准化后才能进行比较。

在临床医学中,设置对照组是非常必要的,但也是比较困难的。对那些可以“自然痊愈”的疾病(如玫瑰糠疹、口腔溃疡)都应设立对照组。

空白对照在基础医学研究的动物实验中可以采用,但是,在临床医学研究中一般不采用。

对照组使用的“安慰剂”只在少数病种中使用,往往用于与心理或精神因素有关的疾病,如疼痛、晕船、情绪不安等。这些安慰剂是外形和真药一样而实质上没有药理作用的物质,它通过医务人员的语言、态度、威信及患者的心理和精神因素而起作用。有的疾病使用“安慰剂”疗效可达35%,有的甚至更高。所以,这类疾病使用“安慰剂”是可行的也是必要的(这样的临床试验必须在有该病种临床试验资质的医院,并且,在患者签订知情同意书的条件下,才可行)。

在临床研究中,往往要判断某种新药或某种新疗法的疗效是否较已知的药物或疗法更好。这时对照组通常是采用已上市的药物中疗效最好的一种。

四、实验设计

在做实验或搞科研的开始,我们就应考虑把实验对象按“完全随机设计”、“配对设计”或“配伍组设计”。设计不同不仅抽样方法不同,在处理其实验结果时使用的统计方法也不同。

(一) 完全随机设计

完全随机设计 (completely random design) 是将受试对象随机分配到各个处理组或对照组中, 或分别从不同总体中随机抽样进行研究。它可以是两样本比较, 也可以是多样本比较; 各样本含量可以相等, 也可以不相等, 但也不宜差别太大。

(二) 配对设计

配对设计 (matching design) 是将受试对象按一定条件配成对子, 称为配对样本, 将它们随机分到两个组中, 分别给以不同的处理。配对条件是指有可能影响实验效应的非主要因素。动物实验中, 常将种属、性别相同, 年(月)龄、体重相近的两个动物配成对子; 人群实验中, 常将性别相同, 年龄、生活劳动条件相近的两个人配成对子。这样就可增加各处理组之间的齐同对比性; 但是, 由于人群的特殊性, 一般很少采用配对设计。

(三) 随机区组设计

随机区组设计 (random block design) 也称为配伍组设计, 它是配对设计的扩展。配对设计的每一“对子”分别随机分到两个处理组中, 而配伍组设计中的每个“配伍组”包含多个受试对象, 要将它们分别随机分到各处理组。

例 1.1 现有 36 个动物, 要将其分到 A、B、C 3 个处理组, 这些动物根据性别、年龄、体重相近原则, 分成 12 个配伍组, 每组 3 个, 第一个配伍组的受试动物编号为 1~3, 第 2 个配伍组的受试动物编号为 4~6, 依此类推。把每个配伍组的 3 个动物随机分到 A、B、C 3 个处理组, 其结果如表 1.3 所示。

表 1.3 配伍组设计

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
A	B	C	C	B	A	B	C	A	A	C	B	B	C	A	C	B	A
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
A	C	B	C	B	A	B	A	C	A	B	C	A	B	C	C	B	A

五、误差

统计上所说的误差 (error) 泛指观测值 (observed value) 与真实值之差, 以及样本指标与总体指标之差, 主要有以下 3 类。

(一) 系统误差

在收集资料的过程中,由于仪器初始状态未调整到零、标准试剂未经校正、医生掌握疗效标准偏高或偏低等原因,使观察结果倾向性的偏大或偏小,这称为系统误差(system error)。系统误差影响原始资料的准确性,必须克服。如果已发生,要尽力查明其原因,予以校正。

(二) 随机测量误差

在收集原始资料的过程中,即使仪器初始状态及标准试剂已经校正,但是,由于各种偶然因素的影响,也会造成同一对象多次测定的结果不完全一致。例如,实验操作人员操作技术不稳定、不同实验操作人员之间的操作差异、电压不稳及环境温度差异等因素造成测量结果的误差。这种误差往往没有固定的倾向,有时高有时低,被称为随机测量误差(random measurement error)。对于这种误差应采取措施,尽最大可能来控制,至少应控制在一定的允许范围内。一般可以用技术培训、指定固定实验操作人员、加强责任感教育及购置一定精度的稳压器、恒温装置等措施,从而达到控制的目的。

(三) 抽样误差

即使消除了系统误差,并把随机测量误差控制在允许范围内,样本均数(或其他统计量)与总体均数(或其他统计量)之间仍可能有差异,这种误差产生的原因是:①个体之间存在变异;②抽样时只能抽取总体中的一部分作为样本。由此样本的数据计算的统计指标(如均数)就会与总体的该指标有差异,由于这种差异是抽样引起的,故称为抽样误差(sampling error),对它要用统计方法进行正确分析。

一般来说,样本含量越大,则抽样误差越小,样本的观察指标与总体的该指标越接近,即越能说明总体的规律。反之,样本含量越小,则抽样误差相应地越大。因此,我们不能仅仅凭观察指标的大小进行简单判断,而应该使用概率与数理统计方法来辨别哪些实验研究的结果是有实际意义的,哪些可能是由抽样误差所造成的,从而得出正确的结论。常用的统计方法是“统计学检验”。

六、概 率

概率(probability)是描写某一事件发生的可能性大小的一个量度。用 A 表示某一事件, P 表示该事件可能发生的概率,可记为 $P(A)$ 。