

生物统计学

第一章 绪论

职晓阳

xyzhi@ynu.edu.cn

云南大学 生命科学学院



會澤百宗 至公天下

- ① 生物统计学的概念
- ② 生物统计学简史
- ③ 学习生物统计学的重要性
- ④ 生物统计学常用术语
- ⑤ 生物统计学的主要任务与作用

① 生物统计学的概念

② 生物统计学简史

③ 学习生物统计学的重要性

④ 生物统计学常用术语

⑤ 生物统计学的主要任务与作用

1.1 生物统计学的概念

统计学 (statistics) 是把数学的语言应用于具体的科学研究领域,
将所研究的问题抽象为数学问题的一门科学;
涉及数据的收集、整理分析 和解释。

1.1 生物统计学的概念

生物统计学 (biostatistics) 是统计学的原理和方法在生物学研究中的应用，旨在揭示生物现象的数量特征及其变化规律。

- ① 生物统计学的概念
- ② 生物统计学简史**
- ③ 学习生物统计学的重要性
- ④ 生物统计学常用术语
- ⑤ 生物统计学的主要任务与作用

1.2 生物统计学简史

Statistics

science dealing with data about the condition of a state or community.

1.2 生物统计学简史

关于国家管理的数据收集和研究，系统性的文字记载初现于 17 世纪的英国。



Sir William Petty (1623-1687)



John Graunt (1620-1674)

1.2 生物统计学简史

古典记录统计学

形成于 17 世纪中叶至 19 世纪中叶。

- 瑞士数学家 Jakob Bernoulli 第一次证明了大数定律。
- 法国数学家 Pierre-Simon Laplace 发展了概率论，提出了拉普拉斯定理（中心极限定理的一部分）。
- 法国数学家 Abraham de Moivre 发现正态分布。
- 德国数学家 Carolus F. Gauss 研究误差的概率分布，提出了“误差分布曲线”。

1.2 生物统计学简史

近代描述统计学

形成于 19 世纪中叶至 20 世纪上半叶。

统计学开始与生物学相结合，并共同发展。

- 英国遗传学家 Francis Galton 建立“人体测量实验室”，引入了中位数、百分位数、以及分布、相关、回归等重要概念。
- Galton, Karl Pearson, Walter F.R. Weldon 创立 *Biometrika*, 提出“生物统计 (biometry)”一词。
- K. Pearson 提出了相关与回归分析问题, χ^2 检验。

1.2 生物统计学简史

现代推断统计学

形成于 20 世纪初至 20 世纪中叶。

从“描述”到“推断”，是统计学发展的一次巨大飞跃。

- William S. Gosset, “Student”; t 检验; 小样本研究理论的先驱。
- Ronald A. Fisher, 发展了显著性检验及参数估计理论; 提出了 F 分布和 F 检验, 创立了方差和方差分析; 随机区组法、拉丁方法和正交试验。
- Jerzy Neyman 和 Egon Pearson, 共同完成假设检验和区间估计理论工作。

- ① 生物统计学的概念
- ② 生物统计学简史
- ③ 学习生物统计学的重要性**
- ④ 生物统计学常用术语
- ⑤ 生物统计学的主要任务与作用

1.3 学习生物统计学的重要性



在最终的分析中，所有知识皆为历史；
在抽象的意义下，所有科学皆为数学；
在理性的世界里，所有判断皆为统计。

Calyampudi R. Rao (1920-2023)

- ① 生物统计学的概念
- ② 生物统计学简史
- ③ 学习生物统计学的重要性
- ④ 生物统计学常用术语**
- ⑤ 生物统计学的主要任务与作用

1.4 生物统计学常用术语

1.4.1 总体与样本

定义 (1.1)

具有相同性质的研究对象/个体所组成的集合，称为**总体** (population)。

定义 (1.2)

从总体抽出的部分研究对象/个体构成的集合，称为**样本** (sample)。

1.4 生物统计学常用术语

1.4.2 参数与统计量

定义 (1.3)

用来描述总体特征的概括性数字度量，称为**参数** (parameter)，通常用希腊字母表示。

定义 (1.4)

用来描述样本特征的概括性数字度量，称为**统计量** (statistic)，通常用英文字母表示。

1.4 生物统计学常用术语

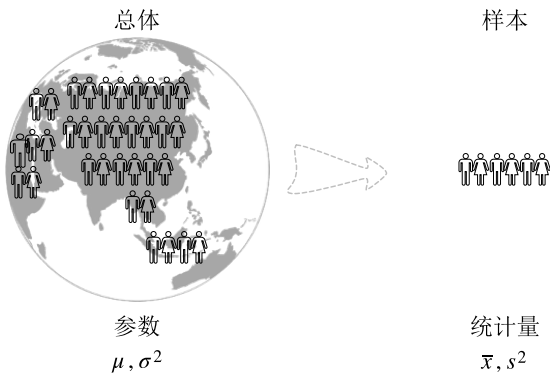


图 1.1 总体与样本、参数与统计量

1.4 生物统计学常用术语

1.4.3 变量与观测值

定义 (1.5)

相同性质的研究对象间可呈现差异的一些特征或属性，称为**变量** (variable)。

定义 (1.6)

变量的取值，称为**观测值** (observed value)。

1.4 生物统计学常用术语

1.4.4 因素与水平

定义 (1.7)

试验中影响试验指标的原因或原因的组合，称为**试验因素** (experimental factor)，或处理因素 (treatment factor)，简称因素，通常用大写英文字母表示。

定义 (1.8)

每个试验因素的不同状态 (处理的某种特定状态或数量上的差别)，称为**因素的水平** (factor level)，简称水平，通常用相应因素的英文字母加下标来表示，如 A_1 和 B_2 。

1.4 生物统计学常用术语

1.4.5 处理与重复

定义 (1.9)

对受试对象 (tested subject, 又称试验单位/单元) 施加的某种外部干预或措施, 称为处理 (experimental treatment)。

定义 (1.10)

在试验中, 将同一个处理实施在两个或多个试验单位上, 称为重复 (replication)。

1.4 生物统计学常用术语

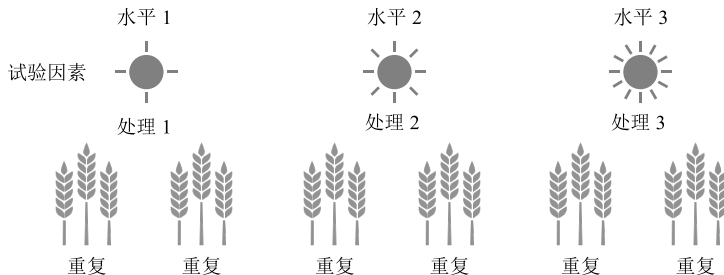


图 1.2 因素与水平、处理与重复

1.4 生物统计学常用术语

1.4.6 效应与交互

定义 (1.11)

试验因素对试验对象产生的、相对独立的作用，称为该因素的**主效应** (main effect)，简称主效或效应。

定义 (1.12)

两个或多个处理因素间相互作用产生的效应，称为**互作效应** (interaction effect)，简称互作或连应。

1.4 生物统计学常用术语

1.4.7 准确性与精确性

定义 (1.13)

在调查或试验中某一试验指标或性状的观测值与真实值接近的程度，称为**准确性** (accuracy)。

定义 (1.14)

调查或试验中同一试验指标或性状的重复观测值彼此接近的程度，称为**精确性** (precision)。

1.4 生物统计学常用术语

1.4.7 准确性与精确性

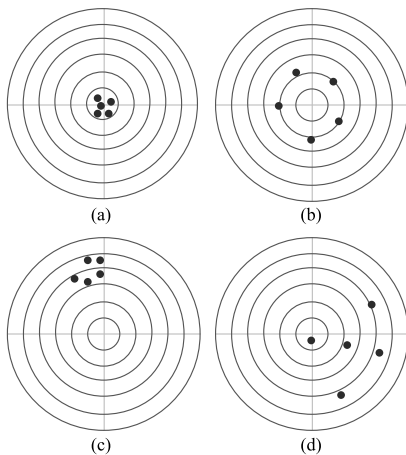


图 1.3 准确性与精确性

1.4 生物统计学常用术语

1.4.8 误差与错误

定义 (1.15)

观测值偏离真实值的差异，称为**误差** (error)。

定义 (1.16)

由于测量方法不正确或违反了操作规则等人为因素引起的差错，称为**错误** (mistake)。

- ① 生物统计学的概念
- ② 生物统计学简史
- ③ 学习生物统计学的重要性
- ④ 生物统计学常用术语
- ⑤ 生物统计学的主要任务与作用**

1.5 生物统计学的主要任务与作用

生物统计学的基本任务，概括来说主要包括：

- 试验设计 (experimental design)
- 统计分析 (statistical analysis)

1.5 生物统计学的主要任务与作用

试验设计是指应用统计学的基本原理和方法制定试验方案、选择试验材料并进行合理分组，使研究者用较少的人力、物质和时间成本获得较多可靠的数据资料。

1.5 生物统计学的主要任务与作用

统计分析是指应用数理统计的原理与方法对数据资料进行分析与推断，主要包括**统计描述** (statistical description) 和**统计推断** (statistical inference)。

技术上，统计分析涉及数据资料的收集和整理、特征数的计算、假设检验、方差分析、回归和相关分析等。

本章小结

- ① 生物统计学的概念
- ② 生物统计学简史
- ③ 学习生物统计学的重要性
- ④ 生物统计学常用术语
- ⑤ 生物统计学的主要任务与作用