



生物统计学统计方法

汇报人：<XXX>

2024-01-25



目录

- 绪论
- 描述性统计
- 概率论基础
- 参数估计
- 假设检验
- 非参数检验
- 回归分析
- 实验设计与方差分析



01

绪论

Chapter





生物统计学的定义与意义



生物统计学定义

生物统计学是应用数理统计学的原理和方法，分析和解释生物学研究中所得数据的一门科学。



生物统计学的意义

在生物学研究中，对实验数据进行统计分析是不可或缺的重要环节。生物统计学提供了科学的方法和工具，帮助研究者从数据中提取有用信息、揭示生物学规律和机制，以及评估研究结果的可靠性和有效性。



统计方法在生物学研究中的应用

通过样本数据推断总体特征，包括参数估计和假设检验等方法，用于评估研究结果的统计显著性和可靠性。

处理多个变量之间的关系，如回归分析、主成分分析、聚类分析等，以揭示复杂的生物学现象和机制。

描述性统计

对数据进行整理和描述，如计算平均值、标准差、绘制图表等，以直观展示数据的分布和特征。

推论性统计

实验设计

合理规划实验方案，减少误差和偏见，提高实验的效率和准确性。

多元统计分析



课程目标与要求

01

掌握基本的统计概念和方法，如概率、假设检验、方差分析等。



02

能够运用适当的统计方法分析和解释生物学数据。



03

了解实验设计和数据收集的基本原则和技巧。



04

培养批判性思维和科学态度，能够客观评估研究结果的可靠性和有效性。





02

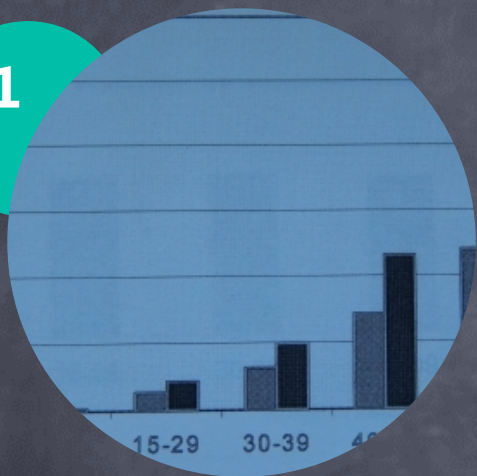
描述性统计

Chapter



数据类型与描述

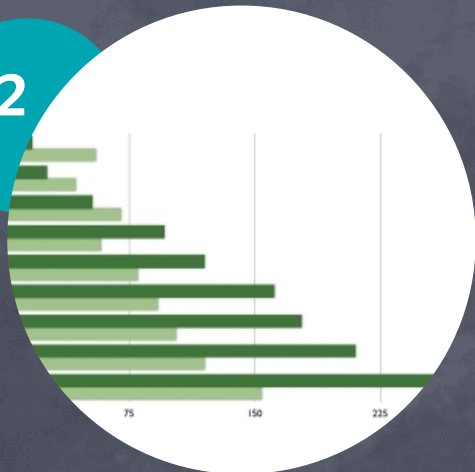
01



定量数据

数值型数据，如身高、体重等，可进行数学运算。

02



定性数据

分类数据，如性别、血型等，用文字或符号表示。

03

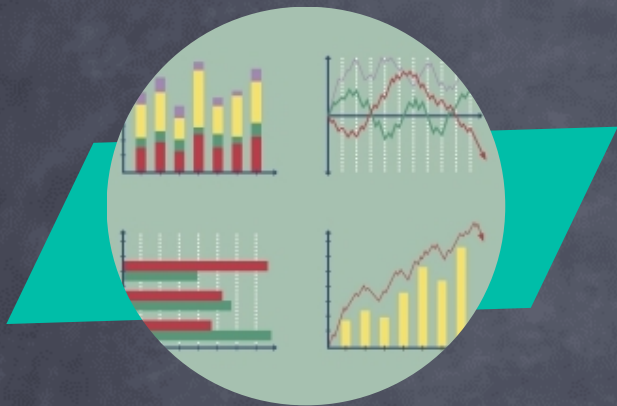


有序数据

具有等级或顺序关系的定性数据，如病情等级、教育程度等。

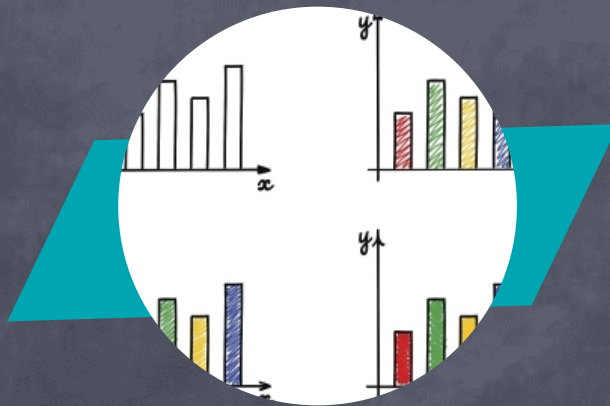


集中趋势的度量



算术平均数

所有观察值之和除以观察值个数，反映数据的平均水平。



中位数

将数据按大小顺序排列后，位于中间位置的数，反映数据的中心位置。



众数

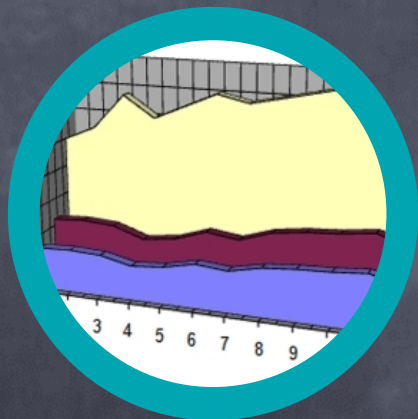
出现次数最多的数，反映数据的集中情况。



离散程度的度量

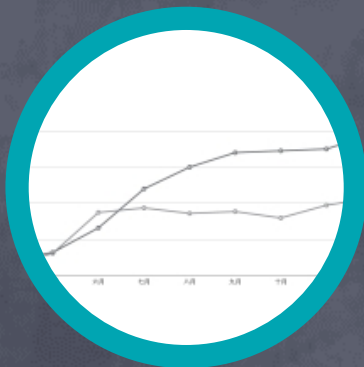
极差

最大值与最小值之差，反映数据的波动范围。



方差

各观察值与平均数之差的平方和的平均数，反映数据的离散程度。



标准差

方差的平方根，用于比较不同数据集的离散程度。





数据分布形态的图形表示

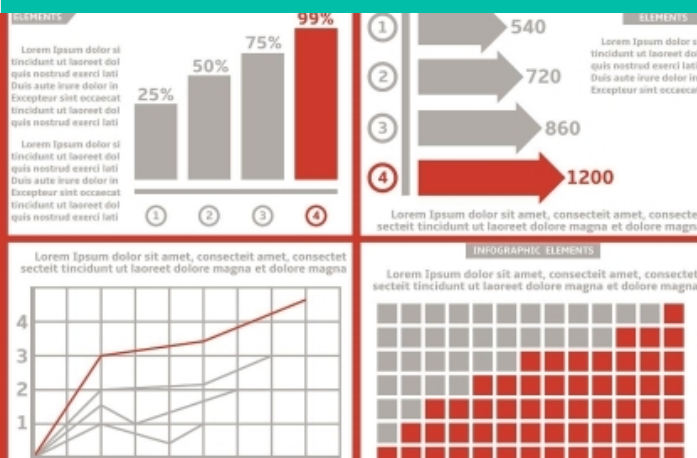
直方图

用矩形的面积表示各组频数，矩形的高度表示频数密度，反映数据的分布情况。



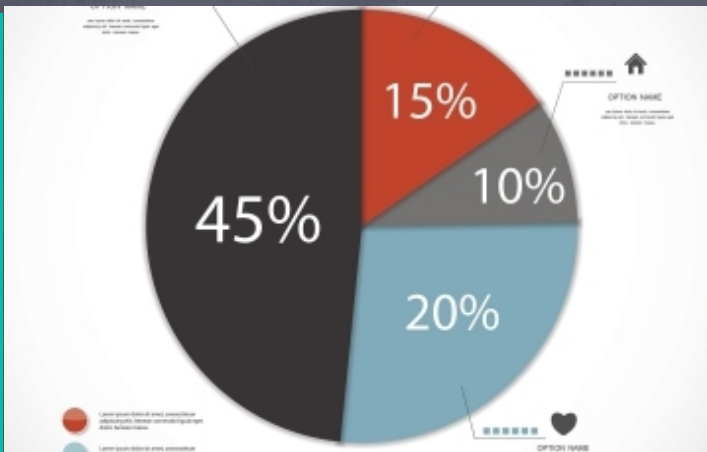
折线图

用线段的升降表示各组频数的变化情况，适用于表示时间序列数据。



箱线图

用箱体、须线和异常点表示数据的分布情况，可直观识别数据中的异常值和偏态情况。





03

概率论基础

Chapter





事件与概率

● 事件

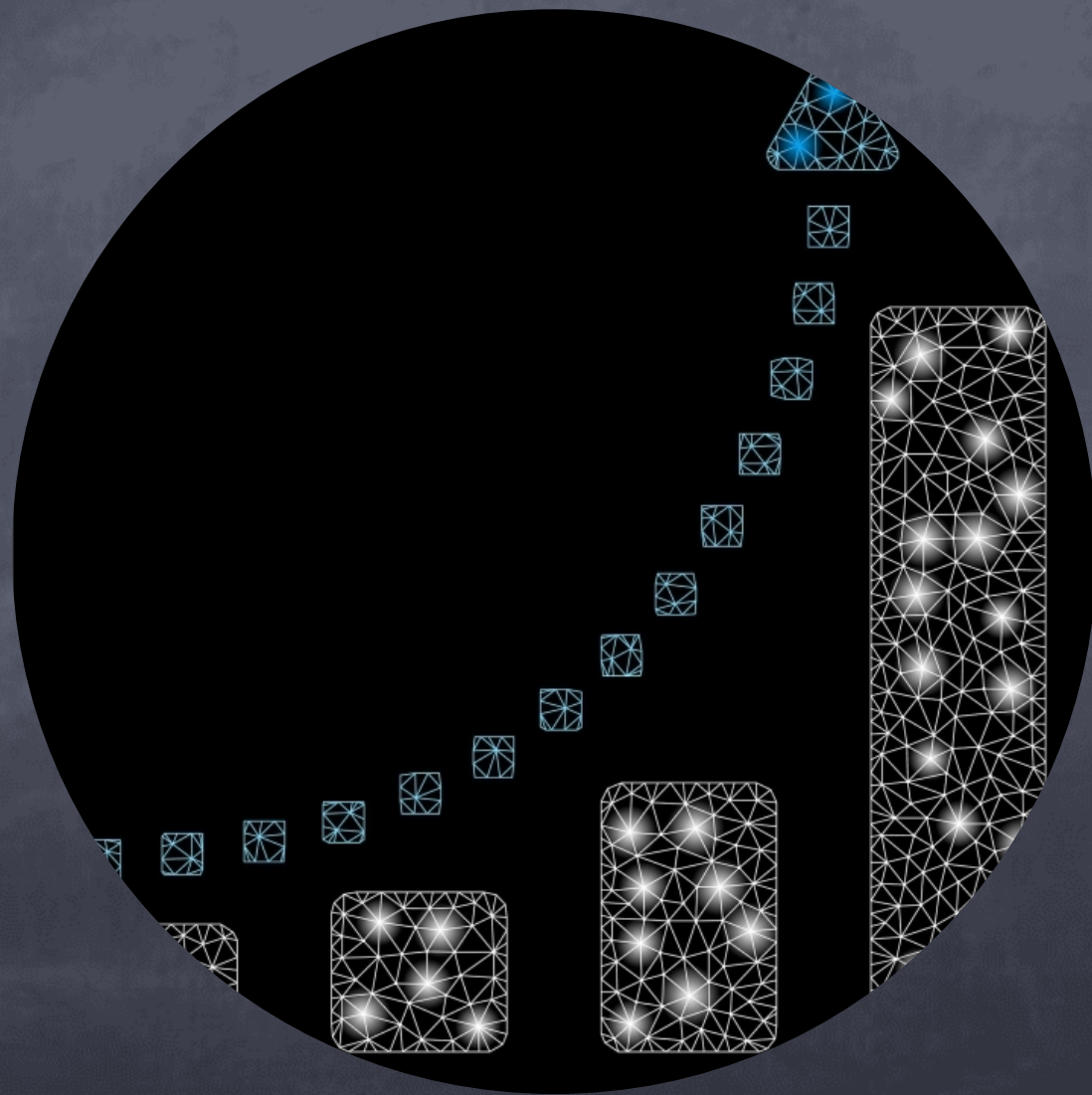
在一定条件下，某现象可能发生或不可能发生，称该现象为随机事件，简称事件。

● 概率

描述事件发生的可能性大小的数值，常用P表示。

● 事件的运算

包括事件的包含、相等、和事件（并）、积事件（交）、差事件及互斥事件等。





随机变量及其分布



01

随机变量

将随机试验的结果数量化，用实数来表示，这样的实数称为随机变量。

02

离散型随机变量及其分布

取值可一一列出的随机变量称为离散型随机变量，其分布可用分布律描述。

03

连续型随机变量及其分布

取值充满某个区间的随机变量称为连续型随机变量，其分布可用概率密度函数描述。



期望与方差

期望

描述随机变量取值的平均水平，是随机变量的重要数字特征之一。对于离散型随机变量，期望是所有可能取值与其对应概率的乘积之和；对于连续型随机变量，期望是概率密度函数与自变量的乘积在整个取值范围内的积分。

方差

描述随机变量取值的离散程度，即各取值与期望的偏离程度。方差越大，说明随机变量的取值越分散；方差越小，说明随机变量的取值越集中。



大数定律与中心极限定理



大数定律

揭示了当试验次数足够多时，频率稳定于概率的现象。即当 n 很大时，事件 A 发生的频率与事件 A 发生的概率有较大偏差的可能性很小。

中心极限定理

揭示了当样本量足够大时，样本均值的分布近似于正态分布的现象。无论总体服从什么分布，只要样本量足够大（通常要求 $n \geq 30$ ），样本均值的分布就近似于正态分布。



04

参数估计

Chapter





点估计

● 矩估计法

利用样本矩来估计总体矩，如用样本均值估计总体均值。

● 最大似然估计法

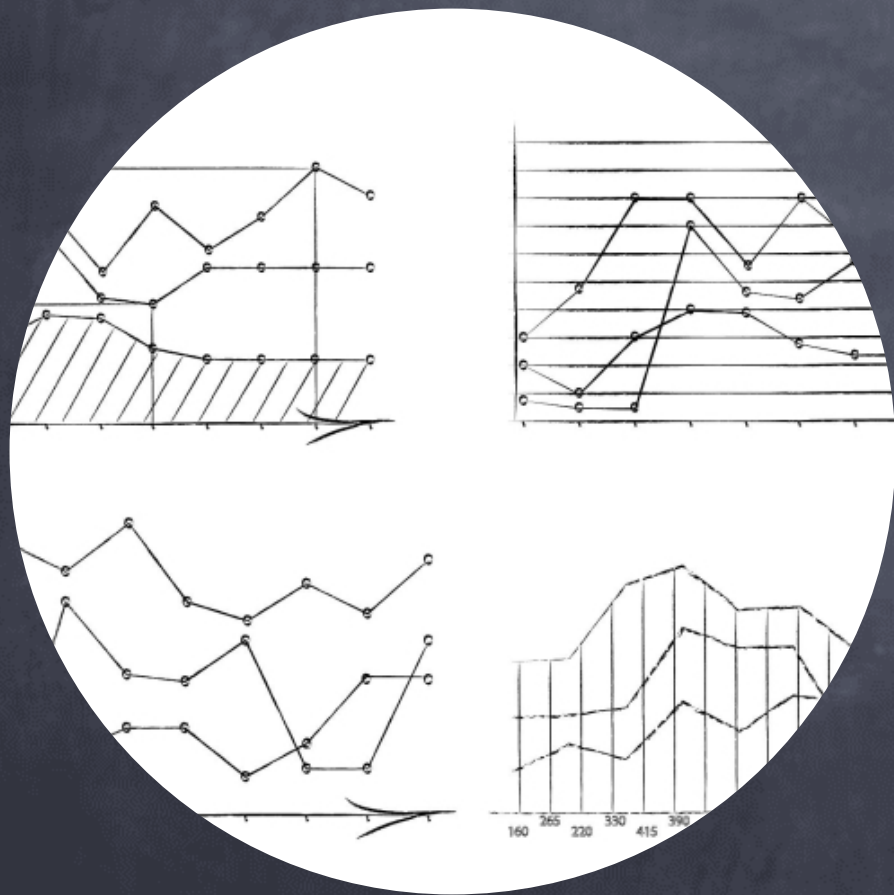
根据样本信息选择使得似然函数达到最大的参数值作为估计值。

● 贝叶斯估计法

基于贝叶斯定理，结合先验信息和样本信息得到后验分布，进而进行参数估计。



区间估计



置信区间

根据样本数据计算得到的一个区间，用于估计未知参数的真实值所在范围，该区间具有一定置信水平。

枢轴量法

构造一个包含未知参数的枢轴量，并根据其分布特性确定置信区间的上下限。

自助法

通过对样本进行重复抽样，生成大量自助样本，进而计算置信区间。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：
<https://d.book118.com/186015145230010130>