

09-12-Statistics

Created on 20220605.

Last modified on 2022 年 6 月 5 日.

目录

Chapter 1 Introduction

数理统计学 a: 抽样理论 (包括抽样分布、抽样调查等), b: 假设检验, c: 非参数统计, d: 方差分析, e: 相关回归分析, f: 统计推断, g: 贝叶斯统计 (包括参数估计等), h: 试验设计, i: 多元分析, j: 统计判决理论, k: 时间序列分析, l: 数理统计学其他学科。

应用统计数学 a: 统计质量控制, b: 可靠性数学, c: 保险数学, d: 统计模拟。

Chapter 2 抽样理论

抽样理论 (包括抽样分布、抽样调查等)

Chapter 3 假设检验

Chapter 4 非参数统计

Chapter 5 方差分析

Chapter 6 相关回归分析

Chapter 7 统计推断

Chapter 8 贝叶斯统计

贝叶斯统计（包括参数估计等）

Chapter 9 试验设计

Chapter 10 多元分析

Chapter 11 统计判决理论

Chapter 12 时间序列分析

Chapter 13 数理统计学其他学科

Chapter 14 统计质量控制

Chapter 15 可靠性数学

Chapter 16 保险数学

Chapter 17 统计模拟

Chapter 18 Information Theory

香农第一定理（可变长无失真信源编码定理。霍夫曼编码也称为最佳编码，由 Huffman 1952 年提出，压缩的极限是不同符号用不同的编码，如果做到每个符号的代码长度等于它出现概率的对数，则编码总长度就是信息熵。）香农第二定理（有噪信道编码定理）香农第三定理（保失真度准则下的有失真信源编码定理）

Shannon, 1948 年 10 月, A Mathematical Theory of Communication, 提出信息熵:

$$H(X) = - \sum_i p_i \log p_i \quad (18.1)$$

提出比特单位。一段信息的信息量是固定的，称为信息熵。无论怎么压缩，信息熵是无失真信源编码的极限值若编码的平均码长小于信息熵值，必然发生差错（也就是有损）。

信道容量 C 计算公式:

$$C = B * \log(1 + \frac{S}{N}) \quad (18.2)$$

B 为信道带宽；S/N 为信噪比，通常用分贝（dB）表示。噪声大，信噪比接近 0，C 结果接近 0。离路由器越远，信噪比越小，网速约下降。

18.1 Models

统计模型

18.1.1 Fitts' Law

Paul Fitts 于 1954 年提出。是一个人机互动以及人体工程学中人类活动的模型。它预测了快速移动到目标区域所需的时间是目标区域的距离和目标区域大小的函数。费兹法则多用于表现指点动作的概念模型。无论是用手或手指进行接触，或是在电脑屏幕上用假想的设备（例如，鼠标）进行虚拟的触碰。

费兹法则可用多种不同公式表达，比较普通用的是，用一维的 Shannon 公式（Mackenzie，约克大学教授提出，因其与香农定律相似而命名）：从一个起始位置移动到一个最终目标所需的时间

由两个参数来决定, 到目标的距离和目标的大小 (上图中的 D 与 W), 用数学公式表达为时间 $T = a + b \log_2(D/W + 1)$ 。 T 是完成动作的时间 a 代表装置开始结束的时间, b 表示该装置的速度, 这些常数可从实测数据进行直线近似取得。 D 是起始位置到目标中心的距离。 w 是目标区域在运动维上的宽度。