

统计理论 All of statistics学习笔记

Contents

前言	1
第一部分：概率论	2
1. 概率论	2
2. 随机变量	3
2. 随机变量	3
3. 期望	3
4. 不等式	3
5. 随机变量的收敛	3
第二部分：统计推断	3
6. 模型，统计推断和学习	3
参数模型和非参数模型	3
统计推断的基本概念	4
8. Bootstrap	4
10. 假设检验和p值	4
第三部分：统计模型和方法	4

前言

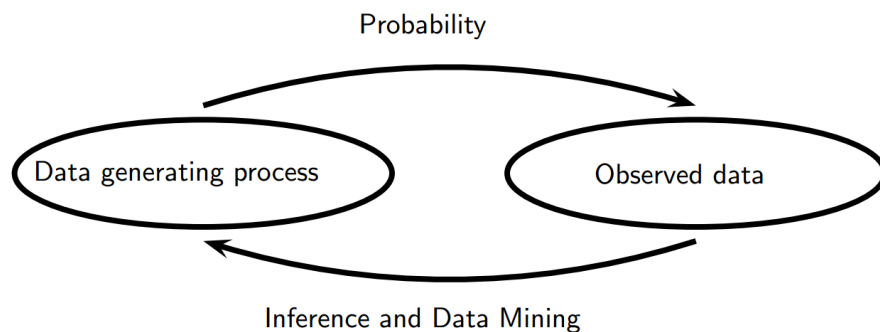
第一部分概率论：不确定性，统计推断的基础

 给一个数据生成的过程，结果的属性是什么？

第二部分统计推断和他的近亲，数据挖掘和机器学习

 给出数据的结果，我们能说数据生成的过程是怎么样的？

预测，分类，聚类，估计是统计推断。数据分析，机器学习和数据挖掘是统计推断的应用



统计、数据挖掘对应含义字典

统计	计算机科学	含义
估计	学习	使用数据去估计未知的属性
分类	监督学习	通过X预测离散的Y
聚类	非监督学习	把数据分组
数据	训练样本	$(X_1, X_1), \dots, (X_n, Y_n)$
协变量	特征	X'_i s
分类器	假设	协变量到结果的映射
假设	—	参数空间 Θ 的子集
置信空间	—	空间包含未知属性
有向非循环图	贝叶斯网络	多维分布给定条件独立关系
贝叶斯推断	贝叶斯推断	统计方法通过数据更新信念
频率推断	—	通过保障频率方法的统计方法
large deviation bounds	PAC学习	uniform bounds on probability of errors

第一部分：概率论

1. 概率论

概率是用于量化不确定性的数学语言

术语小结

Ω	样本空间
ω	结果（点或元素）
A	事件（样本空间的子集）
A^c	A的补充（非A）
$A \cup B$	并（A或者B）
$A \cap B$ or AB	交（A且B）
$A - B$	差（ ω 属于A但不属于B）
$A \subset B$	A属于B

术语小结

\emptyset	假事件（永远为假）
Ω	真事件（永远为真）

概率

定义：函数 P 用一个实数 $P(A)$ 去代表每个事件 A 的概率分布或者概率测量，如果满足一下三个公理

1: $P(A) \geq 0$ 对于每个事件 A

2: $P(\Omega) = 1$

3: 如果 A_1, A_2, \dots 是非连续的，那么有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

有两种方式解释频率 $P(A)$ ，一种是频率，另一种是可信度。分别对应频率派和贝叶斯派。

2. 随机变量

2. 随机变量

3. 期望

4. 不等式

5. 随机变量的收敛

第二部分：统计推断

6. 模型，统计推断和学习

统计推断，在计算机中称为“学习”，一个典型的统计推断的问题是

- 给定样本 X_1, \dots, X_n ，我们如何推断 F ？

在一些情况下我们只推断 F 的特征比如平均值

参数模型和非参数模型

参数模型：通过有限个参数可以确定的模型，比如正态分布

非参数模型：无法通过有限个参数确定的模型，比如CDF

统计推断的基本概念

8. Bootstrap

10. 假设检验和p值

假设检验更多是告诉我们去拒绝某个结论，除非这个结论是唯一的，不然我们无法假定 H_0 就是正确的。

10.8 拟合检验

用卡方拟合检验的严重限制：

如果拒绝 H_0 那么我们可以根据我们不接受这个模型。如果我们不拒绝 H_0 ，我们也不能说这个模型是正确的。因为这个检验没有足够的。

第三部分：统计模型和方法