

我是8位的

I am 8 bits, what about you?

随笔 - 205, 文章 - 0, 评论 - 103, 阅读 - 101万

导航

博客园

首页

新随笔

联系

订阅

管理

<

2022年3月

>

日	一	二	三	四	五	六
27	28	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2
3	4	5	6	7	8	9

公告

你的支持是我的动力

欢迎关注微信公众号“我是8位的”



昵称：我是8位的

园龄：4年7个月

粉丝：288

关注：5

+加关注

盖楼抽奖

#她的梦想在发光#

HWD科技女性故事有奖征集

分享最打动你的科技女性故事

活动时间：2022年3月8日-3月18日

马上参与



搜索

找找看

谷歌搜索

常用链接

我的随笔

我的评论

我的参与

最新评论

我的标签

积分与排名

积分 - 457097

排名 - 1198

概率笔记10——矩估计和最大似然

估计

生活中我们经常估计一些数值，比如从家到学校要走多久？一颗大白菜大概多少斤？凭什么估计出具体数值呢？“估计”不是瞎猜，是根据已有数据计算的。从家到学校往返过多次，手上也拿过无数颗白菜，此时我们会凭借心中的尺度计算出一个大约的数值。

矩估计

矩估计，即矩估计法，也称“矩法估计”，是利用已有样本估计期望值的一种方法。

某个问题的数学期望客观存在的数学特征，是一个具体的数值，只是这个数值计算起来需要知道一些“已知条件”，而这些已知条件在现实世界中并不可知。幸运的是，我们可以随时得到一些随机样本，利用这些样本估计一个数值：

$$E(X) \cong \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

戴帽子的等号表示估计。每个 x_i 都是一个简单随机样本，并且我们认为每个样本都是等可能的，这实际上是真实世界中一种不得已而为之的办法。在大数定律下的作用下，这个估计将会逐渐稳定，逼近真实值。

现在有甲、乙两个射击运动员站在我们面前，他们的平均成绩并没有贴在身上，如何判断他们的成绩呢？

一个符合经验的做法是让他们各打10枪，然后计算均值。比如 x_i 是甲第 i 枪的成绩，那么我们对甲的估计是 $(x_1+x_2+...+x_{10})/10$ 。这里使用的是简单的均值，并没有任何概率参与，原因是我们并不知道甲打出每一环的概率，只好认为是等权平均。数学期望是运动员的真实成绩，我们在计算数学期望时需要已知运动员打出每一环的概率，然而“已知”在并不总是存在于现实世界，因此才退而求其次，使用“估计”。

独立同分布

独立同分布是概率论中的一个概念，即一组数据彼此间互不干扰，在现实环境里随机出现。

独立已经介绍过多次，射击比赛中的每一次射击都是独立的，不会因为本次的结果影响下一枪（抛开运动员心理状态的变化）。如果是从一堆白球中取一个黑球，随着白球的减少，下次取出黑球的概率会不断变大，则不能称每次的取球行为相互独立。

“同分布”的意思是每次都从特定的集合中取结果，比如掷骰子，每次都从1~6中取结果，则称样本是同分布。如果夹杂着几个12面的骰子，则

随笔分类 (211)

- ★★资源下载★★(1)
- Java并发编程(1)
- 程序员的数学(24)
- 单变量微积分(31)
- 多变量微积分(24)
- 概率(24)
- 机器学习(27)
- 软件设计(1)
- 数据分析(6)
- 数据结构与算法(27)
- 随笔(5)
- 线性代数(34)
- 项目管理(2)
- 转载(4)

随笔档案 (205)

- 2021年2月(1)
- 2020年3月(2)
- 2020年2月(6)
- 2020年1月(4)
- 2019年12月(7)
- 2019年11月(15)
- 2019年9月(3)
- 2019年8月(6)
- 2019年7月(1)
- 2019年6月(8)
- 2019年5月(3)
- 2019年4月(5)
- 2019年3月(7)
- 2019年2月(3)
- 2019年1月(7)
- 更多

阅读排行榜

- 1. 使用Apriori进行关联分析 (一) (29768)
- 2. 线性代数笔记12——列空间和零空间 (28772)
- 3. FP-growth算法发现频繁项集 (一)——构建FP树(24430)
- 4. 寻找“最好” (2) ——欧拉-拉格朗日方程(23099)
- 5. 多变量微积分笔记3——二元函数的极值(22772)

评论排行榜

- 1. 隐马尔可夫模型 (一) (8)
- 2. 线性代数笔记12——列空间和零空间 (7)
- 3. 线性代数笔记3——向量2 (点积) (7)
- 4. FP-growth算法发现频繁项集 (一)——构建FP树(5)
- 5. 寻找“最好” (2) ——欧拉-拉格朗日方程(4)

推荐排行榜

- 1. 寻找“最好” (2) ——欧拉-拉格朗日方程(7)
- 2. FP-growth算法发现频繁项集 (一)——构建FP树(7)
- 3. 线性代数笔记3——向量2 (点积) (6)
- 4. FP-growth算法发现频繁项集 (二)——发现频繁项集(5)
- 5. 隐马尔可夫模型 (一) (5)

最新评论

- 1. Re:线性代数笔记3——向量2 (点积)
如果点积小于0，即夹角小于90°，这个写错了。应该是夹角大于90°
--猫猫猫猫大人

样本不是同分布的。

未知的密度函数

在连续型变量中，只要我们知道变量的概率密度 $f(x)$ ，就可以知道它的期望：

$$E(X) = \int_{-\infty}^{\infty} x P(x) = \int_{-\infty}^{\infty} x f(x) dx$$

问题是 $f(x)$ 通常是未知的，只知道它的模型，但不确定具体的模型参数。我们设这个未知的参数是 θ ，概率密度是 $f(x; \theta)$ ，表示 f 受到 θ 的影响，数学期望公式：

$$E(X) = \int_{-\infty}^{\infty} x P(x) = \int_{-\infty}^{\infty} x f(x; \theta) dx$$

实际上 θ 是一个向量，例如：

$$\theta = (\theta_0, \theta_1, \theta_2), \quad f(x; \theta) = \theta_0 + \theta_1 x + \theta_2 x^2$$

示例 设连续型随机变量的概率密度是 $f(x; \theta) = \begin{cases} (\theta + 1)x^\theta, & 0 < x < 1 \\ 0, & \text{others} \end{cases}$ ，求 θ 的矩估计量。

可以先计算出 X 的矩估计：

$$\bar{x} \cong E(X) = \int_{-\infty}^{\infty} x f(x; \theta) dx = \begin{cases} \int_0^1 x (\theta + 1)x^\theta dx, & 0 < x < 1 \\ 0, & \text{others} \end{cases}$$

只有 $0 < x < 1$ 的时候才能计算 θ ：

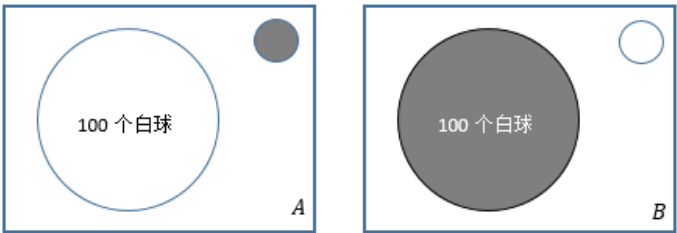
$$\int_0^1 x (\theta + 1)x^\theta dx = (\theta + 1) \int_0^1 x^{\theta+1} dx = (\theta + 1) \frac{x^{\theta+2}}{\theta + 2} \Big|_{x=0}^{x=1} = \frac{\theta + 1}{\theta + 2}$$
$$\frac{\theta + 1}{\theta + 2} \cong \bar{x} \Rightarrow \hat{\theta} = \frac{2\bar{x} - 1}{1 - \bar{x}}$$

最大似然

最大似然估计方法 (Maximum Likelihood Estimate, MLE) 也称为最大似似估计或极大似然估计，是建立在最大似然原理的基础上的一种统计方法。

最大似然的含义

“似然”就是“可能性”的意思。我们经常听到“最大似然”，这个词来源于实际，下图解释了它的含义。



A、B是两个一模一样的箱子，A中有100个白球和1个黑球，B中有100个黑球和1个白球。现在从两个箱子中随意取出一个小球，结果是黑球，这

2. Re:线性代数笔记10——矩阵的LU分解写的很好，不过LU分解的前提是错的，LU分解只需要第三个条件，如果允许行置换就是下面写到的PLU，可以分解所有矩阵

--wiki3D

3. Re:单变量微积分笔记20——三角替换1 (sin和cos) 很nice

--尹保棕

4. Re:线性代数笔记24——微分方程和exp(At) 有些图片挂了呢

--ccchendada

5. Re:寻找“最好” (2) ——欧拉-拉格朗日方程

提个issue，最速降线中

$v = \{2gh\}^{1/2}$ 与配图不一致，建议以起点为原点，向右伸出x轴，向下伸出y轴建立坐标系

--trustInU

个黑球是从哪个箱子中取出的？第一反应是“最有可能从B中取出的”，这符合通常的经验。这里的“最有可能”就是“最大似然”的意思。

似然和似然函数

假设有一个独立同分布的数据集 X ，它的参数是 θ 。现在从 X 中取出一些样本 $x = \{x_1, x_2, \dots, x_n\}$ ， $P(x; \theta)$ 表示给定参数 θ 时，从 X 中取得这些样本 x 的可能性：

$$P(x; \theta) = \begin{cases} \prod_{i=1}^n p(x_i; \theta) & \text{离散型} \\ \prod_{i=1}^n f(x_i; \theta) & \text{连续型} \end{cases}$$

其中 $P(x; \theta)$ 类似于条件概率，但不等于条件概率，因为 θ 只是一个密度函数中的参数，并不是一个事件。

假设现在 θ 有两个取值 θ_1 和 θ_2 ，对于 X 中的一些样本 $x = \{x_1, x_2, \dots, x_n\}$ ，如果 $P(x, \theta_1) > P(x, \theta_2)$ ，就认为 θ_1 对产生 x 的可能性（似然性）要大于 θ_2 ， $P(x, \theta_1)$ 和 $P(x, \theta_2)$ 就是似然，是对参数 θ 产生样本 x 的可能性的度量。

还是以射击为例，假设按运动员的成绩由高到低分为一级、二级、三级，甲打出了10枪 $x = \{9, 9, 10, 10, 8, 9, 9.5, 9.5, 9.5, 9\}$ 。运动员的级别相当于影响成绩参数 θ ，当 θ 等于一级时，甲打出这个成绩的可能性较高。

现在需要根据给定样本 x 来求 $P(x; \theta)$ ，由于样本是已知的，将所有 x 的值代入上面的公式，将得到一个只有 θ 的式子，这个式子称为 θ 的似然函数，记为 $L(x; \theta)$ 或 $L(\theta)$ ：

$$L(\theta) = \begin{cases} \prod_{i=1}^n p(x_i; \theta) & \text{离散型} \\ \prod_{i=1}^n f(x_i; \theta) & \text{连续型} \end{cases}$$

最大似然估计

知道了似然函数，最大似然估计就很容易理解了：对于一个给定的样本集，挑选使得 $P(x; \theta)$ 能够达到最大时的参数 作为 θ 的估计值，使得：

$$L(\hat{\theta}) = \max_{\theta} L(x; \theta)$$

最终将求得 θ 的一个估值，在 时，似然函数的值最大。

极值点通常是在导数等于0的点取得，因此可以通过下式求得 θ ：

$$\frac{dL}{d\theta} = 0$$

如果 θ 是 n 维向量，则：

$$\frac{\partial L}{\partial \theta_1} = 0, \frac{\partial L}{\partial \theta_2} = 0, \dots, \frac{\partial L}{\partial \theta_n} = 0$$

对于一些特殊的密度函数（比如指数密度函数）来说，直接求 $dL/d\theta$ 太过繁琐，由于 L 与 $\ln L$ 在同一 θ 处取到极值，所以也经常使用：

$$\frac{d\ln(L)}{d\theta} = 0$$

示例

示例1

设样本的总体分布率为： $P\{X=x\}=p^x(1-p)^{1-x}$ ，求 p 在观察样本 $\{x_1, x_2, \dots, x_n\}$ 下的最大似然估计量。

这里只不过是把 θ 用 p 表示，现在我们做一下替换，变成熟悉的形式：

$$p^x(1-p)^{1-x} \Rightarrow \theta^x(1-\theta)^{1-x} \Rightarrow p(x_i; \theta) = \theta^{x_i}(1-\theta)^{1-x_i}$$

$$L(\theta) = \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{(x_1+x_2+\dots+x_n)}(1-\theta)^{(1-x_1+1-x_2+\dots+1-x_n)}$$

$L(\theta)$ 是 θ 的指数形式，换成对数更为简单：

$$\begin{aligned}\ln L &= \ln \theta^{(x_1+x_2+\dots+x_n)}(1-\theta)^{(1-x_1+1-x_2+\dots+1-x_n)} \\ &= \ln \theta^{(x_1+x_2+\dots+x_n)} + \ln(1-\theta)^{(1-x_1+1-x_2+\dots+1-x_n)}\end{aligned}$$

根据对数的基本公式继续计算：

$$\begin{aligned}\log_a M^n &= n \log_a M \\ \ln L &= \ln \theta^{(x_1+x_2+\dots+x_n)} + \ln(1-\theta)^{(1-x_1+1-x_2+\dots+1-x_n)} \\ &= \left(\sum_{i=1}^n x_i\right) \ln \theta + \left(\sum_{i=1}^n 1-x_i\right) \ln(1-\theta) \\ \frac{d\ln(L)}{d\theta} &= \left(\sum_{i=1}^n x_i\right) \frac{1}{\theta} - \left(\sum_{i=1}^n 1-x_i\right) \frac{1}{1-\theta} = 0 \\ \text{let } a &= \sum_{i=1}^n x_i, \quad \sum_{i=1}^n 1-x_i = n-a \\ \text{then } \frac{d\ln(L)}{d\theta} &= \frac{a}{\theta} - \frac{(n-a)}{1-\theta} = 0 \\ \Rightarrow \theta &= \frac{a}{n} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}\end{aligned}$$

示例2

总体样本服从参数为 λ 的指数分布， $\{x_1, x_2, \dots, x_n\}$ 是观察样本，求 λ 的最大似然估计值。

总体样本的概率密度是：

$$f(x; \theta) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$L(\lambda) = \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

$$\ln L(\lambda) = \ln \lambda^n e^{-\lambda \sum_{i=1}^n x_i} = \ln \lambda^n + \ln e^{-\lambda \sum_{i=1}^n x_i} = n \ln \lambda - \lambda \sum_{i=1}^n x_i$$

$$\begin{aligned}\frac{d\ln(L)}{d\lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n x_i = \frac{n}{\lambda} - n\bar{x} = 0 \\ \lambda &= \frac{1}{\bar{x}}\end{aligned}$$

作者: 我是8位的

出处: <http://www.cnblogs.com/bigmonkey>

本文以学习、研究和分享为主，如需转载，请联系本人，标明作者和出处，非商业用途！

扫描二维码关注公作者众号“我是8位的”



随笔

分类: 概率

标签: 矩估计, 最大似然, 最大似然估计, 极大似然

[好文要顶](#)[关注我](#)[收藏该文](#)



我是8位的
关注 - 5
粉丝 - 288
[+加关注](#)

2

0

 推荐

 反对

« 上一篇: [概率笔记9——大数定律](#)
» 下一篇: [最小二乘法 \(1\) ——线性问题](#)

posted on 2019-07-01 17:44 我是8位的 阅读(10047) 评论(0) 编辑 收藏 举报

[刷新评论](#) [刷新页面](#) [返回顶部](#)

 登录后才能查看或发表评论，立即 [登录](#) 或者 [逛逛](#) 博客园首页

【推荐】华为 HWD 2022 故事征集，分享最打动你的科技女性故事
【推荐】华为开发者专区，与开发者一起构建万物互联的智能世界



WebForms升级到ASP.NET Core
fineui.com

- 编辑推荐:
- 革命性创新，动画杀手锏 @scroll-timeline
 - 戏说领域驱动设计（十二）—— 服务
 - ASP.NET Core 6框架揭秘实例演示[16]：内存缓存与分布式缓存的使用
 - .Net Core 中无处不在的 Async/Await 是如何提升性能的？
 - 分布式系统改造方案 —— 老旧系统改造篇



最新新闻:

- 乔布斯的创业搭档: 他缺乏工程师才能, 不得不锻炼营销能力来弥补
 - 美国大厂码农薪资曝光: 年薪18万美元, 够养家, 不够买海景房
 - 两张照片就能转视频! Google提出FLIM帧插值模型
 - Android 再推“杀手级”功能, 可回收 60% 存储空间
 - 溺在理财暴雷潮的投资人: 本金63万, 月兑25元不够卖菜
- » 更多新闻...

Powered by:

博客园

Copyright © 2022 我是8位的

Powered by .NET 6 on Kubernetes