

1. 最大似然估计

1. 实例 (引出背景)

实例: 现有黄球, 红球若干个在一个布袋中(无穷, 拿不完)

要求其分布特点.

即

X	$X=\text{黄}$	$X=\text{红}$
p	0	$1-0$

从袋中抽 100 个球 (随机地), 以抽样方式整体分布:

取出: 40 个黄球, 60 个红球

最大似然估计的原则: 存在即合理

为什么会是 40 个黄球, 60 个红球; 而不是其他结果

只能说明是因为抽取 40 个黄球, 60 个红球的组合概率最大
或者说, 参数 0 条件下, 抽 40 个黄球, 60 个红球概率是 最大 的。

上述问题的解为:

$$\text{则 } L(0) = 0^{40} (1-0)^{60} \Rightarrow \ln L(0) = 40 \ln 0 + 60(1-0)$$

$$\text{对 } L(0) \text{ 求导, } \Rightarrow \frac{dL(0)}{d0} = 40 \cdot \frac{1}{0} - \frac{60}{1-0} = 0$$

$$\text{极值点 } \Rightarrow 0 = \frac{2}{5}$$

所以: 当 $0 = \frac{2}{5}$ 时, 抽取 100 个球, 其中 40 个黄球, 60 个红球概率最大;

所以我们就认为 ~~在~~ 原始样本中, $0 = \frac{2}{5}$

因为只有当 $0 = \frac{2}{5}$ 时, 我们才能一次抽样中, 抽 40 个黄球, 60 个红球。

实例 2:

· 实例1: 现要调查一个班级男生身高分布, 已知其服从分布 $\theta = (\mu, \sigma)$

试: 先随机采样, 100个男生, 记录他们的身高。

男生的身高分别为 x_1, x_2, \dots, x_{100}

现在认为: 在 μ 与 σ 的条件下, 抽取出来是 $(x_1, x_2, \dots, x_{100})$ 的概率最大。

而抽取出来 $(x_1, x_2, \dots, x_{100})$ 的概率是:

$$\prod_{i=1}^{100} P(x_i, \theta).$$

如何找到 μ 与 σ 使 $L(\theta) = L(x_1, x_2, \dots, x_{100}; \theta) = \prod_{i=1}^n P(x_i; \theta)$

最大。求, 得极值就知 μ, σ 。

则利用最大似然理论计算出的样本分布特征有: μ, σ 。

2. 似然函数

反映的是在不同参数 θ 取值下, 取得当前样本集的可能性。
因此, 将参数 θ 相对于样本 x 的似然函数, 记为 $L(\theta)$

$$L(\theta) = L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n P(x_i; \theta)$$

现在, 让 $L(\theta)$ 最大, 求个

$$\hat{\theta} = \arg \max (L(\theta))$$



扫描全能王 仓

试: 求导数, 令导数为0, 得到似然方程。
解似然方程, 得到的参数即为所求。

3. 总结

总结：一种统计方法，去估计样本的分布特征值 μ, σ 等。

原因是总体样本过多，我们无法直接统计，所以抽取部分样本出来辅助计算。

现在考虑一个问题：

为什么在原始分布特征下，抽取出的样本是这些，而不是其他



在原始分布特征下，这些样本被同时抽取的概率最大



写出联合概率公式（似然函数），求最大这个概率

↓
最大似然函数



对似然函数求极值，得最优。极值点，就是原始分布中的梯度量值。