

第1章 概率

条件概率

$$P(A|B) = \frac{P(AB)}{P(B)}$$

全概率

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$

贝叶斯

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$$

“两两独立”和“相互独立”

第2章 一维分布

离散分布

01 分布

二项分布

泊松分布

几何分布

连续分布

均匀分布

指数分布

正态分布

第3章 二维分布

联合分布

$X \backslash Y$	-1	0	2
0	0.1	0.3	0
1	0.1	0.05	0.45

边缘分布

则关于X的边缘分布律为:

$$p_{i\cdot} = P\{X = x_i\} = \sum_j p_{ij} \quad (i=1, 2, \dots)$$

关于Y的边缘分布律为:

$$p_{\cdot j} = P\{Y = y_j\} = \sum_i p_{ij} \quad (j=1, 2, \dots)$$

也就是: 当求x的边缘分布则将y看成是无穷的边界

https://blog.csdn.net/fuoleng_103

条件分布

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}, f(y|x) = \frac{f(x,y)}{f_X(x)}.$$

独立性

两个事件互不影响

第4章 数字特征

期望, 均值

即可能取值中的理想平均值

方差

总体差异性

协方差

衡量两个变量之间的总体误差,。一般情况下， 相关系数越大表明相关程度就越高。

相关系数

衡量两个变量的相关程度

矩估计

通过样本估计总体的数字特征。设 A_k 是原点矩， M_k 是中心距

$$a_{nk} = \frac{1}{n} \sum_{i=1}^n X_i^k$$
$$m_{nk} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^k$$

名称	含义
一阶原点矩	平均值
二阶原点矩	平均能量
一阶中心距	0
二阶中心距	方差
三阶中心距	偏度
四阶中心距	峭度

第5章 样本

样本空间

随机试验的所有结果， 条件概率可能会缩小样本空间

统计量

χ^2 分布

T 分布

F 分布

第6章 大数定律以及中心极限定理

大数定律

当样本容量足够大时，样本均值趋向于总体均值

中心极限定理

当样本容量足够大时，样本均值服从正态分布，其均值为总体均值，方差为总体方差的 $1/n$

第7章 参数估计

矩估计

样本原点矩是总体原点矩的无偏估计，样本中心矩经过一些简单的调整也可以使之无偏。

最大似然估计

常常对似然函数取对数

机器学习中常用的 EM 算法就是一种特殊的求极大似然估计的数值方法