连续随机变量下的贝叶斯公式

茆(mao)诗松 《概率论与数理统计》第一版 (十章)

似然: likelihood 可能性 相似,像,源自于最大似然估计

表示:哪个原因最有可能(最像)导致了结果

A班 99男 1女 ,B班,99女 1男

先随机抽一个班,再从此班抽一个人进行观测,结果是女

此女最像是从B班抽出来的

$$P(\underset{因}{状态}\mid\underset{果}{观测}) = \eta P(\underset{果}{观测}\mid\underset{因}{状态})P(状态)$$

后验分布 $P(状态_1\mid观测)$ 似然概率 $P(观测\mid状态_1)$

$P(状态_2\mid观测)$ $P(观测\mid状态_2)$

独立 未必 没有函数关系 $Y=f(X)$ Y与X可能独立,也可能不独立

必然事件 $Y=X+1$ $P(X=1)=1$ $P(Y=2)=1$ $P(X=1,Y=2)=1$ 独立

随机事件 设有一个正态概率分布 $N(\mu,\sigma^2)$ (μ,σ)未知,从此分布中,抽取n个独立的样本,$X_1\cdots X_n$,则$X_1\cdots X_n$独立同分布,则随机变量

$$\bar{X}=\frac{X_1+X_2+\cdots X_n}{n} \qquad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$ 相互独立

连续随机变量的贝叶斯公式

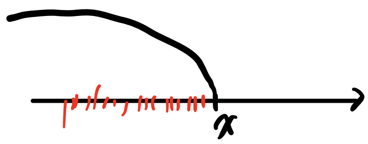
离散 $P(X=x|Y=y)=\dfrac{P(Y=y|X=x)P(X=x)}{P(Y=y)}$

连续 $P(X<x|Y=y)=\dfrac{P(Y=y|X<x)P(X<x)}{P(Y=y)}$? $=0$

贝叶斯公式无法直接运用于连续随机变量

化积分为求和

$X<x \Rightarrow \sum_{u=-\infty}^{x} X=u$



化区间为点的累加

$$\therefore\ P(X<x|Y=y)=\sum_{u=-\infty}^{x}P(X=u|Y=y)=\sum_{u=-\infty}^{x}\frac{P(Y=y|X=u)P(X=u)}{P(Y=y)}$$

$$=\lim_{\varepsilon\to 0}\sum_{u=-\infty}^{x}\frac{P(y<Y<y+\varepsilon|X=u)P(u<X<u+\varepsilon)}{P(y<Y<y+\varepsilon)}$$

$$=\lim_{\varepsilon\to 0}\sum_{u=-\infty}^{x}\frac{(f_{Y|X}(\xi_1|u)\cdot\varepsilon)(f_X(\xi_2)\cdot\varepsilon)}{f_Y(\xi_3)\cdot\varepsilon}$$

其中 $\xi_1\in(y,y+\varepsilon)$

$\xi_2\in(u,u+\varepsilon)$

$\xi_3\in(y,y+\varepsilon)$

$$=\lim_{\varepsilon\to 0}\sum_{u=-\infty}^{x}\frac{f_{Y|X}(y|u)f_X(u)}{f_Y(y)}\cdot\varepsilon$$

$$=\int_{-\infty}^{x}\frac{f_{Y|X}(y|u)f_X(u)}{f_Y(y)}du \Rightarrow \int_{-\infty}^{x}\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}dx$$

连续随机变量的贝叶斯公式

$$P(X<x|Y=y)=\int_{-\infty}^{x}\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}dx$$

$\Downarrow$

$$\int_{-\infty}^{x}f_{X|Y}(x|y)dx \Rightarrow f_{X|Y}(x|y)=\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

连续随机变量的贝叶斯公式

$$= \eta f_{Y|X}(y|x) f_X(x)$$ ✓

$$f_Y(y) = \int_{-\infty}^{+\infty} f(y,x)\,dx$$ （联合概率密度与边缘概率密度的关系）

$$= \int_{-\infty}^{+\infty} f_{Y|X}(y|x) f(x)\,dx = C$$

$$\eta = \frac{1}{\int_{-\infty}^{+\infty} f_{Y|X}(y|x) f_X(x)\,dx}$$ 归一化