

一、最基本的概念 (讨论连续型变量)

① 概率密度

- x 位于区间 (a, b) 的概率 $P(x \in (a, b)) = \int_a^b p(x) dx$
- 概率密度需满足 $\begin{cases} p(x) \geq 0 \\ \int_{-\infty}^{+\infty} p(x) dx = 1 \end{cases}$

② 期望与协方差

- 在概率分布 $p(x)$ 下, 函数 $f(x)$ 的平均值称为 $f(x)$ 的期望 (Expectation)

连续变量: $E[f] = \int p(x) f(x) dx$

离散变量: $E[f] = \sum_x p(x) f(x)$

- 方差 (Variance) 度量了 $f(x)$ 在均值 $E[f(x)]$ 附近变化性的大小.

$$\begin{aligned} \text{Var}[f] &= E[(f(x) - E[f(x)])^2] \\ &= E[f^2(x)] - E[f(x)]^2 \end{aligned}$$

$$\text{Var}[x] = E[x^2] - E[x]^2$$

- 对于两个随机变量 x, y , 定义协方差 (Covariance), 表示在多大程度上 x, y 会共同变化. 如果 x 和 y 相互独立, 那么它们的协方差为 0.

$$\begin{aligned} \text{cov}[x, y] &= E[(x - E(x))(y - E(y))] \\ &= E[xy - xE(y) - yE(x) + E(x)E(y)] \\ &= E(xy) - E(x)E(y) - E(y)E(x) + E(x)E(y) \end{aligned}$$

$$\text{cov}[x, y] = E(xy) - E(x)E(y)$$

③ 贝叶斯概率 (Bayesian)

贝叶斯定理: $P(w|D) = \frac{P(D|w) \cdot P(w)}{P(D)}$

- 其中 $P(w|D)$ 称为后验概率, 在观测到 D 之后估计 w 的不确定性.

- $p(D(w))$ 称为似然函数. 表示在不同参数向量 w 下, 观测数据 y 能出现的概率.