概率论1

目录

为什么要学习概率论？

认识随机变量

- 可以随机取不同值
- 描述状态分布
- 分类：
  - 离散型随机变量
  - 连续型随机变量

概率分布

- 描述随机变量或一簇随机变量在每一个可能取到的状态的可能性的大小
- 离散型随机变量
  - 概率质量函数
- 连续型随机变量
  - 概率密度函数

离散型随机变量与概率质量函数

- 概率质量函数
  - 函数表示P
  - 随机变量x
  - $P(x=x_1)=0$
  - $P(x=x_1)=1$
- 联合概率分布
  - 多个变量的概率分布
  - $P(x=x_1, y=y_1)$

P需满足的条件

- P 的定义域必须是x 所有可能状态的集合。
- $\forall x_1 \in x, 0 \leq P(x) \leq 1$
  - 不可能发生的事件概率为0
  - 能够确保一定发生的事件概率为1
- $\sum_{x_i \in x} P(x = x_i) = 1$
  - 归一化
  - 保证事件发生概率小于1

示例

连续型随机变量与概率密度函数

- 概率密度函数
  - 函数表示 P
  - 随机变量 x
- 联合概率分布
  - 多个变量的概率分布
  - $P(x=x_1, y=y_1)$

P需满足的条件

- P 的定义域必须是x 所有可能状态的集合。
- $\forall x_1 \in x, P(x) \geq 0$
  - 并不要求$P(x) \leq 1$
- $\int p(x)dx = 1$
  - 归一化
  - 保证事件发生概率小于1

连续型 VS 离散型

- 概率密度函数$\mathrm{p}(x)$ 不直接给出特定状态的概率
- 定义落在面积为δx 的无限小的区域内的概率为$\mathrm{p}(x)\delta x$
- 通过概率密度函数求积分来获得点集的真实概率质量
- x 落在区间[a, b] 的概率是$\int_{[a,b]} p(x)dx = 1$

示例

边缘概率

- 已知一组变量的联合概率分布，求子集的概率分布
- 离散型随机变量x 和y，已知$p(x, y)$

$$\forall x \in \mathrm{x}, P(\mathrm{x} = x) = \sum_{y} P(\mathrm{x} = x, \mathrm{y} = y).$$

- 离散型随机变量x 和y，已知$p(x, y)$

$$p(x) = \int p(x, y) dy.$$

条件概率

- 给定$x = x_1$, 则$y = y_1$ 发生的条件概率为 $p(y = y_1|x = x_1)$，其中

$$P(\mathrm{y} = y \mid \mathrm{x} = x) = \frac{P(\mathrm{y} = y, \mathrm{x} = x)}{P(\mathrm{x} = x)}.$$

条件概率的链式法则

- 链式法则：任何多维随机变量的联合概率分布，都可以分解成只有一个变量的条件概率相乘的形式：

$$P(\mathrm{x}^{(1)}, \ldots, \mathrm{x}^{(n)})$$

$$= P(\mathrm{x}^{(1)}) \Pi_{i=2}^{n} P(\mathrm{x}^{(i)} \mid \mathrm{x}^{(1)}, \ldots, \mathrm{x}^{(i-1)}).$$

条件概率的链式法则推导

$$
\begin{aligned}
P(\mathrm{a}, \mathrm{b}, \mathrm{c}) &= P(\mathrm{a} \mid \mathrm{b}, \mathrm{c}) P(\mathrm{b}, \mathrm{c}) \\
P(\mathrm{b}, \mathrm{c}) &= P(\mathrm{b} \mid \mathrm{c}) P(\mathrm{c}) \\
P(\mathrm{a}, \mathrm{b}, \mathrm{c}) &= P(\mathrm{a} \mid \mathrm{b}, \mathrm{c}) P(\mathrm{b} \mid \mathrm{c}) P(\mathrm{c}).
\end{aligned}
$$

随机变量的独立性

- 两个随机变量x 和y，如果它们的概率分布可以表示成两个因子的乘积形式，并且一个因子只包含x另一个因子只包含y，我们就称这两个随机变量是相互独立的

$$\forall x \in \mathrm{x}, y \in \mathrm{y}, p(\mathrm{x} = x, \mathrm{y} = y)$$

$$= p(\mathrm{x} = x) p(\mathrm{y} = y).$$

随机变量的条件独立性

- 如果关于x 和y 的条件概率分布对于z 的每一个值都可以写成乘积的形式，那么这两个随机变量x 和y 在给定随机变量z 时是条件独立的

$$\forall x \in \mathrm{x}, y \in \mathrm{y}, z \in \mathrm{z},$$

$$p(\mathrm{x} = x, \mathrm{y} = y \mid \mathrm{z} = z)$$

$$= p(\mathrm{x} = x \mid \mathrm{z} = z) p(\mathrm{y} = y \mid \mathrm{z} = z)$$

期望

- 函数f(x) 关于某分布P(x) 的期望是指，当x由P 产生，f 作用于x 时，f(x) 的平均值
- 离散型随机变量

$$\mathbb{E}_{\mathrm{x}\sim P}[f(x)] = \sum_x P(x)f(x),$$

- 连续型随机变量

$$\mathbb{E}_{\mathrm{x}\sim p}[f(x)] = \int p(x)f(x)dx.$$

示例

方差

- 对x 依据它的概率分布进行采样时，随机变量x 的函数值会呈现多大的差异：

$$\mathrm{Var}(f(x)) = \mathbb{E}\left[(f(x) - \mathbb{E}[f(x)])^2\right]$$

示例

协方差

- 定义两个变量线性相关性的强度以及这些变量的尺度

$$\mathrm{Cov}(f(x), g(y))$$

$$= \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

示例