条件

xyfJASON

1 条件概率

1.1 条件概率

定义：设 A, B 为两个事件且 $\mathbb{P}(B) > 0$，则称 $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ 为 B 发生的条件下 A 发生的概率，记作 $\mathbb{P}(A \mid B)$，即：

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \mathbb{P}(B) > 0$$

公理化：可以证明，条件概率是一个公理化定义下的概率：

1. 非负性：显然；
2. 规范性：
$$\mathbb{P}(\Omega \mid B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$
3. 可数可加性：
$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{n} A_i \mid B\right) &= \frac{\mathbb{P}\left(\left(\bigcup\limits_{i=1}^{n} A_i\right) \cap B\right)}{\mathbb{P}(B)} \\
&= \frac{\mathbb{P}\left(\bigcup\limits_{i=1}^{n} (A_i \cap B)\right)}{\mathbb{P}(B)} \\
&= \frac{\sum\limits_{i=1}^{n} \mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} \\
&= \sum_{i=1}^{n} \mathbb{P}(A_i \mid B)
\end{aligned}$$

证毕。

因此，**概率的所有性质都适用于条件概率**。

乘法公式：由条件概率定义有：

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A \mid B)$$

推广：

$$\mathbb{P}(A_1A_2\cdots A_n)=\mathbb{P}(A_1)\mathbb{P}(A_2\mid A_1)\mathbb{P}(A_3\mid A_1A_2)\cdots\mathbb{P}(A_n\mid A_1A_2\cdots A_{n-1})$$

1.2 全概率公式

设 $A_1,A_2,\cdots,A_n$ 互不相容，$B\subset A_1\cup A_2\cup\cdots\cup A_n$，则：

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

证：

$$\begin{aligned}\mathbb{P}(B)&=\mathbb{P}\left(B\cap\left(\bigcup_{i=1}^{n}A_i\right)\right)\\&=\mathbb{P}\left(\bigcup_{i=1}^{n}(A_i\cap B)\right)\\&=\sum_{i=1}^{n}\mathbb{P}(A_i\cap B)\\&=\sum_{i=1}^{n}\mathbb{P}(A_i)\mathbb{P}(B\mid A_i)\end{aligned}$$

证毕。

1.3 贝叶斯公式

设 $A_1,A_2,\cdots,A_n$ 互不相容且 $\mathbb{P}(A_i)>0$，$B\subset A_1\cup A_2\cup\cdots\cup A_n$，则：

$$\mathbb{P}(A_i\mid B)=\frac{\mathbb{P}(A_i\cap B)}{\mathbb{P}(B)}=\frac{\mathbb{P}(A_i)\mathbb{P}(B\mid A_i)}{\sum\limits_{j=1}^{n}\mathbb{P}(A_j)\mathbb{P}(B\mid A_j)}$$

关于贝叶斯公式的理解：视事件 A_i 是导致事件 B 发生的原因，我们对于事件 A_i 已有一个先验概率 $\mathbb{P}(A_i)$，现在事件 B 发生了，这必然给我们带了一定的信息，于是我们可以由此修正 A_i 发生的概率，得到 $\mathbb{P}(A_i\mid B)$，即后验概率。

1.4 条件独立

我们知道，两个事件 A,B 相互独立，是指 $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$. 由于条件概率是符合概率公理化定义的概率律，我们可以对条件独立有类似的定义：

给定事件 C，若事件 A,B 满足：

$$\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\mathbb{P}(B \mid C)$$

则称 A, B 在给定事件 C 下条件独立。

注意，A, B 条件独立并不能推出 A, B 独立，反过来亦不成立。

2 条件分布

2.1 条件分布列

离散随机变量 Y 取定某个值 y 后，离散随机变量 X 的条件分布列为：

$$p_{X|Y}(x \mid y) = \mathbb{P}(X = x \mid Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

于是我们有联合分布列、边缘分布列、条件分布列之间的关系：

$$\begin{aligned}
&p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x \mid y) \\
&p_X(x) = \sum_y p_{X,Y}(x,y) = \sum_y p_Y(y)p_{X|Y}(x \mid y)
\end{aligned}$$

2.2 条件概率密度函数

连续随机变量 Y 取定某个值 y 后，连续随机变量 X 的条件概率密度函数为：

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

同样我们有联合概率密度函数、边缘概率密度函数、条件概率密度函数之间的关系：

$$\begin{aligned}
&f_{X,Y}(x,y) = f_Y(y)f_{X|Y}(x \mid y) \\
&f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}y = \int_{-\infty}^{+\infty} f_Y(y)f_{X|Y}(x \mid y)\mathrm{d}y
\end{aligned}$$

3 条件期望

3.1 条件期望

对于离散随机变量，

$$\mathbb{E}[X \mid Y=y]=\sum_{x} x p_{X \mid Y}(x \mid y)$$

对于连续随机变量，

$$\mathbb{E}[X \mid Y=y]=\int_{-\infty}^{+\infty} x f_{X \mid Y}(x \mid y) \mathrm{d} x$$

对于随机变量的函数，我们类似有:

$$\begin{aligned}
\mathbb{E}[g(X) \mid Y=y] &=\sum_{x} g(x) p_{X \mid Y}(x \mid y) && \text{离散} \\
\mathbb{E}[g(X) \mid Y=y] &=\int_{-\infty}^{+\infty} g(x) f_{X \mid Y}(x \mid y) \mathrm{d} x && \text{连续}
\end{aligned}$$

3.2 全期望定理

$$\begin{aligned}
\mathbb{E} X &=\sum_{y} \mathbb{E}[X \mid Y=y] p_{Y}(y) \\
\mathbb{E} X &=\int_{-\infty}^{+\infty} \mathbb{E}[X \mid Y=y] f_{Y}(y) \mathrm{d} y
\end{aligned}$$