第八次习题课

王瑞

April 21, 2022

Statisticians, like artists, have the bad habit of falling in love with their models. – George E. P. Box

1 随机变量的独立性

说明两个随机变量独立的方法有许多, 以下是一个总结

1. 设随机变量 X 和 Y 的联合分布函数为 $F(x,y)$, 边缘分布函数为 $F_X(x)$ 和 $F_Y(y)$, 如果对于任意实数，恒有
$$F(x,y) = F_X(x)F_Y(y) \tag{1}$$
则称随机变量 X 与 Y 独立.

2. 随机变量 X 和 Y 独立的充分必要条件是由 X 生成的任何事件与 Y 生成的任何事件独立, 即, 对于任意的 Borel 集合 A 和 B, 有
$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \tag{2}$$

3. 如果 X 与 Y 独立, 那么对于任意函数 $f(x)$ 和 $g(y)$, $f(X)$ 与 $g(X)$ 独立.

4. 如果 X 和 Y 是离散型随机变量, 其概率分布为 $P(X_i = x_i, Y_j = y_j) = p_{ij}$, 边缘分布分别为 p_i^X 和 p_j^Y, 那么 X 和 Y 独立的充要条件是 $p_{ij} = p_i^X p_j^Y$.

5. 如果 X 和 Y 是连续型随机变量, 其密度函数为 $f(x,y)$, 边缘密度函数分别为 $f_X(x)$ 和 $f_Y(y)$, 那么 X 和 Y 独立的充要条件是 $f(x,y) = f_X(x)f_Y(y)$.

例 1.1 ***(圆盘上的均匀分布)*** 考虑随机变量 X 和 Y 均匀地在圆盘 $S_0 = \{(x,y) : x^2 + y^2 \leq 1\}$ 上取值. 很显然, X 与 Y 不独立（可以有多种说明 X 和 Y 不独立的方法）. 假设我们使用极坐标 $(R, \Theta) \in [0,1] \times [0.2\pi]$ 重新参数化随机向量 (X,Y), 那么可以说明 R 和 Θ 是独立的.

例 1.2 设随机变量 X 和 Y 独立同分布, 且
$$P(X = -1) = P(Y = -1) = P(X = 1) = P(Y = 1) = \frac{1}{2} \tag{3}$$
求 $P(X = Y)$

例 1.3 设随机向量 (X,Y) 的联合密度函数如下, 试问 X 和 Y 是否独立.

1.

$$f(x,y)=\begin{cases}xe^{-x-y} & (x>0,\ y>0)\\ 0, & (\text{其他.})\end{cases} \tag{4}$$

2.

$$f(x,y)=\begin{cases}2 & (0<x<y<1)\\ 0, & (\text{其他.})\end{cases} \tag{5}$$

3.

$$f(x,y)=\begin{cases}24xy & (0<x<1,\ 0<y<1,\ 0<x+y<1)\\ 0, & (\text{其他.})\end{cases} \tag{6}$$

例 1.4 设 $X\sim N(0,1)$, 试说明 X 与 X^2 不独立.

定理 1.1 设随机向量的概率分布或者密度函数为 $f(x,y)$. 那么 X 和 Y 独立的充分必要条件是对任意的 $x\in\mathcal{R}$ 和 $y\in\mathcal{R}$, 存在函数 $g(x)$ 和 $h(y)$, 使得

$$f(x,y)=g(x)h(y) \tag{7}$$

2 条件分布

设 $f(x,y)$ 为联合密度函数或者联合概率分布, $f_Y(y)$ 为边缘密度函数或概率分布, 那么条件分布或条件密度函数的计算为:

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)} \tag{8}$$

可以验证, 固定 y 后 (即把 $f_{X|Y}(x|y)$ 视为 x 的函数),$f_{X|Y}(x|y)$ 是一个密度函数 (或者离散情况下的概率分布).

条件期望的计算:

$$E[g(X)|Y=y]=\int_{\mathcal{R}}g(x)f_{X|Y}(x|y)dx \tag{9}$$

注意 $E[X|Y]$ 和 $E[X|Y=y]$ 的区别.

例 2.1 设 *2* 维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}3x & (0<x<1,\ 0<y<x)\\ 0, & (\text{其他.})\end{cases} \tag{10}$$

试求条件密度函数 $f(y|x)$

例 2.2 设 *2* 维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{21}{4}x^2y & (x^2\le y\le 1)\\ 0, & (\text{其他.})\end{cases} \tag{11}$$

试求条件概率 $P(Y\ge 0.75|X=0.5)$.

例 2.3 设随机向量 (X, Y) 为连续型随机向量, 已知 X 的密度函数 $f_X(x)$ 及对一切 x, 在 $X = x$ 的条件下 Y 的条件密度函数 $f_{Y|X}(y|x)$. 求

1. (X, Y) 的联合密度函数 $f(x, y)$
2. Y 的边缘密度函数 $f_Y(y)$
3. 条件密度函数 $f_{X|Y}(x|y)$

例 2.4 试证明 *(可以假设 (X, Y) 为连续型随机向量))*

$$E[E[X|Y]] = E[X] \tag{12}$$

References

[1] Casella, George, and Roger L. Berger. Statistical inference(2nd ed). Cengage Learning, 2002.

[2] 茆诗松, 程依明, 濮晓龙. " 概率论与数理统计 (第二版)." (2012).

[3] 龙永红. " 概率论与数理统计 (第四版)." (2012).

[4] Michael, Perlman. Probability and mathematical statistics, `https://sites.stat.washington.edu/people/mdperlma/STAT%20512%20MDP%20Notes.pdf`