变量独立性

Guangyao Zhao

2022-08-09

Contents

在观察两个(或多个)变量时只有一个标准,就是定义: $P\{X=x, Y=y\} = P\{X=x\}P\{Y=y\}$

对于离散型变量

假设随机变量 (X, Y) 具有分布律: $P\{X=x, Y=y\} = p^2(1-p)^{x+y-2}, 0<p<1$, 其中 x, y 均为正整数, 问 X, Y 是否独立。

X 的边缘分布如下:

$$
\begin{aligned}
P\{X=x\} &= P\{X=x, Y=1\} + P\{X=x, Y=2\} + \cdots + P\{X=x, Y=+\infty\} \\
&= p^2(1-p)^{x+1-2} + p^2(1-p)^{x+2-2} + \cdots + p^2(1-p)^{x+\infty-2} \\
&= \sum_{y=1}^{+\infty} p^2(1-p)^{x+y-2} \\
&= p^2(1-p)^{x-2}\sum_{y=1}^{+\infty}(1-p)^y \\
&= p^2(1-p)^{x-2}\frac{1-p}{p} \\
&= p(1-p)^{x-1}
\end{aligned}
$$

注意, 此处有个等比数列的小结论:

$$\sum_{k=1}^{+\infty} x^k = \frac{x}{1-x}, 0 < x < 1$$

根据对称性：

$$P\{Y = y\} = p(1-p)^{y-1}$$

综上：

$$P\{X = x\}P\{Y = y\} = p^2(1-p)^{x+y-2}$$

即独立性得以证明。

对于连续型变量

假设二维随机变量 (X, Y) 具有概率密度：

$$f(x,y) = \begin{cases} \quad 2e^{-(2x+y)}, x > 0, y > 0 \\ \quad 0, \text{other} \end{cases}$$

变量 X 的边缘分布：

$$F_X(x) = \int_0^{+\infty} 2e^{-(2x+y)} dy = \begin{cases} 2e^{-2x}, x > 0 \\ 0, \text{other} \end{cases}$$

同理，变量 Y 的边缘分布：

$$F_Y(y) = \int_0^{+\infty} 2e^{-(2x+y)} dx = \begin{cases} e^{-y}, y > 0 \\ 0, \text{other} \end{cases}$$

两者相乘：

$$F_X(x)F_Y(y) = 2e^{-(2x+y)}$$

综上可知 (X, Y) 相互独立。