高中统计学 · 第肆课练习

何濯羽

2024 年 2 月 16 日

1 独立性的定义

随机变量 X 与 Y 的联合分布列如下所示。请问 X 与 Y 相互独立吗?

		Y		
		0	1	
X	0	0.2	0.6	0.8
	1	0.05	0.15	0.2
		0.25	0.75	

2 条件概率与条件期望的计算

		Y		
		0	1	2
	0	0.3	0.15	0.05
X	1	0.2	0.05	0.05
	2	0.1	0.05	0.05

随机变量 X 与 Y 的联合分布列如上所示。

1) 写出 X 和 Y 各自的边际概率质量函数。

2) 写出给定 $Y = 0$ 时 X 的条件概率质量函数。

3) 计算 $E(X|Y = 0)$。

3 相关系数的计算与应用

随机变量 X 与 Y 的联合分布列如下所示。请计算 X 与 Y 的相互系数，并回答 X 与 Y 是否线性相关。

		Y	
		0	2
X	0	0.1	0.3
	2	0.3	0.3

4 变量和的方差【拓展】

求证：$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$。

5 强大的独立性【拓展】

已知 X 与 Y 是一对相互独立的随机变量。它们的数学期望分别为 μ_X 和 μ_Y,方差分别为 σ_X^2 和 σ_Y^2。请用 μ_X、μ_Y、σ_X^2、σ_Y^2 写出 XY 与 Y 相关系数的表达式。

答案

独立性的定义

通过计算，我们发现 $Pr(X=x,Y=y)=Pr(X=x)\cdot Pr(Y=y),\forall x,y$ 成立，所以 X 与 Y 相互独立。

条件概率与条件期望的计算

(1) X 的边际 PMF 为

$$f_X(x)=\begin{cases}0.3+0.15+0.05=0.5 & (x=0)\\ 0.2+0.05+0.05=0.3 & (x=1)\\ 0.1+0.05+0.05=0.2 & (x=2)\end{cases}$$

Y 的边际 PMF 为

$$f_Y(y)=\begin{cases}0.3+0.2+0.1=0.6 & (y=0)\\ 0.15+0.05+0.05=0.25 & (y=1)\\ 0.05+0.05+0.05=0.15 & (y=2)\end{cases}$$

(2)

$$f_{X|Y}(x|y=0)=Pr(X=x|Y=0)=\frac{Pr(X=x,Y=0)}{Pr(Y=0)}=\begin{cases}\frac{0.3}{0.6}=\frac{1}{2} & (x=0)\\ \frac{0.2}{0.6}=\frac{1}{3} & (x=1)\\ \frac{0.1}{0.6}=\frac{1}{6} & (x=2)\end{cases}$$

(3)

$$E(X|Y=0)=\sum_{x=0}^{2}x\cdot Pr(X=x|Y=0)=0\times\frac{1}{2}+1\times\frac{1}{3}+2\times\frac{1}{6}=\frac{2}{3}$$

相关系数的计算与应用

根据 X 与 Y 的联合分布列，我们可以得到 X、Y、XY 的分布列：

X 的取值 x	0	2
$Pr(X=x)$	0.4	0.6

Y 的取值 x	0	2
$Pr(Y=y)$	0.4	0.6

XY 的取值	0	4
概率	0.7	0.3

凭此，我们可以得到

$$E(X) = 1.2, \quad E(Y) = 1.2, \quad E(XY) = 1.2$$

于是，

$$Cov(X, Y) = E(XY) - E(X) \cdot E(Y) = 1.2 - 1.2 \times 1.2 = -0.24$$

我们根据 X、Y 的分布列还可以写出 X^2、Y^2 的分布列：

X^2 的取值	0	4
概率	0.4	0.6

Y^2 的取值	0	4
概率	0.4	0.6

凭此，我们可以得到

$$E(X^2) = 2.4, \quad E(Y^2) = 2.4$$

于是，

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 = 2.4 - (1.2)^2 = 0.96 \\ D(Y) &= E(Y^2) - [E(Y)]^2 = 0.96 \end{aligned}$$

故 X 与 Y 的相关系数为

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{-0.24}{\sqrt{0.96} \times \sqrt{0.96}} = -0.25 < 0$$

据此可知 X 与 Y 是线性负相关的关系。

变量和的方差【拓展】

根据方差和协方差的定义式，即可证明

$$\begin{aligned} Var(X+Y) &= E\left\{[X+Y-E(X+Y)]^2\right\} \\ &= E\Big\{\big([X-E(X)]+[Y-E(Y)]\big)^2\Big\} \\ &= E\big\{[X-E(X)]^2\big\} + E\big\{[Y-E(Y)]^2\big\} + 2E\big\{[X-E(X)][Y-E(Y)]\big\} \\ &= Var(X) + Var(Y) + 2Cov(X, Y) \end{aligned}$$

强大的独立性【拓展】

因为 $E(X) = \mu_X$，$E(Y) = \mu_Y$，$D(X) = E(X^2) - [E(X)]^2 = \sigma_X^2$，$D(Y) = E(Y^2) -$

$[E(Y)]^2 = \sigma_Y^2$，所以

$$E(X^2) = \sigma_X^2 + \mu_X^2$$
$$E(Y^2) = \sigma_Y^2 + \mu_Y^2$$

因为 X 与 Y 相互独立，所以

$$\begin{aligned} Cov(XY, Y) &= E(XY^2) - E(XY) \cdot E(Y) \\ &= E(X) \cdot E(Y^2) - E(X) \cdot E(Y) \cdot E(Y) \\ &= \mu_X(\sigma_Y^2 + \mu_Y^2) - \mu_X\mu_Y^2 \\ &= \mu_X\sigma_Y^2 \end{aligned}$$

且

$$\begin{aligned} D(XY) &= E(X^2Y^2) - [E(XY)]^2 \\ &= E(X^2) \cdot E(Y^2) - [E(X) \cdot E(Y)]^2 \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - (\mu_X\mu_Y)^2 \\ &= \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \mu_X^2\sigma_Y^2 \end{aligned}$$

于是，我们可以求得

$$\begin{aligned} Corr(XY, Y) &= \frac{Cov(XY, Y)}{\sqrt{D(XY) \cdot D(Y)}} \\ &= \frac{\mu_X\sigma_Y^2}{\sqrt{(\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \mu_X^2\sigma_Y^2)\,\sigma_Y^2}} \\ &= \frac{\mu_X\sigma_Y}{\sqrt{\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \mu_X^2\sigma_Y^2}} \end{aligned}$$