知识小料

「电计 2203 班」周常规知识整理共享

22

日期：2024-4-6
学科：概率与统计 A
来源：2021 数 I 考研第 16 题

甲、乙两个盒子中有 2 个红球和 2 个白球，选取甲盒中任意一球，观察颜色后放入乙盒，再从乙盒中任取一球，令 X, Y 分别表示从甲盒和乙盒中取到的红球的个数，则 X 与 Y 的相关系数为__________。

首先列出 X, Y 的**取值**，均为 $0, 1$。由于是离散型随机变量，所以可以考虑列分布列，直接得到数字特征。

要计算 $P(X=0, Y=0)$，可以把取球过程拆成**两个步骤**：

- 在甲盒中取一白球；
- 在已知乙盒中多了一个白球的条件下，在乙盒中取一白球。

因此我们可以借用**概率乘法公式**（也就是分步乘法）：

$$P(X=0, Y=0) = P(X=0) \cdot P(Y=0|X=0) = \frac{1}{2} \times \frac{3}{5} = \frac{3}{10} \tag{1}$$

类似地，可以求出另外的概率值，列出**分布列**：

$$\begin{aligned} P(X=0, Y=1) &= P(X=0) \cdot P(Y=1|X=0) = \frac{1}{2} \times \frac{2}{5} = \frac{1}{5} \\ P(X=1, Y=0) &= P(X=1) \cdot P(Y=0|X=1) = \frac{1}{2} \times \frac{2}{5} = \frac{1}{5} \\ P(X=1, Y=1) &= P(X=1) \cdot P(Y=1|X=1) = \frac{1}{2} \times \frac{3}{5} = \frac{3}{10} \end{aligned} \tag{2}$$

$Y \backslash X$	0	1	$p_{i\bullet}$
0	$\frac{3}{10}$	$\frac{1}{5}$	$\frac{1}{2}$
1	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{2}$
$p_{\bullet j}$	$\frac{1}{2}$	$\frac{1}{2}$	1

表 1: 二维随机变量 (X, Y) 的分布列

有了分布列，就可以计算 (X, Y) 的**相关系数**了。

$$\rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \tag{3}$$

这样就转化为计算 $E(X), E(Y), E(XY), D(X), D(Y)$ 了。

由「**联合确定边缘**」的思想，我们可以写出相关一维随机变量的分布列。其中注意 $P(XY=1)=P(X=1,Y=1)=\dfrac{3}{10}$。

X	0	1
P	$\frac{1}{2}$	$\frac{1}{2}$

(a) X 的分布列

Y	0	1
P	$\frac{1}{2}$	$\frac{1}{2}$

(b) Y 的分布列

XY	0	1
P	$\frac{7}{10}$	$\frac{3}{10}$

(c) XY 的分布列

图 1: 相关一维随机变量的分布列

可以看到，X, Y, XY 三个随机变量均服从二项分布，$X \sim B(1,p)$，$E(X)=p$，$D(X)=p(1-p)$，因此：

- $E(X)=\frac{1}{2}$，$D(X)=\frac{1}{2}\times(1-\frac{1}{2})=\frac{1}{4}$
- $E(Y)=\frac{1}{2}$，$D(Y)=\frac{1}{2}\times(1-\frac{1}{2})=\frac{1}{4}$
- $E(XY)=\frac{3}{10}$

$$\rho_{XY} = \frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{\frac{3}{10}-\frac{1}{2}\times\frac{1}{2}}{\sqrt{\frac{1}{4}}\times\sqrt{\frac{1}{4}}} = \frac{1}{5}$$

【结论】$\dfrac{1}{5}$

【点评】本题以相关系数为切入点，考察了相关系数的公式、协方差的公式、二维离散型随机变量分布列的求解、条件概率与乘法公式的应用等，算是一道综合性较强的问题。要解决此问题，关键是会运用「联合确定边缘」的思想，以及会把抽样过程分成两个步骤求解。