概率论与数理统计

主讲教师：荣腾中

1929
CHONGQING UNIVERSITY

Chapter 1

随机变量的数字特征

1.1 标准化与相关系数

协方差 $cov(\boldsymbol{X},\boldsymbol{Y})$ 是一个二维数字特征，但 $cov(\boldsymbol{X},\boldsymbol{Y})$ 的取值受两个变量 $\boldsymbol{X},\boldsymbol{Y}$ 各自的量纲影响，数字的意义并不明显。

1.1.1 标准化变量

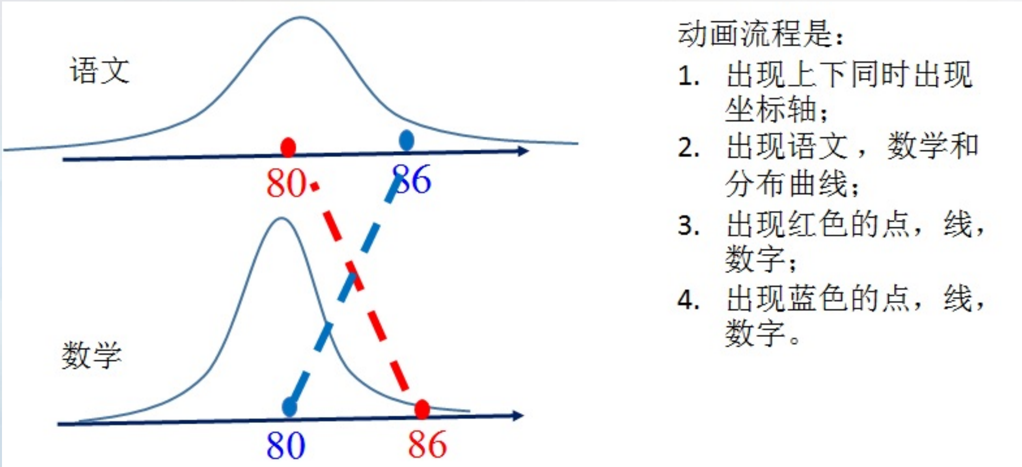
[标准化变量] 如果对随机变量 X 的方差 DX 存在，且 $DX > 0$，称

$$X^* = \frac{X - EX}{\sqrt{DX}}$$

为 X 的标准化变量。

显然，$EX^* = 0$，$DX^* = 1$。

例 1.1.1. 某次考试后，语文和数学的成绩均呈均值 80 的正态分布，但语文方差为 9，数学的方差为 4。甲生语文 80 分，数学 86 分；乙生数学 80 分，语文 86 分。从总分来看，两生总分相同。你觉得这个相同的总分真实地反映了他们的能力了吗？



语文的 86 分对应的标准化成绩 $y^* = \frac{86-80}{3} = 2$。

数学的 86 分对应的标准化成绩 $x^* = \frac{86-80}{2} = 3$。

数学的标准分就比语文的标准分高。

1.1.2 相关系数

[相关系数]　如果 X^*, Y^* 分别为二维随机变量 X, Y 的标准化变量，称

$$\rho(X,Y) = cov(X^*, Y^*)$$

为 X 与 Y 的相关系数。

由协方差的运算性质

$$cov(X^*, Y^*) = cov\left(\frac{X-EX}{\sqrt{DX}}, \frac{Y-EY}{\sqrt{DY}}\right) = \frac{cov(X,Y)}{\sqrt{DXDY}}$$

所以常记为

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{DXDY}}$$

相关系数的性质：

① $|\rho(X,Y)| \leq 1$ ；

证明：令 $X^* = \frac{X-EX}{\sqrt{DX}}$，$Y^* = \frac{Y-EY}{\sqrt{DY}}$，则

$$\begin{aligned} D(X^*+Y^*) &= DX^* + DY^* + 2cov(X^*,Y^*) \\ &= 2 + 2\rho(X,Y) \geq 0 \end{aligned}$$

解不等式有 $\rho(X,Y) \geq -1$。同理，由 $D(X^*-Y^*) \geq 0$，解不等式有 $\rho(X,Y) \leq 1$。

所以

$$|\rho(X,Y)| \leq 1$$

② $\rho(X,Y)=\pm 1$ 的充要条件为 X 与 Y 以概率为 1 完全线性相关，即

$$P\left(Y=\pm\sqrt{\frac{DY}{DX}}X+EY-\sqrt{\frac{DY}{DX}}EX\right)=1$$

证明：由①的证明知

$$\rho(X,Y)=1 \Longleftrightarrow D(X^*-Y^*)=0$$

又因为 $E(X^*-Y^*)=0$，所以 $P\{X^*-Y^*=0\}=1$。于是 $\rho(X,Y)=1$ 的充要条件是

$$P\left(\frac{X-EX}{\sqrt{DX}}=\frac{Y-EY}{\sqrt{DY}}\right)=1$$

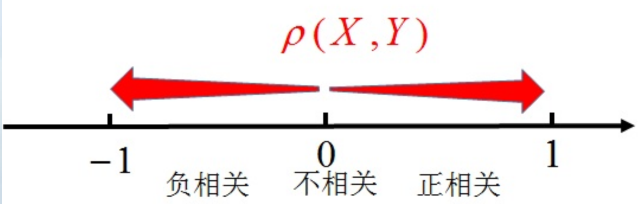
同理

$$\rho(X,Y)=-1 \Longleftrightarrow P\{X^*=-Y^*\}=1$$

结论得证。

[相关] 若 $\rho(X,Y)=0$，则称 X 与 Y 不（线性）相关；若 $\rho(X,Y)\neq 0$，则称 X 与 Y （线性）相关。

- 当 $\rho(X,Y)>0$ 时，X 与 Y 正相关；
- 当 $\rho(X,Y)<0$ 时，X 与 Y 负相关；
- $|\rho(X,Y)|$ 的大小反映了线性相关程度的大小。



动画流程是：
1. 出现坐标轴，0，1，-1三点；
2. 出现r(X,Y)；
3. 显示“负相关，不相关，正相关”；
4. 从0开始向左出现红箭头；
5. 从0开始向左出现红箭头。

③ $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 独立，则 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 不相关，反之则不然。

但对于二维正态随机变量 $(\boldsymbol{X}, \boldsymbol{Y}) \sim \boldsymbol{N}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \boldsymbol{\sigma}_1^2, \boldsymbol{\sigma}_2^2; \boldsymbol{\rho})$，可以证明 $\boldsymbol{\rho} = \boldsymbol{\rho}(\boldsymbol{X}, \boldsymbol{Y})$。从而

$$\boldsymbol{X}, \boldsymbol{Y} \text{ 独立} \Longleftrightarrow \boldsymbol{\rho} = 0 \Longleftrightarrow \boldsymbol{X}, \boldsymbol{Y} \text{ 不相关}$$