

1. 给定如下四个二维数据点：(2, 19)、(9, 6)、(7, 15)、(5, 12)，并定义第一个纬度为 x ，第二个纬度为 y 。

a) 计算 x 和 y 的协方差矩阵，即 $\begin{bmatrix} \text{cov}(x, x) & \text{cov}(x, y) \\ \text{cov}(y, x) & \text{cov}(y, y) \end{bmatrix}$ 。

答案： $\bar{x} = \frac{2+9+7+5}{4} = 5.75$, $\bar{y} = \frac{19+6+15+12}{4} = 13$

$$\text{cov}(x, x) = \frac{1}{4}[(2-5.75)^2 + (9-5.75)^2 + (7-5.75)^2 + (5-5.75)^2] = 6.688$$

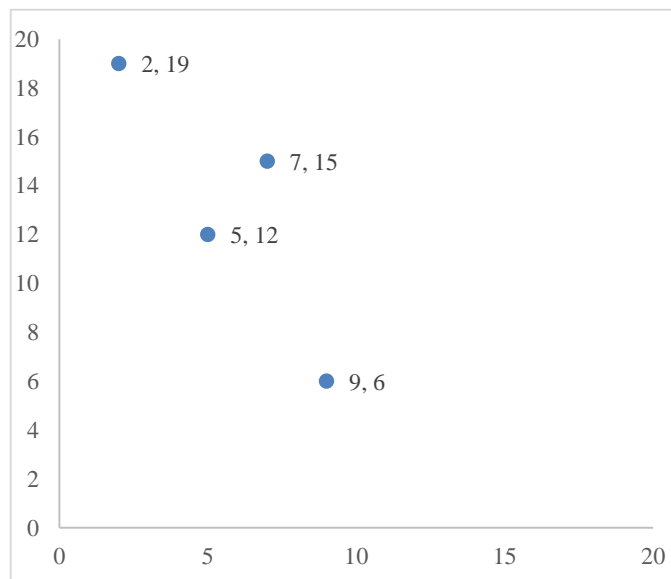
$$\text{cov}(y, y) = \frac{1}{4}[(19-13)^2 + (6-13)^2 + (15-13)^2 + (12-13)^2] = 22.5$$

$$\text{cov}(x, y) = \frac{1}{4}[(2-5.75)(19-13) + (9-5.75)(6-13) + (7-5.75)(15-13) + (5-5.75)(12-13)] = -10.5$$

因此， x 和 y 的协方差矩阵为： $\begin{bmatrix} 6.688 & -10.5 \\ -10.5 & 22.5 \end{bmatrix}$ 。

b) 计算 x 和 y 的相关系数。

答案： x 和 y 的相关系数为： $r_{xy} = \frac{-10.5}{\sqrt{6.688}\sqrt{22.5}} = -0.856$ 。



2. 如果你参加“来做个交易”节目，其规则是这样的：

- ✓ 主持人向你示意三个关闭的大门，然后告诉你每个门后都有一个奖品：一个奖品是一辆车，另外两个是像花生酱和面包这样不值钱的奖品。奖品随机配置，且主持人知道每个奖品在哪。
- ✓ 游戏的目的是要猜哪个门后有车。如果你猜对了就可以拿走。
- ✓ 你先挑选一扇门，我们姑且称之为 A，其他两个称为 B 和 C。
- ✓ 在打开你选中的门 A 之前，为了增加悬念，主持人会打开 B 或 C 中一个没有车的门来增加悬念（如果汽车实际上就是在门 A 后面，那么主持人打开门 B 或门 C 都是安全的，此时他可以随意选择一个；如果汽车在门 C 后面，那么他会打开门 B）。
- ✓ 然后主持人给你一个选择：坚持最初的选择还是换到剩下的未打开的门上。两个选择能够拿走汽车的概率分别是多少？

答案：我们将假设用 H 来表示，数据用 D 来表示。如果数据 D 为：主持人打开了门 B，而且没有汽车在后面，那么可得如下表格：

先验概率 $P(H)$		似然度 $P(D H)$	$P(H)P(D H)$	后验概率 $P(H D)$
假设 1：汽车在门 A 后面	1/3	1/2	1/6	1/3
假设 2：汽车在门 B 后面	1/3	0	0	0
假设 3：汽车在门 C 后面	1/3	1	1/3	2/3

如果数据 D 为：主持人打开了门 C，而且没有汽车在后面，同理可得，假设 1、2、3 的后验概率为 1/3、2/3、0。因此，如果你坚持选择门 A，中奖概率只有 1/3；而如果换到另外一个门，你的机会将是 2/3。