

散点图

一个二维数据点样本 $(x_1, y_1), \dots, (x_n, y_n)$ 的散点图是这些点在 xy 平面上的图形。

当散点图中的数据的 x 和 y 坐标之间存在强相关性时，这种相关性在图中表现得非常明显。例如，下图中右侧的变量显示了明显的线性相关性，而左侧的图则没有明显的相关性。

右侧的图中，线性相关性很明显，但并不完美。点并没有完全落在一条直线上，而是在一条（假想的）直线周围波动。

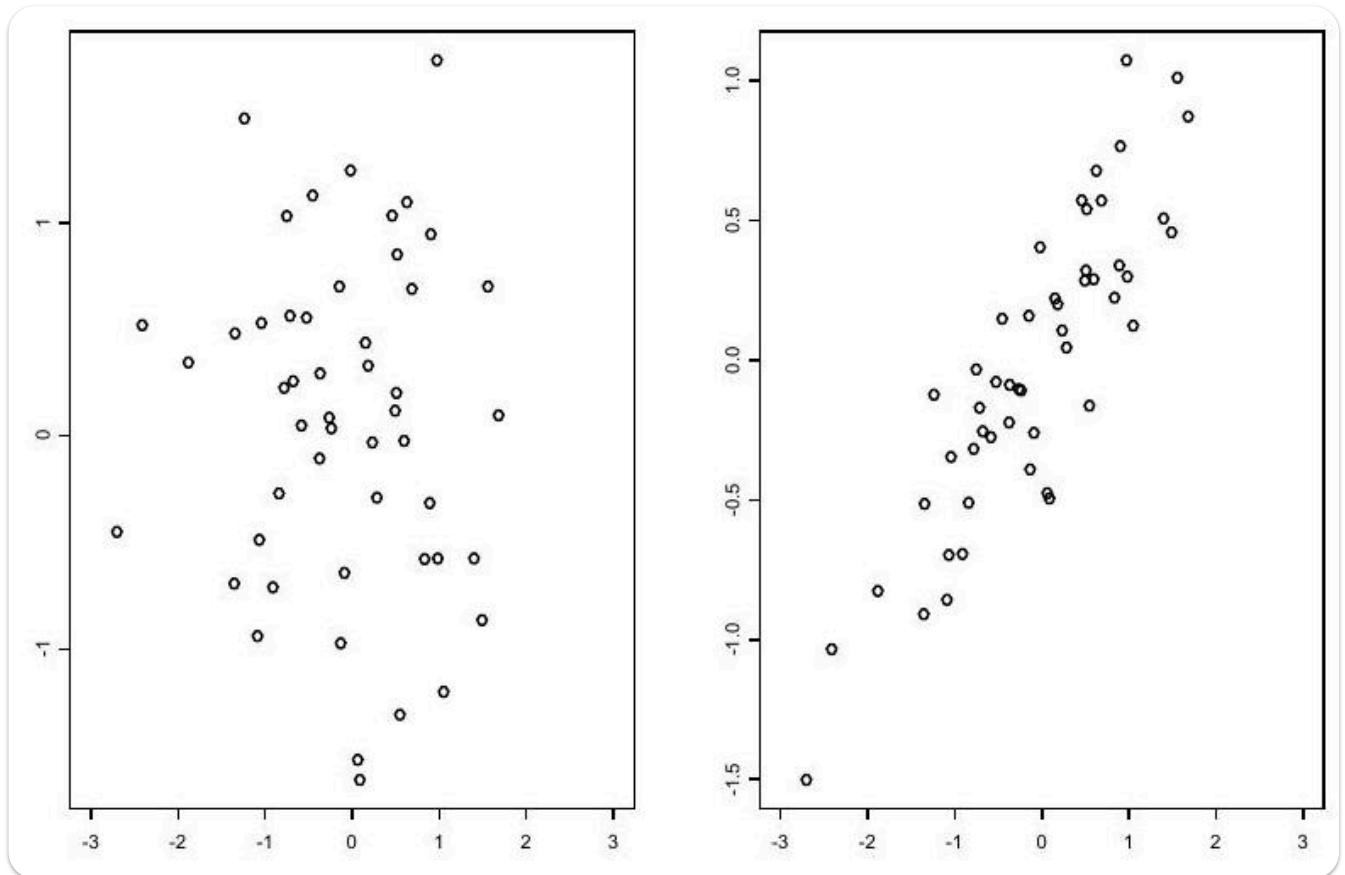


图:两个样本的散点图，每个样本有 50 个点：左侧是独立坐标 ($r_{x,y} = -0.05$)，右侧是具有线性相关性的坐标 ($r_{x,y} = 0.87$)。

散点图还可以用来验证样本 x_1, \dots, x_n 是独立变量实现的常见假设。例如，我们可以绘制点 (x_{2i-1}, x_{2i}) 对于 $i = 1, \dots, \lfloor n/2 \rfloor$ ，或点 (x_i, x_{i+1}) 对于 $i = 1, \dots, n-1$ 。如果假设正确，这些散点图不应显示太多结构。所以我们引入了自相关性的定义。