

第三节

分布形态的测度

1. 偏态系数

2. 标准分数

1. 偏态系数

偏度是指数据分布的偏斜方向和程度，描述的是数据分布对称程度。测度数据分布偏度的统计量称为偏态系数。

$$SK = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{s} \right)^3$$

偏态系数取决于离差三次方的平均数与标准差三次方的比值。

偏态系数的绝对值越大，说明数据分布的偏斜程度越大。

偏态系数取值	0	数据的分布是对称的 		
	正值	数据分布右偏	0 ~ 0.5	轻度右偏
			0.5 ~ 1	中度右偏
			> 1	严重右偏
	负值	数据分布左偏	0 ~ -0.5	轻度左偏
			-0.5 ~ -1	中度左偏
			< -1	严重左偏

【2025年·单选】偏态系数-1.5，该组数据分布为（ ）。

- A. 轻度左偏
- B. 中度右偏
- C. 重度左偏

D. 对称

答案：C

解析：偏态系数为-1.5，绝对值大于1，且为负值，属于重度左偏，故C选项正确。 $|SK| \leq 0.5$ 为轻度偏态， $0.5 < |SK| \leq 1$ 为中度偏态。

2. 标准分数（Z分数）

（1）标准分数（标准分数的平均数为0，标准差为1。）

在统计上，均值和标准差不同时，不同变量的数值是不能比较的，但是每个数值在变量分布中相对于均值的相对位置是可比的，因此可以通过计算标准分数来比较不同变量的取值。

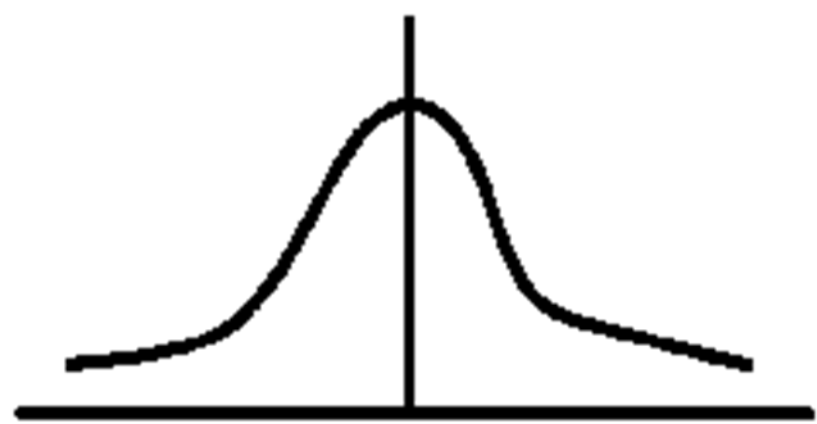
$Z_i = \frac{X_i - \bar{X}}{s}$		
A 地	均值 80	标准差 20
B 地	均值 60	标准差 5

（2）标准分数的运用

当数据服从对称的钟形分布时，经验法则表明：

- ①约有 68% 的数据与平均数的距离在加减 1 个标准差的范围之内；
- ②约有 95% 的数据与平均数的距离在加减 2 个标准差的范围之内；
- ③约有 99% 的数据与平均数的距离在加减 3 个标准差的范围之内。

【口诀】168，295，399



【2023年·单选】一组数据均值为60，标准差20，40是其中一个数值，标准分数为（ ）。

- A. 1
- B. 2
- C. -2
- D. -1

答案：D

解析：标准分数可以给出数值距离均值的相对位置，计算方法是用数值减去均值所得的差除以标准差， $(40-60)/20=-1$ 。

【2016年·多选】某企业客户满意度数据服从对称的钟形分布，均值为75，标准差为5。根据经验法则，关于该企业客户满意度的说法，正确的有（ ）。

- A. 约有68%的客户满意度在[70, 80]范围内
- B. 约有68%的客户满意度在[75, 85]范围内
- C. 约有95%的客户满意度在[75, 95]范围内
- D. 约有95%的客户满意度在[65, 85]范围内
- E. 约有99%的客户满意度在[60, 90]范围内

答案：ADE

解析：经验法则表明：约有68%的数据与平均数的距离在1个标准差之内，约有95%的数据与平均数的距2个标准差之内，约有99%的数据与平均数的距离3个标准差之内。也就是说，对于服从对称的钟形小的标准分数，68%的标准分数在 $[-1, +1]$ 范围内，约有95%的标准分数在 $[-2, +2]$ 范围之内，约有99%的标准分数在 $[-3, +3]$ 范围之内。

【2025年·单选】某城市居民月收入服从对称钟形分布，均值为10,000元，标准差为2,000元。根据经验法则，约有68%的居民月收入在（ ）范围内

- A. [6000, 14000]
- B. [7000, 13000]
- C. [8000, 12000]
- D. [9000, 11000]

答案：C

解析：对称钟形分布即正态分布，约68%的数据落在均值 ± 1 个标准差范围内，即 10000 ± 2000 ，为[8000, 12000]，故C选项正确。

【2025年·单选】某组数据服从对称钟形分布，均值为60，标准差为5。根据经验法则，该组数据中约有95%的数据在（ ）范围内。

- A. [50, 70]

- B. [55, 65]
- C. [40, 80]
- D. [45, 75]

答案：A

解析：95%的数据落在均值±2个标准差范围内，即 $60 \pm 2 \times 5 = 60 \pm 10$ ，为[50, 70]，故A选项正确。

第四节

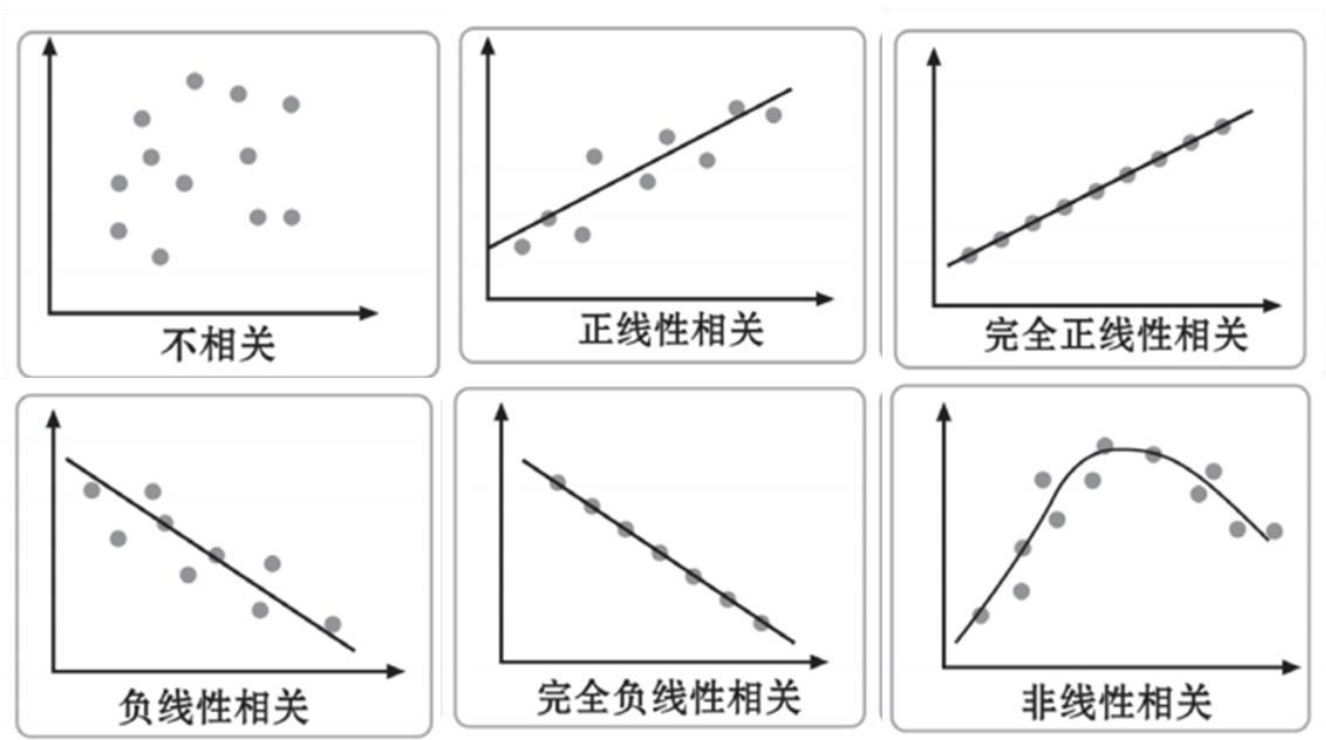
变量间的相关分析

（一）变量间的相关关系★

划分标准	分类	内容
相关程度	完全相关	一个变量的取值变化完全由另一个变量的取值变化所确定。
	不完全相关	两个变量之间的关系介于完全相关和不相关之间，一般相关现象都是指不完全相关。
	不相关	两个变量的取值变化彼此互不影响。
相关方向	正相关	一个变量的取值由小变大，另一个变量的取值也由小变大。
	负相关	一个变量的取值由小变大，另一个变量的取值相反地由大变小。
相关形式	线性相关	两个相关变量之间的关系大致呈现为线性关系。
	非线性相关	两个相关变量之间，近似于某种曲线方程的关系。

（二）散点图★

两个变量间的关系可以用散点图来展示。在散点图中，每个点代表一个观测值，横纵坐标值分别代表两个变量相应的观测值。



【2025年·单选】某医院长期记录患者年龄（岁）和住院天数（天）的数据，适用于分析这两个变量相关关系的是（ ）。

- A. 散点图
- B. 直方图
- C. 饼图
- D. 箱线图

答案：A

解析：散点图用于展示两个定量变量之间的相关关系，故A正确。B用于单变量分布，C用于构成比例，D用于分布特征。

（三）相关系数★★★

1. 相关系数的取值范围

相关系数的取值范围在+1和-1之间，即 $-1 \leq r \leq 1$ 。

相关系数	变量 X 和 Y 之间的关系	
$0 < r \leq 1$	正线性相关关系	
$-1 \leq r < 0$	负线性相关关系	
$r = 1$	完全正线性相关	当 $ r = 1$ 时，变量 Y 的取值完全依赖于 X
$r = -1$	完全负线性相关	

相关系数	变量 X 和 Y 之间的关系
$r = 0$	不存在线性相关关系，但并不说明变量之间没有任何关系，应结合散点图合理解释。

2. 两变量之间的相关程度

在说明两个变量之间的线性关系强弱时，根据相关系数的绝对值大小，可将相关程度分为：【口诀】358

$ r $ 的取值	两变量之间的相关程度
$ r < 0.3$	相关程度极弱，可视为无线性相关关系
$0.3 \leq r < 0.5$	可视为低度相关
$0.5 \leq r < 0.8$	可视为中度相关
$ r \geq 0.8$	可视为高度相关

【2023年·单选】两个变量的Pearson相关系数为0.03则这两个变量间可视为（ ）。

- A. 低度线性相关
- B. 无线性相关关系
- C. 高度线性相关
- D. 中度线性相关

答案：B

解析：①当 ≥ 0.8 时，视为高度相关；②当 $0.5 \leq < 0.8$ 时，视为中度相关；③当 $0.3 \leq < 0.5$ 时，视为低度相关；④当 < 0.3 时，变量之间的相关程度极弱，视为无线性相关关系。

【2024年·多选】关于Pearson相关系数的说法，正确的有（ ）。

- A. Pearson相关系数适用于分类变量间的相关分析
- B. Pearson相关系数 r 大于0，说明两个变量间存在高度相关关系
- C. Pearson相关系数 r 等于0，说明两个变量间不存在相关关系
- D. Pearson相关系数度量的是两个变量间的线性相关关系
- E. Pearson相关系数取值范围在-1和+1之间

答案：DE

解析：Pearson相关系数度量线性相关关系，取值范围在-1和+1之间，故D、E正确。A错误，适用于数值型变量；B错误， $r > 0$ 仅说明正相关，高度相关需 $|r| \geq 0.8$ ；C错误， $r = 0$ 说明不存在线性相关，可能存在非线性相

关。

框架导航

