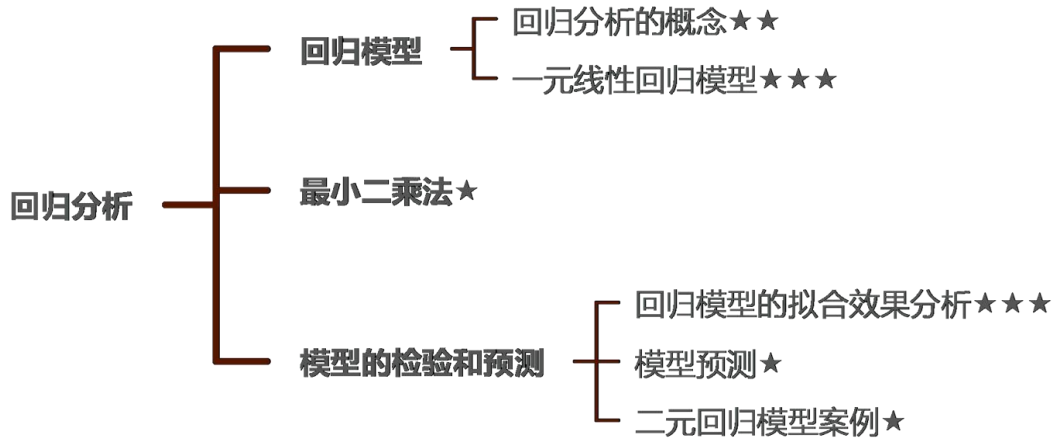


第二十六章

回归分析

框架导航



第一节 回归模型

（一）回归分析的概念★★

1. 回归分析与相关分析的关系

（1）联系

①相关分析需要依靠**回归分析**来表明现象数量**相关的具体形式**，而回归分析则需要依靠**相关分析**来表明现象数量变化的**相关程度**。

②只有当变量之间存在着**高度相关**时，进行回归分析寻求其相关的**具体形式**才有意义。

（2）区别

①相关分析

研究变量之间**相关的方向**和**相关的程度**，不能指出变量间相互关系的具体形式，也无法从一个变量的变化来推测另一个变量的变化情况。

②回归分析

研究变量之间相互关系的**具体形式**，它对具有相关关系的变量之间的数量联系进行测定，确定一个相关的**数学方程式**，根据这个数学方程式可以从已知量来**推测未知量**，从而为估算和预测提供了一个重要的方法。

2. 因变量与自变量

被预测或被解释的变量称为**因变量**，一般用 Y 表示；

用来预测或解释因变量的变量称为**自变量**，一般用 X 表示。

【2023 年 · 多选】关于相关分析和回归分析的说法正确的有

()。

- A. 相关分析可以从一个变量的变化来推测另一个变量的变化
- B. 相关分析和回归分析中都需要确定因变量和自变量
- C. 相关分析和回归分析具有相同的研究方法和研究目的
- D. 相关分析需要依靠回归分析来表明现象数量相关的具体形式
- E. 回归分析需要依靠相关分析来表明现象数量变化的相关程度

答案：DE

解析：相关分析和回归分析在研究目的和方法上有明显区别：（1）相关分析①研究变量之间相关的方向和相关的程度。选项 A 错误；②相关分析不能指出变量间相互关系的具体形式，也无法从一个变量的变化来推测另一个变量的变化情况。（2）回归分析选项 C 错误 DE 正确①研究变量之间相互关系的具体形式，对具有相关关系的变量之间的数量联系进行测定，确定一个相关的数学方程式。②根据这个数学方程式可以从已知量来推测未知量，从而为估算和预测提供了一个重要方法。综上所述本题选择 DE。

（二）一元线性回归模型★★★

1. 回归模型的概念及分类

回归模型可以用描述因变量 Y 如何依赖自变量 X 和误差项 ε 的方程来表示。

分类依据	类型
自变量的多少	一元回归模型、多元回归模型
是否是线性	线性回归模型、非线性回归模型

2. 一元线性回归

（1）回归模型

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

β_0 和 β_1 为模型的参数。

① $\beta_0 + \beta_1 X$ 反映了由于自变量X的变化而引起的因变量Y的线性变化。

②误差项 ε 是个随机变量，表示除X和Y的线性关系之外的随机因素对Y的影响，是不能由X和Y的线性关系所解释的Y的变异性。

(2) 回归方程

$$E(Y) = \beta_0 + \beta_1 X$$

β_0 是回归直线的截距, β_1 是回归直线的斜率, 表示 X 每变动一个单位时, $E(Y)$ 的变动量。

【2025 年·单选】一元线性回归方程的形式为 ()。

- A. $E(Y) = \beta_0 + \beta_1 X$
- B. $E(Y) = \beta_0 + \beta_1 X + \varepsilon$
- C. $Y = \beta_0 + \varepsilon$
- D. $Y = \beta_0 + \beta_1 X$

答案: A

解析: 一元线性回归方程形式为 $E(Y) = \beta_0 + \beta_1 X$, 故 A 正确。B 是包含随机误差项的模型表达式, D 缺少期望符号。

【2024 年·多选】关于估计的一元线性回归模型 $Y=6000+0.7X$ 的说法, 正确的有 ()。

- A. X 每增加 1%, Y 平均增长 0.7%
- B. X 每增加 1 个单位, Y 平均增长 0.7 个单位
- C. X 每减少 1 个单位, Y 平均增长 0.7 个单位
- D. 当 X 为 20000 时, Y 约为 20000
- E. X 每减少 1%, Y 平均增长 0.7%

答案: BD

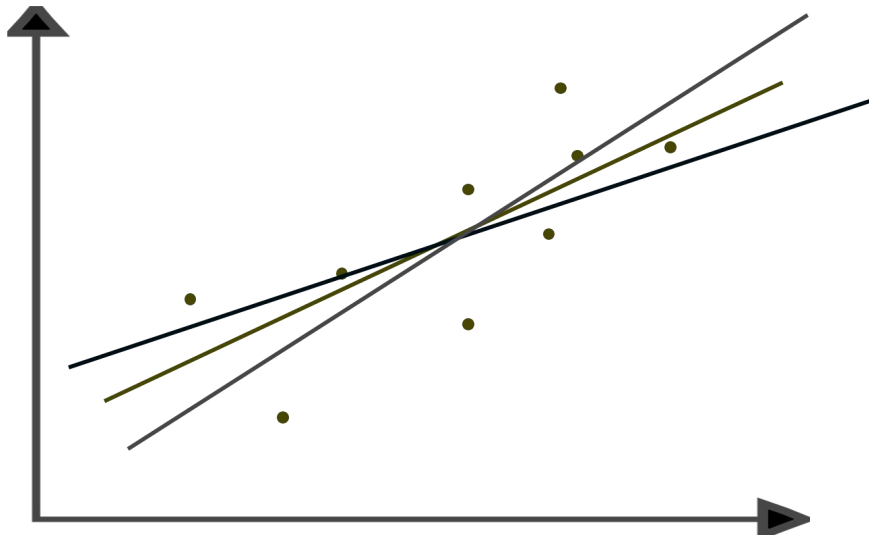
解析: 回归系数 0.7 表示 X 每增加 1 个单位, Y 平均增长 0.7 个单位 (B 正确); 当 $X=20000$ 时, $Y=6000+0.7 \times 20000=20000$ (D 正确)。A、E 百分比表述错误; C 方向错误。

第二节 最小二乘法★

回归分析是根据相关关系的具体形态, 选择一个合适的数学模型, 来近似地表达变量间的依赖关系。

最小二乘法就是使得因变量的观测值 y_i 与估计值 $\hat{y}_i$ 之间的离差 (又称残差) 平方和最小来估计参数 β_0 和 β_1 的方法。

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$



【2022 年 · 单选】最小二乘法的原理是（ ）。

- A. 因变量的观测值和均值之间的离差平方和最小
- B. 因变量的观测值和估计值之间的离差平方和最小
- C. 因变量的均值和估计值均值之间的离差平方和最小
- D. 自变量的观测值和均值之间的离差平方和最小

答案：B

解析：最小二乘法的原理是使得因变量的观测值与估计值之间的离差（残差）平方和最小。

【2023 年 · 单选】在回归分析中常用的估计方法是（ ）。

- A. 有偏估计法
- B. 无偏估计法
- C. 最小二乘法
- D. 最大二乘法

答案：C

解析：对回归模型进行估计的方法称为最小二乘法。