



第二十六章

回归分析



第一节 回归模型

一、回归分析的概念

1、定义

所谓回归分析就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。



第一节 回归模型

2、相关分析与回归分析的关系

(1) 联系

它们不仅具有共同的研究对象，而且在具体应用时，常常必须互相补充。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。



第一节 回归模型

(2) 区别

相关分析和回归分析在研究目的和研究方法上有明显区别：

相关分析：研究变量之间相关的方向和相关程度。

回归分析：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，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。



第一节 回归模型

- 3、进行回归分析时，首先确定因变量和自变量
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。



第一节 回归模型

【例-多选题】关于相关分析和回归分析的说法，正确的有（ ）。

- A. 相关分析研究变量间相关的方向和相关程度
- B. 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
- C. 回归分析研究变量间相互关系的具体形式
- D. 相关分析和回归分析在研究方法和研究目的上有明显区别
- E. 相关分析中需要明确自变量和因变量

答案：ACD



第一节 回归模型

二、一元线性回归模型

1、回归模型的分类：

根据自变量的多少，回归模型可以分为一元回归模型和多元回归模型。

根据回归模型是否线性，回归模型可以分为线性回归模型和非线性回归模型。



第一节 回归模型

2、一元线性回归是描述两个变量之间相关关系的最简单的回归模型。

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。只涉及一个自变量的一元线性回归模型可以表示为：

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

式中 β_0 和 β_1 为模型的参数

y是x的线性函数（ $\beta_0 + \beta_1 X$ ）加上误差项 ε 。



第一节 回归模型

$\beta_0 + \beta_1 X$ 反映了由于x的变化而引起的y的变化；

误差项 ε 是随机变量，反映了除x和y之间的线性关系之外的随机因素对y的影响，是不能由x和y之间的线性关系所解释的Y的变异性。



第一节 回归模型

【例-单选题】线性回归模型中 $Y = \beta_0 + \beta_1 X + \varepsilon$ 误差 ε 的含义是（ ）。

- A. 回归直线的截距
- B. 除 x 和 Y 线性关系之外的随机因素对 Y 的影响
- C. 回归直线的斜率
- D. 观测值和估计值之间的残差

答案：B

解析：误差项 ε 是随机变量，反映了除 x 和 y 之间的线性关系之外的随机因素对 y 的影响，是不能由 x 和 y 之间的线性关系所解释的 Y 的变异性。



第一节 回归模型

3、描述因变量Y的期望 $E(Y)$ 如何依赖自变量X的方程称为回归方程。一元线性回归方程的形式为：

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

一元线性回归方程的图示是一条直线， 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。