

回归分析

一：回归模型

(一) 回归分析的概念

1.回归分析：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。

2.回归分析vs相关分析（研究目的和方法有明显区别）

- (1) 相关分析需要依赖回归分析表明现象数量相关的具体形式。
- (2) 回归分析依赖相关分析表明现象数量变化的相关程度。
- (3) 相关分析无法从一个变量的变化来推测另一个变量变化情况，而回归分析的数学方程式可以。
- (4) 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。

3.进行回归分析时，首先需要确定因变量和自变量

- (1) 因变量：被预测或被解释的变量，一般用y表示；
- (2) 自变量：用来预测或解释因变量的变量，一般用X表示。

(二) 一元线性回归模型

1.回归模型分类

- (1) 根据自变量的多少：一元回归模型和多元回归模型。
- (2) 根据回归模型是否线性：线性回归模型和非线性回归模型

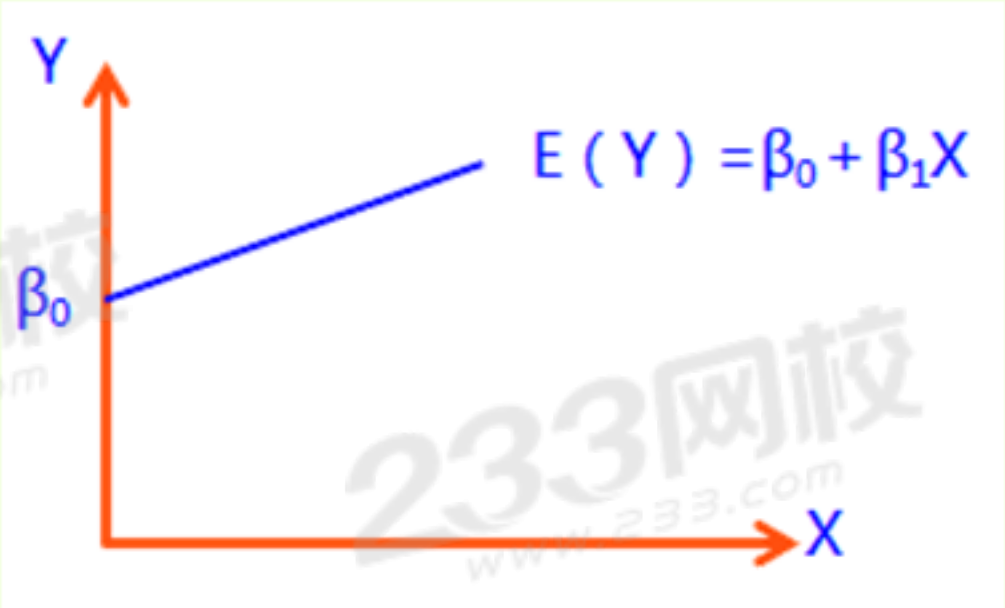
【注意2】误差项ε：随机变量，表示除X和Y的线性关系之外的随机因素对Y的影响；不能由X和Y的线性关系所解释的Y的变异性。

2.一元线性回归模型：描述两个变量之间相关关系的最简单的回归模型。一元线性回归是只涉及一个自变量的回归问题。

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

- 【注意1】一元线性回归方程的图示是一条直线。
- 【注意2】β0是回归直线的截距，β1是回归直线的斜率，表明x每变动一个单位，E(Y)的变动量。

一元线性回归方程：E(Y) = β0 + β1X



(三) 回归方程：描述因变量Y的期望E(Y)如何依赖自变量X的方程。

二：最小二乘法

(一) 最小二乘法：使得因变量的观测值yi与估计值yi^之间的离差平方和最小来估计参数β0和β1的方法。

(二) 最小二乘法是对回归模型进行参数估计的方法

三：模型的检验与预测

(一) 回归模型的拟合效果分析：

1.在使用估计的回归方程之前，需要对模型进行检验

- (1) 结合经济理论和经验分析回归系数的经济含义是否合理；
- (2) 分析估计的模型对数据的拟合效果如何；
- (3) 对模型进行假设检验。

2.决定系数（R2）：测度回归直线对样本数据的拟合程度。说明回归模型所能解释的因变量变化占因变量总变化的比例

3.决定系数的取值：在0到1之间

- R2 = 1，说明回归直线可以解释因变量的所有变化。
- R2 = 0，说明回归直线无法解释因变量的变化，因变量的变化与自变量无关。

决定系数VS相关系数VS偏态系数

决定系数	$0 \leq R^2 \leq 1$	越接近 1 效果越好
相关系数	$-1 \leq r \leq 1$	0.3, 0.5, 0.8（无关、低度、中度、高度）
偏态系数	SK	0.5, 1（轻度、中度、严重）

(二) 模型预测：回归分析的一个重要应用就是预测，即利用估计的回归模型预估因变量数值。

(三) 二元回归模型案例