中央财经大学
Central University of Finance and Economics

计量经济学

中国财政发展协同创新中心
陈怡心
cyx@cufe.edu.cn

第2章 一元线性回归

陈怡心
中央财经大学中国财政发展协同创新中心

一元线性回归模型：例子

- 为什么在青少年时期要选择上学?

  好奇心、求知欲及个人成长

  教育→未来的收入水平

- 如何从理论上解释教育投资的回报率(returns to schooling)?

- Mincer (1958)提出基于效用最大化的理性选择模型：
- 个体选择多上一年学，则需推迟一年挣钱(另需交学费)；为弥补其损失，市场均衡条件要求给予受教育多者更高的未来收入。
- 由此可得工资对数与教育年限的线性关系：

$$lnw = \alpha + \beta s$$

lnw 为工资对数，s为教育年限(schooling)，而α与β为参数。

- α为截距项，表示当教育年限为 0 时的工资对数水平，因为 $lnw = \alpha + \beta \cdot 0 = \alpha$
- β为斜率，表示教育年限对工资对数的边际效应，即每增加一年教育，将使工资增加百分之几，因为对方程两边求导可得

$$\beta = \frac{dlnw}{ds} = \frac{\frac{dw}{w}}{ds} \approx \frac{\frac{\Delta w}{w}}{\Delta s}$$

例：如果$\beta = 0.05$，$\Delta s = 1$年，则$\frac{\Delta w}{w} \approx \beta \cdot \Delta s = 0.05 = 5\%$

- 显然，教育年限只是影响工资对数的因素之一，故严格来说，方程应写为

$$lnw = \alpha + \beta s + \text{其他因素}$$

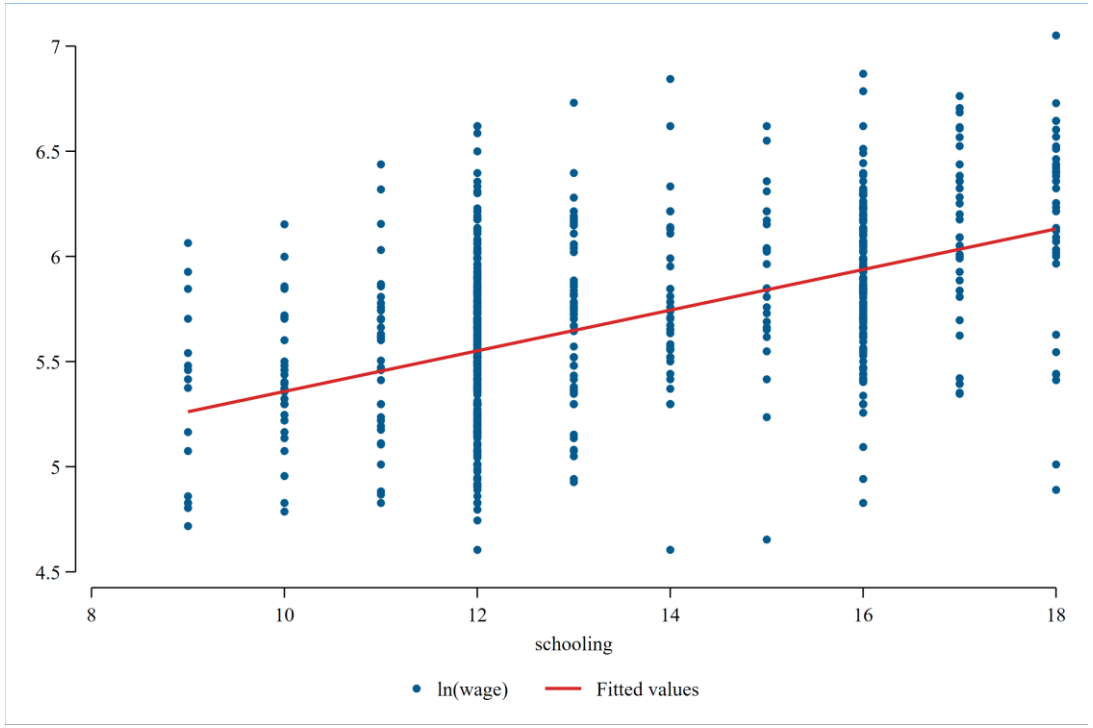
- 将其他因素记为ε,则有

$$lnw = \alpha + \beta s + \varepsilon$$

即为劳动经济学(labor economics)中著名的明瑟方程(the Mincer equation)的基本形式(Mincer, 1974)。

- 但多上一年学，究竟能使未来收入提高百分之几?
- 这取决于参数β的取值。明瑟模型并未提供关于α与β具体取值的信息。
- 对于这种定量问题(quantitative question)，只有通过数据才能给 出定量回答(quantitative answer)。
- 需要用计量经济学方法，通过样本数据来估计未知参数α与β。

- 样本数据包括 758 位美国年轻男子的教育投资回报率数据。
- 为了考察工资对数与教育年限的关系，画二者的散点图，并在图上画出离这些样本点最近的“回归直线”。



工资对数与教育年限正相关，似乎存在线性关系，在形式上与明瑟方程相一致。

一个财政学领域的例子

- “中国经济增长奇迹”

- 地方分权的制度是中国经济增长的制度关键（钱颖一、许成钢）
- 发展了Tiebout模型等财政分权理论，将财政分权与地方政府的激励、经济转型和经济增长联系起来。形成了第二代财政分权理论。

- 强调地方的分权，促进了地方竞争（包括以下五个方面）
- 促进和维护了市场机制的发展
- 促进了乡镇企业的发展
- 促进了城市化和基础设施的建设
- 导致了改革实验的发生和模仿
- 促进了外商直接投资的流入，从而促进了经济增长

- 如何建立财政分权与经济增长的计量模型呢?
- 不妨假设两者之间存在一种线性关系

$$lnpgdp = \alpha + \beta dec$$

$lnpgdp$为人均GDP的对数形式，dec为财政分权程度，α和β为待估参数，α为截距项，β为斜率。

财政分权只是影响经济增长的因素之一故严格来说，方程应写为

$$lnpgdp = \alpha + \beta dec + \text{其他因素}$$

- 将其他因素记为ε,则有

$$lnpgdp = \alpha + \beta dec + \varepsilon$$

财政分权程度的增加，究竟能使人均GDP提高百分之多少呢?
这取决于参数β的取值。
通过样本数据估计未知参数，α和β

一元线性回归模型

• 更一般地，假设从总体随机抽取n位个体，则一元线性回归模型可写为

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (i = 1, \ldots, n)$$

• y_i：被解释变量(dependent variable, regressand)

• x_i：解释变量(explanatory variable, independent variable, regressor)

y	x
因变量	自变量
被解释变量	解释变量
回归子	回归元
响应变量	控制变量

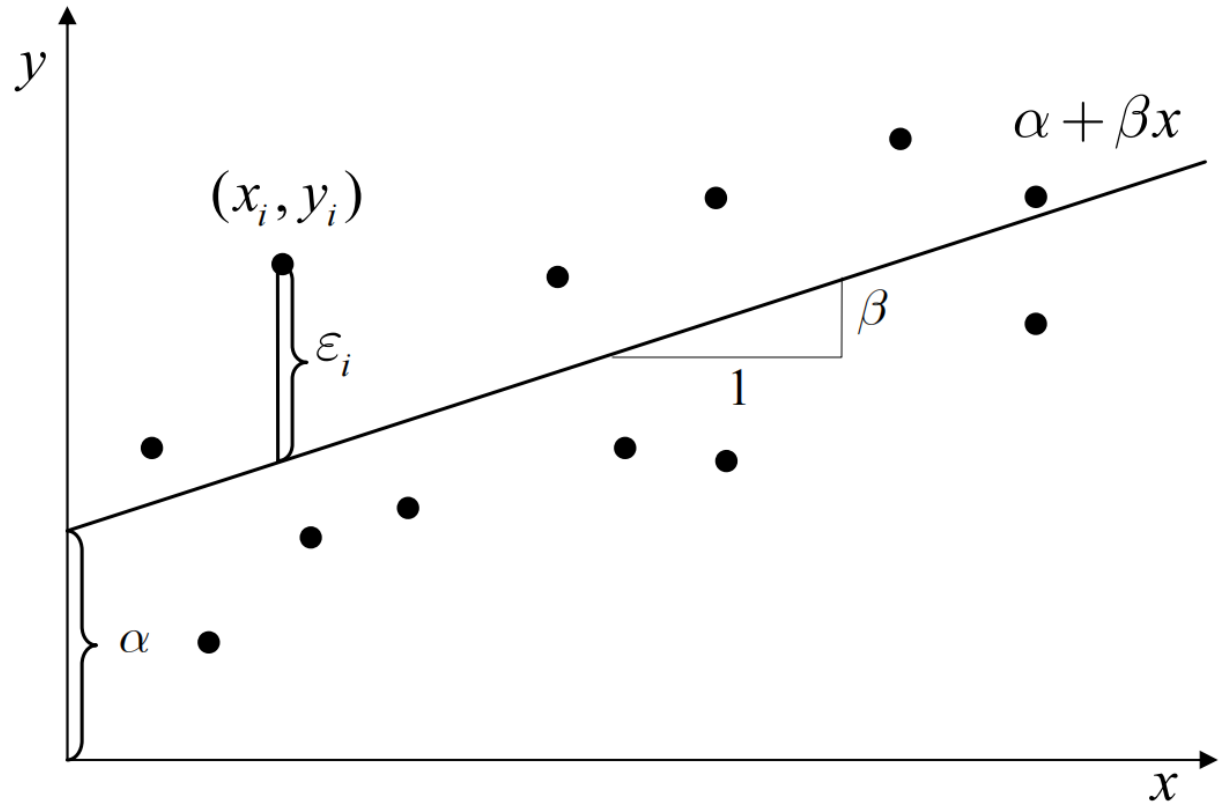
- α：截距项(intercept)或常数项(constant)
- β：斜率(slope)
- α与β：统称“回归系数”(regression coefficients)或“参数”(parameters)
- ε_i：“误差项”(error term)或“扰动项”(disturbance)，包括遗漏的其他因素、变量的测量误差、回归函数的设定误差(比如，忽略了非线性项)以及人类行为的内在随机性等。
- 除x_i以外，影响y_i的所有其他因素都在ε_i中。
- 下标 i 表示个体 i，比如第 i 个人，第 i 个企业，第 i 个国家等。
- i 的取值为1, …, n，其中 n 为“样本容量”(sample size)。

- 方程右边的确定性部分为 $\alpha + \beta x_i$ ，称为 **总体回归线** (population regression line)或**总体回归函数**(population regression function，简记PRF)。

- 假设总体回归函数为线性，可视为一阶近似(忽略二次项及高阶项)。

数据生成过程

- 模型$y_i = \alpha + \beta x_i + \varepsilon_i$ 也称为**数据生成过程**(Data Generation Process，简记DGP)。
- 从数据生成的角度来看， 随机变量x_i与ε_i首先从相应的概率分布 中抽取观测值(observation)。
- 确定x_i 与 ε_i的取值后，根据方程$y_i = \alpha + \beta x_i + \varepsilon_i$ 生成y_i的取值。由于ε_i通常无法观测(unobservable)，故研究者只知道(x_i, y_i)。
- 计量经济学的主要任务之一就是通过数据$\{x_i, y_i\}_{i=1}^n$来获取关于总体参数(α, β)的信息。


y
$\alpha + \beta x$
(x_i, y_i)
ε_i
β
1
α
x

OLS估计量的推导

- 如何根据观测值$\{x_i, y_i\}_{i=1}^n$来估计总体回归线$\alpha + \beta x_i$?
- 我们希望在(x,y)平面
  找到一条回归线。

- 希望找到一条什么样的线更好?
- 怎么定义?

OLS估计量的推导

- 离所有这些点（观测值）最近的直线

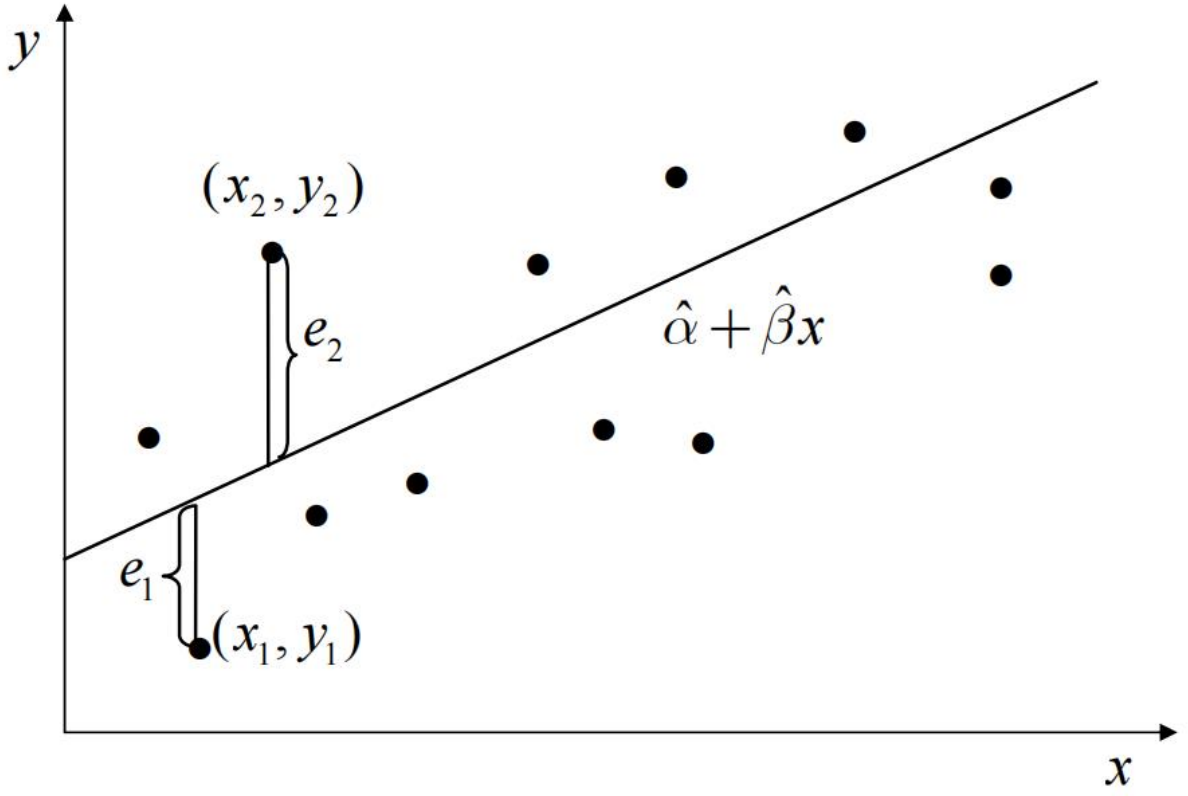
- 在此平面上， 任意给定一条直线

$y_i = \hat{\alpha} + \hat{\beta} x_i$

- 每个点(观测值)到这条线的距离

$e_i \equiv y_i - \hat{\alpha} - \hat{\beta} x_i$

称为“残差”(residual)。



如直接把残差加起来，$\sum_{i=1}^{n} e_i$,会出现正负相抵的现象。

使用绝对值，$\sum_{i=1}^{n}|e_i| = \sum_{i=1}^{n}|y_i - \hat{\alpha} - \hat{\beta}x_i|$

使用平方和，

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)^2$$

- 称为“残差平方和”(Sum of Squared Residuals，简记SSR；或Residual Sum of Squares，简记RSS)

“普通最小二乘法”(Ordinary Least Squares，简记OLS)就是选择$\hat{\alpha}$和$\hat{\beta}$，使得残差平方和最小化。

- 可将OLS 的目标函数写为

$$\min_{\hat{\alpha},\hat{\beta}} \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{\alpha} - \hat{\beta} x_i\right)^2$$

OLS 是线性回归模型的基本估计方法。

- 此最小化问题的一阶条件为

$$\begin{cases} \dfrac{\partial}{\partial \hat{\alpha}} \displaystyle\sum_{i=1}^{n} e_i^2 = -2 \sum_{i=1}^{n} \left(y_i - \hat{\alpha} - \hat{\beta} x_i \right) = 0 \\ \dfrac{\partial}{\partial \hat{\beta}} \displaystyle\sum_{i=1}^{n} e_i^2 = -2 \sum_{i=1}^{n} \left(y_i - \hat{\alpha} - \hat{\beta} x_i \right) x_i = 0 \end{cases}$$

消去方程左边的“-2”可得

$$\begin{cases} \sum_{i=1}^{n}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)=0 \\ \sum_{i=1}^{n}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)x_i=0 \end{cases}$$

对上式各项分别求和，并移项可得

$$\begin{cases} n\hat{\alpha}+\hat{\beta}\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i \\ \hat{\alpha}\sum_{i=1}^{n}x_i+\hat{\beta}\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_iy_i \end{cases}$$

这是有关估计量$\hat{\alpha}$和 $\hat{\beta}$ 的二元一次线性方程组，称为“正规方程组”(normal equations)。

- 从方程组第1个方程可得

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

- 其中，$\bar{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i$ 为y的样本均值，$\bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$ 为x的样本均值。
- 代入方程组的第2个方程可得，

$$\left(\bar{y} - \hat{\beta}\bar{x}\right)\sum_{i=1}^{n} x_i + \hat{\beta}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

合并同类项，并移项可得

$$\hat{\beta}\left(\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i y_i - \bar{y}\sum_{i=1}^{n} x_i$$

- 使用关系式 $\sum_{i=1}^{n} x_i = n\bar{x}$, 求解$\hat{\beta}$可得

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

可写为更为直观的离差形式

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

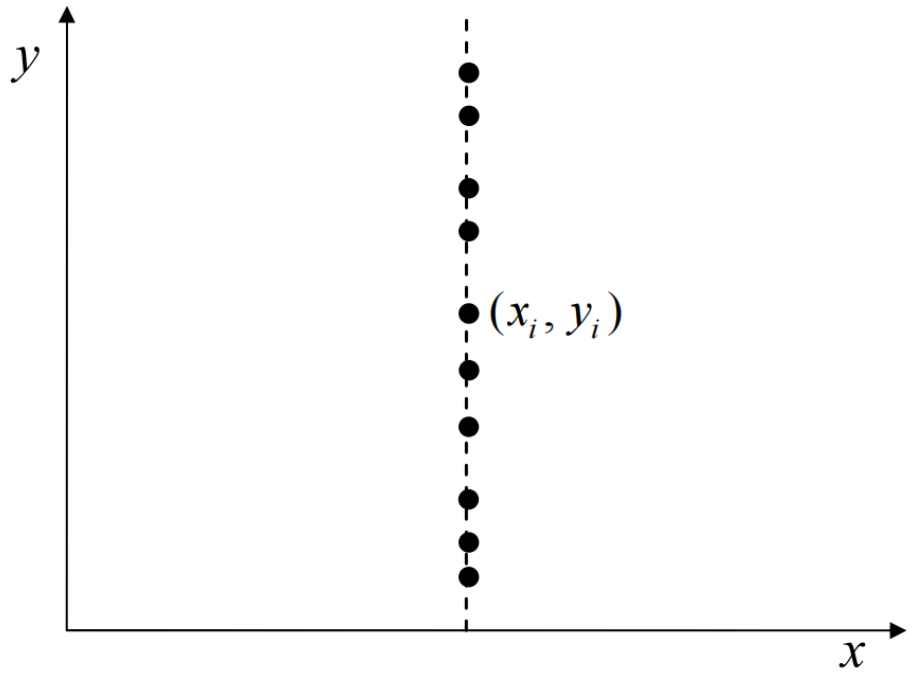
- 综上，斜率α和截距β的OLS估计量分别为

- $\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$
- $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

- 估计的截距$\hat{\alpha}$和斜率$\hat{\beta}$是利用X和Y的n组样本观测值计算得到的。它们分别为总体斜率α和截距β的估计。

- 需要x_i有变化，不能是常数。

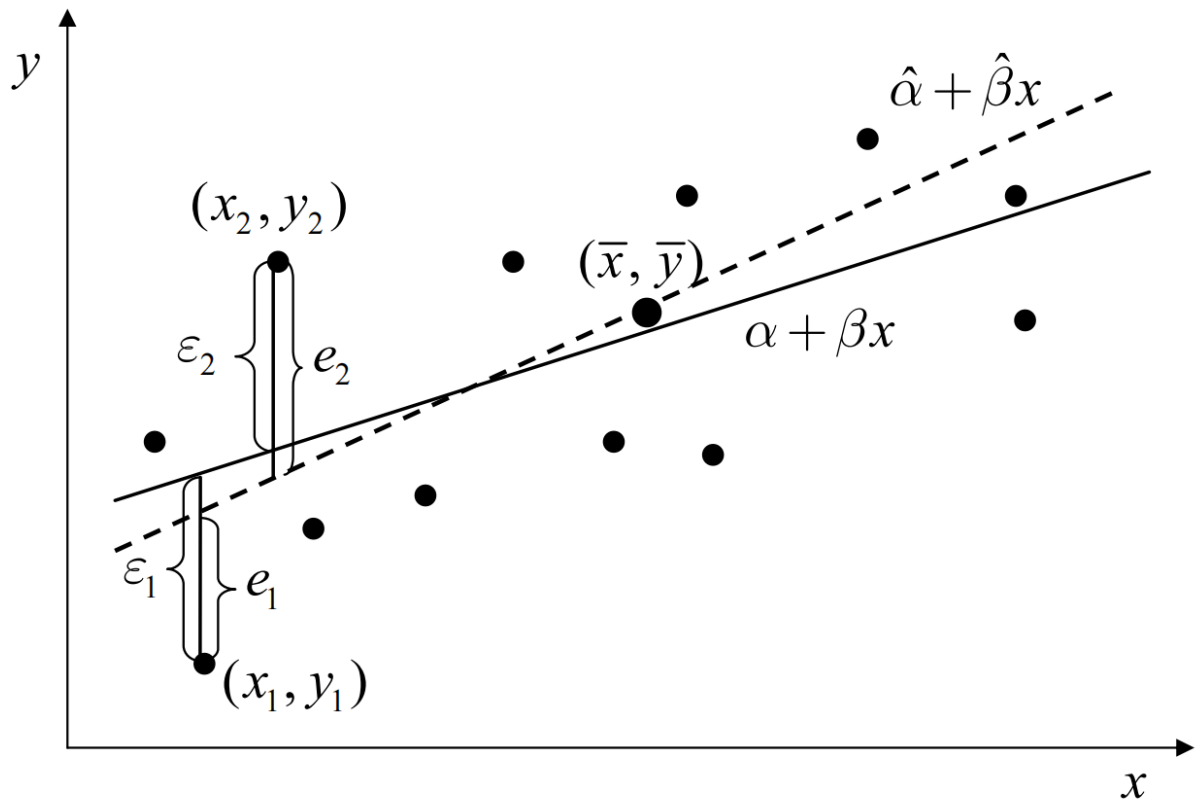
- 思考：如何x_i没有变化?
- $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0$



- 求解OLS估计量$\hat{\alpha}$和$\hat{\beta}$，可得到

$$\hat{y} = \hat{\alpha} + \hat{\beta} x_i$$

- 称为**样本回归线**(sample regression line)
  或**样本回归函数**(sample regression function，简记 SRF)。
- 样本回归线一定经过$(\bar{x}, \bar{y})$



矩估计

- 随机误差项的性质
- 零期望假设：

$$E(\varepsilon_i) = 0, \forall i, or \;\; E(\boldsymbol{\varepsilon}) = \mathbf{0}$$

- 注释：这不是一个严格的假设，通常可以通过调整截距项α从而标准化地获得$\mathrm{E}(\varepsilon) = 0$
- 零条件期望假设：

$$E(\varepsilon_i|x_i) = 0, \forall i, or \;\; E(\boldsymbol{\varepsilon}|X) = \mathbf{0}$$

- 注释：这是一个更强的假设，u对于所有的x实现值来说，其均值都是0，这就意味着它们是不相关的，严格独立的。
- 在模型$y = \alpha + \beta x + \varepsilon$两边同时取条件期望后，得到总体回归方$E(y|x) = \alpha + \beta x$

矩估计

由零条件期望假设可得

$$E(\varepsilon_i|x_i)=0 \Longrightarrow \begin{cases} E(\varepsilon_i)=0 \\ cov(x_i,\varepsilon_i)=E(x_i\varepsilon_i)=0 \end{cases}, \varepsilon_i = y_i-\alpha-\beta x_i, \forall i$$

- 那么有：

$$E(y_i-\alpha-\beta x_i)=0$$
$$E[x_i(y_i-\alpha-\beta x_i)]=0$$

- 样本近似（矩条件）：

$$\frac{1}{n}\Sigma_{i=1}^{n}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)=0 \qquad (1)$$
$$\frac{1}{n}\Sigma_{i=1}^{n}x_i\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)=0 \qquad (2)$$

矩估计

- 由公式（1）：

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

- 于是公式（2）变为：

$$\frac{1}{n}\Sigma_{i=1}^{n} x_i\left(y_i - \left(\bar{y} - \hat{\beta}\bar{x}\right) - \hat{\beta}x_i\right) = 0$$

$$\Sigma_{i=1}^{n} x_i(y_i - \bar{y}) = \hat{\beta}\Sigma_{i=1}^{n} x_i(x_i - \bar{x})$$

$$\Sigma_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \hat{\beta}\Sigma_{i=1}^{n} (x_i - \bar{x})^2$$

$$\hat{\beta} = \frac{\Sigma_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\Sigma_{i=1}^{n} (x_i - \bar{x})^2}$$

矩估计

- 截距和斜率的矩估计量分别为：

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = \frac{\Sigma_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\Sigma_{i=1}^{n}(x_i - \bar{x})^2}$$

- OLS估计量能够很好地近似总体的性质。

OLS的正交性

定义被解释变量y_i的“拟合值”(fitted value)或“预测值”(predicted value)为

$$\hat{y}_i \equiv \hat{\alpha} + \hat{\beta} x_i$$

可将残差写为

$$e_i = y_i - \left(\hat{\alpha} + \hat{\beta} x_i\right) = y_i - \hat{y}_i$$

故可将被解释变量分解为

$$y_i = \hat{y}_i + e_i$$

正规方程组

$$\begin{cases} \sum_{i=1}^{n}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)=0 \\ \sum_{i=1}^{n}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)x_i=0 \end{cases}$$

可写为

$$\begin{cases} \sum_{i=1}^{n} e_i=0 \\ \sum_{i=1}^{n} x_i e_i=0 \end{cases}$$

将上式写为向量内积的形式：

$$(1 \quad \cdots \quad 1)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = 0,\ (x_1 \quad \cdots \quad x_n)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = 0$$

定义常数向量、残差向量、解释向量以及拟合值向量为

$$\mathbf{1} \equiv \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{n\times 1},\ \boldsymbol{e} \equiv \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix},\ \boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ \hat{\boldsymbol{y}} = \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix}$$

则正规方程可写为

$$\mathbf{1}'\boldsymbol{e} = 0, \quad \boldsymbol{x}'\boldsymbol{e} = 0$$

故残差向量$\boldsymbol{e}$与常数向量$\mathbf{1}$正交，而且$\boldsymbol{e}$也与解释向量$\boldsymbol{x}$正交。

- 将常数项视为取值都为1的解释变量，而α为此变量的系数。
- 残差向量与所有解释变量(包括$\mathbf{1}$与$\boldsymbol{x}$)正交。
- 残差向量$\boldsymbol{e}$也与拟合值向量 $\hat{\boldsymbol{y}}$正交，因为

$$\hat{\mathbf{y}}'\mathbf{e} \equiv \begin{pmatrix}\hat{y}_1 & \cdots & \hat{y}_n\end{pmatrix}\begin{pmatrix}e_1\\ \vdots \\ e_n\end{pmatrix} = \sum_{i=1}^{n}\hat{y}_i e_i = \sum_{i=1}^{n}\left(\hat{\alpha}+\hat{\beta}x_i\right)e_i = \hat{\alpha}\underbrace{\sum_{i=1}^{n}e_i}_{=0} + \underbrace{\hat{\beta}\sum_{i=1}^{n}x_i e_i}_{=0} = 0$$

- OLS 残差与解释变量及拟合值的正交性是 OLS 的重要特征，为推导证明提供了方便。

- 比如，考虑方程$e_i = y_i - \hat{y}_i$，将两边对i加总，并除以n可得：

$$0 = \frac{1}{n}\sum_{i=1}^{n} e_i = \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{1}{n}\sum_{i=1}^{n} \hat{y}_i = \bar{y} - \bar{\hat{y}}$$

- 其中，$\bar{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n} \hat{y}_i$，为拟合值$\hat{y}_i$的均值。由上式可知，被解释变量的均值恰好等于拟合值的均值，即

$$\bar{y} = \bar{\hat{y}}$$

平方和分解公式

- 被解释变量可分解为相互正交的两个部分，即$y_i = \hat{y_i} + e_i$

- 如果回归方程有常数项（通常都有），则被解释变量的离差平方和$\sum_{i=1}^{n}(y_i - \bar{y})^2$(Total Sum of Squares,TSS)可分解为

$$\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})^2}_{\text{TSS}} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}_{\text{ESS}} + \underbrace{\sum_{i=1}^{n}e_i^2}_{\text{RSS}}$$

- 此方程称为“平方和分解公式”

- 将$\sum_{i=1}^{n}(y_i-\bar{y})^2$分解为两部分。
- 右边第一项为$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$，由于$\bar{y}=\bar{\hat{y}}$(被解释变量的均值等于拟合值的均值)，故可写为$\sum_{i=1}^{n}\left(\hat{y}_i-\bar{\hat{y}}\right)^2$,即可由模型解释的部分，称为Explained Sum of Squares (ESS)
- 右边第二项为残差平方和$\sum_{i=1}^{n}e_i^2$ (Residual Sum of Squares,RSS)，是模型所无法解释的部分。
- 平方和分解公式能够成立，正是由于OLS 的正交性。

- 证明： 将离差$(y_i - \bar{y})$写为$(y_i - \hat{y_i} + \hat{y_i} + \bar{y})$，则可将TSS 写为

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n}(e_i + \hat{y}_i - \bar{y})^2$$
$$= \sum_{i=1}^{n} e_i^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + 2\sum_{i=1}^{n} e_i(\hat{y}_i - \bar{y})$$

只需证明交叉项$\sum_{i=1}^{n} e_i(\hat{y}_i - \bar{y}) = 0$即可，而这又由OLS的正交性所保证：

$$\sum_{i=1}^{n} e_i(\hat{y}_i - \bar{y}) = \sum_{i=1}^{n} e_i\hat{y}_i - \bar{y}\sum_{i=1}^{n} e_i = 0 - 0 = 0$$

如果没有常数项，则无法保证$\sum_{i=1}^{n} e_i = 0$，故平方和分解公式在无常数项的情况下不再成立。

拟合优度

- OLS 的样本回归线为离所有样本点最近的直线。
- 但究竟离这些样本点有多近?
- 希望有绝对的度量，以衡量样本回归线对数据的拟合优良程度。
- 在有常数项的情况下，根据平方和分解公式，可将被解释变量的离差平方和分解为模型可以解释与不可解释的部分。
- 如果模型可以解释的部分所占比重越大，则样本回归线的拟合程度越好。

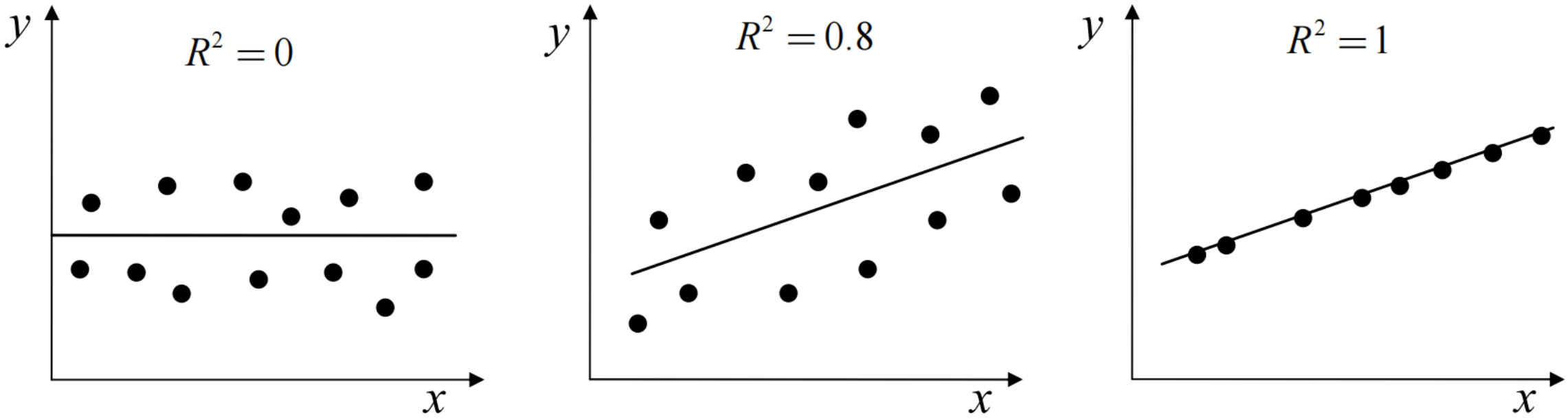
• 定义 **拟合优度**(goodness of fit) R^2为

$$R^2 \equiv \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

• 拟合优度R^2也称为**可决系数**(coefficient of determination)

• 在有常数项的情况下，拟合优度等于被解释变量y_i与拟合值$\hat{y}_i$之间相关系数的平方，即$R^2 = [Corr(y_i, \hat{y}_i)]^2$，故记为$R^2$。

• 显然，$0 \leq R^2 \leq 1$。



- R^2越高，则样本回归线对数据的拟合程度越好。
- 如果R^2 =1，则解释变量x可以完全解释y的变动。
- 如果R^2 =0，则解释变量x对于解释y没有任何帮助。
- 如果0 < R^2 < 1，则为介于以上两种极端的中间情形，即x可以解释y的一部分，但无法解释其余部分。

- 注意：
- R^2 只是反映拟合程度的好坏，除此外并无太多意义。
- 评估回归方程是否显著，应使用 F 检验(R^2 与 F 统计量也有联系)。

- 初学者的误区：

（1）为了提高R^2，放入很多解释变量。

（2）样本容量很小，通常会得到很高的R^2。（极端情况，样本容量为2）

无常数项的回归

- 偶尔也进行无常数项的回归，或许是经济理论的要求，也可能在模型变换时消去常数项。
- 无常数项的回归必然经过原点，也称为“经过原点的回归”(regression through the origin)。
- 此时， 一元线性回归模型可写为

$$y_i = \beta x_i + \varepsilon_i \quad (i = 1, \ldots, n)$$

- 依然进行OLS 估计，最小化残差平方和：

$$\min_{\hat{\beta}} \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta} x_i\right)^2$$

- 一阶条件为

$$\frac{d}{d\hat{\beta}} \sum_{i=1}^{n} e_i^2 = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta} x_i\right) x_i = 0$$

- 消去方程左边的“-2”可得

$$\sum_{i=1}^{n} \left(y_i - \hat{\beta} x_i\right) x_i = 0$$

- 求解$\hat{\beta}$可得

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

- 该方程与有常数项回归的表达式类似。

- 但即使没有常数项，OLS 也仍满足正交性，因为正规方程 (组) 的表达式基本不变。
- 记$e_i = y_i - \hat{\beta} x_i$，则正规方程可写为

$$\sum_{i=1}^{n} x_i e_i = 0$$

- 记拟合值$\hat{y}_i \equiv \hat{\beta} x_i$，容易证明残差仍与拟合值正交

$$\sum_{i=1}^{n} \hat{y}_i e_i = \sum_{i=1}^{n} \hat{\beta} x_i e_i = \hat{\beta} \sum_{i=1}^{n} x_i e_i = \hat{\beta} \cdot 0 = 0$$

- 因此，仍可利用OLS的正交性将$\sum_{i=1}^{n}\hat{y}_i^2$分解为：

$$\sum_{i=1}^{n}y_i^2=\sum_{i=1}^{n}(\hat{y}_i+e_i)^2=\sum_{i=1}^{n}\hat{y}_i^2+2\underbrace{\sum_{i=1}^{n}\hat{y}_i e_i}_{=0}+\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}\hat{y}_i^2+\sum_{i=1}^{n}e_i^2$$

- $\sum_{i=1}^{n}\hat{y}_i^2$为可由模型解释的部分，而$\sum_{i=1}^{n}e_i^2$为模型不可解释的部分。

- 定义 **非中心R^2** (uncentered R^2)

$$R_{uc}^2=\frac{\sum_{i=1}^{n}\hat{y}_i^2}{\sum_{i=1}^{n}y_i^2}$$

  - 如果无常数项， Stata 汇报的R^2正是R_{uc}^2
  - R_{uc}^2与R^2的定义不同，二者不具有可比性，但在 Stata 中都称为“R-squared”(在无常数项回归时汇报R)。

谢谢！