



计量经济学(Econometrics)

胡华平

西北农林科技大学

经济管理学院数量经济教研室

huhuaping01@hotmail.com

2021-09-07

西北农林科技大学

第5章：一元回归：模型形式扩展

5.1 过原点回归

5.2 尺度与测量单位

5.3 标准化变量回归

5.4 对数线性模型

5.5 半对数模型

5.6 倒数模型

5.7 函数模型的选择

5.1 过原点回归



过原点回归的模型形式

过原点回归 (regression through the origin) : 没有截距项的线性模型

在实践中, 双变量PRM过原点回归采取如下的形式:

$$Y_i = \beta_2 X_i + u_i$$

适用于这种模型的例子:

- 弗里德曼的持久收入假说(permanent income hypothesis);
- 资本资产定价模型(the capital Asset Pricing Model, CAPM)等。



资本资产定价模型 (CAPM)

资本资产定价模型(the capital Asset Pricing Model, CAPM):

$$(ER_i - r_f) = \beta_i (ER_m - r_f)$$

其中:

- ER_i 证券 i 的期望回报率;
- ER_m 市场证券组合的期望回报率 (如标准普尔S&P500综合股票指数);
- r_f 为无风险回报率 (90天国债回报率)。
- β_i 为系数, 表明第 i 种证券回报率与市场互动程度的度量。(注:不要把这个 β_i 和双变量回归的斜率系数 β_2 混同起来。)

一个大于1的 β_i 意味着证券 i 是一种易波动或进攻型证券; 一个小于1的 β_i 意味着证券 i 是一种防御型证券。



资本资产定价模型 (CAPM)

$$R_i - r_f = \beta_i (R_m - r_f) + u_i$$

$$R_i - r_f = \alpha_i + \beta_i (R_m - r_f) + u_i$$

- 如果CAPM成立，则预期 α_i 为0。
- 这样的模型如何估计呢？



资本资产定价模型 (CAPM)

这类模型的SRM可以写成：

$$Y_i = \hat{\beta}_2 X_i + e_i$$

OLS方法下求解回归系数：

$$\sum e_i^2 = \sum (Y_i - \hat{\beta}_2 X_i)^2$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_2} = 2 \sum (Y_i - \hat{\beta}_2 X_i) (-X_i) = 0$$

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2} = \frac{\sum X_i (\beta_2 X_i + u_i)}{\sum X_i^2} = \beta_2 + \frac{\sum X_i u_i}{\sum X_i^2}$$

$$E(\hat{\beta}_2) = \beta_2$$



资本资产定价模型 (CAPM)

OLS方法下求解得到的方差:

$$\begin{aligned} \text{Var}(\hat{\beta}_2) &= E(\hat{\beta}_2 - \beta_2)^2 \\ &= E\left[\frac{\sum X_i u_i}{\sum X_i^2}\right]^2 \\ &= \frac{\sigma^2}{\sum X_i^2} \end{aligned}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-1}; \quad E(\hat{\sigma}^2) = \sigma^2$$



资本资产定价模型 (CAPM)

OLS估计量对比：无截距和有截距的差异：

$$Y_i = \hat{\beta}_2 X_i + e_i$$

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-1}$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}$$

第一，对有截距项的模型来说，总有 $\sum e_i = 0$ ；对无截距项的模型来说， $\sum e_i = 0$ 不一定成立，只有 $\sum e_i X_i = 0$ 成立。

第二，对有截距项的模型，判定系数 $r^2 \geq 0$ ；但是，对无截距模型来说， r^2 时可能出现负值。



资本资产定价模型 (CAPM)

过原点回归的判定系数 r^2 的计算公式如下：

$$TSS = \sum y_i^2 = \sum Y_i^2 - n\bar{Y}^2$$

$$RSS = \sum e_i^2 = \sum Y_i^2 - \hat{\beta}_2^2 \sum X_i^2$$

$$r^2 = 1 - \frac{RSS}{TSS} > 0; \quad r^2 = 1 - \frac{RSS}{TSS} < 0;$$

因此，对于无截距模型，我们给出拟合优度指标为毛判定系数 (Raw r^2)：

$$Raw \quad r^2 = \frac{\sum (X_i Y_i)^2}{\sum X_i^2 \sum Y_i^2}$$



资本资产定价模型 (CAPM)

启示:

- 第一, 尽管模型含有截距项, 但若该项的出现是统计上不显著的(即统计上等于零), 则从任何实际方面考虑, 都可认为这个结果是一个过原点回归模型。
- 第二, 如果在模型中确实有截距, 而我们却执意拟合一个过原点回归, 我们就犯了设定错误。



资本资产定价模型 (CAPM) : 数据

1980.01-1999.12年间104 种股票构成的一个指数的超额回报率 $Y_t(\%)$ 和英国总体股票指数的超额回报率 $X_t(\%)$ 的月度数据共 $n=240$ 个月观测。其中超额回报率指的是超过无风险资产回报率的部分。

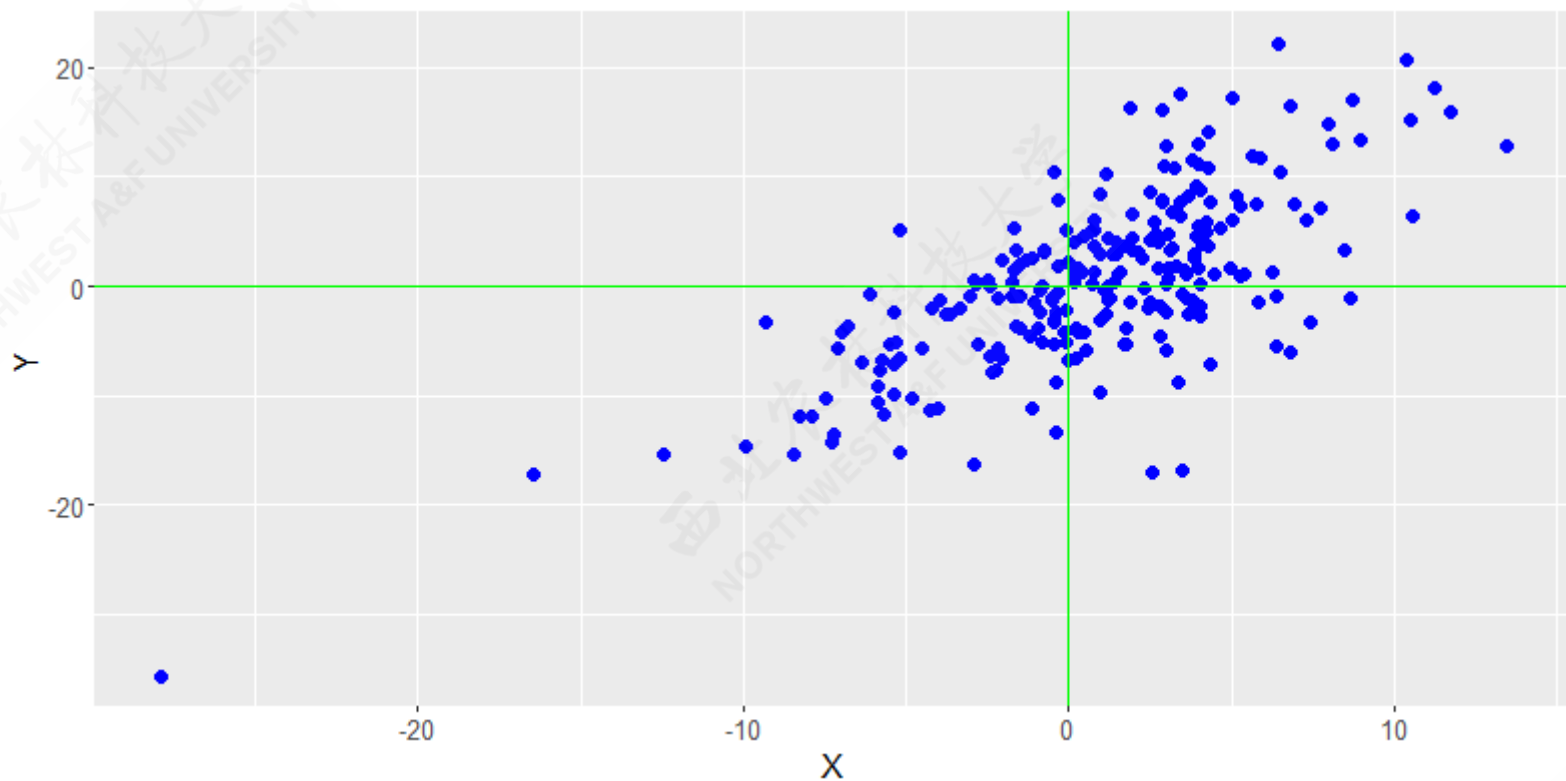
year	month	X	Y
1980	1	7.2634	6.0802
1980	2	6.3399	-0.9242
1980	3	-9.2852	-3.2862
1980	4	0.7933	5.2120
1980	5	-2.9024	-16.1642
1980	6	8.6132	-1.0547
1980	7	3.9821	11.1724
1980	8	-1.1502	-11.0633

Showing 1 to 8 of 240 entries



资本资产定价模型 (CAPM) : 散点图

下面先直接给出二者的散点图：





资本资产定价模型 (CAPM) : 回归结果

两类模型回归结果对比:

无截距模型:

$$Y_i = \hat{\beta}_2 X_i + e_i$$

$$\hat{Y} = + 1.16X$$

$$(t) \quad (15.5320)$$

$$(se) \quad (0.0744)$$

$$(\text{fitness}) R^2 = 0.5023; \bar{R}^2 = 0.5003$$

$$F^* = 241.24; p = 0.0000$$

有截距模型:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

$$\hat{Y} = - 0.45 + 1.17X$$

$$(t) \quad (-1.2329) \quad (15.5350)$$

$$(se) \quad (0.3629) \quad (0.0754)$$

$$(\text{fitness}) R^2 = 0.5035; \bar{R}^2 = 0.5014$$

$$F^* = 241.34; p = 0.0000$$

5.2 尺度与测量单位



案例数据

回归分析中，因变量Y和解释变量X的测量单位不同会造成回归结果的差异吗？

Year	GPDIB	GPDIM	GDPB	GDPM	GPDIB_std	GDPB_std
1990	886.6	886600	7112.5	7112500	-1.2942	-1.3459
1991	829.1	829100	7100.5	7100500	-1.4624	-1.3550
1992	878.3	878300	7336.6	7336600	-1.3185	-1.1768
1993	953.5	953500	7532.7	7532700	-1.0986	-1.0287
1994	1042.3	1042300	7835.5	7835500	-0.8389	-0.8002

Showing 1 to 5 of 16 entries

Previous

1

2

3

4

Next

>

- GPDIB = 以2000年10亿(Billions)美元计国内私人总投资； GPDIM = 以2000年百万(millions)美元计国内私人总投资；



尺度变换

把某一测量单位下的回归模型，变换为另一测量单位的回归模型：

$$\begin{aligned} Y_i &= \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i \\ Y_i^* &= \hat{\beta}_1^* + \hat{\beta}_2^* X_i + e_i^* \end{aligned}$$

尺度因子： $\omega_1; \omega_2$ 分别表示为Y和X的尺度因子。

$$Y_i^* = \omega_1 Y_i \quad X_i^* = \omega_2 X_i$$

如果 $(Y_i; X_i)$ 都是以10亿(billion)美元计量的，我们把它们改为用百万(million)美元去度量，就会有：

$$Y_i^* = 1000Y_i; \quad X_i^* = 1000X_i; \quad \omega_1 = \omega_2 = 1000$$



OLS估计

进行数据转换，新模型的OLS估计量如下：

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

$$Y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* X_i + e_i^*$$

$$Y_i^* = \omega_1 Y_i; \quad X_i^* = \omega_2 X_i; \quad e_i^* = \omega_1 e_i$$

$$\begin{aligned} \hat{\beta}_2^* &= \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}} & \Leftarrow \text{var}(\hat{\beta}_2^*) &= \frac{\sigma^{*2}}{\sum x_i^{*2}} \\ \hat{\beta}_1^* &= \bar{Y}^* - \hat{\beta}_2^* \bar{X}^* & \Leftarrow \text{var}(\hat{\beta}_1^*) &= \frac{\sum X_i^{*2}}{n \sum x_i^{*2}} \cdot \sigma^{*2} \\ \hat{\sigma}^{*2} &= \frac{\sum e_i^{*2}}{n - 2} \end{aligned}$$



OLS估计

进行数据转换，两个模型下OLS估计量有如下关系：

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

$$Y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* X_i + e_i^*$$

$$Y_i^* = \omega_1 Y_i; \quad X_i^* = \omega_2 X_i; \quad e_i^* = \omega_1 e_i$$

$$\hat{\beta}_2^* = \left(\frac{\omega_1}{\omega_2} \right) \hat{\beta}_2 \quad \Leftarrow \text{var}(\hat{\beta}_2^*) = \left(\frac{\omega_1}{\omega_2} \right)^2 \text{var}(\hat{\beta}_2)$$

$$\text{var}(\hat{\beta}_2^*) = \left(\frac{\omega_1}{\omega_2} \right)^2 \text{var}(\hat{\beta}_2) \quad \Leftarrow \text{var}(\hat{\beta}_1^*) = \omega_1^2 \text{var}(\hat{\beta}_1)$$

$$\hat{\sigma}^{*2} = \omega_1^2 \hat{\sigma}^2$$

$$r_{xy}^2 = r_{x^*y^*}^2$$



相关结论

$$\begin{aligned} Y_i &= \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i \\ Y_i^* &= \hat{\beta}_1^* + \hat{\beta}_2^* X_i + e_i^* \\ Y_i^* &= \omega_1 Y_i; \quad X_i^* = \omega_2 X_i; \quad e_i^* = \omega_1 e_i \end{aligned}$$

模型对比，得出如下主要结论：

- $\omega_1 = \omega_2$ ，即尺度因子相等时，斜率系数及其标准误不受尺度从 (Y_i, X_i) 到 (Y_i^*, X_i^*) 的影响。截距及其标准误却放大或缩小至 ω_1 倍。
- X_i 尺度不变 $\omega_2 = 1$ ， Y_i 尺度因子 ω_1 变化，那么，斜率和截距系数以及它们各自的标准误都要乘以同样的因子 ω_1 。
- Y_i 尺度不变 $\omega_1 = 1$ ，而 X_i 尺度因子 ω_2 变化，那么，斜率系数及其标准误都要乘以因子 $1/\omega_1$ ，而截距系数及其标准误不变。



案例分析结果

GPDI和GDP都以十亿美元计算：

$$\begin{aligned}\widehat{GPDIB} &= -926.09 + 0.25GDPB \\ (t) & \quad (-7.9590) \quad (19.5824) \\ (se) & \quad (116.3577) \quad (0.0129) \\ (fitness) & \quad R^2 = 0.9648; \bar{R}^2 = 0.9623 \\ & \quad F^* = 383.47; p = 0.0000\end{aligned}$$

GPDI和GDP都以百万美元计算：

$$\begin{aligned}\widehat{GPDIM} &= -926090.39 + 0.25GDPM \\ (t) & \quad (-7.9590) \quad (19.5824) \\ (se) & \quad (116357.6965)(0.0129) \\ (fitness) & \quad R^2 = 0.9648; \bar{R}^2 = 0.9623 \\ & \quad F^* = 383.47; p = 0.0000\end{aligned}$$

GPDI以十亿美元，而GDP以百万美元：

$$\begin{aligned}\widehat{GPDIB} &= -926.09 + 0.00GDPM \\ (t) & \quad (-7.9590) \quad (19.5824) \\ (se) & \quad (116.3577) \quad (0.0000) \\ (fitness) & \quad R^2 = 0.9648; \bar{R}^2 = 0.9623 \\ & \quad F^* = 383.47; p = 0.0000\end{aligned}$$

GPDI以百万美元而GDP以十亿美元：

$$\begin{aligned}\widehat{GPDIM} &= -926090.39 + 253.52GDPB \\ (t) & \quad (-7.9590) \quad (19.5824) \\ (se) & \quad (116357.6965)(12.9465) \\ (fitness) & \quad R^2 = 0.9648; \bar{R}^2 = 0.9623 \\ & \quad F^* = 383.47; p = 0.0000\end{aligned}$$

5.3 标准化变量回归



标准化变量回归

假设如下双变量回归：

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

对Y和X作如下标准化变换，得到相应的标准化变量：

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}; \quad X_i^* = \frac{X_i - \bar{X}}{S_X}$$

- 标准化变量的特征是：其均值总是0 和标准差总是1。



标准化变量回归

得到如下新的双变量回归模型:

$$\begin{aligned} Y_i^* &= \hat{\beta}_1^* + \hat{\beta}_2^* X_i + e_i^* \\ &= \hat{\beta}_2^* X_i + e_i^* \end{aligned}$$

- 对标准化的回归子和回归元做回归，截距项总是零!
- 实际上变成了过原点回归模型!



标准化变量回归

模型比较与结论：

- 第一，由于标准化回归本质上是一个过原点回归，而我们在已经指出通常过原点回归的不能使用 r^2 ，所以我们就没有给出其 r^2 值。
- 第二，传统模型的系数与这里的系数之间存在一种有趣的关系。在双变量情形中，这种关系如下（证明过程略：自学练习题！）：

$$\hat{\beta}_2^* = \frac{S_X}{S_Y} \hat{\beta}_2$$

- 第三，在多元回归中，变量标准化可以去除多个自变量之间数量尺度(量纲)的差别，因而具有一定的优点！



标准化数据变换

下面我们对以十亿美元计的GPDIB和GDPB进行标准化数据变换：

Year	GPDIB	GDPB	GPDIB_std	GDPB_std
1990	886.6	7112.5	-1.2942	-1.3459
1991	829.1	7100.5	-1.4624	-1.3550
1992	878.3	7336.6	-1.3185	-1.1768
1993	953.5	7532.7	-1.0986	-1.0287
1994	1042.3	7835.5	-0.8389	-0.8002
1995	1109.6	8031.7	-0.6421	-0.6521
1996	1209.2	8328.9	-0.3508	-0.4277
1997	1320.6	8703.5	-0.0250	-0.1450

Showing 1 to 8 of 16 entries

Previous

1

2

Next



OLS比较

GPGDI和GDP都以十亿美元计算:

$$\begin{aligned} \widehat{GPDIB} &= -926.09 + 0.25GDPB \\ (t) & \quad (-7.9590) \quad (19.5824) \\ (se) & \quad (116.3577) \quad (0.0129) \\ (fitness) & \quad R^2 = 0.9648; \bar{R}^2 = 0.9623 \\ & \quad F^* = 383.47; p = 0.0000 \end{aligned}$$

标准化变量后的模型估计:

$$\begin{aligned} \widehat{GPDIB}_{std} &= +0.98GDPB_{std} \\ (t) & \quad (20.2697) \\ (se) & \quad (0.0485) \\ (fitness) & \quad R^2 = 0.9648; \quad \bar{R}^2 = 0.9624 \\ & \quad F^* = 410.86; \quad p = 0.0000 \end{aligned}$$

回归模型的函数形式

对数线性模型

半对数模型

倒数模型

5.4 对数线性模型



对数线性模型的形式

指数回归模型 (exponential regression model)

$$Y_i = \beta_1 X_i^{\beta_2} e^{u_i}$$

可化为：

$$\ln Y_i = \ln \beta_1 + \beta_2 \ln X_i + u_i$$

$$\ln Y_i = \alpha + \beta_2 \ln X_i + u_i \quad \Leftarrow \alpha = \ln \beta_1$$

这种模型被称为对数-对数(log-log)，双对数(double-log)或对数一线性(log-linear)模型。进而有：

$$Y_i^* = \alpha + \beta_2 X_i^* + u_i \quad \Leftarrow [Y_i^* = \ln Y_i; \quad X_i^* = \ln X_i]$$

从而可用OLS方法可以得到BLUE估计量：

$$Y_i^* = \hat{\alpha} + \hat{\beta}_2 X_i^* + e_i$$



对数线性模型：学会如何测度弹性

双数线性模型：

$$\ln Y_i = \alpha + \beta_2 \ln X_i + u_i$$

$$Y_i^* = \hat{\alpha} + \hat{\beta}_2 X_i^* + e_i \quad \Leftarrow \hat{\alpha} = \ln \hat{\beta}_1$$

$$\beta_2 = \frac{d(\ln Y)}{d(\ln X)} = \frac{\frac{1}{Y} dY}{\frac{1}{X} dX} = \frac{dY/Y}{dX/X}$$

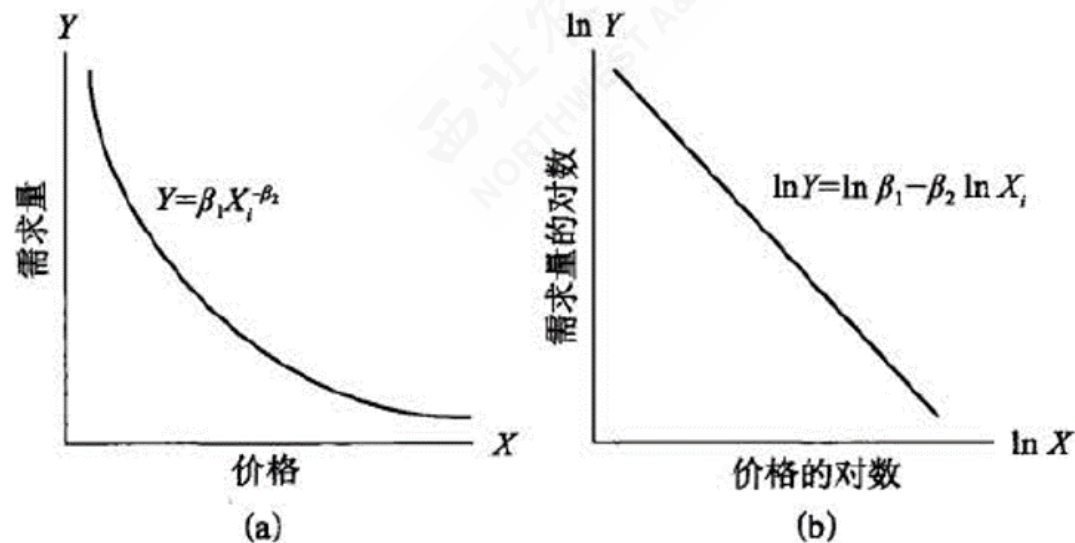
斜率就是Y对X的弹性！如果Y代表商品需求量Q，X代表商品价格P，则表示该商品的需求价格弹性。



学会如何测度弹性

双数线性模型有如下性质：

- Y对X的弹性在整个研究范围内是常数，一直为 β_2 ，因此这种模型也称为不变弹性模型(constant elasticity model)。
- 虽然 $\hat{\alpha}$ 和 $\hat{\beta}_2$ 是无偏估计量，但是进入原始模型的参数 β_1 的估计值 $\hat{\beta}_1$ 却是有偏估计，而且 $\beta_1 = \text{antilog} \hat{\alpha}$ 。





耐用品消费案例

耐用品支出与个人消费总支出的关系：

obs	EXPDUR	PCEXP	ln_expdur	ln_pcexp
2003-I	971.4	7184.9	6.8787	8.8797
2003-II	1009.8	7249.3	6.9175	8.8887
2003-III	1049.6	7352.9	6.9562	8.9029
2003-IV	1051.4	7394.3	6.9579	8.9085
2004-I	1067	7479.8	6.9726	8.9200

Showing 1 to 5 of 15 entries

Previous

1

2

3

Next

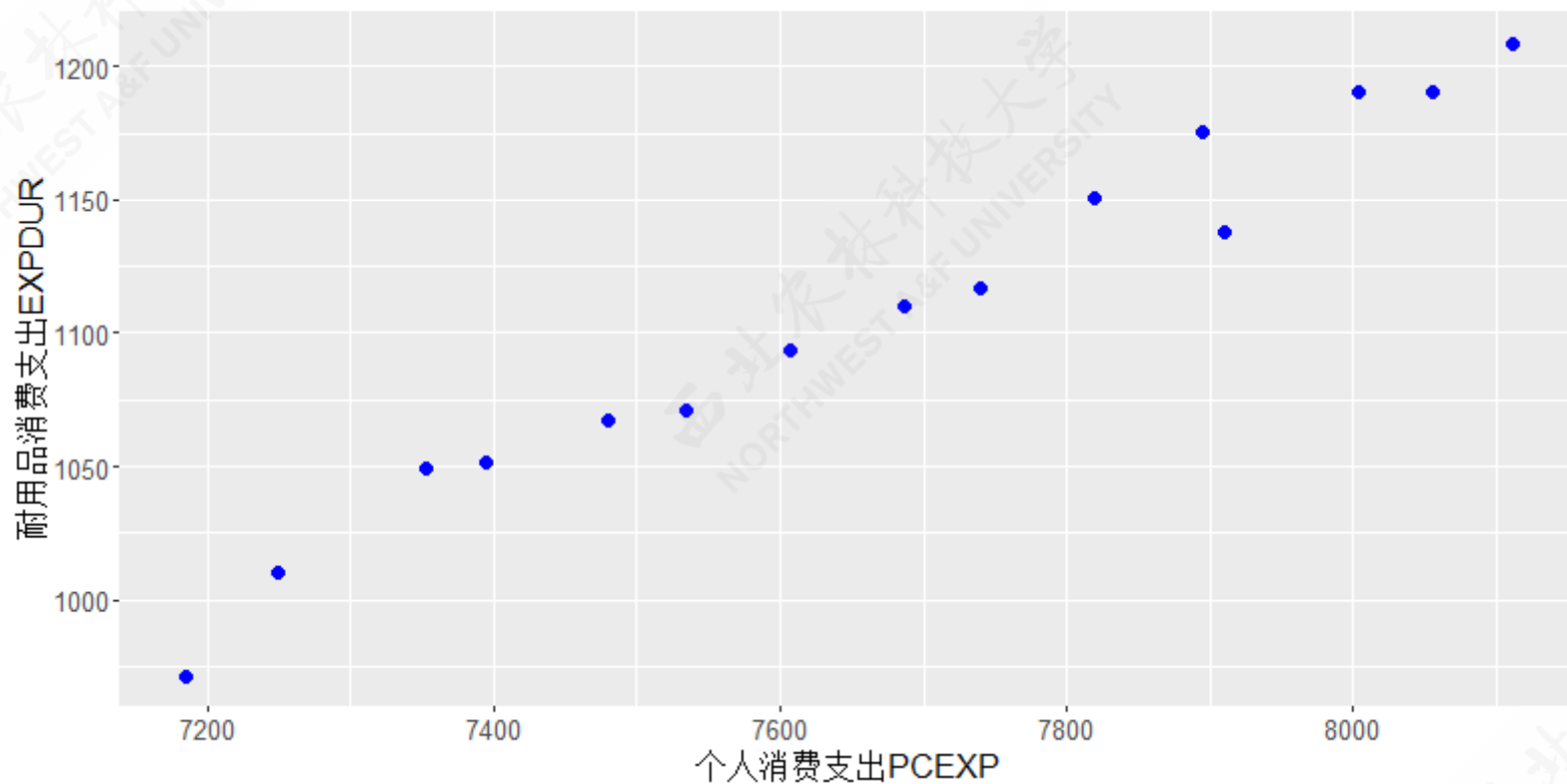
其中：PCEXP=个人消费支出, EXPDUR=耐用品消费支出，单位10亿美元（按2000年价格计）



耐用品消费案例

假设我们想求出耐用品支出对个人消费总支出的斜率。

将耐用品支出相对于个人消费总支出做散点图：

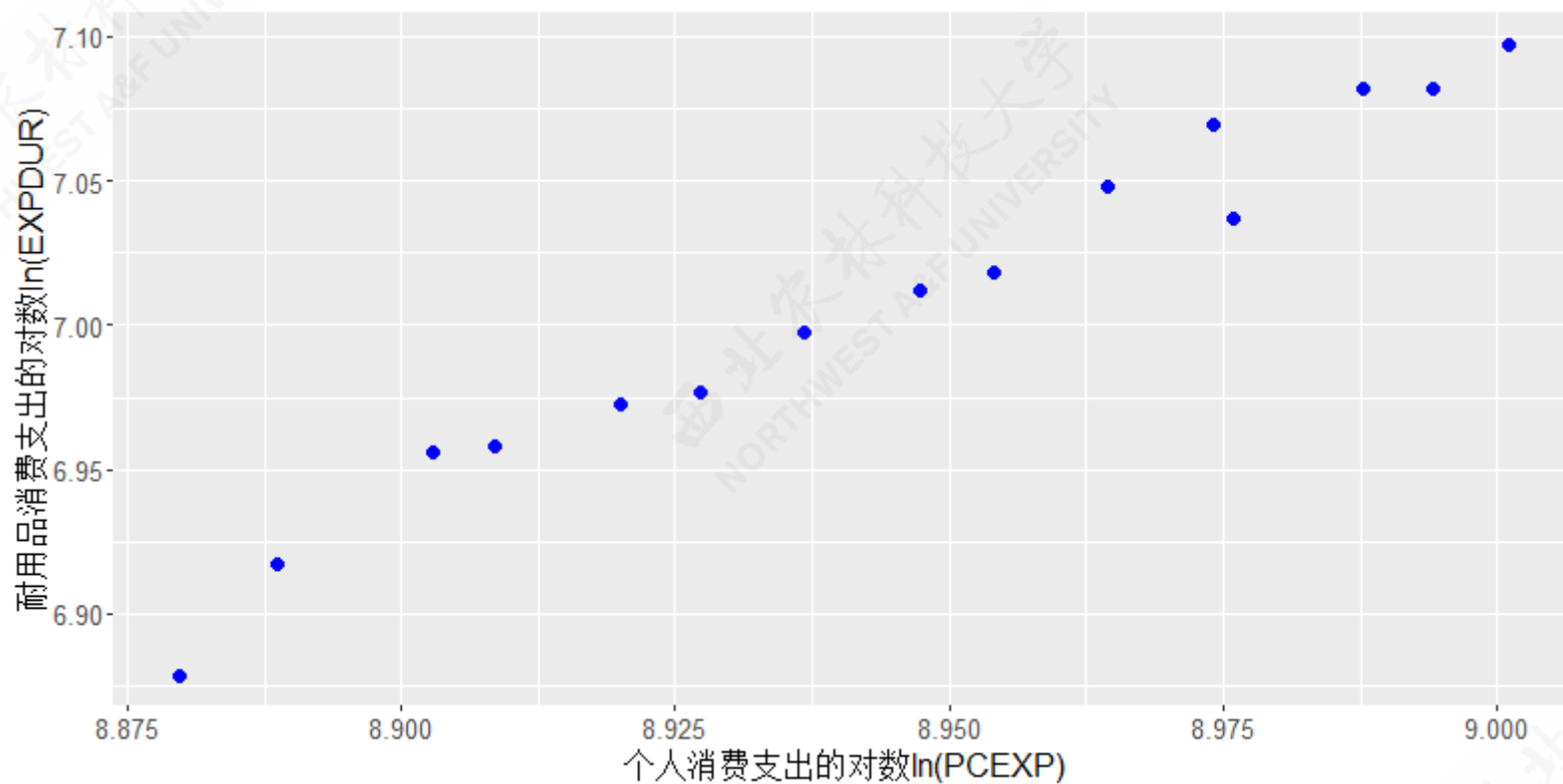




耐用品消费案例

假设我们想求出耐用品支出对个人消费总支出的弹性。

将耐用品支出的对数相对于个人消费总支出的对数做散点图：





耐用品消费案例

耐用品消费案例中，我们可以实证得到如下的双对数模型：

$$\begin{aligned} \log(\widehat{EXPDUR}) &= -7.54 && + 1.63\log(PCEXP) \\ (t) &&& (-10.5309) \quad (20.3152) \\ (se) &&& (0.7161) \quad (0.0801) \\ (fitness) &&& R^2 = 0.9695; \bar{R}^2 = 0.9671 \\ &&& F^* = 412.71; p = 0.0000 \end{aligned}$$

5.5 半对数模型



线性到对数模型

怎样测量增长率？经济学家、企业人员与政府常常对于求出某些经济变量的增长率感兴趣，如人口、GNP、货币供给、就业、生产力、贸易赤字等。

$$Y_t = Y_0(1 + r)^t$$

Y_t = 时期t的劳务实际支出； Y_0 = 劳务实际支出的初始值(为2002年第四季度末的值)； r是Y的复合增长率。

劳务支出数据

obs	t	EXPSERVICES	ln_expservices
2003-I	1	4143.3	8.3292
2003-II	2	4161.3	8.3336
2003-III	3	4190.7	8.3406
2003-IV	4	4220.2	8.3476

Showing 1 to 4 of 15 entries

Previous

1

2

3

4

Next



线性到对数模型

半对数模型(semilog models):

- 线性到对数模型(log-lin model): 只有回归子Y取对数
- 对数到线性模型(lin-log model): 只有回归元X取对数



线性到对数模型

半对数模型的形式：

$$Y_t = Y_0(1 + r)^t$$

$$\ln Y_t = \ln Y_0 + t \ln(1 + r)$$

$$\ln Y_t = \beta_1 + \beta_2 t \quad \Leftarrow [\beta_1 = \ln Y_0; \quad \beta_2 = \ln(1 + r)]$$

$$\ln Y_t = \beta_1 + \beta_2 t + u_t$$

斜率 β_2 的经济学含义：

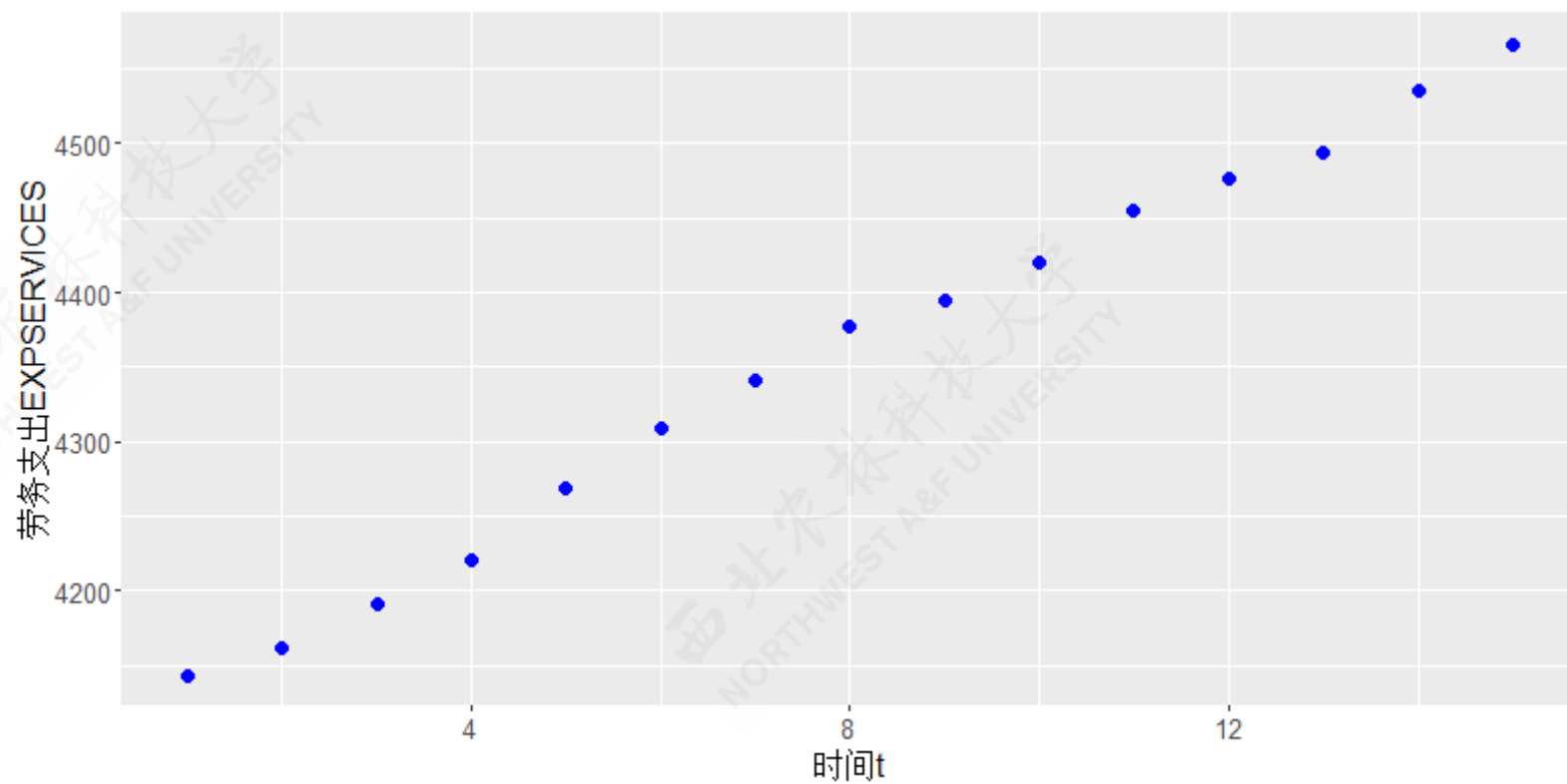
$$\beta_2 = \frac{d \ln Y}{dt} = \frac{dY/Y}{dt}$$

恒定相对增长率模型：上述模型描述了因变量Y的恒定相对增长率

- 恒定相对增长模型： $\beta_2 > 0$
- 恒定相对衰减模型： $\beta_2 < 0$

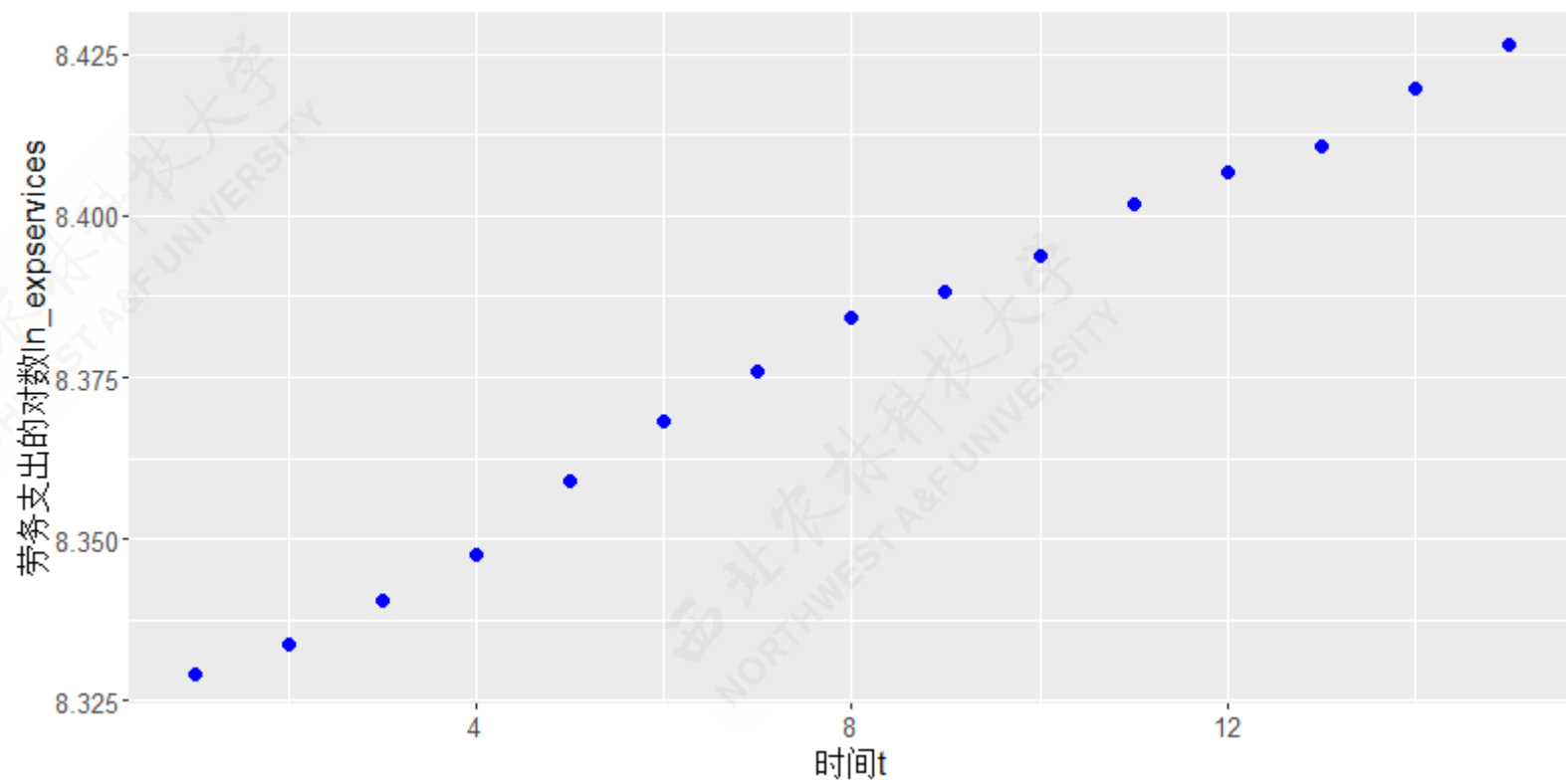


散点图1





散点图2





半对数模型：线性到对数模型

$$\ln Y_t = \beta_1 + \beta_2 t + u_t \quad \Leftrightarrow [\beta_1 = \ln Y_0; \quad \beta_2 = \ln(1 + r)]$$

$$\begin{aligned} \log(\widehat{EXP SERVICES}) &= + 8.32 & + 0.01t \\ (t) & (5186.2999) & (39.9648) \\ (se) & (0.0016) & (0.0002) \\ (fitness) & R^2 = 0.9919; \bar{R}^2 = 0.9913 \\ & F^* = 1597.18; p = 0.0000 \end{aligned}$$

$$\begin{aligned} \hat{\beta}_2 &= \ln(1 + r) = 0.00705 \\ r &= \text{antilog}(\hat{\beta}_2) - 1 \\ &= \text{antilog}(0.00705) - 1 \\ &= 0.00708 \end{aligned}$$

- $\hat{\beta}_2 = 0.00705$ 表示瞬时增长率
- $r = 0.00708$ 表示复合增长率



回归结果比较

下面做一个对比模型。线性趋势模型：Y直接对时间t回归：

$$Y_t = \beta_1 + \beta_2 t + u_t$$

$$\begin{aligned} \widehat{EXPSE\text{RVICES}} &= + 4111.54 & + 30.67t \\ (t) & (655.5628) & (44.4671) \\ (se) & (6.2718) & (0.6898) \\ (fitness) & R^2 = 0.9935; \bar{R}^2 = 0.9930 \\ & F^* = 1977.32; p = 0.0000 \end{aligned}$$

解释如下：在2003年第1季度至2006年第3季度期间，劳务支出以每季度约300 亿美元的绝对速度(注意不是相对速度)增加，即劳务支出有上涨的趋势。



对数到线性模型 (lin-log model)

如果我们的目的是测量X的一个百分比变化时，Y的绝对变化量，则要用对数到线性模型 (lin-log model)。

$$Y_i = \beta_1 + \beta_2 \ln X_i + u_i$$

$$\beta_2 = \frac{dY}{d \ln X} = \frac{dY}{dX/X} = \frac{\Delta Y}{\Delta X/X}$$

$$\Delta Y = \beta_2 \frac{\Delta X}{X}$$

例如：恩格尔支出(Engel expenditure) 模型：

- “用于食物的总支出以算术级数增加，而总支出以几何级数增加。”



家庭食物支出案例

食物支出（foodexp）与家庭总支出（totalexp）的关系：

obs	foodexp	totalexp	ln_totalexp
1	217	382	5.9454
2	196	388	5.9610
3	303	391	5.9687
4	270	415	6.0283
5	325	456	6.1225

Showing 1 to 5 of 55 entries

Previous

1

2

3

4

5

...

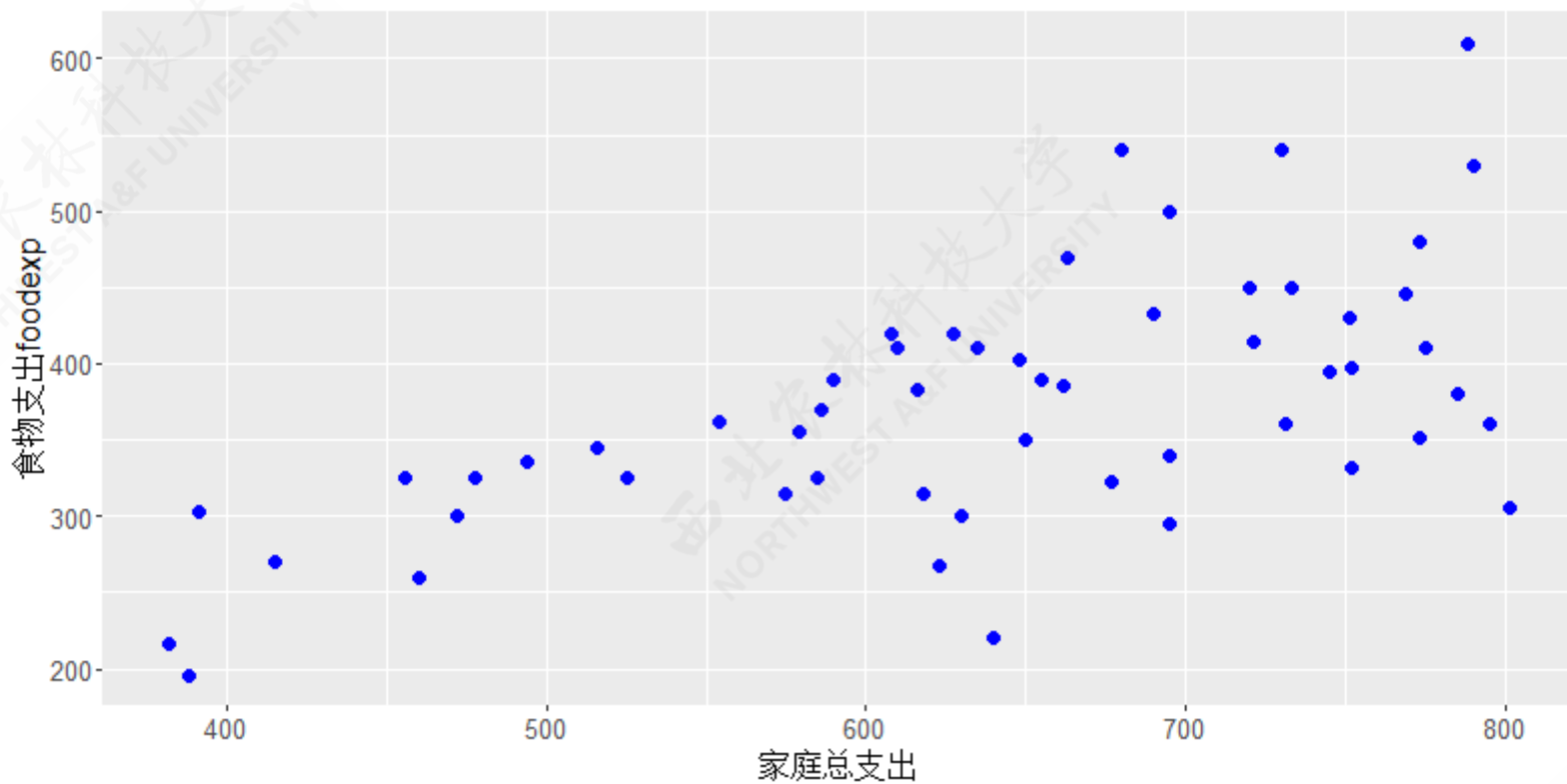
11

Next



家庭食物支出案例

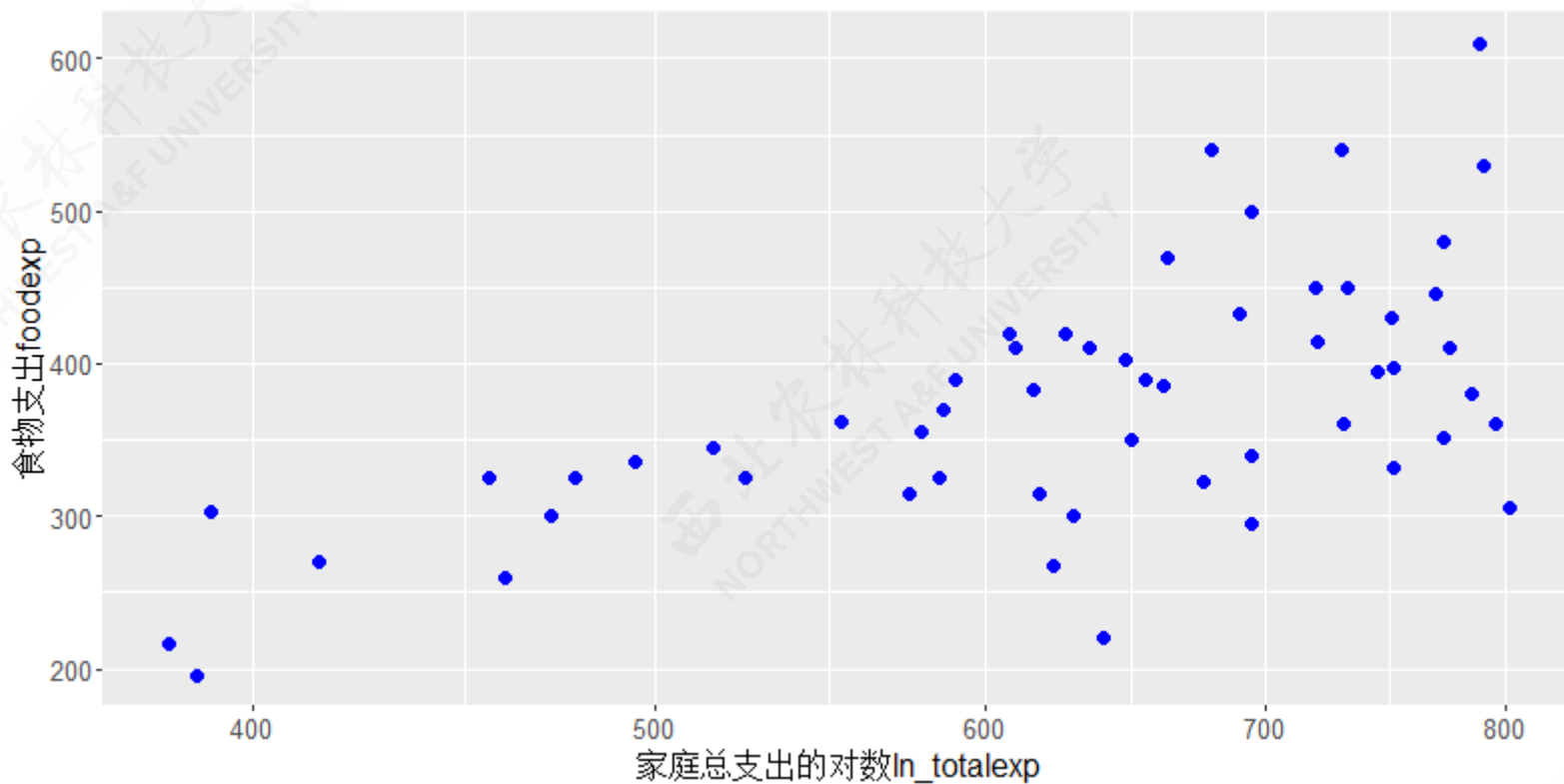
原始数据作散点图：





家庭食物支出案例

对家庭总支出取对数 $\ln(\text{totalexp})$ ，再做散点图：





家庭食物支出案例

构建如下对数到线性模型：

$$Y_t = \beta_1 + \beta_2 \ln X_i + u_i$$

家庭食物支出案例的OLS估计结果如下：

$$\begin{aligned} \widehat{foodexp} &= -1283.91 + 257.27 \log(totalexp) \\ (t) & \quad (-4.3848) \quad (5.6625) \\ (se) & \quad (292.8105) \quad (45.4341) \\ (fitness) & \quad R^2 = 0.3769; \bar{R}^2 = 0.3652 \\ & \quad F^* = 32.06; p = 0.0000 \end{aligned}$$

对比构建如下经典线性模型及其OLS估计结果：

$$Y_t = \beta_1 + \beta_2 X_i + u_i$$

$$\begin{aligned} \widehat{foodexp} &= +94.21 + 0.44 totalexp \\ (t) & \quad (1.8524) \quad (5.5770) \\ (se) & \quad (50.8563) \quad (0.0783) \\ (fitness) & \quad R^2 = 0.3698; \bar{R}^2 = 0.3579 \\ & \quad F^* = 31.10; p = 0.0000 \end{aligned}$$

5.6 倒数模型



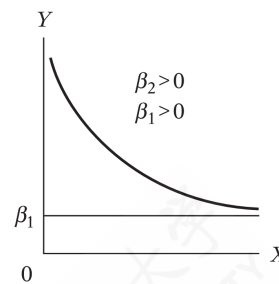
倒数模型

形式:

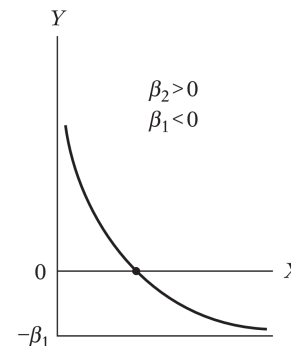
$$Y_i = \beta_1 + \beta_2 \left(\frac{1}{X_i} \right) + u_i$$

特征: 总有一条内在的渐近线!

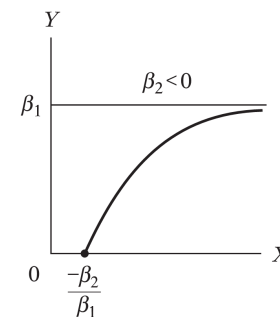
- a. 平均固定成本(AFC)曲线
- b. 菲利普斯曲线 (Phillips curve)
- c. 恩格尔曲线 (the Engel expenditure curve)



(a)



(b)



(c)

$$X \rightarrow \infty; \quad \beta_2 \left(\frac{1}{X} \right) \rightarrow 0; \quad Y \rightarrow \beta_1$$



儿童死亡率案例

儿童死亡率（CM，千分数）与人均GNP（PGNP，1980年的人均GNP）的关系：

obs	CM	PGNP	rep_PGNP
1	128	1870	0.0005
2	204	130	0.0077
3	202	310	0.0032
4	197	570	0.0018
5	96	2050	0.0005

Showing 1 to 5 of 64 entries

Previous

1

2

3

4

5

...

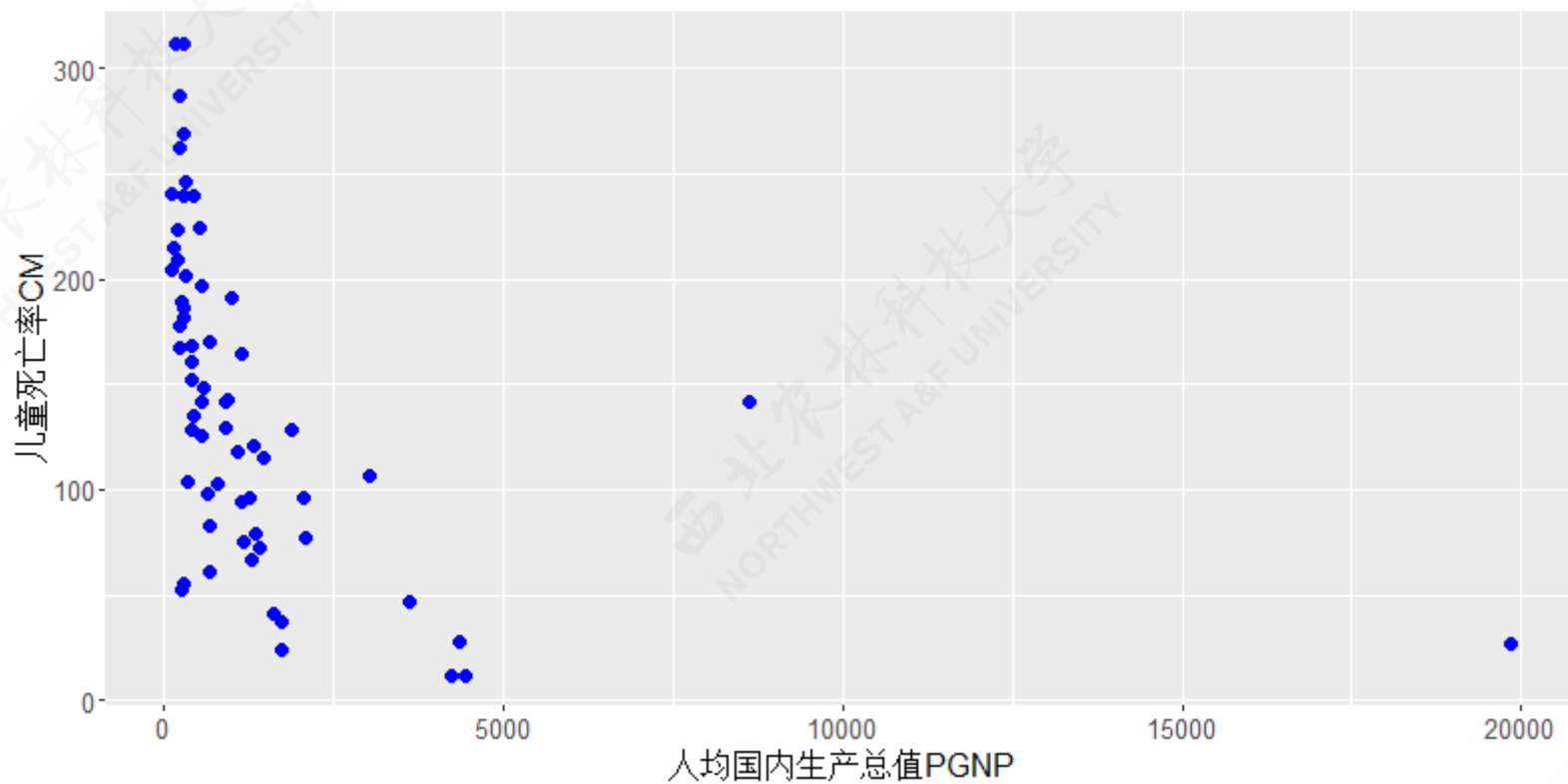
13

Next



儿童死亡率案例

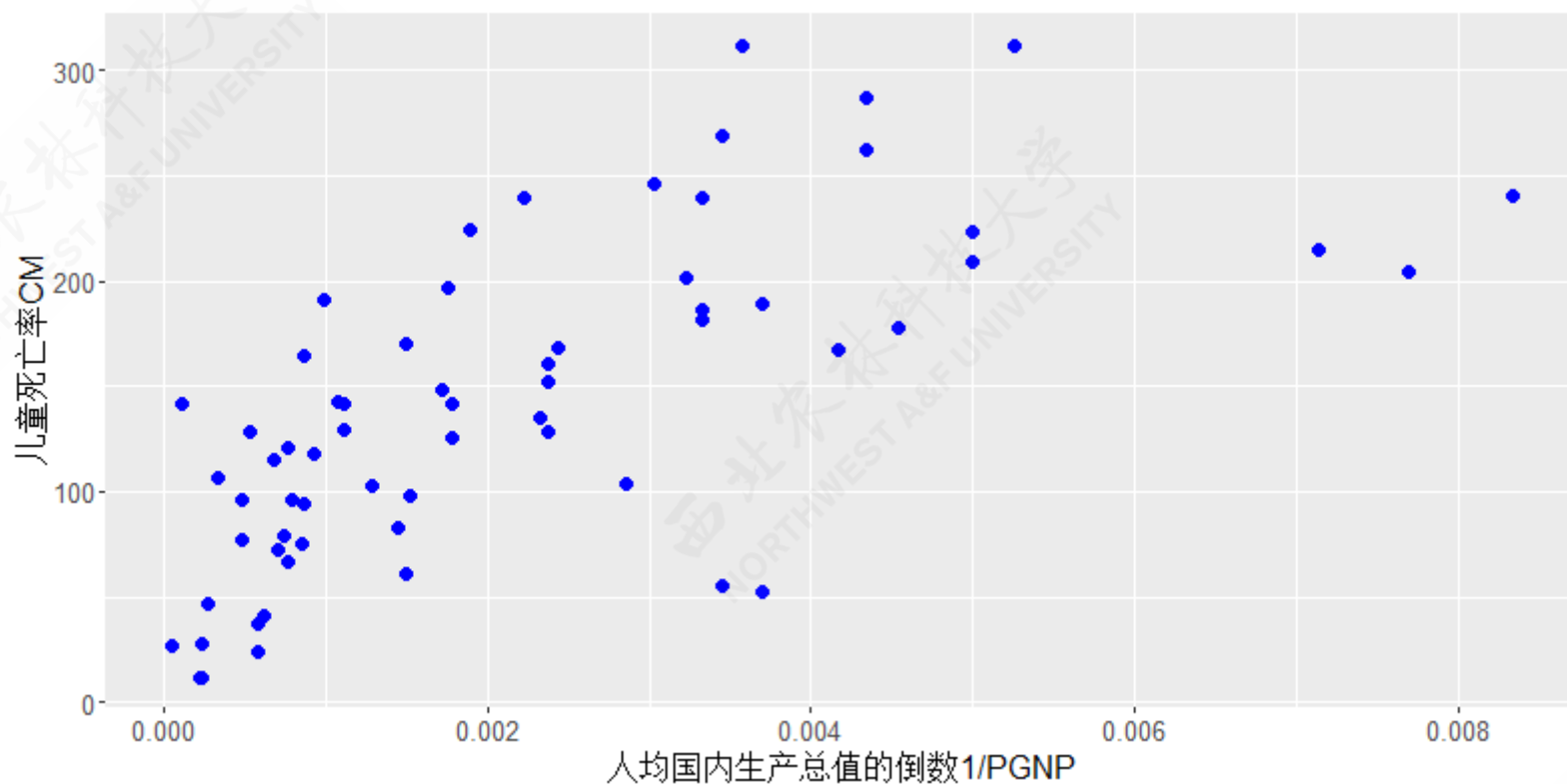
原始数据作散点图：





儿童死亡率案例

把PGNP取倒数 $1/PGNP$ 再作散点图：





儿童死亡率案例

构建如下倒数模型：

$$Y_i = \beta_1 + \beta_2 \left(\frac{1}{X_i} \right) + u_i$$

儿童死亡率案例倒数模型的OLS估计结果如下：

$$\begin{aligned} \widehat{CM} &= +81.79 & +27273.17 rep_{PGNP} \\ (t) & (7.5511) & (7.2535) \\ (se) & (10.8321) & (3759.9992) \\ (fitness) R^2 &= 0.4591; \bar{R}^2 = 0.4503 \\ F^* &= 52.61; p = 0.0000 \end{aligned}$$

对比构建如下经典线性模型：

$$Y_t = \beta_1 + \beta_2 X_i + u_i$$

其OLS估计结果如下：

$$\begin{aligned} \widehat{CM} &= +157.42 & -0.01 PGNP \\ (t) & (15.9893) & (-3.5157) \\ (se) & (9.8456) & (0.0032) \\ (fitness) R^2 &= 0.1662; \bar{R}^2 = 0.1528 \\ F^* &= 12.36; p = 0.0008 \end{aligned}$$



菲利普斯曲线

通货膨胀率（infrate，%）与失业率（unrate，%）的关系：

year	infrate	unrate	rep_unrate
1960	1.7182	5.5	0.1818
1961	1.0135	6.7	0.1493
1962	1.0033	5.5	0.1818
1963	1.3245	5.7	0.1754
1964	1.3072	5.2	0.1923

Showing 1 to 5 of 47 entries

Previous

1

2

3

4

5

...

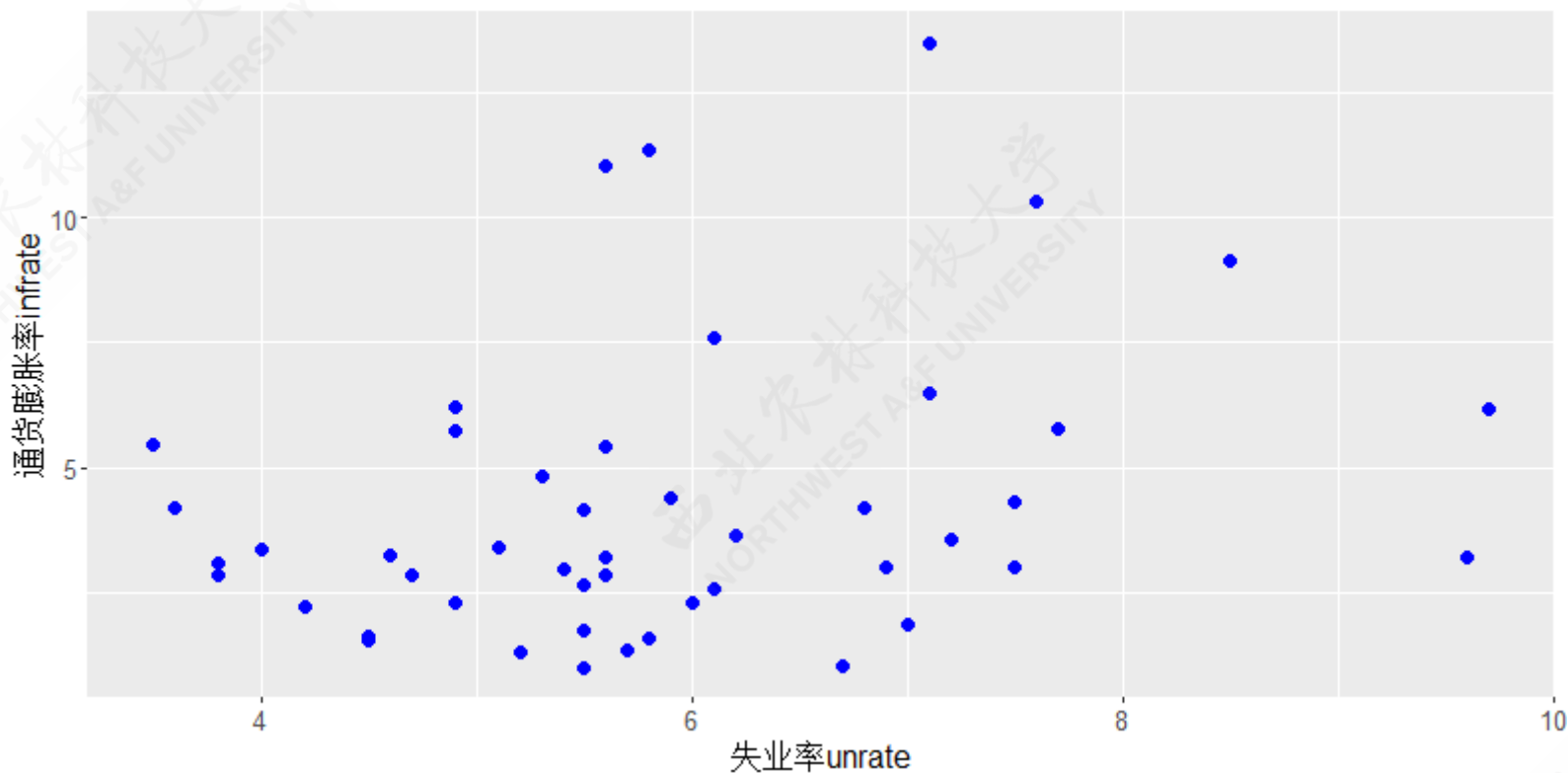
10

Next



菲利普斯曲线

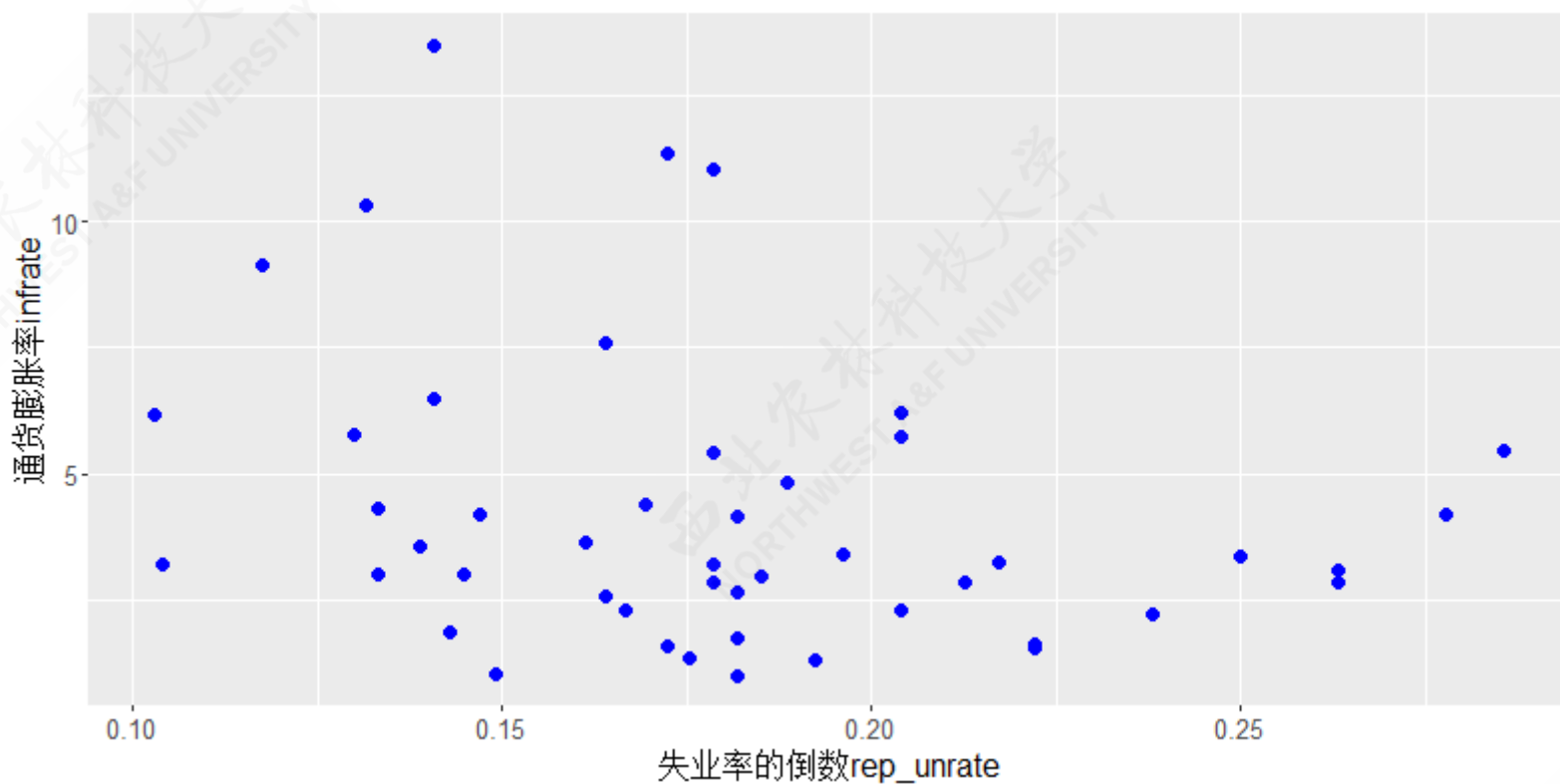
原始数据作散点图：





菲利普斯曲线

把失业率 `unrate` 取倒数 $1/unrate$ 再作散点图：





菲利普斯曲线

构建如下倒数模型：

$$Y_i = \beta_1 + \beta_2 \left(\frac{1}{X_i} \right) + u_i$$

菲利普斯曲线案例倒数模型的OLS估计结果如下：

$$\begin{aligned} \widehat{infrate} &= + 7.37 & - 17.37 rep_{unrate} \\ (t) & (4.1723) & (-1.8212) \\ (se) & (1.7670) & (9.5364) \\ (fitness) & R^2 = 0.0686; \bar{R}^2 = 0.0479 \\ & F^* = 3.32; p = 0.0752 \end{aligned}$$

对比构建如下经典线性模型及其OLS估计结果：

$$Y_t = \beta_1 + \beta_2 X_i + u_i$$

$$\begin{aligned} \widehat{infrate} &= + 0.81 & + 0.59 unrate \\ (t) & (0.4642) & (2.0377) \\ (se) & (1.7347) & (0.2874) \\ (fitness) & R^2 = 0.0845; \bar{R}^2 = 0.0641 \\ & F^* = 4.15; p = 0.0475 \end{aligned}$$

5.7 函数形式的选择



技巧和经验

选择适当模型时，需要一些技巧和经验：

1. 模型背后的理论(如菲利普斯曲线)可能给出了一个特定的函数形式。
2. 最好能求出回归子相对回归元的变化率〈即斜率〉和回归子对回归元的弹性(见下页ppt)。
3. 所选模型的系数应该满足一定的先验预期。
4. 有时多个模型都能相当不错地拟合一个给定的数据集。
5. 通常不应该过分强调这个指标
6. 在有些情形中，确定一个特定的函数形式不是那么容易，此时我们或许可以使用所谓的博克斯-考克斯变换(Box-Cox transformations)



计算表一览

模型	方程	斜率	点弹性	平均弹性
models	eq	$\frac{dY}{dX}$	$\frac{dY}{dX} \cdot \frac{X_i}{Y_i}$	$\frac{dY}{dX} \cdot \frac{\bar{X}}{\bar{Y}}$
M_1 线性模型	$Y_i = \beta_1 + \beta_2 X_i + u_i$	β_2	$\beta_2 X_i / Y_i$	$\beta_2 \bar{X} / \bar{Y}$
M_2 过原点模型	$Y_i = \beta_2 X_i + u_i$	β_2	$\beta_2 X_i / Y_i$	$\beta_2 \bar{X} / \bar{Y}$
M_3 双对数模型	$\ln(Y_i) = \beta_1 + \beta_2 \ln(X_i) + u_i$	$\beta_2 Y_i / X_i$	β_2	β_2
M_4 线性到对数模型	$\ln(Y_i) = \beta_1 + \beta_2 X_i + u_i$	$\beta_2 Y_i$	$\beta_2 X_i$	$\beta_2 \bar{X}$
M_5 对数到线性模型	$Y_i = \beta_1 + \beta_2 \ln(X_i) + u_i$	β_2 / X_i	β_2 / Y_i	β_2 / \bar{Y}
M_6 倒数模型	$Y_i = \beta_1 + \beta_2 / X_i + u_i$	$-\beta_2 / X_i^2$	$-\beta_2 / (X_i Y_i)$	$-\beta_2 / (\bar{X} \bar{Y})$
M_7 对数倒数模型	$\ln(Y_i) = \beta_1 - \beta_2 / X_i + u_i$	$\beta_2 Y_i / X_i^2$	β_2 / X_i	β_2 / \bar{X}

本章結束

