

计量经济学 II

复习提纲

教材和题型：

1. 学习和考试教材

本次考试，以古扎拉蒂的《计量经济学基础》（上下册）为准，电子版请登陆到网络教学平台下载学习。

Gujarati, D. and D. Porter. Basic Econometrics 5ed [M], McGraw-Hill Education, 2008.

2. 考试题型

- 整个测试只有一个题型：分析计算题！
- 整个测试会围绕 2-3 个案例数据展开，分大题分小题，一直提问下去……！

3. 不要紧张，没有想象的那么难！——换一种测试风格，研究生学习才更精彩！

下面，就好好看复习要点吧！———》》》》》

第一部分：多元回归分析（回顾本科阶段计量知识）

（一）约定的记号和写法

课程要求：熟练、准确地书写基本的模型和函数形式；

1. 总体回归模型（PRM）：

理论表达式：形如 $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

2. 总体回归函数（PRF）：

理论表达式：形如 $E(Y_i|X_i) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i}$

3. 样本回归模型（SRM）：

理论表达式：形如 $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + e_i$

数值表达式：形如 $Y_i = 20.5 + 0.6X_{2i} + 1.4X_{3i} + e_i$

4. 样本回归函数（SRF）：

理论表达式：形如 $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$

数值表达式为：形如 $\hat{Y}_i = 20.5 + 0.6X_{2i} + 1.4X_{3i}$

(二) 阅读统计软件给出的分析报告

课程要求：会熟练、正确阅读统计软件给出的各类分析报告，理解其中的关键信息和内涵。这些分析报告包括：传统的多元回归分析报告；时间序列分析报告；联立方程分析报告；logistics 回归分析报告；以及各种计量检验的辅助分析报告（如异方差 white 检验报告）等。

根据统计软件的不同（stata；Eview；R），各种分析报告呈现形式略有差异，但基本要素和信息都大抵一致。

起码要了解下面一些报告形式：

1. 整理好的**精炼报告**（形式 1：多行方程表达法）。

根据统计软件的原始报告，往往是选取最关键的信息，经过整理并以多行方程的形式呈现，**精炼报告**的形式一般为：

$$\begin{array}{llll} \widehat{lprice} = & +0.65 & -0.10lquan & -0.08mon \\ (t) & (1.4753) & (-1.9541) & (-0.7961) \\ (se) & (0.4408) & (0.0501) & (0.1042) \\ (cont.) & -0.08tue & -0.07wed & +0.05thu \\ (t) & (-0.7979) & (-0.6801) & (0.5341) \\ (se) & (0.1048) & (0.1077) & (0.1008) \\ (cont.) & +0.29stormy & +0.09cold & -0.15change \\ (t) & (3.6030) & (1.2080) & (-2.0013) \\ (se) & (0.0815) & (0.0713) & (0.0738) \\ (fitness) & n = 111; & R^2 = 0.2610; & \overline{R^2} = 0.2030 \\ & F^* = 4.50; & & p = 0.0001 \end{array} \quad (2)$$

2. 整理好的**精炼报告**（形式 2：表格列示法）。

根据统计软件的原始报告，往往是选取最关键的信息，经过整理以表格形式呈现，**表格列示法**的形式一般为：

表 8 回归分析的表格列示法

term	estimate	std.error	statistic	p.value
(Intercept)	0.6504	0.4408	1.4753	0.1432
lquan	-0.0978	0.0501	-1.9541	0.0534
mon	-0.0830	0.1042	-0.7961	0.4278
tue	-0.0836	0.1048	-0.7979	0.4268
wed	-0.0732	0.1077	-0.6801	0.4980
thu	0.0538	0.1008	0.5341	0.5944
stormy	0.2937	0.0815	3.6030	0.0005
cold	0.0861	0.0713	1.2080	0.2299
change	-0.1478	0.0738	-2.0013	0.0480

3. 原始报告——传统多元回归分析的 R 软件分析报告

一般呈现为:

Call:

```
lm(formula = mod_mult, data = fultonfish)
```

Residuals:

```
  Min   1Q Median   3Q   Max
-0.835 -0.240  0.027  0.252  0.726
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.6504   0.4408   1.48 0.14321
lquan      -0.0978   0.0501  -1.95 0.05343 .
mon        -0.0830   0.1042  -0.80 0.42780
tue        -0.0836   0.1048  -0.80 0.42677
wed        -0.0732   0.1077  -0.68 0.49801
thu         0.0538   0.1008   0.53 0.59441
stormy      0.2937   0.0815   3.60 0.00049 ***
cold        0.0861   0.0713   1.21 0.22986
change     -0.1478   0.0738  -2.00 0.04801 *
```

Signif. codes:

```
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.341 on 102 degrees of freedom

Multiple R-squared: 0.261, Adjusted R-squared: 0.203

F-statistic: 4.5 on 8 and 102 DF, p-value: 0.000104

4. 原始报告——时间序列回归分析的 R 软件分析报告

形式一般为（以ARMA(2,0)模型为例）：

```
Series: fultonfish$price
ARIMA(2,0,0) with non-zero mean

Coefficients:
    ar1    ar2    mean
 0.908 -0.198 -0.186
s.e. 0.095 0.095 0.079

sigma^2 estimated as 0.0625: log likelihood=-2.55
AIC=13.1  AICc=13.47  BIC=23.93

Training set error measures:
      ME  RMSE  MAE  MPE  MAPE  MASE
Training set 0.002292 0.2466 0.1993 Inf  Inf 0.9593
      ACF1
Training set -0.002839
```

5. 原始报告——联立方程回归分析的 R 软件报告

形式一般为（以两个方程的联立方程组回归分析为例。大家重点关注回归部分，诊断部分不需要关注！）：

```
systemfit results
method: 2SLS

      N DF SSR detRCov OLS-R2 McElroy-R2
system 222 213 110  0.107 0.094  -0.598
```

```
      N DF SSR MSE RMSE  R2 Adj R2
eq1 111 105 52.1 0.496 0.704 0.139 0.098
eq2 111 108 57.5 0.533 0.730 0.049 0.032
```

The covariance matrix of the residuals

```
      eq1 eq2
eq1 0.496 0.396
eq2 0.396 0.533
```

The correlations of the residuals

```
      eq1 eq2
eq1 1.000 0.771
eq2 0.771 1.000
```

2SLS estimates for 'eq1' (equation 1)

Model Formula: $\text{lquan} \sim \text{lprice} + \text{mon} + \text{tue} + \text{wed} + \text{thu}$

Instruments: $\sim \text{mon} + \text{tue} + \text{wed} + \text{thu} + \text{stormy}$

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.5059	0.1662	51.19	<2e-16 ***
lprice	-1.1194	0.4286	-2.61	0.010 *
mon	-0.0254	0.2148	-0.12	0.906
tue	-0.5308	0.2080	-2.55	0.012 *
wed	-0.5664	0.2128	-2.66	0.009 **
thu	0.1093	0.2088	0.52	0.602

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.704 on 105 degrees of freedom

Number of observations: 111 Degrees of Freedom: 105

SSR: 52.09 MSE: 0.496 Root MSE: 0.704

Multiple R-Squared: 0.139 Adjusted R-Squared: 0.098

2SLS estimates for 'eq2' (equation 2)

Model Formula: lquan ~ lprice + stormy

Instruments: ~mon + tue + wed + thu + stormy

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.62835	0.38897	22.18	<2e-16 ***
lprice	0.00106	1.30955	0.00	1.00
stormy	-0.36325	0.46491	-0.78	0.44

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.73 on 108 degrees of freedom

Number of observations: 111 Degrees of Freedom: 108

SSR: 57.522 MSE: 0.533 Root MSE: 0.73

Multiple R-Squared: 0.049 Adjusted R-Squared: 0.032

(三) 理解多元回归系数的含义、计算拟合值或预测值、拟合优度

课程要求：会正确理解回归系数的经济学含义；给定已知条件下，会利用回归方程结果，计算出Y的拟合值或预测值；掌握拟合优度的具体指标。

- 系数的含义：

双对数模型 $\ln(Y_i) = \hat{\beta}_1 + \hat{\beta}_2 \ln(X_i) + e_i$ 中， $\hat{\beta}_1$ 表示Y对X的弹性。

- 计算 Y_i 的拟合值或预测值：

模型 $\ln(Y_i) = 20 + 0.6X_i + e_i$ ，在给定 $X_i = 10$ 时， Y_i 的拟合值 $\hat{Y}_i = e^{(20+0.6*10)} = 1.957310^{\{11\}}$

- 拟合优度的主要指标：

在OLS估计法下的多元回归，其拟合优度指标有可决系数 R^2 或调整可决系数 \bar{R}^2 。

(注意，在Logistics回归里，采用的是ML估计法，其拟合优度测度另有相应指标！)

(四) 主要的统计学检验和计量经济学检验

课程要求：熟练掌握假设检验的基本原理和过程；根据不同检验目的和手段，会进行基本的检验和判断。

1. 假设检验的基本过程

- step1: 提出原假设 H_0 和备择假设 H_1
- step2: 构造合适的样本统计量（一般是 Z^* 统计量、 t^* 统计量、 F^* 统计量、 χ^{2*} 统计量等）

- step3: 给定显著性水平（如 $\alpha = 0.05$ ）和自由度（ f ）下查出**理论分布值**，如 $Z_{0.975}(f_z)$; $t_{0.975}(f_t)$; $F_{0.95}(f_1, f_2)$; $\chi^2_{0.95}(f_\chi)$ 等。（注意：对称性分布如Z分布和t分布，采用双尾查表，得到右侧正值；非对称的F分布和 χ^2 分布），采用单尾查表。其次，不同情况下，自由度f的取自要视情况而定，需要牢记自由度f!）
- step4: 比较**样本统计量**与**理论分布值**的大小关系，得出检验结论。如果**样本统计量大于理论分布值**，则表明在 $1 - \alpha = 95\%$ 置信度下是**显著的**，应**显著的拒绝原假设 H_0** ；如果**样本统计量小于理论分布值**，则表明在 $1 - \alpha = 95\%$ 置信度下是**不显著的**，也即不能**显著的拒绝原假设 H_0** ，所以暂时接受原假设 H_0 。

注意 1：原假设和备择假设，在不同检验目的下会不相同。如何才能更好地记忆呢？一条基本的**法则**就是：原假设 H_0 则一般是希望被拒绝的，从而希望得到备择假设 H_1 。——也就是说，原假设 H_0 永远只是“稻草人”，我们希望它被打到，真正的醉翁之意实际上是备择假设 H_1 ！

注意 2：除了使用 step4 的办法，我们还可以直接看**样本统计量**对应的 p 值，也可以判断是显著还是不显著！

2. 重要的几类检验：

- t 检验：主要是回归系数显著性检验.....
- F 检验：广泛用于模型整体显著性检验、约束条件模型检验等.....
- χ^2 检验：用于异方差怀特检验、logit 回归的模型整体显著性检验（对数似然 LR 检验）.....

（五）虚拟变量回归

课程要求：掌握定性变量如何用一组虚拟变量来表达；有截距模型和无截距模型的差异；给定初始条件，计算 Y 的拟合值或预测值。

- 给定条件下 Y 期望值的表达和计算:

给定{quan=2.3026, mon=1, tue=0, wed=0, thu=0, stormy=1, cold=1, change=0}下,

$E(\ln price | \dots)$ 可以表达和计算为:

$$\begin{aligned}
 E(\ln price \mid & \text{quan} = 2.3; \text{mon} = 1.0; \text{tue} = 0.0; \text{wed} = 0.0; \text{thu} = 0.0; \\
 & \text{stormy} = 1.0; \text{cold} = 1.0; \text{change} = 0.0) \\
 = & +\beta_1 + \beta_2(2.3) + \beta_3(1) + \beta_4(0) + \beta_5(0) \\
 & +\beta_6(0) + \beta_7(1) + \beta_8(1) + \beta_9(0) \\
 = & +\beta_1 + 2.3\beta_2 + \beta_3 \\
 & +\beta_7 + \beta_8
 \end{aligned}$$

(六) 多重共线性问题

课程要求: 掌握多重共线性问题的主要诊断方法, 尤其是方差膨胀因子(VIF)。

(七) 异方差问题

课程要求: 掌握异方差问题的主要诊断方法, 尤其是观测残差图初步识别法和怀特异方差检验(White test)法。

怀特异方差检验(White test)法:

- 需要注意怀特辅助诊断方程的设置特点!
- 怀特检验将会采用 χ^2 统计量进行假设检验!
- 熟记原假设和备择假设!

第二部分：probit 模型和 logit 模型

课程要求：理解二分类因变量回归模型的构建原理；掌握 probit 模型和 logit 模型的异同；重点要会解读 logit 回归模型的分析报告，尤其是回归系数的含义。

重要的知识点包括：

1. 识记 logit 累积概率函数和 probit 累积概率函数：

能写出二者的函数表达式；能识别 probit 累积概率密度曲线和 logit 累积概率密度曲线（见书本第 15 章,pg:571）。

2. 理解 logit 模型及表达：

$$\text{令 } Z_i = \beta_0 + \beta_1 X_i$$

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = Z_i = \beta_0 + \beta_1 X_i$$

3. 掌握 logit 模型的拟合优度指标（见书本第 15 章,pg:562）

计数 $R^2 = \frac{Y_i \text{ 的正确预测数}}{Y_i \text{ 的总观测数 } n}$ （重点掌握）。

麦克法登 R^2 （Mcfadden R^2 ）。

4. logit 回归方程的解读，尤其是回归系数的含义（见书本第 15 章,pg:562-563）

主要有三个层面的解读和计算：

- 简单系数解读（没多大意义）。大概过程是：（假定 X_2 也是二分类变量，且 X_2 前的回归系数 $\hat{\beta}_2 = 0.6$ ），若 X_2 由 0 变为 1 的话，则（保持其他因素不变下）Logit 估计值 L_i 将变为 0.6.....

- **机会比率**解读（有那么点实际意义）。大概过程是：（假定 Y 为二分类变量， X_2 也是二分类变量，且 X_2 前的回归系数 $\hat{\beta}_2 = 0.6$ ），若 X_2 由 0 变为 1 的话，则（保持其他因素不变下） $Y = 1$ 与 $Y = 0$ 的机会比率为 1.8221.....
- 给定条件下 $Y=1$ 的实际发生概率。计算稍微复杂，请自行训练和脑部吧！

第三部分：时间序列分析

课程要求：

1. 时间序列的平稳性检验——定量方法

（见书本第 21 章,pg:764）

对于时间序列变量，首先应该检验它是否平稳。

我们需要重点关注采用**增广迪基-富勒检验法（Augmented Dickey-Fuller Test, ADF）**检验时间序列变量是否平稳。

- 要注意**增广迪基-富勒检验法（Augmented Dickey-Fuller Test）**的基本检验原理
- 要会看**增广迪基-富勒检验法（Augmented Dickey-Fuller Test）**的检验分析报告，并正确得出平稳性检验的结论。

很遗憾，书中只给了一个不太实用的分析报告——估计你们也还是看不懂。

下面给出 R 软件的增广迪基-富勒检验法 (Augmented Dickey-Fuller Test) 检验报告, 你们自己要学会看懂!! (这个检验法经常要用到, 所以还是咬咬牙搞定它吧!!)

```
Augmented Dickey-Fuller Test

data: fultonfish$lprice
Dickey-Fuller = -4.8, Lag order = 4, p-value =
0.01
alternative hypothesis: stationary
```

2. 时间序列的波动模式——初步的图形方法

如果时间序列变量是平稳的, 那么我们就可以很快对其进行平稳性序列建模并预测啦!

但是....., 平稳性序列建模有很多种形式啊——AR(p)模型, 或者 MA(q)模型, 或者 ARMA(p,q)模型, 或者 ARIMA(p,d,q)模型.....。你选哪一种呢? (抓狂)。

还好, 我们先给一个简单一点的图形判断规则吧:

直接看该时序变量的自相关图(PAC)和偏自相关图(PAC)吧。什么图形会匹配什么样的模型, 也是有一个经验法则的 (好好参看书本第 22 章, pg:790 页吧!!!!)。

至于自相关图(ACF)和偏自相关图(PAC), 大概长成这个样子喽!

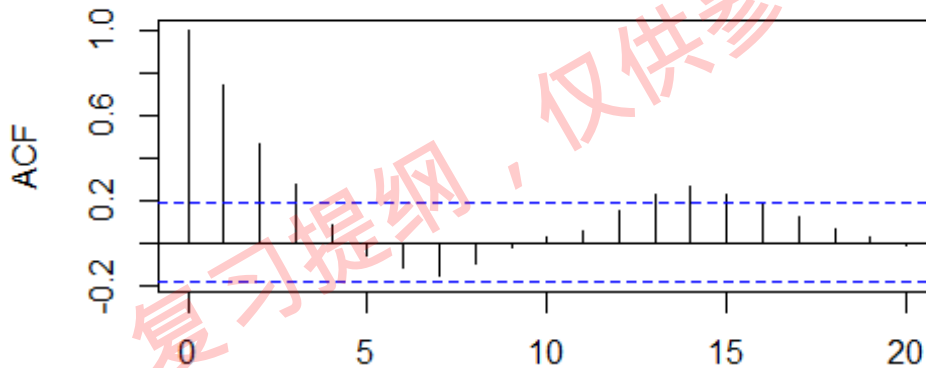


图 10 对数化价格变量 lprice 的自相关图

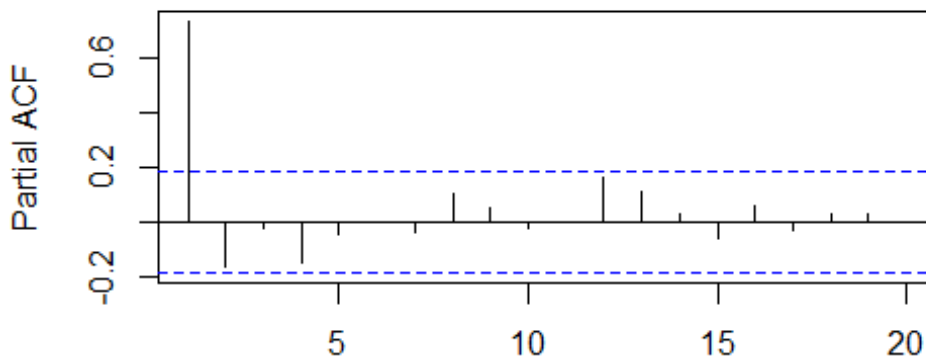


图 11 对数化价格变量 lprice 的偏自相关图

自己去对照法则看图说话吧。这个还是很简单, 也是很基本的“技能”!

- **后话 1:** 实际上自相关图(ACF)和偏自相关图(PAC)也可以初步判明一个序列是不是平稳。这一点, 我们往往可以用来查看回归残差序列。或者换句话说, 如果残差没有明显的特定模式, 那么就说明回归模型很理想喽!!
- **后话 2:** 当然, 以上通过图形方法来识别还是很难人的(只是个经验法则而已), 其实也可以用更复杂和更高深的定量方法来识别究竟适合做哪一类平稳模型——AR(p)模型, 或者 MA(q)模型, 或者 ARMA(p,q)模型, 或者 ARIMA(p,d,q)模型.....。我还是不说了吧, 感兴趣的请自行脑补!!

3. 能读懂 ARMA(p, q)模型的回归分析报告!!!

例如, 能根据回归分析报告, 正确写出样本回归函数形式。

下面也给出 R 软件的回归分析结果:

```
Series: fultonfish$price
ARIMA(2,0,2) with non-zero mean

Coefficients:
  ar1  ar2  ma1  ma2  mean
 1.473 -0.604 -0.572 -0.074 -0.191
s.e. 0.186 0.144 0.216 0.166 0.063

sigma^2 estimated as 0.0627: log likelihood=-1.7
AIC=15.41 AICc=16.22 BIC=31.66

Training set error measures:
      ME  RMSE  MAE  MPE  MAPE  MASE
Training set 0.002642 0.2446 0.1984 Inf Inf 0.9546
      ACF1
Training set -0.01111
```

ar? ma? mean? 分别是对应什么? 请自行复习!

第四部分：联立方程模型

这部分内容请自己看上课 ppt 和讲义。

考察重点主要是：

- 联立方程转换成约简形式
- 联立方程的识别规则（重点还是阶条件识别法则）
- 两阶段最小二乘法（2SLS）估计联立方程
- 联立方程回归结果的经济学解释

后话 1：大家还是需要利用经济学知识，多关注回归结果的经济学解释。例如，弹性的概念（又提到它了！）；利用供给方程结果和需求方程结果，在给定条件下，计算市场均衡价格和均衡数量！！——经济学的烂账还是要补回来的，欠的账迟早要还啊！

后话 2：识别的秩条件还是有点复杂和啰嗦，大家知道它跟阶条件的关系就行！