

《中级计量经济学》2018 年秋试题

第一部分：鳕鱼供需案例分析计算（共 3 大题，共 75 分）得分：
_____ 分

案例背景介绍：研究者收集了纽约市 **Fulton** 海鲜市场 1991-12-02 日到 1992-05-08 日期间共 111 天的相关日度数据。研究者主要关注鳕鱼产品的市场价格和数量，以及其他市场相关变量，如工作日（周一~周五）、气候条件、鳕鱼存货量变化情况。具体变量及含义见下表：

表 1: 鳕鱼案例的变量说明

变量	含义	备注
price	鳕鱼市场价格 (美元/千克)	连续变量
quan	鳕鱼的数量 (吨)	连续变量
lprice	(自然) 对数化的鳕鱼市场价格	连续变量
lquan	(自然) 对数化的鳕鱼数量	连续变量
mon	周一	虚拟变量:1= 周一; 0= 其他
tue	周二	虚拟变量:1= 周二; 0= 其他
wen	周三	虚拟变量:1= 周三; 0= 其他
thu	周四	虚拟变量:1= 周四; 0= 其他
stormy	暴风雨天气	虚拟变量:1= 暴风雨; 0= 其他
cold	寒冷天气	虚拟变量:1= 寒冷; 0= 其他
change	鳕鱼存货变化情况	虚拟变量:1= 变化大; 0= 变化小

请根据以上案例背景和数据，分析计算并回答下面一系列问题：

1. (共计 30 分，共 6 小题) A 同学重点考察鳕鱼价格的影响因素，并构建如下的多元回归价格模型：

$$lprice_t = \beta_1 + \beta_2 lquan_t + \beta_3 mon_t + \beta_4 tue_t + \beta_5 wed_t + \beta_6 thu_t + \beta_7 stormy_t + \beta_8 cold_t + \beta_9 change_t + u_t \quad (\text{价格模型})$$

A 同学利用统计软件对上述价格模型进行 OLS 回归分析，得到如下报告摘要：

```
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
```

请根据以上回归结果，回答如下问题：

(1) (5 分) 根据 A 同学的回归分析报告，请你写出相应的样本回归模型（要求写出理论表达式和数值表达式）。

(2) (2 分) 关于变量 *lquan* 的回归系数，它有什么样的经济学含义（请进行简单的经济学解释）？其符号和大小是否符合理论上的预期（请简要说明理由）？

(3) (5 分) A 同学的统计分析报告为：回归系数的 t 检验结果为显著（给定 $\alpha = 0.1$ ）的变量包括哪些？（请简要说明判断依据）？回归方程的整体显著性 F 检验结果如何？（请简要说明判断依据）？回归方程的判定系数 R^2 和调整判定系数 $\bar{R}^2$ 分别是多少？

(4) (5 分) 如果 A 同学给出一些已知条件 { 暴风雨；寒冷；鳕鱼存量变化小；鳕鱼的市场数量为 **54.5996** 吨 }，也即 { $stormy = 1; cold = 1; change = 0; quan = 54.5996$ }。请计算在给定条件下，Fulton 生鲜市场在星期五的鳕鱼期望价格是多少？（要求：请写出理论表达式，并计算出数值结果；对数计算请以自然数 $e = 2.7183$ 为底；计算结果保留 2 位小数）。

(5) (5 分) 为了诊断价格模型是否存在多重共线性问题，A 同学利用统计软件得到了下面表 2 所示的方差膨胀因子计算表。

表 2: 回归方程的方差膨胀因子 VIF

X	lquan	mon	tue	wed	thu	stormy	cold	change
VIF	1.305	1.592	1.723	1.697	1.594	1.302	1.212	1.106

根据表 2 的信息，请判断回归方程是否存在多重共线性问题？（请简要说明判断依据。）

(6) (小计 8 分，共 2 小问) 为了诊断价格模型是否存在异方差问题，A 同学利用统计软件对上述回归方程的残差 e_t 进行绘图，得到如下时序图（见图 1）：

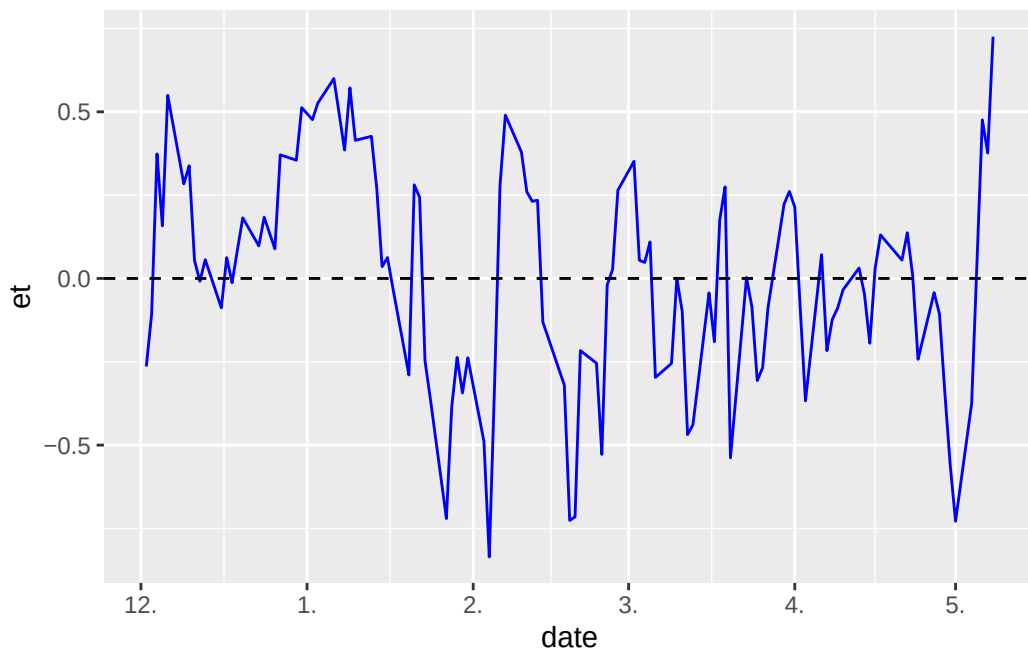


图 1: 残差 e_t 的时序图

同时，为量化地考察价格模型是否存在异方差问题。A 同学决定进行怀特异方差诊断（White Test），对残差 e_t 构建如下辅助回归模型：

$$\begin{aligned}
 e_t^2 = & +\alpha_1 + \alpha_2 lquan + \alpha_3 mon + \alpha_4 tue \\
 & +\alpha_5 wed + \alpha_6 thu + \alpha_7 stormy + \alpha_8 cold \\
 & +\alpha_9 change + \alpha_{10}(lquan)^2 + v_t
 \end{aligned}$$

并得到辅助回归模型的分析结果为：

$$\begin{aligned}
\hat{e}_t^2 = & -1.09 & +0.30lquan & -0.03mon & -0.05tue \\
(t) & (-0.8705) & (0.9923) & (-0.7829) & (-1.0942) \\
(se) & (1.2536) & (0.3054) & (0.0441) & (0.0438) \\
(cont.) & -0.07wed & -0.05thu & +0.02stormy & -0.00cold \\
(t) & (-1.4756) & (-1.0744) & (0.7264) & (-0.0555) \\
(se) & (0.0450) & (0.0421) & (0.0342) & (0.0298) \\
(cont.) & +0.00change & -0.02(lquan)^2 & & \\
(t) & (0.1213) & (-1.0006) & & \\
(se) & (0.0310) & (0.0185) & & \\
(fitness) & n = 111; & R^2 = 0.0429; & \bar{R}^2 = -0.0424 & \\
& F^* = 0.50; & p = 0.8691 & &
\end{aligned}$$

请根据以上分析结果，对价格模型是否存在异方差问题进行如下回答：

a) (3 分) 观测残差的时序图 (图 1)，请初步判断价格模型是否存在明显的异方差问题？(请简要说明理由)。

b) (5 分) 利用上述怀特异方差检验的辅助回归模型的回归结果，对异方差问题进行 χ^2 显著性检验。(给定 $\alpha = 0.05$ 水平下，理论查表值有 $\chi_{0.95}^2(9) = 16.92$ ； $\chi_{0.95}^2(10) = 18.31$ 。)(要求：写出原假设 H_0 和备择假设 H_1 ；计算出卡方统计量 $\chi^{*2} = n \cdot R^2 \sim \chi^2(f)$ ；确定正确的理论查表值 $\chi_{1-\alpha}^2(f)$ ；明确给出价格模型是否存在异方差的诊断结论。)

2. (共计 20 分，共 3 小题) B 同学重点关注鳕鱼对数化价格变量 $lprice_t$ 的平稳性和预测问题。

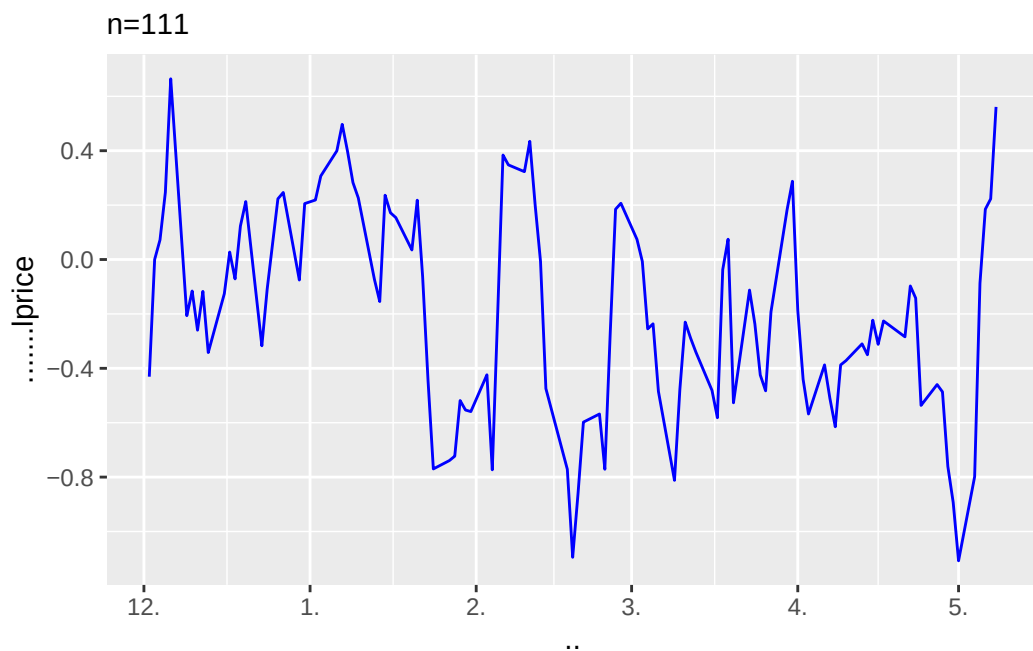


图 2: 鳕鱼对数化价格变量 lprice 的时序图

(1) (5 分) 为了分析对数化价格变量 lprice_t 的平稳性问题, B 同学采用增广迪基-富勒检验法 (**Augmented Dickey-Fuller Test, ADF**) 进行验证。B 同学利用统计软件分析, 得出如下表 (表 3) 所示的分析结果:

表 3: 对数化价格变量 lprice 的增广迪基-富勒检验 (ADF)

项目	细节
检验方法	Augmented Dickey-Fuller Test
备择假设 H_1	alternative hypothesis:stationary (平稳)
检验变量	lprice
迪基-富勒统计量	Dickey-Fuller = -4.76
滞后阶数	Lag order = 4
概率值	p-value = 0.01

请根据以上信息, 判断对数化价格变量 lprice_t 是不是平稳时间序列? (请简要给出理由。)

(2) (5 分) 为了识别对数化价格变量 $lprice_t$ 的波动模式, B 同学进一步绘制出了 $lprice_t$ 序列的自相关图 (ACF, 见图 3) 和偏自相关图 (PACF, 见图 4)

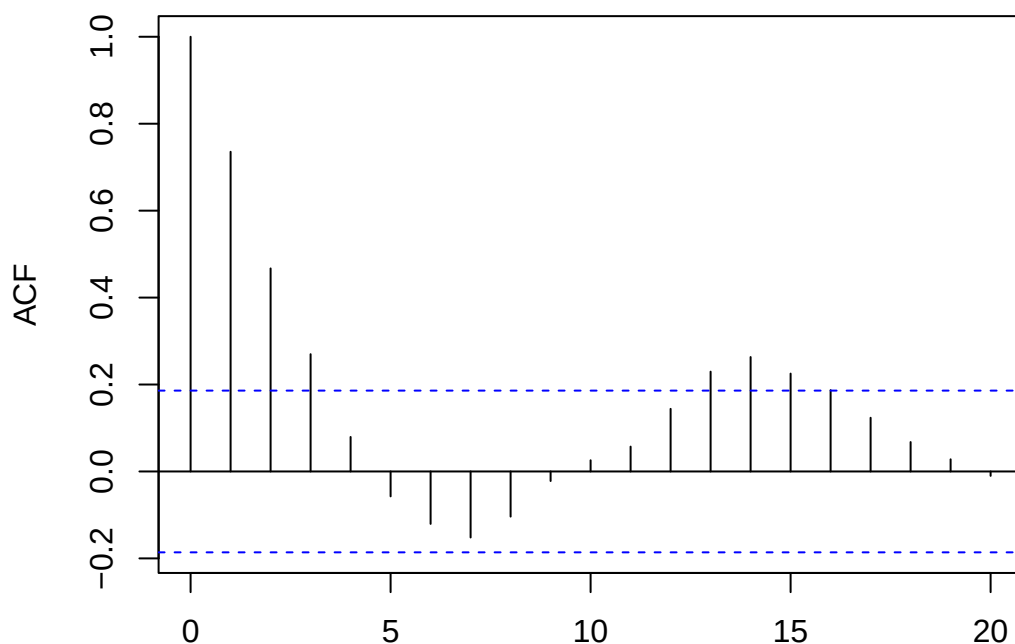


图 3: 对数化价格变量 $lprice$ 的自相关图

请根据以上图形信息, 初步判断对数化价格变量 $lprice_t$ 的时间序列过程是属于哪一种模式? (要求: 指明是 AR 过程/MA 过程/ARMA 过程中的哪一种, 不需要指明阶数, 并给出简要理由。)

(3) (小计 10 分, 共 2 小问) 为进一步对变量 $lprice_t$ 进行准确 $ARMA(p, q)$ 建模, B 同学尝试进行 $ARMA(2, 0)$ 建模分析。B 同学使用统计软件, 得到如下分析结果:

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

Coefficients:
      ar1      ar2      mean
      0.75    0.45    0.00
```

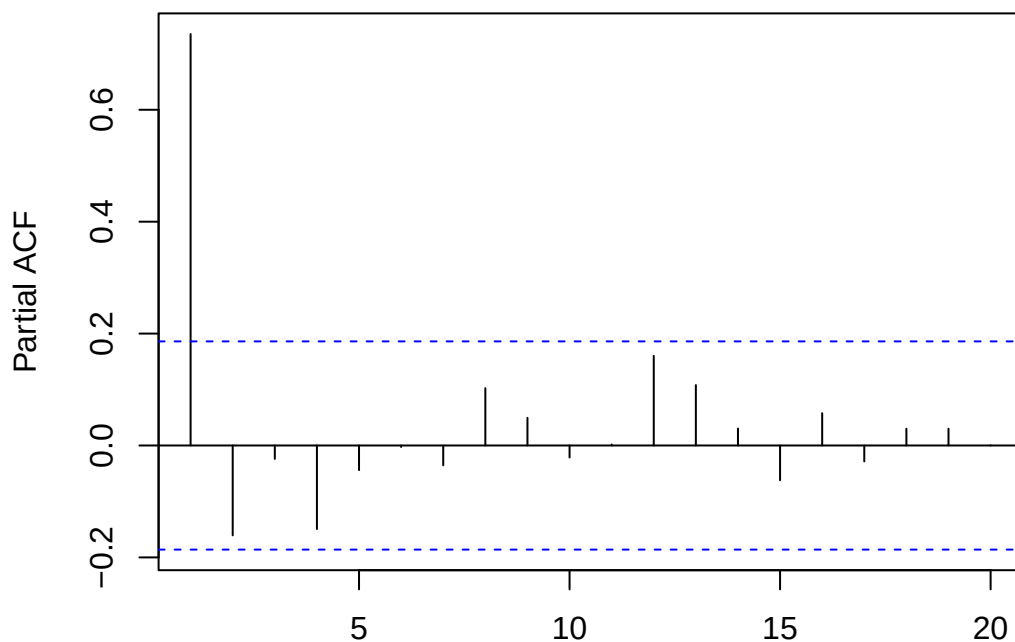


图 4: 对数化价格变量 lprice 的偏自相关图

```

      0.908   -0.198   -0.186
s.e.  0.095    0.095    0.079

```

```

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

```

同时, B 同学进一步对以上 $ARMA(2,0)$ 建模结果的残差绘图, 得到回归残差的自相关图 (ACF, 见图 5) 和偏自相关图 (PACF, 见图 6)

请回答:

a) (6 分) 根据分析报告摘要信息, 写出 $ARMA(2,0)$ 模型的样本回归方程形式。

b) (4 分) 根据残差绘图结果 (图 5 和图 6), 判断 B 同学的 $ARMA(2,0)$ 建模是否合理, 并给出简要理由。

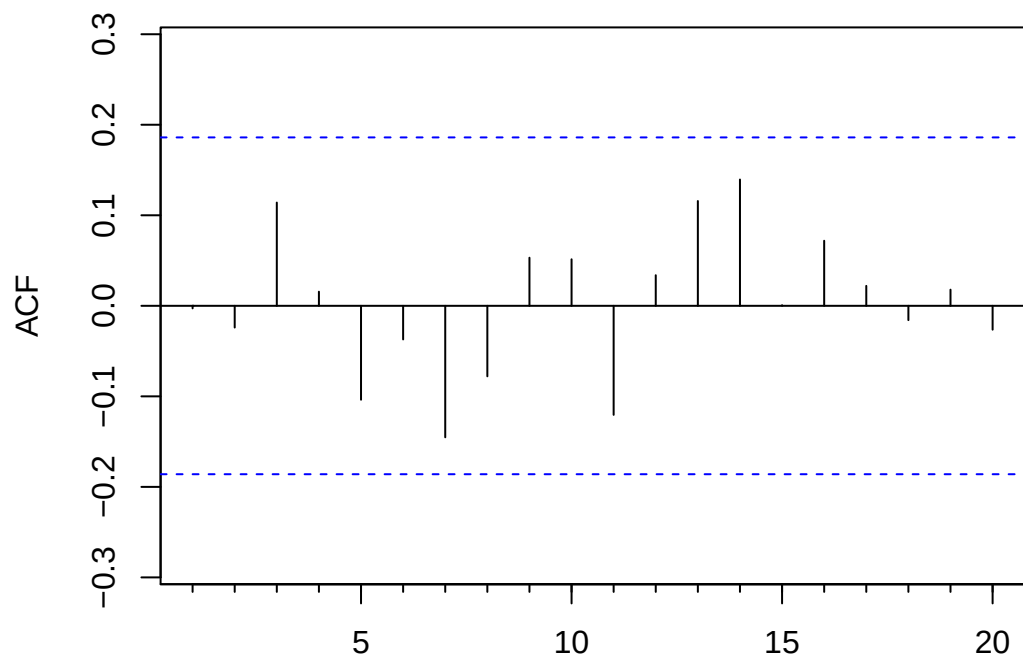


图 5: 残差的自相关图

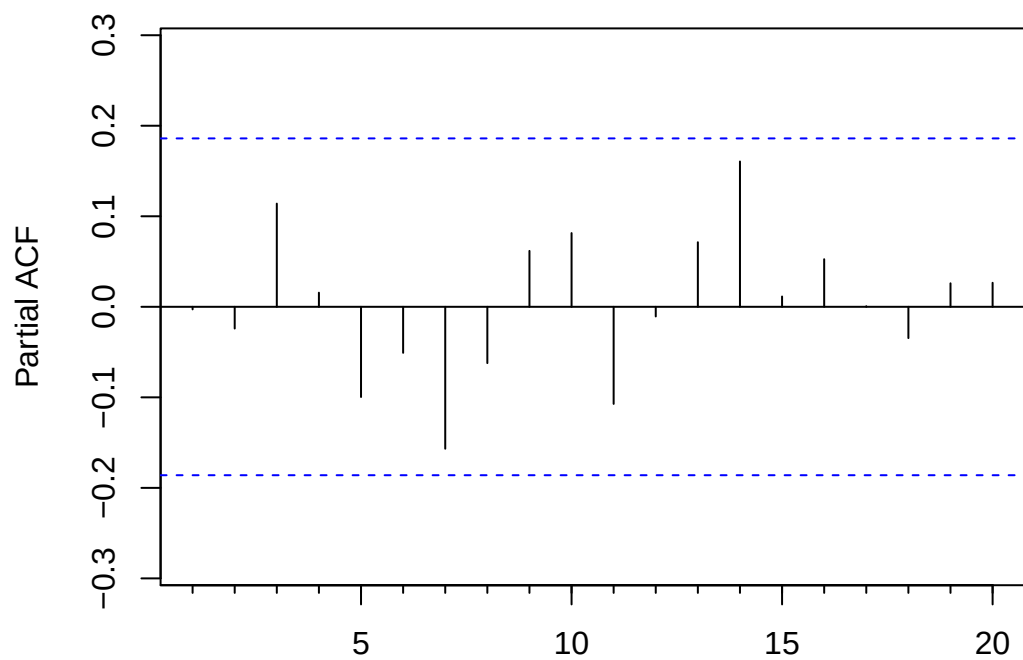


图 6: 残差的偏自相关图

3. (共计 25 分, 共 5 小题) C 同学重点关注鳕鱼供需关系的联立方程建模分析。

对此, C 同学构建了如下的联立方程:

$$lquan_t = \alpha_0 + \alpha_1 lprice_t + \alpha_2 mon_t + \alpha_3 tue_t + \alpha_4 wen_t + \alpha_5 thu_t + u_{1,t} \quad (\text{需求方程})$$

$$lquan_t = \beta_0 + \beta_1 lprice_t + \beta_3 stormy_t + u_{2,t} \quad (\text{供给方程})$$

(1) (3 分) 对于以上联立方程模型, 需求方程中变量 $lprice_t$ 的参数 α_1 , 以及供给方程中变量 $lprice_t$ 的参数 β_1 , 理论上这两个参数的大小和符号应该分别有什么要求?

(2) (5 分) 请根据阶条件规则, 对上述联立方程模型中需求方程和供给方程分别作出识别性问题的判断, 并请说明判断依据。

(3) (5 分) 请将以上需求方程和供给方程分别表达成约简方程形式。(要求: 不需要写出结构系数与约简系数的关系式。数量方程 ($lquan$) 约简系数记为 $\pi_{1,p}$ ($p = 1, 2, \dots$), 随机干扰项记为 $v_{1,t}$; 价格方程 ($lprice$) 约简系数记为 $\pi_{2,q}$ ($q = 1, 2, \dots$), 随机干扰项记为 $v_{2,t}$ 。)

(4) (小计 6 分, 共 2 小问) C 同学直接使用普通最小二乘法 (OLS) 对前述第 (3) 问中的约简方程分别进行回归, 并得到如下分析结果。

对数化鳕鱼数量 ($lquan$) 的约简方程 OLS 回归结果为：

$$\begin{aligned} \widehat{lquan} &= +8.81 && +0.10mon_i - 0.48tue_i \\ (s) && (0.1470) & (0.2065) & (0.2011) \\ (t) && (+59.92) & (+0.49) & (-2.41) \\ (cont.) && -0.55wed_i + 0.05thu_i - 0.39stormy_i \\ (s) && (0.2058) & (0.2010) & (0.1437) \\ (t) && (-2.69) & (+0.27) & (-2.70) \end{aligned}$$

对数化鳕鱼价格 ($lprice$) 的约简方程 OLS 回归结果为：

$$\begin{aligned} \widehat{lprice} &= -0.27 && -0.11mon_i - 0.04tue_i \\ (s) && (0.0764) & (0.1073) & (0.1045) \\ (t) && (-3.56) & (-1.05) & (-0.39) \\ (cont.) && -0.01wed_i + 0.05thu_i + 0.35stormy_i \\ (s) && (0.1069) & (0.1045) & (0.0747) \\ (t) && (-0.11) & (+0.48) & (+4.64) \end{aligned}$$

a) (4 分) 如果已知条件为 { 周一; 暴风雨 }, 也即 { $mon = 1$ $stormy = 1$ }。在给定条件下, 根据上述约简方程结果, 请计算 Fulton 生鲜市场的鳕鱼价格 $price$ 和鳕鱼数量 $quan$ 分别是多少? (要求: 对数计算请以自然数 $e = 2.7183$ 为底; 计算结果保留 2 位小数)。

b) (2 分) C 同学坚持认为: 以上步骤计算得到的鳕鱼价格 $price$ 和鳕鱼数量 $quan$ 就是 Fulton 生鲜市场在周一且天气为暴风雨时的市场均衡价格和市场均衡数量。你认为 C 同学的观点是否正确? 请简要说明你的理由。

(5) (小计 6 分, 共 2 小问) C 同学随后又采用了两阶段最小二乘法 (2SLS) 对联立方程模型进行了回归分析, 得到如下表所示的分析结果 (见表 4):

表 4: 两阶段最小二乘法 (2SLS) 回归结果

eq	vars	Estimate	Std. Error	t value	Pr(> t)
eq1	(Intercept)	8.5059	0.1662	51.1890	0.0000
eq1	lprice	-1.1194	0.4286	-2.6115	0.0103
eq1	mon	-0.0254	0.2148	-0.1183	0.9061
eq1	tue	-0.5308	0.2080	-2.5518	0.0122
eq1	wed	-0.5664	0.2128	-2.6620	0.0090
eq1	thu	0.1093	0.2088	0.5233	0.6018
eq2	(Intercept)	8.6284	0.3890	22.1826	0.0000
eq2	lprice	0.0011	1.3095	0.0008	0.9994
eq2	stormy	-0.3632	0.4649	-0.7813	0.4363

a) (2 分) 基于以上分析报告, 请你计算出鳕鱼的需求价格弹性 η_1 和供给价格弹性 η_2 分别是多少?

b) (4 分) 基于以上分析报告, 请你帮 C 同学计算出 Fulton 生鲜市场在周一且天气为暴风雨时 (也即 { **mon=1; stormy=1** }) 鳕鱼的市场均衡价格和市场均衡数量。(计算结果保留 2 位小数。)

第二部分: 乳腺癌案例分析计算 (共 5 小题, 共 25 分) 得分:

_____ 分

案例背景介绍: 一项关于女性乳腺癌患病情况的控制性研究, 收集了共 178 名女性的相关资料。研究者主要关注二分类变量是否患乳腺癌 (*Cancer*), 以及与其他相关变量的关系, 如年龄、受教育年数、是否接受过正规检查、生育第一胎时的年龄、流产次数、体重等。具体变量及含义见下表所示 (表 5)

表 5: 女性乳腺癌案例的变量说明

变量	含义	备注
Cancer	是否患乳腺癌	虚拟变量:1= 患有乳腺癌; 0= 未患乳腺癌
Age	年龄 (岁)	数值变量
HIGD	受教育年数 (年)	数值变量
CHK	是否接受过正规检查	虚拟变量:1= 接受过正规检查; 0= 未接受过正规检查
AGPI	生育第一胎时的年龄	数值变量
Miscar	流产次数 (次)	数值变量
Births	生育子女数 (个)	数值变量
Weight	体重 (磅)	数值变量

对变量是否患有乳腺癌(*Cancer*), 构造如下的总体回归模型 (见式 1):

$$\begin{aligned}
 \widehat{Cancer} = & + 0.20 & - 0.00Age_i & - 0.01HIGD_i \\
 (s) & (0.2975) & (0.0032) & (0.0132) \\
 (t) & (+0.68) & (-0.22) & (-0.93) \\
 (cont.) & - 0.20CHK_i & + 0.02AGPI_i & + 0.05Miscar_i \\
 (s) & (0.0618) & (0.0082) & (0.0306) \\
 (t) & (-3.20) & (+2.91) & (+1.54) \\
 (cont.) & + 0.04Births_i & - 0.00Weight_i & \\
 (s) & (0.0197) & (0.0010) & \\
 (t) & (+2.20) & (-3.00) &
 \end{aligned} \tag{1}$$

研究者对上述模型 (式 1) 进行 logit 形式的极大似然估计, 统计软件的分析结果如下表所示 (见表 6):

表 6: Logit 回归估计结果

term	estimate	std.error	statistic	p.value
(Intercept)	-0.5273	2.3190	-0.2274	0.8201
Age	0.0058	0.0219	0.2660	0.7902
HIGD	-0.0615	0.0860	-0.7144	0.4750
CHK	-1.5960	0.5286	-3.0191	0.0025
AGPI	0.1339	0.0562	2.3844	0.0171
Miscar	0.3415	0.2395	1.4257	0.1540
Births	0.2923	0.1302	2.2455	0.0247
Weight	-0.0294	0.0100	-2.9310	0.0034

基于以上的 logit 回归分析，研究者统计了乳腺癌患病预测情况，具体见下面表 7：

表 7: logit 回归分析预测情况统计

真实值	预测值	频次
0	0	130
1	0	28
0	1	8
1	1	12

请回答下面的一系列问题：

(1) (5 分) 给定 Y_i 为二分类变量， X_i 为连续变量，如果分别构建 probit 模型和 logit 模型，并令概率 $P(Y_i = 1|X) = \Phi(\beta_0 + \beta_1 X_i)$ ，且令 $Z_i = \beta_0 + \beta_1 X_i$ 。研究人员分别绘制下图所示的累积概率曲线（见图 7）。

```
Warning in geom_text(aes(x = 1, y = 0.88, label = "a")): All aesthetics have 1 rows.
```

```
i Please consider using `annotate()` or provide this layer with data containing a single row.
```

```
Warning in geom_text(aes(x = 1, y = 0.69, label = "b")): All aesthetics have 1 rows.
```

```
i Please consider using `annotate()` or provide this layer with data containing a single row.
```

请指出图 7 中，probit 累积概率曲线和 logit 累积概率密度曲线分别是哪一条？并请写出 logit 累积概率函数表达式。

(2) (5 分) 根据 Logit 回归估计结果（见表 7），请正确写出 logit 样本回归函数。（要求：logit 变量用 L_i 表达；回归系数保留 4 位小数）。

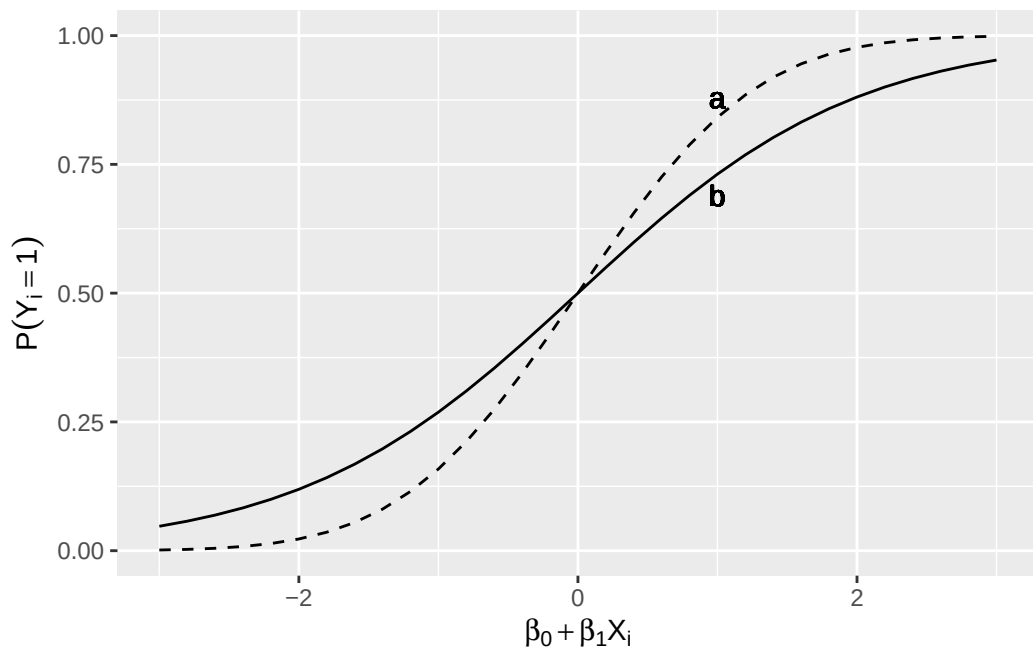


图 7: probit 模型和 logit 模型下的累积概率曲线

(3) (5 分) 根据 logit 回归的预测情况统计 (见表 7), 请计算拟合优度指标——计数 R^2 值。(要求: 写出理论计算公式; 计算结果保留 2 位小数)

(4) (5 分) 如果已经计算得到似然比统计量 $LR^* = \chi^{2*} = 25.97$, 请对 logit 回归进行模型整体显著性卡方检验。(给定 $\alpha = 0.05$ 水平下, 卡方分布的理论查表值有 $\chi^2_{0.95}(7) = 14.07$; $\chi^2_{0.95}(8) = 15.51$ 。)(要求: 写出原假设 H_0 和备择假设 H_1 ; 确定正确的理论查表值 $\chi^2_{1-\alpha}(f)$; 明确给出模型整体显著性检验结论。)

(5) (5 分) 根据 logit 回归结果, 请计算一个接受过正规检查 (也即 $CHK = 1$) 与一个未接受过正规检查 (也即 $CHK = 0$) 的女性, 两种情形下罹患乳腺癌的机会比率值是多少 (假定其他条件不变)?