

应用回归分析课程

论文（报告、案例分析）

院 系 信息学院

专 业 应用统计学

班 级 142132001 班

学生姓名 王晓璠

学 号 1421320022

任课教师 张方风

2016 年 12 月 26 日

基于非线性回归模型对价格与需求的统计研究

14 级应用统计学 1 班 1421320022 王晓璠

摘要:

本文搜集了 10 组关于价格与需求量的样本数据, 结合经济学知识, 即需求曲线和需求价格弹性的概念来研究价格与需求量之间的关系, 以价格作为 x , 需求量作为 y , 绘制价格与需求量的散点图, 观察趋势为非线性回归问题, 利用函数曲线拟合价格与需求量的曲线方程, 再将非线性回归问题转化为线性回归问题进行解决。

本次回归分析, 建立不同模型研究因变量(需求量)和自变量(价格)之间的关系, 用 AIC 最小信息量准则进行最优配置指标的比较, 从而选出反映价格与需求量之间关系的最优模型。这种统计分析方法可用于研究经济问题, 观察需求曲线从左上方向右下方倾斜的经济变动趋势, 以及回归方程结果的曲线特征, 可进一步论证经济规律, 即价格上升, 需求量减少; 价格下降, 需求量增加。另外, 还可以帮助政府更好地调控市场价格, 制定更符合市场规律的经济政策, 从而推动经济发展。

关键字: 非线性回归; 价格; 需求量; AIC 准则;

一、研究背景与研究意义

为了建立价格与需求量的函数关系式, 先介绍一下需求曲线和需求价格弹性的概念。

假设知道某种商品在某一时期内价格和需求量的情况, 把这些一一对应的数字在直角坐标图上绘出若干点, 把这些点连接起来, 就可以得到一条曲线。这条曲线就叫做需求曲线。由于价值规律的作用, 价格与需求量总是成反比关系。因此, 需求曲线总是有向下倾斜的趋势。

而需求价格弹性反映了不同的产品, 或者同一产品在不同的条件下, 由于价格变化对市场需求量的影响程度不同, 需求曲线向下倾斜的角度也各不相同。由此可见, 价格在影响需求的因素中起着决定的作用, 受需求函数的制约, 价格的改变必然引起需求量的改变, 而需求量的改变影响着企业收益的改变。企业经常想通过调节价格来增加收益。为了有的放矢的减少风险, 就要充分考虑该商品在市场的价格弹性。

因此，分析价格和需求量之间的关系有着重要的理论意义和经济意义，首先要建立价格与需求量的函数关系式进行拟合。本文利用最小二乘法进行对价格与需求量回归方程的参数估计。通过 *AIC* 最小信息量准则进行最优配置指标的比较，从而选出反映价格与需求量之间关系的最优模型。

二、研究方法

（一）曲线估计

在变量间相关关系的分析中，变量之间的关系并不总是表现出线性关系，非线性关系也是极为常见的。对于非线性关系通常无法直接通过线性回归来分析，无法直接建立线性模型。可通过绘制散点图的方式大致观察非线性关系。变量之间的非线性可以划分为本质线性关系和非本质线性关系。本质线性关系是指变量关系形式上虽然呈非线性关系（如二次函数曲线），但可以通过变量变换化为线性关系，并可以最终进行线性回归分析建立线性模型。而非本质线性关系是指变量关系不仅形式上呈非线性关系，而且也无法通过变量变换转化为线性关系，最终无法进行线性回归分析建立线性模型。本文的曲线估计解决本质线性关系问题的，建立的模型如表 2-1 所示。

表 2-1 本质线性模型

模型名称	回归方程	变量变换后的线性方程
二次曲线	$y = \beta_0 + \beta_1 x + \beta_2 x^2$	$y = \beta_0 + \beta_1 x + \beta_2 x_1$ (其中 $x_1 = x^2$)
复合曲线	$y = \beta_0 \beta_1^x$	$\ln(y) = \ln(\beta_0) + \ln(\beta_1)x$
增长曲线	$y = e^{\beta_0 + \beta_1 x}$	$\ln(y) = \beta_0 + \beta_1 x$
对数曲线	$y = \beta_0 + \beta_1 \ln(x)$	$y = \beta_0 + \beta_1 x_1$ (其中 $x_1 = \ln(x)$)
三次曲线	$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$	$y = \beta_0 + \beta_1 x + \beta_2 x_1 + \beta_3 x_2$ (其中 $x_1 = x^2, x_2 = x^3$)
S 曲线	$y = e^{\beta_0 + \beta_1 / x}$	$\ln(y) = \beta_0 + \beta_1 x_1$ (其中 $x_1 = \frac{1}{x}$)

指数曲线	$y = \beta_0 e^{\beta_1 x}$	$\ln(y) = \ln(\beta_0) + \beta_1 x$
逆函数	$y = \beta_0 + \beta_1 / x$	$y = \beta_0 + \beta_1 x_1$ (其中 $x_1 = \frac{1}{x}$)
幂函数	$y = \beta_0 (x^{\beta_1})$	$\ln(y) = \ln(\beta_0) + \beta_1 x_1$ (其中 $x_1 = \ln(x)$)
逻辑函数	$y = \frac{1}{1/\mu + \beta_0 \beta_1 x}$	$\ln(\frac{1}{y} - \frac{1}{\mu}) = \ln(\beta_0 + \ln(\beta_1)x)$

在曲线估计中，首先，在不能明确究竟哪种模型更接近样本数据时可在上述多种可选择的模型中选择集中模型；然后，可用 *Eviews* 完成模型的参数估计，并输出回归方程显著性检验的 F 值和概率 P 值、判定系数 R^2 等统计量；最后，以判定系数为主要依据选择其中的最优模型，并进行 *MAPE* 预测精度分析，从而使模型用于经济问题的预测研究中。

(二) 回归方程的统计检验

通过样本数据建立回归方程后一般不能立即用于对实际问题的分析和预测，通常要进行各种统计检验，主要包括回归方程的拟合优度检验、回归方程的显著性检验、回归系数的显著性检验、残差分析和模型修正等。

1. 回归方程的拟合优度检验

回归方程的拟合优度检验是检验样本数据点聚集在回归线周围的密集程度，从而评价回归方程对样本数据的代表程度。

拟合优度检验从对被解释变量 y 取值变化的成因分析入手，分析 y 的各个观察值之间的差异（与其均值的差异）主要有两方面原因造成：一是解释变量 x 取值的不同；二是其它随机因素的影响。等价方程表示为：

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

即为：总离差平方和(SST) = 回归平方和(SSA) + 剩余平方和(SSE)。说明回归方程能够解释的变差所占比例越大，那么回归方程的拟合优度会越高。

2. 回归方程的显著性检验

线性回归方程能够较好地反映被解释变量和解释变量之间统计关系的前提是：被解释变量和解释变量之间确实存在显著的线性关系。回归方程的显著性检验正是要检验被解释变量与所有解释变量之间的线性关系是否显著，用线性模型来描述它们之间的关系是否恰当。

3. 回归系数的显著性检验

回归系数的显著性检验主要目的是检验回归方程中的每个解释变量和被解释变量之间是否存在显著地线性关系，也就是研究解释变量是否能有效地解释被解释变量的线性变化，它们能否保留在线性回归方程中。

回归系数显著性检验是围绕回归系数估计值的抽样分布展开的，由此造成从某种理论分布的检验统计量，并进行检验。

4. 残差分析

所谓残差是指由回归方程计算所得的预测值和实际样本之间的差距，定义为：

$$e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_p x_{pi})$$

它是回归模型中 ε_i 的估计值，由多个 e_i 形成的序列成为残差序列。残差分析是回归方程检验中的重要组成部分，其出发点是，如果回归方程能够较好地反映被解释变量的特征和变化规律，那么残差序列中应不包含明显的规律性和趋势性。

(1) 异方差检验

残差的异方差性是相对于同方差而言的。所谓同方差，是为了保证回归参数统计量具有良好的统计性质，经典线性回归模型的一个重要假定：总体回归函数中的随机误差项满足同方差性，即它们都有相同的方差。如果不满足这一假定，那么线性回归模型存在异方差性。

若线性回归模型存在异方差性，则用传统的最小二乘法估计模型，得到的参数估计量不是有效估计量，此时无法对模型参数进行有关显著性检验。对存在异方差性的模型可以采用加权最小二乘法进行估计，从而修正模型。

(2) 自相关检验

如果随机误差项的各个期望值之间存在着相关关系，则称随机误差项之间存在自相关性。本文通过 LM 检验来作回归方程的自相关检验。若方程存在自相关，可用广义差分方程进行模型修正。

（三）AIC 准则---选出相对最优模型

AIC 准则是由日本统计学家赤池弘次（*Akaike*）于 1973 年提出，*AIC* 全称是最小信息量准则（*Akaike info criterion*）。*AIC* 准则是一种考评综合最优配置的指标，它是拟合精度和参数未知个数的加权函数，计算公式为：

$$AIC = n \ln \left(\frac{SSE}{n} \right) + 2p$$

使 *AIC* 函数达到最小值的模型被认为是最优模型。

三、基于非线性回归模型对价格与需求量的统计研究

为了拟合价格与需求量的回归模型，本文搜集了 10 组数据，以价格作为 *x*，需求量作为 *y*，见下表 3-1。绘制需求量 *y* 关于价格 *x* 的散点图，如图 3-2 所示。

表 3-1 价格与需求量的原始数据

价格 <i>x</i> （元）	需求量 <i>y</i> （千克）
1	58
2	50
3	44
4	38
5	34
6	30
7	29
8	26
9	25
10	24

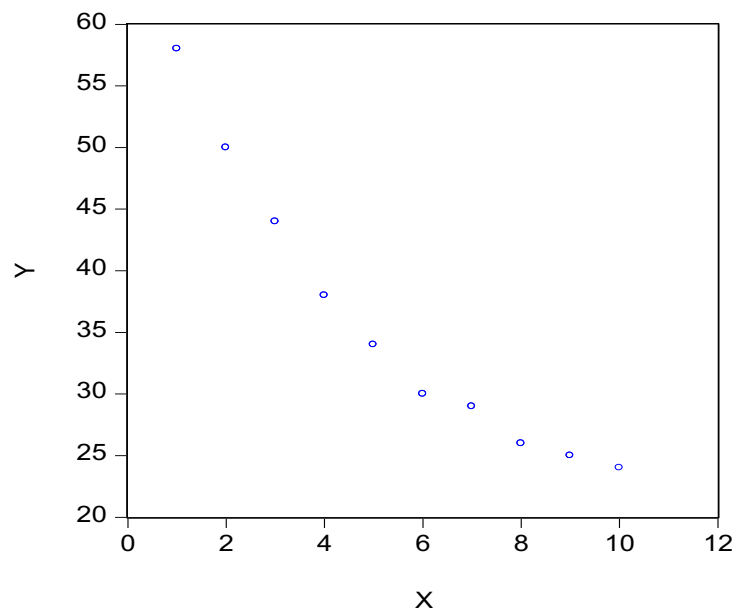


图3-2需求量 y 随价格 x 变动程度的散点图

从图3-2中可以看出，价格 x 与需求量 y 之间呈非线性关系，可尝试选择二次、三次曲线、对数函数模型、指数函数模型和幂函数模型，初步进行这五个模型的曲线估计，并作这五个模型的拟合回归线。确定这五个模型是否符合散点图的基本变化趋势。如表3-3-1、图3-3-2所示。

表3-3-1需求量 y 的曲线估计结果

模型汇总和参数估计值

因变量:: y 自变量为: x

方程	模型汇总					参数估计值			
	R 方	F	df1	df2	Sig.	常数	b1	b2	b3
对数	.991	919.216	1	8	.000	59.506	-15.695		
二次	.997	1303.940	2	7	.000	65.450	-8.440	.436	
三次	.999	1589.253	3	6	.000	67.467	-10.229	.823	-.024
幂	.970	257.993	1	8	.000	63.736	-.411		
指数	.964	216.912	1	8	.000	59.130	-.099		

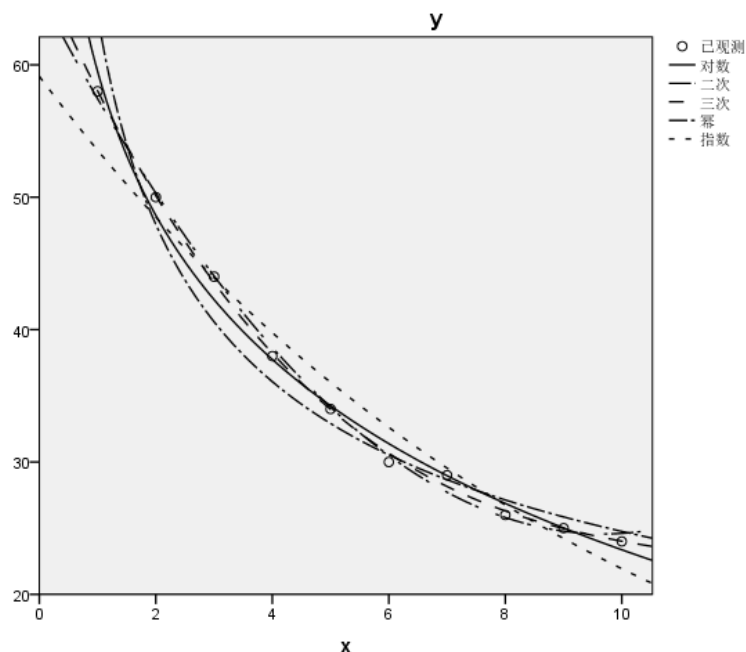


图3-3-2需求量 y 的各模型拟合回归线

由表3-3-1可知，对数函数模型、二次、三次曲线、幂函数模型和指数函数模型的可决系数 R^2 分别为0.991、0.997、0.999、0.970、0.964，五个模型的拟合优度都比较高，而且图3-3-2所示的五个模型的拟合趋势线也基本符合散点图的分布轨迹。因此，本文分别对五个回归模型进行拟合以及模型的检验、修正。

（一） 对数函数模型的建立、检验、修正

1.模型检验

由已知对数曲线方程： $y = \beta_0 + \beta_1 \ln(x)$ ，将对数曲线方程转化为线性方程为： $y = \beta_0 + \beta_1 x_1$ （其中 $x_1 = \ln(x)$ ），将 x_1 作解释变量、 y 作被解释变量，建立 y 、 x_1 的回归方程。回归结果如表 3-4-1 所示。

表3-4-1 对数函数模型的回归结果

Dependent Variable: Y
Method: Least Squares
Date: 12/27/16 Time: 14:39
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	-15.69491	0.517666	-30.31858	0.0000
C	59.50624	0.860795	69.12944	0.0000
R-squared	0.991372	Mean dependent var		35.80000
Adjusted R-squared	0.990294	S.D. dependent var		11.55470
S.E. of regression	1.138386	Akaike info criterion		3.273956
Sum squared resid	10.36737	Schwarz criterion		3.334473
Log likelihood	-14.36978	Hannan-Quinn criter.		3.207569
F-statistic	919.2164	Durbin-Watson stat		1.557639
Prob(F-statistic)	0.000000			

2.模型检验

(1) 对数模型的拟合优度检验:

从表3-4-1中可以看出模型的拟合优度情况,模型的可决系数 R^2 为0.991,调整后 R^2 为0.990。可见 R^2 和调整后 R^2 都在98%以上。说明该模型回归的贡献很大,表示对数函数回归模型拟合的结果十分理想。

(2) 对回归方程进行假设检验--- F 检验:

从表3-4-1中可以看出模型的假设检验结果,表中 F 值为919.216,概率值为0.000,远小于0.05,所以 F 检验通过。说明在对数函数回归模型中 y 与 x_1 (其中 $x_1 = \ln(x)$) 的线性回归高度显著,这与相关系数的检验结果是一致的。

(3) 对回归系数进行假设检验--- t 检验:

从表3-4-1中可知模型的回归系数 β_1 为-15.695,常数项 β_0 为59.506,因而可以得到回归方程为 $\hat{y} = 59.506 - 15.695 \ln(x)$,回归系数 β_1 检验的 t 值为-30.319,概率值为0.0000;而常数项 β_0 的 t 值为69.129,概率值为0.0000。所以, t 检验通过,说明在对数函数回归模型中 y 与 x_1 (其中 $x_1 = \ln(x)$) 的线性假设成立,与 F 检验和相关系数的检验结果一致。

(4) 残差检验，如图3-4-2所示：

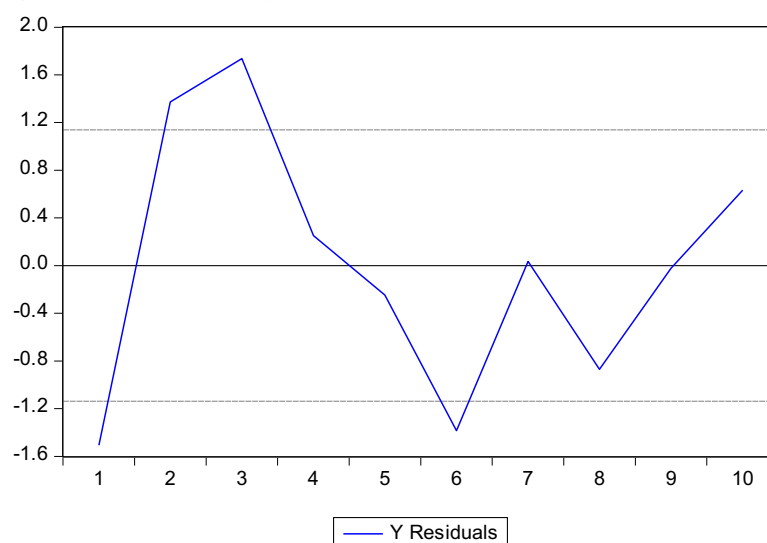


图3-4-2 指数函数模型的残差图

如图3-4-2所示，从指数函数模型的残差图中可以看出，指数函数模型的残差在0附近波动，且上下到0轴间距的绝对值不超过2，可见残差项通过检验。

(5) 自相关检验——*LM* 检验，输出结果如表3-4-3所示：

表3-4-3 *LM* 检验结果

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.226082	Prob. F(2,6)	0.8042
Obs*R-squared	0.700794	Prob. Chi-Square(2)	0.7044

从表3-4-3中可以看出，*LM* 值检验结果的概率值为0.7044，远大于显著性水平 $\alpha = 0.05$ ，因此接受 *LM* 检验原假设，可以认为残差序列不存在自相关。

(6) 异方差检验——*White* 检验，输出结果如表3-4-4所示：

表3-4-4 *White* 异方差检验结果

Heteroskedasticity Test: White

F-statistic	2.797531	Prob. F(2,7)	0.1280
Obs*R-squared	4.442266	Prob. Chi-Square(2)	0.1085
Scaled explained SS	1.515109	Prob. Chi-Square(2)	0.4688

从表3-4-4中可以看出，其中F-statistic是辅助方程整体显著性的F统计量；Obs*R-squared是 *White* 检验统计量 NR^2 ，其概率值为0.1085，明显大于显著性水平 $\alpha = 0.05$ ，因此接受 *White* 检验原假设，认为原方程的残差序列不存在异方差。

3. 小结

对选取的 10 组关于价格与需求量的数据建立了对数函数模型：

$$y = \beta_0 + \beta_1 \ln(x)$$

在应用过程中，由于将曲线方程转化为线性方程进行回归分析，因而需将回归方程 $\hat{y} = 59.506 - 15.695x_1$ （其中 $x_1 = \ln(x)$ ）转化为对数函数方程：

$$y = 59.506 - 15.695 \ln(x)$$

因此此模型通过了所有的假设检验，所以该模型可以反映价格与需求量的关系，并且可以对关于价格与需求量的经济问题作统计预测。

（二）二次多项式模型的建立、检验、修正

1. 模型建立

由已知二次曲线方程： $y = \beta_0 + \beta_1 x + \beta_2 x^2$ ，将二次曲线方程转化为线性方程为： $y = \beta_0 + \beta_1 x + \beta_2 x_1$ （其中 $x_1 = x^2$ ），将 x 、 x_1 作解释变量、 y 作被解释变量，建立 y 、 x 、 x_1 的回归方程。回归结果如表 3-5-1 所示。

表3-5-1 二次多项式模型的回归结果

Dependent Variable: Y				
Method: Least Squares				
Date: 12/27/16 Time: 14:43				
Sample: 1 10				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	-8.440152	0.332983	-25.34711	0.0000
X2	0.435606	0.029501	14.76580	0.0000
C	65.45000	0.797292	82.09037	0.0000
R-squared	0.997323	Mean dependent var		35.80000
Adjusted R-squared	0.996558	S.D. dependent var		11.55470
S.E. of regression	0.677882	Akaike info criterion		2.303638
Sum squared resid	3.216667	Schwarz criterion		2.394413
Log likelihood	-8.518188	Hannan-Quinn criter.		2.204057
F-statistic	1303.940	Durbin-Watson stat		2.064429
Prob(F-statistic)	0.000000			

2.模型检验

(1) 二次多项式模型的拟合优度检验：

从表3-5-1中可以看出模型的拟合优度情况，模型的可决系数 R^2 为0.997，调整后 R^2 为0.9966。可见 R^2 和调整后 R^2 都在98%以上。说明该模型回归的贡献很大，表示二次多项式回归模型拟合的结果十分理想。

(2) 对回归方程进行假设检验--- F 检验：

从表3-5-1中可以看出模型的假设检验结果，表中 F 值为1303.940，概率值为0.000，远小于0.05，所以 F 检验通过。说明在二次多项式回归模型中 y 与 x 、 x_1 （其中 $x_1 = x^2$ ）的线性回归高度显著，这与相关系数的检验结果是一致的。

(3) 对回归系数进行假设检验--- t 检验：

从表3-5-1中可以得到回归方程为 $\hat{y} = 65.450 - 8.44x + 0.436x^2$ ，回归系数 β_1 检验的 t 值为-25.347，概率值为0.0000； β_2 检验的 t 值为14.766，概率值为0.0000；而常数项 β_0 的 t 值为82.090，概率值为0.0000。所以， t 检验通过，说明在二次多项式模型中 y 与 x 、 x_1 的线性假设成立，与 F 检验和相关系数的检验结果一致。

(4) 残差检验，如图3-5-2所示：

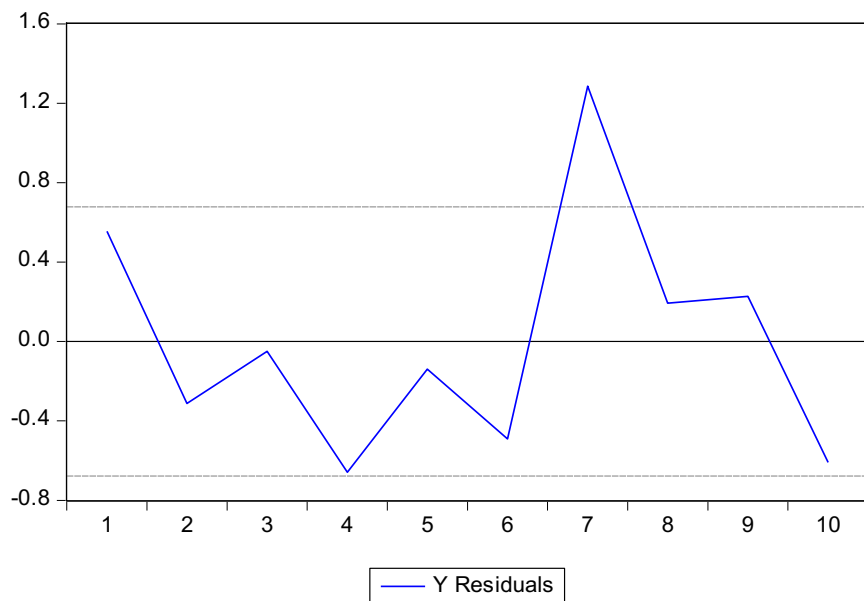


图3-5-2 二次多项式模型的残差图

如图3-5-2所示，从二次多项式模型的残差图中可以看出，二次多项式模型的残差在0附近波动，且上下到0轴的间距的绝对值不超过2，可见残差项通过检验。

(5) 自相关检验——*LM* 检验，输出结果如表3-5-3所示：

表 3-5-3 *LM* 检验结果

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.094257	Prob. F(2,5)	0.9116
Obs*R-squared	0.363331	Prob. Chi-Square(2)	0.8339

从表3-5-3中可以看出，*LM* 值检验结果的概率值为0.8339，远大于显著性水平 $\alpha = 0.05$ ，因此接受 *LM* 检验原假设，可以认为残差序列不存在自相关。

(6) 异方差检验——*White*检验，输出结果如表3-5-4所示：

表3-5-4 *White*异方差检验结果

Heteroskedasticity Test: White

F-statistic	0.299332	Prob. F(4,5)	0.8671
Obs*R-squared	1.932007	Prob. Chi-Square(4)	0.7483
Scaled explained SS	1.004760	Prob. Chi-Square(4)	0.9091

从表3-5-4中可以看出，其中F-statistic是辅助方程整体显著性的F统计量；Obs*R-squared是 *White* 检验统计量NR²，其概率值为0.7483，明显大于显著性水平 $\alpha = 0.05$ ，因此接受 *White* 检验原假设，认为原方程的残差序列不存在异方差。

3. 小结

对选取的 10 组关于价格与需求量的数据建立了二次多项式模型：

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

在应用过程中，由于将曲线方程转化为线性方程进行回归分析，因而需将回归方程 $\hat{y} = 65.450 - 8.44x + 0.436x_1$ （其中 $x_1 = x^2$ ）转化为二次多项式方程：

$$y = 65.450 - 8.44x + 0.436x^2$$

因为此模型通过了所有的假设检验，所以该模型可以反映价格与需求量的关系，并且可以对关于价格与需求量的经济问题作统计预测。

（三）三次多项式模型的建立、检验、修正

1. 模型建立

由已知三次曲线方程： $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ ，将三次曲线方程转化为线性方程为： $y = \beta_0 + \beta_1 x + \beta_2 x_1 + \beta_3 x_2$ （其中 $x_1 = x^2$ ， $x_2 = x^3$ ），将 x 、 x_1 、 x_2 作解释变量、 y 作被解释变量，建立 y 、 x 、 x_1 、 x_2 的回归方程。回归结果如表 3-5-1 所示。

表3-5-1 三次多项式模型的回归结果

Dependent Variable: Y

Method: Least Squares

Date: 12/27/16 Time: 15:52

Sample: 1 10

Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	-10.22883	0.729840	-14.01517	0.0000
X2	0.823427	0.150541	5.469776	0.0016
X3	-0.023504	0.009027	-2.603706	0.0405
C	67.46667	0.973707	69.28847	0.0000
R-squared	0.998743	Mean dependent var		35.80000
Adjusted R-squared	0.998115	S.D. dependent var		11.55470
S.E. of regression	0.501706	Akaike info criterion		1.747571
Sum squared resid	1.510256	Schwarz criterion		1.868605
Log likelihood	-4.737857	Hannan-Quinn criter.		1.614797
F-statistic	1589.253	Durbin-Watson stat		3.319320
Prob(F-statistic)	0.000000			

2.模型检验

(1) 三次多项式模型的拟合优度检验:

从表3-6-1中可以看出模型的拟合优度情况,模型的可决系数 R^2 为0.999,调整后 R^2 为0.998。可见 R^2 和调整后 R^2 都在98%以上。说明该模型回归的贡献很大,表示三次多项式模型拟合的结果十分理想。

(2) 对回归方程进行假设检验--- F 检验:

从表 3-6-1 中可以看出模型的假设检验结果,表中 F 值为 1589.253,概率值为 0.000,远小于 0.05,所以 F 检验通过。

说明在三次多项式模型 $y = \beta_0 + \beta_1 x + \beta_2 x_1 + \beta_3 x_2$ (其中 $x_1 = x^2$, $x_2 = x^3$) 中 y 与 x 、 x_1 、 x_2 的线性回归高度显著,这与相关系数的检验结果是一致的。

(3) 对回归系数进行假设检验--- t 检验:

从表3-6-1中可以得到回归方程为 $\hat{y} = 67.467 - 10.229x + 0.823x^2 - 0.024x^3$,回归系数 β_1 检验的 t 值为-14.015,概率值为0.0000; β_2 检验的 t 值为5.470,概率值为0.0000; β_3 检验的 t 值为-2.604,概率值为0.0000;而常数项 β_0 的 t 值为67.467,概率值为0.0000。所以, t 检验通过,说明在三次多项式模型中 y 与 x 、 x_1 、 x_2 的线性假设成立,与 F 检验和相关系数的检验结果一致。

(4) 残差检验，如图3-6-2所示：

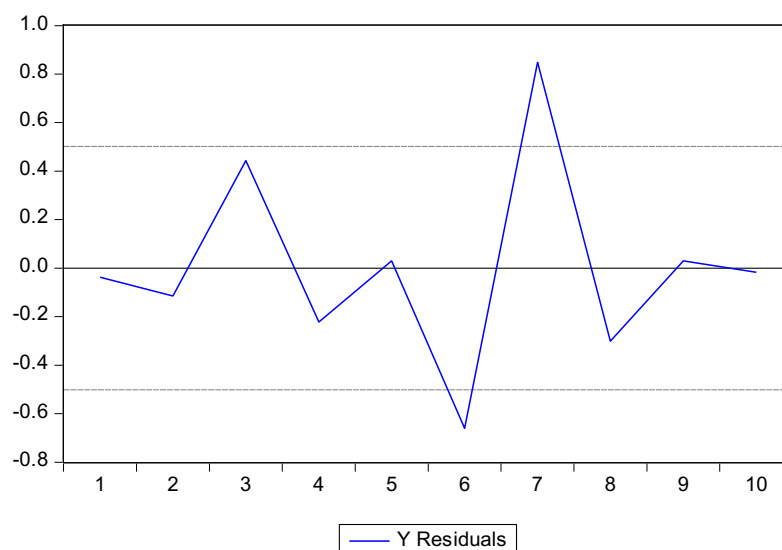


图3-6-2 三次多项式模型的残差图

如图3-6-2所示，从三次多项式模型的残差图中可以看出，三次多项式模型的残差在0附近波动，且上下到0轴的间距的绝对值不超过1，可见残差项通过检验。

(5) 自相关检验——*LM* 检验，输出结果如表3-6-3所示：

表3-6-3 *LM* 检验结果

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.842543	Prob. F(2,4)	0.2709
Obs*R-squared	4.795114	Prob. Chi-Square(2)	0.0909

从表3-6-3中可以看出，*LM* 值检验结果的概率值为0.0909，大于显著性水平 $\alpha = 0.05$ ，因此接受 *LM* 检验原假设，可以认为残差序列不存在自相关。

(6) 异方差检验——*White* 检验，输出结果如表3-6-4所示：

表3-6-4 *White* 异方差检验结果

Heteroskedasticity Test: White

F-statistic	1.137140	Prob. F(6,3)	0.4973
Obs*R-squared	6.945893	Prob. Chi-Square(6)	0.3259
Scaled explained SS	2.912721	Prob. Chi-Square(6)	0.8197

从表3-6-4中可以看出，其中F-statistic是辅助方程整体显著性的F统计量；Obs*R-squared是 *White* 检验统计量 NR^2 ，其概率值为0.3259，明显大于显著性水平 $\alpha = 0.05$ ，因此接受 *White* 检验原假设，认为原方程的残差序列不存在异方差。

3. 小结

对选取的 10 组关于价格与需求量的数据建立了三次多项式模型：

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

在应用过程中，由于将曲线方程转化为线性方程进行回归分析，因而需将回归方程 $\hat{y} = 67.467 - 10.229x + 0.823x_1 - 0.024x_2$ （其中 $x_1 = x^2$ ， $x_2 = x^3$ ）转化为三次多项式方程：

$$y = 67.467 - 10.229x + 0.823x^2 - 0.024x^3$$

因为此模型通过了所有的假设检验，所以该模型可以反映价格与需求量的关系，并且可以对关于价格与需求量的经济问题作统计预测。

（四）幂函数模型的建立、检验、修正

1. 模型建立

由已知幂函数表达式： $y = \beta_0(x^{\beta_1})$ ，将幂函数曲线方程转化为线性方程为： $\ln(y) = \ln(\beta_0) + \beta_1 x_1$ （其中 $x_1 = \ln(x)$ ），将 x_1 作解释变量、 $y_1(y_1 = \ln(y))$ 作被解释变量，建立 y_1 、 x_1 的回归方程。回归结果如表 3-7-1 所示。

表3-7-1 幂函数模型的回归结果

Dependent Variable: Y1				
Method: Least Squares				
Date: 12/27/16 Time: 16:12				
Sample: 1 10				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	-0.410697	0.025569	-16.06217	0.0000
C	4.154745	0.042517	97.71878	0.0000
R-squared	0.969924	Mean dependent var		3.534412
Adjusted R-squared	0.966165	S.D. dependent var		0.305682
S.E. of regression	0.056228	Akaike info criterion		-2.741931
Sum squared resid	0.025293	Schwarz criterion		-2.681414
Log likelihood	15.70965	Hannan-Quinn criter.		-2.808318
F-statistic	257.9934	Durbin-Watson stat		1.094874
Prob(F-statistic)	0.000000			

2.模型检验

(1) 幂函数模型的拟合优度检验:

从表3-7-1中可以看出模型的拟合优度情况,模型的可决系数 R^2 为0.9699,调整后 R^2 为0.966。可见 R^2 和调整后 R^2 都在95%以上。说明该模型回归的贡献很大,表示二次多项式回归模型拟合的结果比较理想。

(2) 对回归方程进行假设检验--- F 检验:

从表 3-7-1 中可以看出模型的假设检验结果,表中 F 值为 257.993,概率值为 0.000,远小于 0.05,所以 F 检验通过。

说明在幂函数回归模型 $\ln(y) = \ln(\beta_0) + \beta_1 x_1$ (其中 $x_1 = \ln(x)$) 中 $\ln(y)$ 与 x_1 的线性回归高度显著,这与相关系数的检验结果是一致的。

(3) 对回归系数进行假设检验--- t 检验:

从表3-7-1中可以得到回归方程为 $\ln(y) = 4.155 - 0.411\ln(x)$,回归系数 β_1 检验的 t 值为-16.062,概率值为0.0000;而常数项 β_0 的 t 值为97.719,概率值为0.0000。所以, t 检验通过,说明在幂函数回归模型中 $\ln(y)$ 与 x_1 (其中 $x_1 = \ln(x)$) 的线性假设成立,与 F 检验和相关系数的检验结果一致。

(4) 自相关检验--- LM 检验,输出结果如表3-7-2所示:

表3-7-2 LM 检验结果

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.284169	Prob. F(2,6)	0.7622
Obs*R-squared	0.865268	Prob. Chi-Square(2)	0.6488

从表3-7-2中可以看出, LM 值检验结果的概率值为0.6488,远大于显著性水平 $\alpha = 0.05$,因此接受 LM 检验原假设,可以认为残差序列不存在自相关。

(5) 异方差检验--- $White$ 检验,输出结果如表3-7-3所示:

表3-7-3 $White$ 异方差检验结果

Heteroskedasticity Test: White

F-statistic	7.434855	Prob. F(2,7)	0.0186
Obs*R-squared	6.799226	Prob. Chi-Square(2)	0.0334
Scaled explained SS	2.538658	Prob. Chi-Square(2)	0.2810

从表3-7-3中可以看出,其中F-statistic是辅助方程整体显著性的F统计量;Obs*R-squared是 $White$ 检验统计量 NR^2 ,其概率值为0.0334,小于显著性水平 $\alpha = 0.05$,因此拒绝 $White$ 检验原假设,认为原方程的残差序列存在异方差,需要对原模型进行相应的修正以解决异方差的影响。

3.模型修正---运用加权最小二乘法修正异方差

(1) 进行加权最小二乘法估计，输出结果如表3-7-4所示：

表3-7-4 加权最小二乘法下模型估计结果

Dependent Variable: Y1

Method: Least Squares

Date: 12/27/16 Time: 16:30

Sample (adjusted): 2 10

Included observations: 9 after adjustments

Weighting series: W

Weight type: Inverse standard deviation (EViews default scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.243802	0.023062	184.0197	0.0000
X1	-0.456013	0.017222	-26.47913	0.0000

Weighted Statistics				
R-squared	0.990115	Mean dependent var	3.575155	
Adjusted R-squared	0.988703	S.D. dependent var	2.007987	
S.E. of regression	0.030860	Akaike info criterion	-3.925578	
Sum squared resid	0.006666	Schwarz criterion	-3.881750	
Log likelihood	19.66510	Hannan-Quinn criter.	-4.020158	
F-statistic	701.1442	Durbin-Watson stat	1.833925	
Prob(F-statistic)	0.000000	Weighted mean dep.	3.686113	

Unweighted Statistics				
R-squared	0.989034	Mean dependent var	3.475964	
Adjusted R-squared	0.987467	S.D. dependent var	0.258245	
S.E. of regression	0.028911	Sum squared resid	0.005851	
Durbin-Watson stat	1.584344			

如表3-7-4所示为加权最小二乘法估计结果，包括了加权最小二乘法下参数的估计结果，以及加权与未加权方法下模型估计统计量的比较。

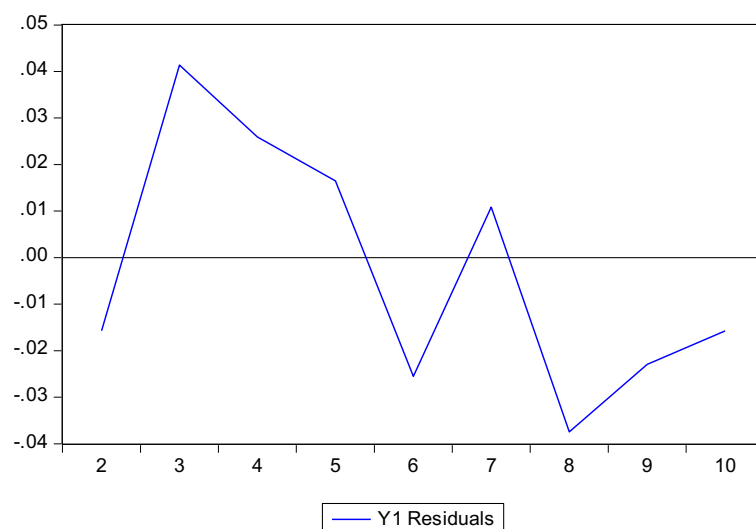


图3-7-5 修正后幂函数模型的残差图

如图3-7-5所示，从加权最小二乘法修正后的残差图中可以看出，修正后的幂函数模型的残差在0附近波动，且上下间距的绝对值不超过1，可见残差项通过检验。

(2) 对加权的模型进行异方差检验

表3-7-6加权模型的异方差检验结果

Heteroskedasticity Test: White

F-statistic	4.455881	Prob. F(2,6)	0.0651
Obs*R-squared	5.378698	Prob. Chi-Square(2)	0.0679
Scaled explained SS	2.285457	Prob. Chi-Square(2)	0.3189

如表3-7-6所示的加权模型的异方差检验结果。可以看到Obs*R-squared统计量，其概率值为0.0679，大于显著性水平 $\alpha = 0.05$ ，因此接受原假设即不存在异方差。此时可以将加权最小二乘法估计的参数值带入原模型，得到修正后回归模型为： $\ln(y) = 4.244 - 0.456\ln(x)$ 。

4. 小结

对选取的 10 组关于价格与需求量的数据建立了幂函数模型：

$$y = \beta_0(x^{\beta_1})$$

由于该模型进行拟合时存在异方差，因而采用了加权最小二乘法消除异方差，作了模型修正。修正后的模型通过检验。

在应用过程中，由于将曲线方程转化为线性方程进行回归分析，因而需将回归方程 $\ln(y) = 4.244 - 0.456\ln(x)$ 转化为幂函数方程：

$$y = 62.778x^{-0.456}$$

因为此模型通过了所有的假设检验，所以该模型可以反映价格与需求量的关系，并且可以对关于价格与需求量的经济问题作统计预测。

（五）指数函数模型的建立、检验、修正

1.模型建立

由已知指数函数表达式： $y = \beta_0 e^{\beta_1 x}$ ，将指数函数曲线方程转化为线性方程为： $\ln(y) = \ln(\beta_0) + \beta_1 x$ ，将 x 作解释变量、 $y_1(y_1 = \ln(y))$ 作被解释变量，建立 y_1 、 x 的回归方程。回归结果如表 3-8-1 所示。

表3-8-1 指数函数模型的回归结果

Dependent Variable: Y1

Method: Least Squares

Date: 12/27/16 Time: 20:10

Sample: 1 10

Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	-0.099152	0.006732	-14.72793	0.0000
C	4.079747	0.041772	97.66618	0.0000
R-squared	0.964430	Mean dependent var		3.534412
Adjusted R-squared	0.959984	S.D. dependent var		0.305682
S.E. of regression	0.061148	Akaike info criterion		-2.574167
Sum squared resid	0.029913	Schwarz criterion		-2.513650
Log likelihood	14.87084	Hannan-Quinn criter.		-2.640554
F-statistic	216.9118	Durbin-Watson stat		0.590934
Prob(F-statistic)	0.000000			

2.模型检验

（1）指数函数模型的拟合优度检验：

从表3-8-1中可以看出模型的拟合优度情况，模型的可决系数 R^2 为0.964，调整后 R^2 为0.960。可见 R^2 和调整后 R^2 都在95%以上。说明该模型回归的贡献很大，表示对数函数回归模型拟合的结果比较理想。

（2）对回归方程进行假设检验--- F 检验：

从表3-8-1中可以看出模型的假设检验结果，表中 F 值为216.912，概率值为0.000，远小于0.05，所以 F 检验通过。

说明在指数函数回归模型 $\ln(y) = \ln(\beta_0) + \beta_1 x$ 中 $\ln(y)$ 与 x 的线性回归高度显著，这与相关系数的检验结果是一致的。

（3）对回归系数进行假设检验--- t 检验：

从表3-8-1中可以得到回归方程为 $\ln(y) = 4.080 - 0.099x$ ，回归系数 β_1 检验的 t 值为-14.728，概率值为0.000；而常数项 β_0 的 t 值为97.667，概率值为0.000。所以， t 检验通过，说明在指数函数回归模型中 $\ln(y)$ 与 x 的线性假设成立，与 F 检验和相

关系数的检验结果一致。

(4) 残差检验，如图3-8-2所示：

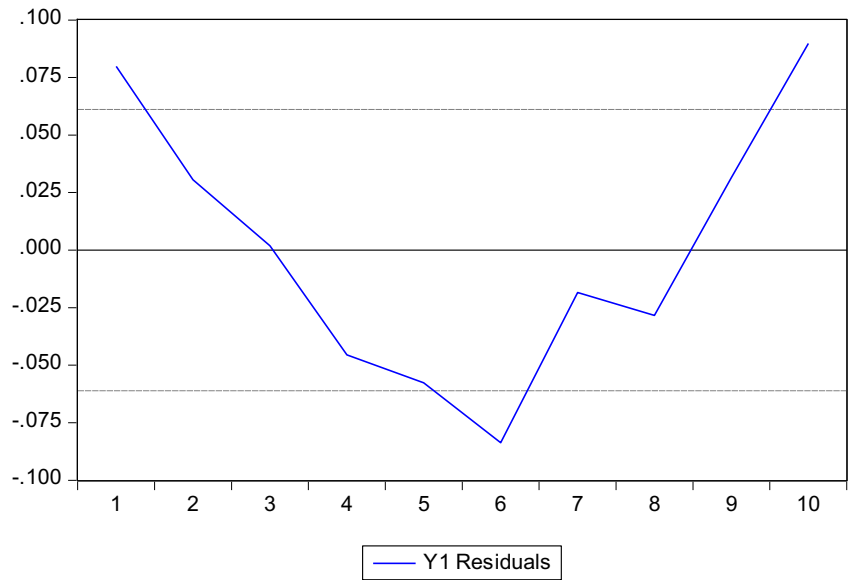


图3-8-2 指数函数模型的残差图

如图3-8-2所示，从指数函数模型的残差图中可以看出，指数函数模型的残差在0附近波动，且上下间距的绝对值不超过1，可见残差项通过检验。

(5) 自相关检验——*LM* 检验，输出结果如表3-8-3所示：

表3-8-3 *LM* 检验结果

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	1.692368	Prob. F(2,6)	0.2613
Obs*R-squared	3.606640	Prob. Chi-Square(2)	0.1648

表3-8-3中可以看出，*LM* 值检验结果的概率值为0.1648，大于显著性水平 $\alpha = 0.05$ ，因此接受 *LM* 检验原假设，可以认为残差序列不存在自相关。

(6) 异方差检验——*White* 检验，输出结果如表3-8-4所示：

表3-8-4 *White* 异方差检验结果

Heteroskedasticity Test: White			
F-statistic	0.699628	Prob. F(2,7)	0.5284
Obs*R-squared	1.665929	Prob. Chi-Square(2)	0.4348
Scaled explained SS	0.495496	Prob. Chi-Square(2)	0.7806

从表3-8-4中可以看出，其中F-statistic是辅助方程整体显著性的F统计量；Obs*R-squared是 *White* 检验统计量 NR^2 ，其概率值为0.4348，明显大于显著性水

平 $\alpha = 0.05$ ，因此接受 *White* 检验原假设，认为原方程的残差序列不存在异方差。

3. 小结

对选取的 10 组关于价格与需求量的数据建立了指数函数模型：

$$y = \beta_0 e^{\beta_1 x}$$

在应用过程中，由于将曲线方程转化为线性方程进行回归分析，因而需将回归方程 $\ln(y) = 4.080 - 0.099x$ 转化为指数函数方程：

$$y = 59.13e^{-0.099x}$$

因为此模型通过了所有的假设检验，所以该模型可以反映价格与需求量的关系，并且可以对关于价格与需求量的经济问题作统计预测。

四、模型评价

下面将对价格与需求量的各个拟合模型的实际值与预测值的拟合结果进行效果比较，从而选出表示价格与需求量关系的最优模型。本文用 *AIC* 准则进行模型的精度评价，*AIC* 准则是一种考评综合最优配置的指标，它是拟合精度和参数未知个数的加权函数，计算公式为：

$$AIC = n \ln \left(\frac{SSE}{n} \right) + 2p$$

使 *AIC* 函数达到最小值的模型被认为是最优模型。各模型比较结果如表 4-1 所示。

表 4-1 价格与需求量模型精度的评价

模型名称	模型函数表达式	拟合优度 R^2	<i>AIC</i>
对数函数模型	$y = 59.506 - 15.695 \ln(x)$	0.991	3.274
二次多项式模型	$y = 65.450 - 8.44x + 0.436x^2$	0.997	2.304
三次多项式模型	$y = 67.467 - 10.229x + 0.823x^2 - 0.024x^3$	0.999	-2.742
幂函数模型	$y = 62.778x^{-0.456}$	0.990（修正后）	-3.926
指数函数模型	$y = 59.13e^{-0.099x}$	0.964	-2.574

从表 4-1 中，可以得出结论：在 *AIC* 准则模型最优配置指标的评价中，对数函数、二次多项式、三次多项式、幂函数、指数函数模型的 *AIC* 值分别为 3.274、2.304、-2.742、-3.926、-2.574，这五个模型的值 *AIC* 均比较小，说明这五个模型都可以较好地对价格与需求量间的关系进行拟合和预测。

但综合比较后，进行加权最小二乘修正后的幂函数模型的 AIC 值为-3.926，是五个模型里最小的，所以，在进行价格与需求量之间关系的拟合和预测时，选取修正后的幂函数模型为最优模型，而三次多项式和指数函数模型的 AIC 值分别为-2.742、-2.574，也比较小，故这两个模型可作备选模型。

综上所述，幂函数模型为本文选取的最优模型，可用于对价格与需求量之间的关系的回归分析和预测。

五、总结

本文选取了价格与需求量的样本数据，为研究价格与需求量之间的关系，建立了多种回归模型进行模型拟合与模型优化，分别是对数函数、二次多项式、三次多项式、幂函数、指数函数模型。再得到五个函数模型的回归方程结果之后，又分别对其进行预测精度的比较与分析。通过比较五个模型的 AIC 值，以及模型的拟合效果，最终选取用加权最小二乘法修正后的幂函数模型作为本文研究出的最优拟合与预测模型。方程如下所示：

$$\text{幂函数模型: } y = 62.778x^{-0.456}$$

本次回归分析，建立不同模型研究因变量（需求量）和自变量（价格）之间的关系，实现对价格与需求量之间关系的回归拟合。这种统计分析方法可用于研究经济问题，观察需求曲线从左上方向右下方倾斜的经济变动趋势，以及回归方程结果的曲线特征，可进一步论证经济规律，即价格上升，需求量减少；价格下降，需求量增加。另外，还可以帮助政府更好地调控市场价格，制定更符合市场规律的经济政策，从而推动经济发展。

参考文献

- [1]何晓群. 应用回归分析（第四版）. 北京：中国人民大学出版社，2015