物资计划与市场

目　　录

第八章　中长期物资供需预测及其数量模型的建立

党的十二届三中全会《关于经济体制改革的决定》中指出："计划工作的重点要转到中期或长期计划上来，适当简化年度计划，并相应改革计划方法，充分重视经济信息和预测，提高计划的科学性"。在物资计划中，搞好预测，特别是搞好中、长期物资供需预测工作，乃是实现国家的经济与社会发展战略目标，保持发展的稳定性和连续性的重要手段。

第一节　预测分类及适用于中长期计划的预测方法

目前，有统计的预测方法不下一百种。根据研究任务的不同，按照一定的特征可把预测方法划分为不同的类别：

一、按照预测对象范围大小不同，可以分为宏观预测和微观预测

（一）宏观预测。宏观预测是对整个国民经济发展前景的预测。它以社会经济发展的总图景作为考核对象，研究经济中各个有关的总量指标、相对数指标和平均数指标之间的联系和发展趋势。如研究国民经济发展水平、速度和规模等都属于宏观预测。

（二）微观预测。微观预测是指对一个企业，一个部门的

某一经济活动发展前景的预测。它以单一的经济活动前景作为考察对象，研究经济发展中各有关经济指标和它们之间的发展趋势。

二、按预测的时间长短不同，可分为长期、中期和近期预测

（一）长期预测。长期预测是指对五年以上经济发展前景的预测，包含物资供求平衡的预测。它是制订经济发展十年或十年以上远景规划和规定长期发展任务的依据。

（二）中期预测。中期预测是指对一年以上、五年或五年以下经济发展前景的预测。它是制订经济发展五年计划和规定五年发展任务的依据。

（三）近期预测。近期预测是指一年或一年以内经济发展前景的预测。近期预测是制订年度计划、季度计划和月度计划等，执行和修订近期经济发展具体任务的依据。

三、按预测的属性不同，可分为定性预测和定量预测

（一）定性预测。定性预测是指对事物的性质作出描述的预测。它一般是判断事物的性质，大致确定出量的概念。因此，定性预测需要更多地依靠业务专家，凭他们的经验作出判断。

（二）定量预测。定量预测一般是从历史数据入手，然后建立数量模型，最后推导出预测值的方法。

定量预测在第二次世界大战后的西方国家已经得到广泛应用和迅速发展。这是因为：第一，西方国家在三十年代以后逐步建立了经济核算体系，为定量预测提供了基础资料；第二，电子计算机的迅速发展与普及，解决了定量预测中的许多复杂疑难的计算问题。同时，数量模型的建立，为定量预测

提供了良好的数学方法；第三，定量预测的成果已成功地为计划与决策服务，收到了良好的效果。所以，目前不仅在西方国家，而且在社会主义国家也引起了极大重视。

四、按预测方法的性质不同，可分为调查预测、因果预测和历史资料延伸预测

（一）调查预测。调查预测是指预测者深入实际，深入群众，走访有关业务专家，进行调查研究并结合自己的实践经验进行判断，然后作出预测的方法。

（二）因果预测。因果预测是把所要预测的经济对象同其他有关因素联系起来进行分析，建立起揭示因果关系的回归数量模型，然后根据回归数量模型进行预测的方法。

（三）历史资料延伸预测。历史资料延伸预测是应用现有的历史统计资料，将过去的发展趋势延伸到未来，并建立数量模型，进行预测的方法。历史资料延伸预测方法把时间的顺序当作预测的对象的趋势因素。因此，它是因果预测的一种，仍可采用回归数量模型进行预测。

调查预测比较简单易行，可利用预测者的丰富实践经验和专业知识作出量的判断，但其缺点是包含的主观因素较多，往往对同一问题，不同的预测者会得出不同的预测结论，很难作出准确的仲裁。因果预测和历史资料延伸预测以系统、准确的统计资料为依据，应用统计方法进行预测的，所以常常把两者统称为统计预测。统计预测充分考虑了数量的和时间的因果关系，可以采用回归方法定量预测较长时期经济发展和变动的前景。但其缺点在于它是以影响的因素变化较稳定的假设条件为依据，一旦因素发生突然意外的变化，预测结果就会大大偏离实际。为了使预测结果比较接近

实际，通常要技术与艺术相结合，与调查预测结果相对照，分析差异的原因，进行必要的调整，求得合理的预测方案。这是取得预测良好效果的重要途径。

通过以上对预测方法的分类，现将主要定量预测方法的性质归纳于表8—1。本章选择适于中、长期预测方法作重点介绍，并配上电子计算机源程序，力求使这些预测方法在物资计划工作中得以广泛应用。

表8—1　主要定量预测方法性质比较

方法＼性质			适应时间范围			所属类型			
			近期	中期	长期	时间序列型	因果型	统计型	非统计型
定量预测方法	线性平滑法	移动平均法	√			√			√
定量预测方法	线性平滑法	指数平滑法	√			√			√
定量预测方法	随机时间序列法		√	√		√	√	√	
定量预测方法	回归分析法	一元线性回归法	√	√	√	√	√	√	
定量预测方法	回归分析法	多元线性回归法	√	√	√		√	√	
定量预测方法	回归分析法	非线性回归法	√	√	√	√	√	√	
定量预测方法	结构分析法	经济计量模型法	√	√	√		√	√	
定量预测方法	结构分析法	投入产出模型法		√	√		√	√	

第二节　一元线性回归预测及其电子计算机源程序设计

根据事物之间关系的复杂程度，回归预测可分成多种类

型：当相关关系的统计规律呈线性关系时，称为线性回归；反之称为非线性回归。在线性回归中，自变量只有一个，称为一元线性回归；自变量在两个或两个以上的，称为多元线性回归。本节先介绍一元线性回归。

一、一元线性回归预测步骤

第一，收集资料，确定变量。根据预测目的和要求，收集有关资料，找出对预测目标(即预测值)影响最大的因素作为自变量，而将影响较小的因素归为随机误差，使问题得到简化；

第二，绘制散布图，将收集的资料画在坐标平面图上，便可直观地看出因变量(Y_i)和自变量(X_i)变化的趋势，据以确定回归方程的形式(线性的或非线性的，非线性时是指数曲线还是抛物线等。这里假定是直线)；

第三，计算回归参数(也称回归系数)。根据收集的资料，采用最小二乘法推导的公式，计算出一元线性回归方程$\hat{Y}_i = b_0 + b_1X_i$的两个参数b_0和b_1将其代入，于是建立起回归数量模型；

第四，进行统计检验。用回归数量模型表示的$\hat{Y}_i$与X_i的关系是线性的，它和统计的实际值总会有误差。为了确知回归数量模型是否符合客观实际，还必须对其进行各种统计检验；

第五，应用回归数量模型进行预测。只要给出一个自变量X_i值，便可以通过模型计算出一个因变量$\hat{Y}_i$值。这个值就是我们需要的预测值。

二、一元线性回归数量模型的建立

一元线性回归方程一般表示为：

$$\hat{Y}_i = b_0 + b_1X_i + \delta_i \qquad (i = 1, 2, \cdots, m)$$

式中：X_i是自变量。它是影响$\hat{Y}_i$的主要因素；

$\hat{Y}_i$是因变量。它随着自变量X_i的变化而变化；

δ_i是随机误差。它表示除X_i对$\hat{Y}_i$起主要作用外，其余的非主要因素或偶然因素的影响；

b_0、b_1是回归参数，即待定未知常数。它们的大小代表了X_i对$\hat{Y}_i$的影响的程度。其中b_0是直线截矩，表示X_i为0时$\hat{Y}_i$的某一具体值；b_1是直线的斜率，表示X_i每变化一个单位时的$\hat{Y}_i$变化值。

假定暂不考虑自变量X_i以外因素的影响，则一元线性回归方程为：

$$\hat{Y}_i = b_0 + b_1X_i$$

为了使预测误差最小，同时又不使预测误差正负抵消，故追求预测误差平方和最小，即$\sum(Y_i - \hat{Y}_i)^2 = \min$。用最小二乘法导出$b_0$、$b_1$两个参数的计算公式为：

$$b_0 = \overline{Y} - b_1\overline{X}$$

$$b_1 = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})^2} = \frac{LXY}{LXX}$$

式中：$\overline{Y}$为Y_i的平均值即$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$（m为样本容量）

$\overline{X}$为X_i的平均值即$\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$

$\Sigma(X_i-\overline{X})(Y_i-\overline{Y})$是$X_i$和$Y_i$的离差乘积之和，简写为“LXY”。

$\Sigma(X_i-\overline{X})^2$是$X_i$的离差平方和，简写为“LXX”。

只要求出b_0，b_1，将其代入回归方程，即可建立起一元线性回归数量模型。

三、拟合优度检验

为了考查预测回归模型与统计值之间的线性拟合程度，这就要求有一个数量指标来加以判别。这个数量指标就是样本值的判别系数，习惯上常称为相关系数γ。它可以判别出回归值线性拟合优度。γ的计算公式是：

$$\gamma=\frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum(X_i-\overline{X})^2\cdot\sum(Y_i-\overline{Y})^2}}=\frac{LXY}{\sqrt{LXX\cdot LYY}}$$

式中：

$\sum(Y_i-\overline{Y})^2$是Y_i的离差平方和，俗称“总变差”，简写为“LYY”。

相关系数γ值摆动在$+1\sim-1$之间。它的绝对值$|\gamma|$大小反映了变量Y_i与X_i的线性相关程度(详见图8—1)。

从图8—1中可以看出：只有当γ的绝对值接近1时，才能用回归直线近似地描述变量Y_i和X_i的线性关系。这时，说明拟合性能好，有实用价值。当γ接近于0时，可以认为X_i与Y_i不存在线性关系，更谈不上拟合优度，只能采用非线性回归模型描述。

γ值的大小不仅与各点的离散程度有关，而且也与数据点的个数有关。为了保证线性关系对γ的最低要求值(即临界值)，可以查阅相关系数检验表(详见表8—2)。表中的数

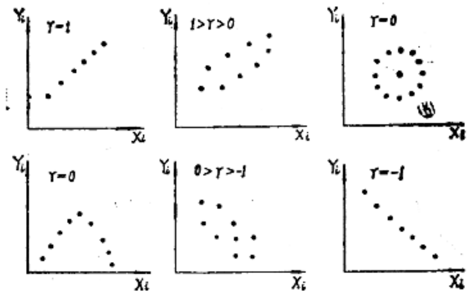

图 8 — 1　相关程度示意

值是相关性检验的标准数值。其中α称为显著水平。一般有两种水平$\alpha = 0.05$和$\alpha = 0.01$，常用的是$\alpha = 0.05$。f称为自由度。在一元线性回归中，它等于样本数量m减去回归系数个数所得的差。为了确定显著水平，可以从表中查出相应的临界值γ_α，然后将计算出来的相关系数γ值与临界值γ_α相比较。若$|\gamma| > \gamma_\alpha$，则可认为Y_i与X_i存在线性关系，可以用回归数量模型近似地进行预测。否则，没有实用价值。

显著水平α的含义可以这样理解：当$|\gamma| > \gamma_\alpha$时，其显著水平为$100 \times (1-\alpha)\%$。例如取$\alpha = 0.05$，则可认为Y_i与X_i线性相关的显著水平为95%。

这里要注意：γ值的大小只能说明两个变量之间存在线性关系的密切程度，而不能说明非线性关系的密切程度；另外，也不能把存在线性关系就看成一定是因果关系。

四、预测精度检验

利用回归数量模型进行预测时，给定一个自变量值，可

表 8 — 2　相关系数检验表

f m − 2	α=0.05	α=0.01	f m − 2	α=0.05	α=0.01	f m − 2	α=0.05	α=0.01
5	0.754	0.874	15	0.482	0.606	25	0.381	0.487
6	0.707	0.874	16	0.468	0.590	30	0.349	0.449
7	0.666	0.798	17	0.456	0.575	35	0.325	0.418
8	0.632	0.765	18	0.444	0.561	40	0.304	0.393
9	0.603	0.735	19	0.433	0.549	45	0.288	0.372
10	0.576	0.708	20	0.423	0.537	50	0.273	0.354
11	0.553	0.684	21	0.413	0.526	100	0.195	0.254
12	0.532	0.661	22	0.404	0.515	200	0.138	0.181
13	0.514	0.641	23	0.396	0.505	300	0.113	0.148
14	0.497	0.632	24	0.388	0.496	1000	0.062	0.081

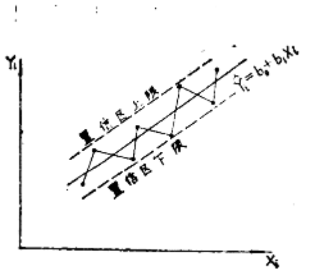

图8—2　置信区间

以计算出一个因变量确定的值。但这个确定的值仅仅是一个估计值，在实际中对预测有意义的不一定是个确定值，而往往是一个范围和区间。同时，还要确定未来的实际值在这个区间内的可能性有多大。这个区间就称为置信区间(详见图8—2)。这个区间越小，说明预测的精度越高。

每个X_i值对应一个$\hat{Y}_i$值。它们总是按一定的分布规律在回归直线附近波动。只要按照最小二乘法原理建立的回归数量模型，当样本数量足够多时，其波动规律一般都属于正态分布。也就是说，$\hat{Y}_i$实际是正态分布的随机变量。这个随机变量上下波动程度，用一个标准量S来表示。它的计算公式是：

$$S=\sqrt{\frac{\sum(Y_i-\hat{Y}_i)^2}{m-2}}$$

式中：

$\sum(Y_i-\hat{Y}_i)^2$为剩余平方和。因为它是自变量X_i对因变量$\hat{Y}_i$的线性影响以外的一切因素作用的结果，所以它也称为预测误差平方和。m－2为剩余平方和的自由度。

根据正态分布性质，预测值在$\hat{Y}_i-S\sim\hat{Y}_i+S$区间的概率为68.27%(即有68.27%的可能性使预测值落入$\hat{Y}_i-S\sim\hat{Y}_i+$

S范围内)；在$\hat{Y}_i-2S\sim\hat{Y}_i+2S$区间的概率为95.45%(即有95.45%的可能性使预测值落入$\hat{Y}_i-2S\sim\hat{Y}_i+2S$范围内)；在$\hat{Y}_i-3S\sim\hat{Y}_i+3S$区间的概率为99.73%(即有99.73%的可能性使预测值落入$\hat{Y}_i-3S\sim\hat{Y}_i+3S$范围内)。因此，标准差越小，置信区间越小，预测的精度越高。一般多取$\hat{Y}_i\pm 2S$作为回归预测区间的上、下限(即置信区间)。

五、举例说明

例8—1，根据过去某种物资的需求量Y_i和工业产值X_i的九组统计资料(详见表8—3)，试建立相应的回归数量模型并预测工业产值6.5万元的需求量。

表8—3　过去的九组统计数据

需求量Y_i(万吨)	4.8	5.7	7.0	8.3	10.9	12.4	13.1	13.6	15.3
工业产值X_i(万元)	1.5	1.8	2.4	3.0	3.5	3.9	4.4	4.8	5.0

因为该例已给出因果条件，所以从绘制散布图开始确定因果相关性质。

1. 绘制散布图(详见图8—3)。

从图中可判断X_i与Y_i的因果关系可以近似地用直线表示，即可用一元线性回归方程：$\hat{Y}_i=b_0+b_1X_i$来描述。

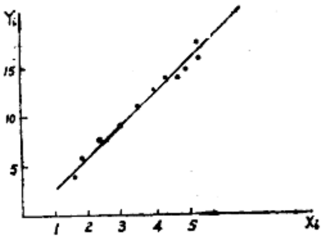

图8—3　散布图

2. 计算回归参数b_0、b_1，建立一元线性回归数

表 8 — 4　手工计算表

序号	Y_i	X_i	$X_i-\bar{X}$	Y_i-Y	$(X_i-X_i)(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$	$(Y_i-Y)^2$
1	4.8	1.5	−1.8667	−5.3222	9.9350	3.4846	28.3258
2	5.7	1.8	−1.5667	−4.4222	6.9283	2.4545	19.5559
3	7.0	2.4	−0.9667	−3.1222	3.0182	0.9345	9.7481
4	8.3	3.0	−0.3667	−1.8222	0.6682	0.1345	3.3204
5	10.9	3.5	0.1333	0.7778	0.1037	0.0178	0.6050
6	12.4	3.9	0.5333	2.2778	1.2148	0.2844	5.1884
7	13.1	4.4	1.0333	2.9778	3.0770	1.0677	8.8673
8	13.6	4.8	1.4333	3.4778	4.9847	2.0543	12.0951
9	15.3	5.0	1.6333	5.1778	8.4569	2.6677	26.8096
Σ	91.1 $\bar{Y}=10.1222$	30.3 $\bar{X}=3.3667$			38.3868	13.1000	114.5156

量模型。

手工计算时，为了避免出现错误，常常把计算的数据列在一张表上（详见表8—4）。则：

$$b_1=\frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sum(X_i-\overline{X})^2}=\frac{38.3868}{13.1000}=2.93$$

$$b_0=\overline{Y}-b_1\overline{X}=10.1222-2.93\times3.3667=0.26$$

于是得一元线性回归数量模型为：

$$\hat{Y}_i=0.26+2.93X_i$$

3. 计算相关系数，检验拟合优度

用表8—4数据，则：

$$\gamma=\frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum(X_i-\overline{X})^2\cdot\sum(Y_i-\overline{Y})^2}}=\frac{38.3868}{\sqrt{13.1000\times114.5156}}$$

$$=0.991$$

令显著水平$\alpha=0.05$，因其自由度为$9-2=7$，查表8—2相关系数检验表知：$\gamma_\alpha=0.666$。显然，$\gamma>\gamma_\alpha$。这说明需求量与工业产值的线性关系密切，拟合优度高。故所建立的一元线性回归数量模型$\hat{Y}_i=0.26+2.93X_i$，至少有95%〔$100\times(1-0.05)=95\%$〕的把握。

4. 计算标准偏差，检验预测精度

将X_i值分别代入模型计算出$\hat{Y}_i$，然后计算出剩余平方和$\sum(Y_i-\hat{Y}_i)^2$（详见表8—5）。

则：

$$S=\sqrt{\frac{\sum(Y_i-\hat{Y}_i)^2}{m-2}}=\sqrt{\frac{2.0320}{9-2}}=0.539$$

表8—5　手工计算表

序号	Y_i	X_i	$\hat{Y}_i$	$Y_i-\hat{Y}_i$	$(Y_i-\hat{Y}_i)^2$
1	4.8	1.5	4.655	0.145	0.0210
2	5.7	1.8	5.534	0.166	0.0276
3	7.0	2.4	7.292	−0.292	0.0853
4	8.3	3.0	9.050	−0.750	0.5625
5	10.9	3.5	10.515	0.385	0.1482
6	12.4	3.9	11.687	0.713	0.5084
7	13.1	4.4	13.152	−0.052	0.0027
8	13.6	4.8	14.324	−0.724	0.5242
9	15.3	5.0	14.910	0.390	0.1521
Σ					2.0320

按两个标准差$2S=2\times0.539=1.078$计算，则预测的置信区间为：

上限：$\hat{Y}_{上}=0.26+2.93X_i+1.078=1.338+2.93X_i$

下限：$\hat{Y}_{下}=0.26+2.93X_i-1.078=-0.818+2.93X_i$

故，预测精度在这一区间出现的可能性为95.45%。

5. 进行预测

将$X=6.5$万元代入上述模型，得预测值为：

$$\hat{Y}_i=0.26+2.93\times6.5=19.305(\text{万吨})$$

因为标准差为0.539，所以预测精度在18.227～20.383万吨之间出现的可能性为95.45%。或者说有4.55%的风险度超出这一区间范围。