

运用多元线性回归模型和 微型计算机预测专门人才需求量

周 镇 江

(化工系)

一、引 言

人才规划是一种战略性规划。人才预测是一个复杂的社会系统工程。在进行预测专门人才的需求时,通常采用传统经验法和数学模型预测法两大类。前者是一种直观性预测,它从事物的外部,通过主观判断和集体思维去预测未来。后者是一种探索性预测,主要研究事物发展的可能趋势和可能途径,它是建筑在充分经验和相当可靠的理论基础上,有重要的实用价值。

根据教育部、国家计委、劳动人事部1983年召开的全国专门人才规划工作会议的精神和《关于进行全国专门人才现状调查和需求预测的通知》的要求,“在预测和规划中,传统方法和现代方法结合,并尽可能采用现代的科学预测规划的方法。”为此,我们在轻业工部领导下,在洗涤用品工业的全国专门人才的预测工作中,除采用传统的经验预测法外,还应用现代的数学模型预测法,并运用微型计算机进行解算。

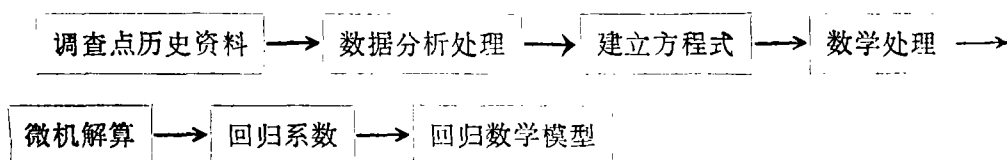
本文将两种方法预测专门人才的结果作一比较,重点介绍回归数学模型预测法。

二、模型的建立

根据六个典型厂的调查发现,一个厂的专门人才(专门人才是指:研究生、本科、大专和中专生四个学历层次)的发展有一定的规律,它与生产的发展速度、规模、劳动力总数、投资总额、利润、税金以及技术措施投资等因素密切相关。就全国来说,整个经济的发展状况、科学技术水平都会对专门人才的需求产生很深的影响。鉴于这些复杂的因素,结合典型厂的具体状况,我们采用多元线性回归模型和微型计算机去预测未来对专门人才的需求量。

多元线性回归方程属于经验归纳模型,其特点是立足于数据,得出模型较简便,不足之处是缺乏明确的理论基础,对深入认识问题有困难。由于我们通过对典型厂的调查,收集到较完整的历史统计数据,因此,可以通过运用相关分析的统计方法得出回归方程。用多元线性回归模型预测专门人才,若考虑的因素多些(当然不能丢掉主要因素),则预测的准确性会高些,这是因为专门人才不但与各因素之间有“纵”的关系,而且在诸因素之间又存在着各种“横”的关系。

用下列方块图示意模型建立的程序:



1. 原理

以 y 表示专门人才数, $x_i (i=1,2,3,\dots)$ 表示职工总数、劳动生产率、总产值、固定资产、利润、税金、技术措施投资等因素。若 y 与 x 之间存在着较明显的线性关系, 则有

$$y = f(x_1, x_2, x_3, \dots)$$

设方程形式如下:

$$y = \beta_0 + \beta_1(x_1 - \bar{x}_1) + \beta_2(x_2 - \bar{x}_2) + \dots + \beta_7(x_7 - \bar{x}_7)$$

其中 $x_i (i=1,2,\dots,7)$ 表示第 i 个因子的取值; $\bar{x}_i$ 表示第 i 个因子取值的平均值; y 为对应一组 $(x_1, x_2, \dots, x_7)$ 的效应; $\beta_0, \beta_1, \dots, \beta_7$ 为所需确定的常数。

由典型厂得到的统计数据, 用最小二乘法拟得方程 $\beta = (X'X)^{-1}X'Y$ 其中

$$X = \begin{pmatrix} 1 & x_{11} - \bar{x}_1 & x_{21} - \bar{x}_2 & \dots & x_{m1} - \bar{x}_m \\ 1 & x_{12} - \bar{x}_1 & x_{22} - \bar{x}_2 & \dots & x_{m2} - \bar{x}_m \\ 1 & x_{13} - \bar{x}_1 & x_{23} - \bar{x}_2 & \dots & x_{m3} - \bar{x}_m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} - \bar{x}_1 & x_{2n} - \bar{x}_2 & \dots & x_{mn} - \bar{x}_m \end{pmatrix}$$

X' 为 X 的转置矩阵; $(X'X)^{-1}$ 为 $X'X$ 的逆阵;

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

由方程 $\beta = (X'X)^{-1}X'Y$ 可解得 $\beta_0, \beta_1, \dots, \beta_7$ 的估计值 $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_7$, 从而可得到含所有因子的回归方程。

以国营洗涤剂 D_2 型厂为例。该厂有下列历史统计数据(表1)。

由表1可见, 从1976至1983年间, 专门人才数量的变化与其七个因素之间存在着较明显的线性关系, 经过数学处理可得到相应的多元线性回归方程。

2. 回归方程

由表1可得到下列数据:

$$\begin{aligned} \bar{x}_1 &= 261.13, & x_2 &= 58264.38, & \bar{x}_3 &= 1600.55, & \bar{x}_4 &= 123.06, \\ \bar{x}_5 &= 113.44, & \bar{x}_6 &= 289.19, & x_7 &= 11.61 \end{aligned}$$

表 1

项 目 年 份	职 工 人 数 (人)	劳 动 生产率 (元/人)	工 业 总产值 (万元)	固 定 资 产 (万元)	年税金 (万元)	年利润 (万元)	技 措 投 资 (万元)	专 门 人 才 (人)
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	y_i
76	218	32146	700.8	111.70	45.8	80.6	8.23	7
77	226	34353	776.4	113.19	49.9	84.1	2.89	9
78	243	35885	872.0	115.88	53.8	116.4	2.69	11
79	249	49104	1222.7	120.67	66.7	223.7	16.84	14
80	275	44120	1213.3	124.58	77.0	216.0	12.30	15
81	280	61307	1716.6	129.42	137.0	382.7	10.50	19
82	299	90388	2702.6	134.05	208.3	540.0	12.46	22
83	299	118812	3600.0	135.00	269.0	670.0	27.00	29

因而可用下列矩阵形式表示:

$$X = \begin{pmatrix} 1 & -43.13 & -26118.38 & -899.75 & -11.36 & -67.64 & -208.59 & -3.38 \\ 1 & -35.13 & -23911.38 & -824.15 & -9.87 & -63.54 & -205.09 & -8.72 \\ 1 & -18.13 & -22379.38 & -728.55 & -7.18 & -59.64 & -172.79 & -8.92 \\ 1 & -12.13 & -9160.38 & -377.85 & -2.39 & -46.74 & -65.49 & 5.23 \\ 1 & 13.87 & -14144.38 & -387.25 & 1.52 & -36.44 & -73.19 & 0.69 \\ 1 & 18.87 & 3042.62 & 116.05 & 6.36 & 23.56 & 93.51 & -1.11 \\ 1 & 37.87 & 32123.62 & 1102.05 & 10.99 & 94.86 & 250.81 & 0.85 \\ 1 & 37.87 & 60547.62 & 1999.45 & 11.94 & 155.56 & 380.81 & 15.39 \end{pmatrix}$$

$$Y = \begin{pmatrix} 7 \\ 9 \\ 11 \\ 14 \\ 15 \\ 19 \\ 22 \\ 29 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \end{pmatrix}.$$

上列矩阵经过数学处理和运用微型计算机计算,求得下列回归系数:

$$\begin{aligned} \hat{\beta}_0 &= 15.7572, & \hat{\beta}_1 &= 1.10845, & \hat{\beta}_2 &= 5.01448E-03, \\ \hat{\beta}_3 &= -0.178232, & \hat{\beta}_4 &= -2.14004, & \hat{\beta}_5 &= 0.544506, \\ \hat{\beta}_6 &= -0.115717, & \hat{\beta}_7 &= 0.580344 \end{aligned}$$

将回归系数代入,得到回归方程

$$\hat{y} = -51.66 + 1.108x_1 + 0.005x_2 - 0.178x_3 - 2.14x_4 + 0.545x_5 \\ - 0.116x_6 + 0.58x_7$$

三、回归方程的验算

1. 计算回归值 $\hat{y}_i$

将得到的回归方程以 1976 年至 1983 年的历史统计数据代入, 即可得到 $\hat{y}_i (i=1, 2, \dots, 8)$ 。今以 1976 年为例, 其历史统计资料为

$$x_1 = -43.13, \quad x_2 = -26118.38, \quad x_3 = -899.75, \quad x_4 = -11.36, \\ x_5 = -67.64, \quad x_6 = -208.59, \quad x_7 = -3.38$$

将各 x_i 值代入回归方程, 得

$$\begin{aligned} \hat{y}_{1(1976)} &= -51.66 + 1.108(-43.13) + 0.005(-26118.38) \\ &\quad - 0.178(-899.75) - 2.14(-11.36) + 0.545(-67.64) \\ &\quad - 0.116(-208.59) + 0.58(-3.38) \\ &= 7.215 \text{ 人} \end{aligned}$$

其余各年, 类此计算, 其数据列入表 2。

表 2

年 份	1976	1977	1978	1979	1980	1981	1982	1983
人才数								
$\hat{y}_i$	7.215	9.2	11.185	14.044	15.111	18.96	21.75	28.538
y_i	7	9	11	14	15	19	22	29

2. 计算复相关系数 R

公式:

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

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i \quad (i=1, 2, \dots, 8) \\ &= 15.75 \end{aligned}$$

$$\sum_{i=1}^8 (\hat{y}_i - \bar{y})^2 = 349.769$$

$$\sum_{i=1}^8 (y_i - \bar{y})^2 = 373.504$$

则

$$R^2 = \frac{349.769}{373.504} = 0.9365$$

$$R = \pm 0.968$$

因 $R = \pm 0.968$ 接近 1, 表明该回归方程是相当正确的。

四、两种预测方法的比较

1. 经验预测

通过领导、科技人员和有关干部对该厂历史和现状的分析、综合, 以及对今后发展远景作出的判断, 经验预测 1990 年该厂对专门人才的需求量为 112 人。同 1983 年现状比较, 其变化见表 3。

表 3

项 目 年 份	职 工 总 数 (人)	专门人才			专门人才比例 本科 : 大专 : 中专	专门人才 占职工比 (%)	生产能力 (万吨/年)
		本 科	大 专	中 专			
1983年现状	299	共 计 29				9.70	1.35
1990年预测 (经验法)	620	10	34	68	1 : 3.4 : 6.8	18.06	3.35

2. 模型预测

模型预测是运用得到的多元线性回归方程, 参照该厂历史上诸因素对专门人才的影响资料, 并作出有依据的适当调整后再以 1983 年为比较基准进行的。预计到 1990 年对各因素的影响如表 4。

表 4

因 素	x_1	x_2	x_3	x_4	x_5	x_6	x_7
比1983年增长倍率	2.08	1.50	2.00	1.00	2.00	2.50	3.60

将表 4 中 x_1 、 x_2 、 x_3 、 $\dots$ 、 x_7 和其倍率的关系代入回归方程, 预测得到 1990 年该厂专门人才的需求量为

$$\begin{aligned}
 \hat{y}_{1990} &= -51.66 + 1.108(621.92) + 0.005(178218) \\
 &\quad - 0.178(7200) - 2.14(135) + 0.545(538) \\
 &\quad - 0.116(1675) + 0.58(97.2) \\
 &= 113.31 \text{ 人 (取 113)}
 \end{aligned}$$

同理,我们以1990年为比较基准,参照该厂的历史资料和发展趋势,预计到2000年对各因素的影响如表5。

表5

因 素	x_1	x_2	x_3	x_4	x_5	x_6	x_7
比1990年增长倍率	1.5	1.7	1.8	1.3	2.0	2.0	3.8

从而得到2000年该厂对专门人才的需求量为

$$\begin{aligned}\hat{y}_{2000} &= -51.66 + 1.108(621.92 \times 1.5) + 0.005(178218 \times 1.7) \\ &\quad - 0.178(7200 \times 1.8) - 2.14(135 \times 1.3) \\ &\quad + 0.545(538 \times 2) - 0.116(1675 \times 2) \\ &\quad + 0.58(97.2 \times 3.8) \\ &= 226.42 \text{ 人(取 226)}\end{aligned}$$

经过专家讨论、比较和审定,认为用多元线性回归法预测专门人才的需求量是可行的。

五、覆 盖

覆盖是以典型调查点的预测为依据,然后扩展到全国同类型企业到达1990年和2000年两个目标年需求的专门人才数。目前我国洗涤剂 and 肥皂企业数量相当多,且在生产规模、产品种类等方面不尽一致,因此,如仅选择一个典型厂去覆盖全国,则代表性显得不足,覆盖后的精度差,但又不可能也不需要选择很多的典型厂去覆盖全国同类型的企业。为此,我们将全国洗涤用品工业分为洗涤剂和肥皂两大类,每类选择两个有代表性的厂作为典型调查点,对它们除进行经验预测外,同时进行了模型预测。现将四个典型厂的数学预测模型介绍如下。

1. 洗涤剂 D_1 型厂的数学模型

$$\begin{aligned}\hat{y}_{D1} &= -25580.922 + 19.038 x_1 + 4252.94 x_2 - 3.382 x_3 \\ &\quad + 5.729 x_4 - 1.848 x_5 - 0.509 x_6\end{aligned}$$

2. 洗涤剂 D_2 型厂的数学模型

$$\begin{aligned}\hat{y}_{D2} &= -51.66 + 1.108 x_1 + 0.005 x_2 - 0.178 x_3 - 2.14 x_4 \\ &\quad + 0.545 x_5 - 0.116 x_6 + 0.58 x_7\end{aligned}$$

3. 肥皂 S_1 型厂的数学模型

$$\begin{aligned}\hat{y}_{S1} &= 407.318 + 0.0127 x_1 - 0.0034 x_2 - 0.0093 x_3 \\ &\quad - 0.0307 x_4 + 0.0009 x_5 + 0.0729 x_6\end{aligned}$$

4. 肥皂 S_2 型厂的数学模型

$$\begin{aligned}\hat{y}_{S2} &= 18.062 + 0.0327 x_1 - 0.0004 x_2 + 0.0076 x_3 \\ &\quad + 0.0179 x_4 - 0.1703 x_5 + 0.0875 x_6 - 0.028 x_7\end{aligned}$$

在进行全国同类型企业的覆盖时,为使覆盖的结果精确度高些,我们考虑以下两点:①凡接近或类似上述四个典型厂之一的都朝该点靠拢;②现有(1983年)全国同类型企业到1990年能达到或接近达到上述四个典型厂的有多少。后一点,我们是通过通过对现有(1983年)全国所有洗涤剂 and 肥皂企业的排队分析,预计发展趋势和可能,提出一个初步数目,然后请主管部门审定。今以洗涤剂工厂为例,全国现有(1983年)企业约74家,到1990年估计可折合典型厂的数目如表6。

表 6

项 目	年 份	1983 年		1990年	合 计
		规模(万吨/年)	工厂数	折合典型厂数	
工 厂 类 别	D ₂ 型	<1	58	14	24
		<2	6	4	
		<3	7	6	
	D ₁ 型	<5	2	1	2
		>5	1	1	

根据表6的数据,我们就可以以D₁型和D₂型厂为基准,覆盖全国洗涤剂行业至1990年需求的专门人才数量。表7为1990年全国洗涤剂行业实际需要补充各类专门人才数量(扣除退休的专门人才)。

同样,我们可以用上述方法覆盖全国肥皂行业至1990年实际需要补充的各类专门人才的数量。

把洗涤剂和肥皂二个行业叠加起来,就可得到1990年全国洗涤用品工业各类专门人才的实际需要补充数。

表 7 1990年全国洗涤用品工业各类专门人才需要补充数(D型厂)

项 目	年 份	1983年				1990年								各类专门人才需要补充总数	专门人才层次比例 本:大:中 科:专:专
		规 模 (万吨/年)	数 量 (个)	折 合 典型厂 (个)	调查点需要补充数				全国同类型企业需要补充数						
					研究 生	本科	大专	中专	研究 生	本科	大专	中专			
工 厂 类 别	D ₂ 型	<1	58	14	—	9	25	59	—	126	350	826	2700	1:2.5:5.7	
		<2	6	4	—	9	25	59	—	36	100	236			
		<3	7	6	—	9	25	59	—	54	150	354			
	小 计	71	24	—	27	75	177	—	216	600	1416				
	D ₁ 型	<5	2	1	3	38	60	133	3	38	60	133			
		>5	1	1	3	38	60	133	3	38	60	133			
	小 计	3	2	6	76	120	266	6	76	120	266				
合 计		74	26	6	103	195	443	6	292	720	1682				

如果需要预测 2000 年时所需专门人才数量, 则同样可以运用上述数学模型进行计算。已如上述, D_2 型洗涤剂厂对专门人才的模型预测为 $\hat{y}_{1990} = 113$ 人, 而 $\hat{y}_{2000} = 226$ 人。

同理, 运用 D_1 型洗涤剂厂也可用数学模型求出 $\hat{y}_{1990}$ 和 $\hat{y}_{2000}$ 的专门人才需求量。

由此, 根据 D_1 型和 D_2 型洗涤剂厂可预测出整个洗涤剂行业到 2000 年时的专门人才需求量。

推而广之, 由 D_1 型和 D_2 型洗涤剂厂与 S_1 型和 S_2 型肥皂厂可覆盖全国洗涤用品工业到 2000 年时各类专门人才的需求量。

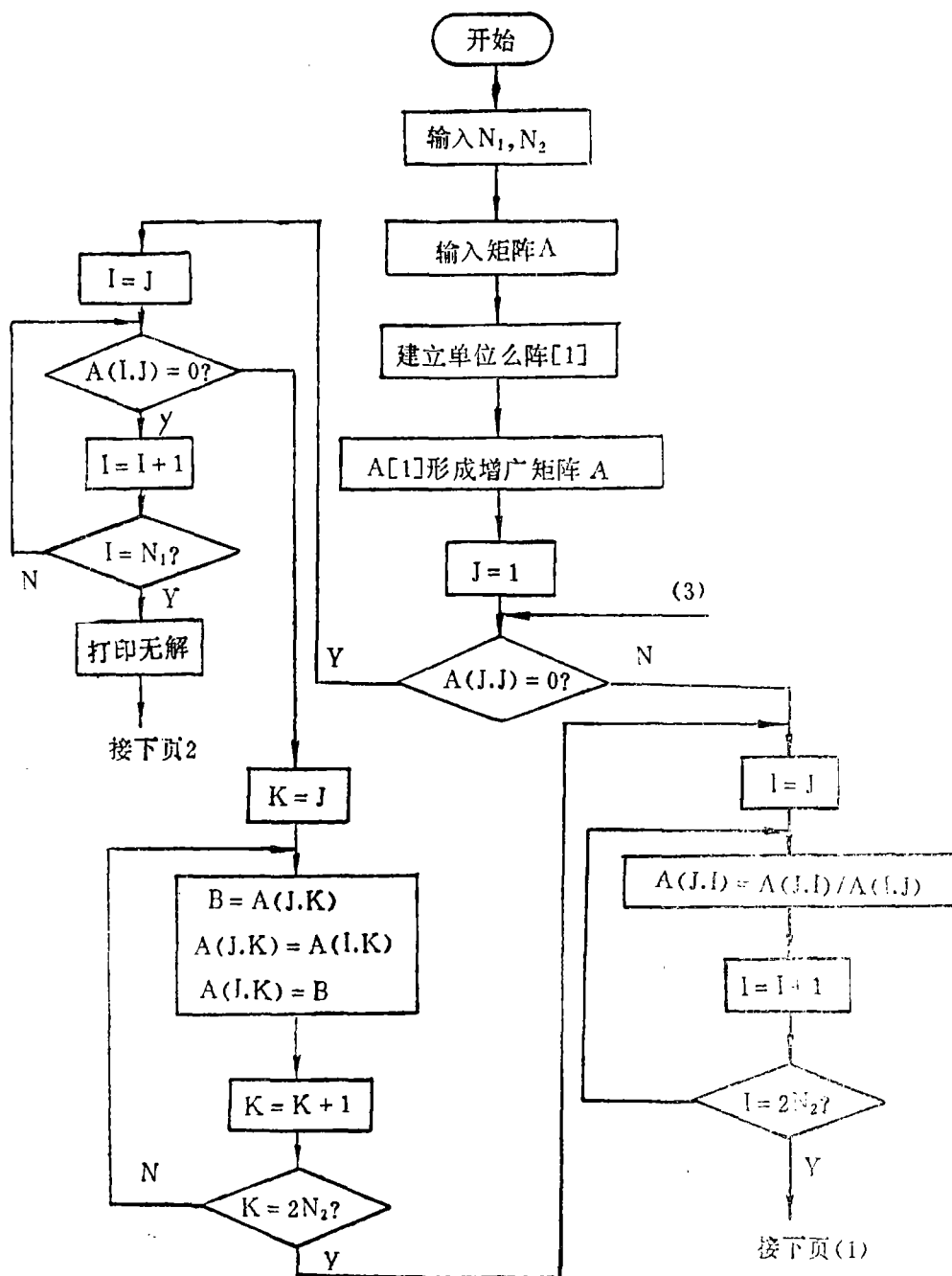
致 谢

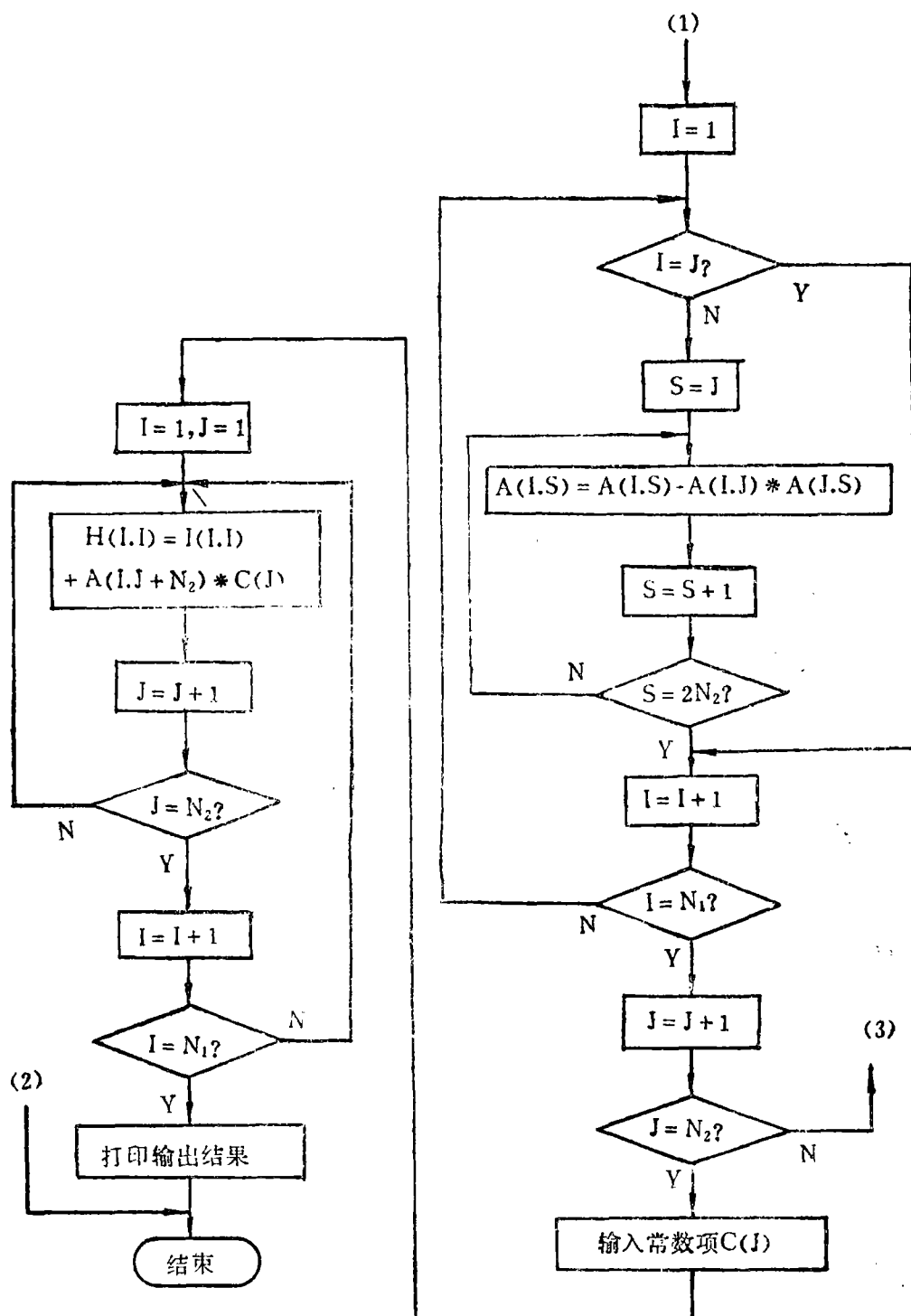
本文能得到完成, 是与轻工业部塑化局、教育司领导的关怀和支持, 与参加调查组的薛珩(组长)、吴群、高曼珠、刘德等同志的努力分不开的; 在写作过程, 北京轻工业学院陈永毅副教授给予热情的鼓励, 无锡轻工业学院曹传书老师对全文进行审查, 黄学祥老师在微机上进行解算, 胡惠良老师给予具体的帮助, 在此, 一并致谢。

参 考 文 献

- [1] Douglas C. M., Design and Analysis of Experiments, Wiley 1976, P304~339.
- [2] 北京大学数学力学系概率统计组编, 《正交设计法》, 石油化学工业出版社, 1975, P 109~120

附录 求解巨阵流程图





84022

九点格式椭圆型差分方程的 Monte Carlo 解法 《无锡轻工业学院学报》, 1984 年, 第 3 卷, 第 3 期

关键词 九点格式椭圆型差分方程, Monte Carlo 解法, 平均游动步数。

摘要 本文给出了九点格式椭圆型差分方程的第一、第二和第三边值问题的 Monte Carlo 解法及平均游动步数的估计。

作者: 张建华 费荣昌

84023

运用多元线性回归模型和微型计算机预测专门人才需求量 《无锡轻工业学院学报》, 1984 年, 第 3 卷, 第 3 期

关键词 人才预测, 微电子计算机, 数学模型。

摘要 本文运用数学回归法和微型计算机建立了一种预测专门人才的数学方法。得到了四种不同类型的数学模型。利用数学模型可预测未来(1990—2000 年)对专门人才的需求量。

作者: 周镇江

84024

用阳极溶出伏安法测定啤酒中铜、铝的含量 《无锡轻工业学院学报》, 1984 年, 第 3 卷, 第 3 期

关键词 啤酒, 铜铝测定, 阳极溶出伏安法。

摘要 本文主要介绍了阳极溶出伏安法在啤酒成分中痕量铜铝离子的分析方法, 用玻璃电极为工作电极, 以醋酸—醋酸钠缓冲液为底液, 对铜铝离子进行了本方法的条件试验, 选择了理想的电解电位、电积时间、扫描速率等条件。酒样用标准加入法, 所测铜铝的实验数据有较好的重现性, 铜铝测定范围可达 10—360 ppb, 回收率在 90—110% 之内, 变异系数 $< \pm 10\%$ 。

作者: 余槐 陆军

84025

蛋白质分散指数 PDI 测定实践 《无锡轻工业学院学报》, 1984 年, 第 3 卷, 第 3 期

关键词 大豆, 蛋白质, 指数。

摘要 测定 PDI 在衡量大豆制品中蛋白质的变性程度及评价其质量起着重要作用, 近来, 我国不少单位也进行了这项测定, 但缺乏统一的仪器和方法。本文介绍的是采用国产仪器进行 PDI 测定的一种方法, 此法较接近 A.O.C.S. 规定, 测得数据重复性较好, 可使测定 PDI 成为一般油厂和实验室的常规分析手段。

作者: 张大煜 陆时凤

84024

DETERMINATION of COPPER and LEAD in BEER by ANODIC STRIPPING VOLTAMMETRY «Journal of the Wuxi Institute of Light Industry», Vol.3, No.3, 1984

KEYWORDS beer, copper lead determination, anodic stripping Voltammetry. ABSTRACT The results of the determination, obtained by Standard Addition method are fairly good, the detection range is from 10 to 360 ppb, Recovery is about 90—110%. The Variation coefficient of copper and lead is lower than $\pm 10\%$. The application is studied of Anodic stripping Voltammetry in the determination of traces of copper and lead in beer, in which the glassy carbon electrode is Used. Various working conditions such as deposition potential, deposition time and scanning rate are selected.

Author: Yu Hual Lu Jun

84022

THE MONTE CARLO SOLUTIONS for NINE-POINT SCHEME ELLIPTIC DIFFERENCE EQUATIONS «Journal of the Wuxi Institute of Light Industry», Vol.3, No.3, 1984

KEYWORDS nine-point scheme elliptic difference equations, Monte Carlo solutions, mean steps of random walk.

ABSTRACT In this paper, the Monte Carlo Solutions for the first, Second and third boundary value problems of nine-point scheme elliptic difference equations and estimation about its mean steps of random walk are given.

Author: Zhang Jianhua Fei Rongchang

84025

THE PRACTICE on DETERMINATION of P.D.I «Journal of the Wuxi Institute of Light Industry», Vol.3, No.3, 1984

KEYWORDS soybean, protein. index.

ABSTRACT The determination of the P.D.I. of soybean products plays an important rule in evaluating the degree of their denaturation and their quality. Recently, this kind of determination has been carried out by some institutes in our country. But the unification of instruments and consistency of methods are still lacking. A method for the determination of P.D.I. using instruments made in our country is introduced by this paper. The method approaches that designed by A.O.C.S. The data determined by this method is better in reproducibility. Thus, this method for the determination P.D.I. may be suggested as a conventional analytical method in oil mills or in laboratories.

Author: Zhang Daiyu Lu Jifon

84023

THE APPLICATION of MULTIPLE LINEAR REGRESSION MODEL and MICROCOMPUTERS to FORECAST the DEMAND of SPECIALIST «Journal of the Wuxi Institute of Light Industry», Vol.3, No.3, 1984

KEYWORDS personnel forecast, microcomputer, mathematical model.

ABSTRACT This paper explores the problem of forecasting the number of required specialists. A mathematical method has been studied for forecasting the demand of specialists using regression method and microcomputers. Four different Kinds of regression models are given. By using mathematical models we may forecast the number of Specialists required in future.

Author: Zhou Zhenjiang