

农业机械化 程序编制及应用

高焕文 主编

北京农业机械化学院

中央农业管理干部学院北京农业机械化分院

523

农业机械化程序编制及应用

主 编 高焕文

副主编 梁汉业

北京农业机械化学院
中央农业管理干部学院北京农业机械化分院

前 言

计算机在农业机械化领域中的应用有着广阔的前景。

国内外农机化发展的经验证明，实现生产与管理的科学化，为科学的预测与决策提供可靠的依据，革新数据处理及计划报表编制的手段，以及解算生产与科研上的各种难题等等，离开计算机的辅助都是难以成功的。

近年来，计算机在我国农机化行业中的应用虽然有所发展，但与世界先进水平相比，与国内其它先进行业相比，差距还不小。为了加速农机化事业的发展，在配备计算机的同时，大力开发农机化程序软件，培养计算机人才，特别是培训广大农机化工作者，使他们尽快掌握专业程序的编制及应用技术，已成为刻不容缓的一大任务。为此，我们编写了这本书，以期完成这项艰巨而具有重要意义的任务谨献菲薄之力。

《农业机械化程序编制及应用》一书，提供了22个农机化方面的实用程序，这些程序都是完整的、能独立运行的。读者可以根据本单位本部门的实际情况，直接应用这些程序或少许加以修改解决问题。除介绍程序及其应用外，本书对有关的专业理论、基础知识、数学模型和编程技巧也作了必要的阐述，以帮助读者获得改编和开发新程序的初步能力。在内容的选择上，既考虑到涉及农机化的各个方面，又尽可能包容着各种类型的编程，如计算与画图、单一程序与组合程序、线性方法与非线性方法、英文输出与汉字输出等。

根据目前农机化行业的计算机拥有及应用情况，本书中的程序均采用 Basic 语言，以在 PC-1500型和APPLE-Ⅱ型微型计算机上的使用为主。

本书可作为农机化管理、科研和生产部门的计算机程序库，适用作专业培训教材和农机、农林院校机械化专业的教材或参考书，同时可供具有中等文化程度的农机化工作者自学参考。

由于我国农机化行业应用计算机尚处在发展阶段，书中内容不可能是完善的，有待进一步补充和提高；加之编著及出版时间仓促，书中难免有缺点和错误，恳请读者批评指正。

本书由北京农业机械化学院农业机械化系和中央农业管理干部学院北京农业机械化分院联合出版，由北京农业机械化学院农业机器运用教研室高煥文任主编，梁汉业任副主编。参加编著的有高煥文（第一、二、三、六、九、十和十一章），梁汉业（第七、八章），董维寿（第四章），王耀发（第五章）和傅泽田（第十二章）。

本书由北京农业机械化学院学报编辑部负责编辑，责任编辑李兴昌。

目 录

第一篇 农机化数据处理

第一章 单一变量的统计处理	(1)
§1 抽样.....	(1)
§2 单一变量统计处理程序.....	(2)
一、数学模型.....	(2)
二、计算机程序.....	(2)
三、计算举例.....	(4)
§3 程序编制说明.....	(5)
练习.....	(6)
第二章 两变量回归及其预报与控制	(7)
§1 直线回归及其预报与控制.....	(7)
一、直线回归方程.....	(7)
二、相关检验.....	(9)
三、直线回归程序.....	(10)
四、直线回归举例.....	(12)
五、程序2-2的画图设计.....	(13)
六、预报与控制.....	(14)
练习.....	(15)
§2 两变量曲线回归.....	(16)
一、幂函数回归.....	(16)
二、指数函数回归.....	(19)
§3 多方案回归处理.....	(22)
练习.....	(24)
附表1 相关系数检验表.....	(25)
第三章 多元回归及其应用	(27)
§1 多元线性回归.....	(27)
一、基本思路与数学模型.....	(27)
二、相关检验.....	(28)
三、多元线性回归程序.....	(29)
四、应用举例.....	(32)
§2 多元非线性回归.....	(35)
一、多元多次回归.....	(35)
二、多项式回归.....	(36)

三、多元相乘幂函数回归·····	(38)
四、多元相乘幂函数回归程序·····	(39)
练习·····	(42)
附表 2 F 检验的临界值 (F_{α}) 表·····	(44)
第四章 农业机器技术状态测试数据处理 ·····	(48)
§1 发动机调速特性试验数据处理·····	(48)
一、计算功率、扭矩和油耗的数学模型·····	(48)
二、程序编制与使用·····	(50)
§2 拖拉机牵引特性试验数据处理·····	(52)
一、计算牵引特性的数学模型·····	(52)
二、拖拉机牵引特性计算程序·····	(52)
三、程序编制与使用·····	(54)
练习·····	(55)

第二篇 农业机器计划管理

第五章 农业机器作业成本计算 ·····	(57)
§1 机械作业成本计算·····	(57)
一、机械作业成本分项费用的计算·····	(57)
二、机械作业成本的计算程序·····	(62)
练习·····	(68)
§2 机械作业收费标准计算·····	(68)
一、机械作业成本的确定·····	(68)
二、盈利的确定·····	(69)
三、机械作业收费标准计算程序·····	(70)
练习·····	(73)
第六章 线性规划法与机械作业计划 ·····	(74)
§1 线性规划及单纯形法·····	(74)
一、线性规划的概念·····	(74)
二、单纯形法·····	(77)
三、单纯形法计算机程序·····	(78)
四、应用举例——运输计划的制定·····	(81)
§2 机械作业计划的制定·····	(82)
练习·····	(90)
第七章 油料计划和技术保养计划 ·····	(92)
§1 油料计划的制定·····	(92)
一、主燃油需要量的计算公式·····	(92)
二、程序编制·····	(93)
三、程序的操作步骤·····	(94)

§2 拖拉机技术保养计划的制定	(98)
一、年度技术保养计划的计算模型	(98)
二、程序编制	(99)
三、程序的操作步骤	(101)
练习	(102)
第八章 油库库存管理及农机管理报表的数据计算	(104)
§1 油库库存管理	(104)
一、程序的设计	(108)
二、程序的操作步骤	(110)
§2 农机管理报表的数据计算	(114)
一、程序的设计	(121)
二、程序的操作步骤	(122)
练习	(125)

第三篇 农业机械的配备与更新

第九章 线性规划配备法	(126)
§1 田间作业机器系统的配备准备	(128)
§2 数学模型的建立	(129)
一、约束方程组	(129)
二、目标函数	(132)
§3 模型中部分系数的计算	(132)
一、机器可作业班次数	(132)
二、拖拉机进行田间作业应分摊的年固定费	(133)
三、农具的年固定费	(133)
四、机组作业可变费用	(133)
§4 计算程序与计算结果	(134)
§5 单一机器的配备	(142)
一、设变量	(142)
二、建立约束方程	(143)
三、建立目标函数	(143)
四、计算结果及分析	(144)
练习	(145)
第十章 “机器-时间”系统配备法	(148)
§1 配备原理及数学模型	(148)
一、作业时间损失	(148)
二、作业总成本	(149)
三、经济功率计算	(150)
§2 自走式农业机械配备程序说明	(150)

一、程序框图·····	(150)
二、程序特点与使用说明·····	(150)
§3 自走式农业机器配备程序·····	(156)
§4 配备举例·····	(166)
练习·····	(170)
第十一章 农业机械的更新期计算·····	(171)
§1 合理更新期计算的基本原理·····	(171)
§2 大中型拖拉机更新期计算模型·····	(172)
§3 大中型拖拉机更新期计算程序·····	(174)
§4 计算举例·····	(176)
练习·····	(178)

第四篇 农业与农机专业户经营管理决策

第十二章 农业与农机专业户经营决策优化·····	(179)
§1 专业户机器选择和生产决策优化的基本思想·····	(179)
一、基本思想·····	(179)
二、计算模型的一般特点·····	(180)
三、优化模型及程序的具体特点·····	(180)
§2 数学模型·····	(181)
一、目标函数·····	(181)
二、约束条件·····	(182)
§3 农村专业户机器选择和生产决策优化的计算程序·····	(184)
一、程序的特点·····	(184)
二、程序使用说明·····	(185)
三、程序运行说明·····	(186)
四、程序的输出结果说明·····	(188)
五、专业户经营决策优化程序·····	(189)
§4 计算举例·····	(194)
§5 计算结果分析·····	(197)
练习·····	(199)
主要参考书目·····	(200)

第一篇 农机化数据处理

大量的农机化数据（各类机器数量、作业量、作业成本、耗油量、维修费、作业时间、地块大小、土壤比阻、作物产量、机械化程度……）为我们探寻农机化发展的内在规律，研究农机化与社会的、自然的各种因素之间的相互关系，预测未来的发展等一系列工作提供了丰富的材料和科学的依据，这是十分宝贵的。但是，要从繁多的数据中找到有用的东西，得到可靠的结论，还必须对数据进行认真的检验和整理，去粗取精、去伪存真、抓住事物的本质和规律。

因此，数据的收集和处理可以说是科学地解决农机化问题的基础，是管理工作实现量化、系统化的前提。但是用手工进行数据处理是十分费事和缓慢的，有时甚至完不成任务。而用计算机进行数据处理，则可明显地节省人力、物力，大大提高处理速度和结果的精度。为此，可以断定，利用计算机进行的数据处理将在农机化领域中发挥重要的作用。

本篇介绍农机化中常用的数据处理程序和所用的数学知识（主要是概率论与数理统计）。

第一章 单一变量的统计处理

第一节 抽 样

为了获得某一农机化参数的可靠数据，必须进行调查、测试或试验。如果能对研究对象进行全面调查和试验，当然十分可靠。但这常常是不可能的。如要知道某地区土壤的比阻，不可能对该地区所有地块所有面积上的土壤都进行比阻测试。又如拖拉机大修费，虽然有可能把数万台拖拉机的大修费收集起来，计算其平均值，但在人力、物力、财力和时间上的耗费是相当大的，很不合算。因而只有抽样调查才是一种切实可行的方法。

抽样时必须遵循随机取样的原则，即不能带有主观动机专门选取大值或选取小值。为了保证一定的代表性，在农机化部门中常常采用机械抽样的办法。例如，要了解某县600个大队的农机化程度，拟抽30个大队调查。可将全县大队编号为1, 2, 3, ..., 100, 101, ..., 599, 600，然后每隔20号抽取一个样，如大队的抽样调查编号为5, 25, 45, 65, ..., 205, ..., 585等。

单一变量统计处理以及后两章的回归处理，实际上都是对抽样的观测数据进行处理，用抽样的结果近似地描述所要知道的总体规律。

第二节 单一变量统计处理程序

一、数学模型

设为得到某县轻壤土的比阻，抽取了 n 个地块进行测试，测得比阻值分别为 $X_1, X_2, X_3, \dots, X_n$ 。它们的平均值 $\bar{X}$ ，方差 S^2 或标准差 S ，变异系数 V 的计算公式如下。

1. 平均值

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1-1)$$

一般用抽样的平均值代替总体平均值，它是描述总体特征的一个主要指示，对本例来说即代表该县轻壤土比阻。

2. 方差

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (1-2)$$

3. 标准差

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (1-3)$$

标准差表明观测变量对平均值的波动情况，它是描述总体规律的另一个重要指标。根据概率论的正态分布规律，可知变量中有：

68%的点落在值 $(\bar{X} \pm S)$ 之间

95%的点落在值 $(\bar{X} \pm 2S)$ 之间

99.7%的点落在值 $(\bar{X} \pm 3S)$ 之间

掌握这个规律，对变量预测非常有用。

4. 变异系数

$$V = \frac{S}{\bar{X}} \times 100\% \quad (1-4)$$

变异系数是观测变量波动的比例，由于它排除了平均值大小的影响，故便于用于不同参数间的比较。

为了使数据分布情况一目了然，可绘制数据分布直方图。

二、计算机程序

程序1-1是根据上述计算内容和绘图要求而编制的计算机程序。

程序1-1 单一变量统计处理 程序调用名①：“STAPRB”

①所谓程序调用名，是程序名的英文缩写，用作储存或调用程序的代号

```

10 : REM STATISTIC PROGRAM
20 : CLEAR
25 : INPUT "NUMBER=" ; N
30 : DIM X(N), Q(5)
35 : M=10E7
40 : FOR J=1 TO N
50 : INPUT "X(I)=" ; X(I)
65 : IF D<X(I) LET D=X(I)
68 : IF M>X(I) LET M=X(I)
70 : T=T+X(I)
80 : NEXT I
85 : D1=(D-M)/5 : D1=INT (D1 *
      1000+0.5)/1000
90 : X1=T/N
100 : FOR I=1 TO N
105 : IF X(I)>(4*D1+M) GOTO 145
110 : IF X(I)>(3*D1+M) GOTO 140
115 : IF X(I)>(2*D1+M) GOTO 135
120 : IF X(I)>D1+M GOTO 130
125 : C1=C1+1 : Q( 1 )=C1/N * 100 :
      GOTO 150
130 : C2=C2+1 : Q( 2 )=C2/N * 100 :
      GOTO 150
135 : C3=C3+1 : Q( 3 )=C3/N * 100 :
      GOTO 150
140 : C4=C4+1 : Q( 4 )=C4/N * 100 :
      GOTO 150
145 : C5=C5+1 : Q( 5 )=C5/N * 100
150 : P=(X(I)-X1) * (X(I)-X1)
155 : S=S+P
160 : NEXT I
165 : S1=S/(N-1)
170 : S2=√ S1
175 : V=S2/X1 * 100
180 : CSIZE 1, COLOR 0
185 : FOR I=1 TO N STEP 2
190 : LPRINT TAB 1;"X( ";I;" )=" ;
      X(I); TAB 18; "X ( " ;I+1;" )
      =" ; X(I+1)
195 : NEXT I
200 : CSIZE 2 : COLOR 3
205 : LPRINT
210 : LPRINT "X1="; X1
220 : LPRINT "S2=" ; INT (S1 * 1000
      +0.5)/1000
230 : LPRINT "S="; INT (S2 * 1000+
      0.5)/1000
240 : LPRINT "V="; INT(V*10+0.5)
      /10; "% "
245 : LPRINT
250 : GRAPH
260 : L=0
270 : GLCURSOR (10,-300) : SORGN
280 : LINE (0,0)-(200,300),0,1,B
290 : FOR J=1 TO 5
300 : Q=Q(J) : H=4 * Q : Q=INT (Q *
      10+0.5)/10
310 : GOSUB 500
320 : L=L+40
330 : NEXT J
335 : L=0
340 : FOR J=1 TO 5
350 : GOSUB 600
355 : L=L+40 : M=M+D1
360 : NEXT J
370 : GLCURSOR (0, -200)
380 : TEXT
390 : END
500 : LINE (L,0)-(L+40,H),0,3,B
520 : GLCURSOR (L,H+10) : COLOR
      0
525 : CSIZE 1
530 : LPRINT Q : "% "
540 : RETURN
600 : ROTATE 1
610 : GLCURSOR (L+15, -5)
615 : IF J=1LPRINT M; "-";M+D1:
      GOTO 630

```

620: LPRINT ">"; M; "-"; M+D1

640: RETURN

630: ROTATE 0

该程序运算的简单操作过程为：检查输入的计算机程序，确认无误后，变换计算机为运行状态。键盘键入“RUN”指令，计算机执行运算；计算机显示“NUMBER=?”，寻问要处理的数据有多少，操作者键入欲进行统计处理的数据个数；计算机显示“X(I)=”时，操作者依次键入各数据值，至此输入结束。计算机运算并打印出计算结果。

三、计算举例

【例1-1】为了掌握某地区轻壤土的比阻，根据随机抽样原则，共测得20个数据如表1-1。试求其平均值、方差、标准差和变异系数，并分五组画出百分比直方图。

表1-1 某地轻壤土比阻测定值（单位：公斤/厘米²）

0.37	0.36	0.35	0.40	0.38
0.39	0.23	0.43	0.34	0.38
0.34	0.42	0.25	0.30	0.33
0.41	0.29	0.35	0.28	0.34

解：调用程序1-1，输入表中20个数据（不管输入的先后次序）。计算机打印结果如下，绘制的直方图见图1-1。

X1=0.347
S2=0.003
S=0.055
V=15.8%

其中：X1——平均值
S2——方差
S——标准差
V——变异系数

从计算结果可知：（1）该地区轻壤土平均比阻为0.347公斤/厘米²。（2）方差0.003或标准差为0.055公斤/厘米²，说明大部分土壤比阻在0.347±0.055即0.292—0.402公斤/厘米²的范围内。（3）变异系数为15.8%。

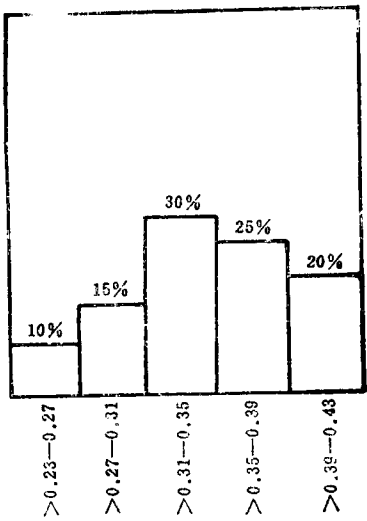


图1-1 计算机绘制的土壤比阻分布直方图

直方图反映了不同比阻地块的分布情况，比阻在0.31—0.35之间的地块最多，占总数的30%，其次为比阻在0.35—0.38之间的，占25%。

结果分析：根据计算结果，一般情况下用0.347作为该地区轻壤土的比阻即可。在要求更精确时，如进行拖拉机编组计算，则应考虑波动范围。满足68%的土地情况，比阻范围是0.292—0.402公斤/厘米²；满足95%情况，比阻为0.237—0.457公斤/厘米²。比较好的办法是结合本地具体情况选取，对本地土质偏硬或板结的情况取较大的比阻值，如0.40公斤/厘米²；

反之取较小值，以使编组的拖拉机动力既满足要求，又不致浪费。

【例1-2】对16台东方红-75型拖拉机机组的耕地生产率进行了测定，其结果如表1-2。试求其平均耕地生产率与变异情况，并绘制直方分布图。

表1-2 16台拖拉机的耕地生产率

车 号	耕 地 生 产 率 (亩/小时)	车 号	耕 地 生 产 率 (亩/小时)
1	8.8	9	10.5
2	12.8	10	11.2
3	11.2	11	10.0
4	8.0	12	11.5
5	12.1	13	10.2
6	11.9	14	9.3
7	11.3	15	8.9
8	11.2	16	13.0

解：用程序1-1计算，结果如下，直方图如图1-2。

$$\bar{X}=10.74 \text{ (亩/小时)}$$

$$S^2=2.113$$

$$S=1.454 \text{ (亩/小时)}$$

$$V=13.5\%$$

直方图表明，生产率在11—12亩/小时之间的拖拉机最多，占总数的37.5%；在8—9亩/小时和12—13亩/小时之间的次之，各为18.8%。

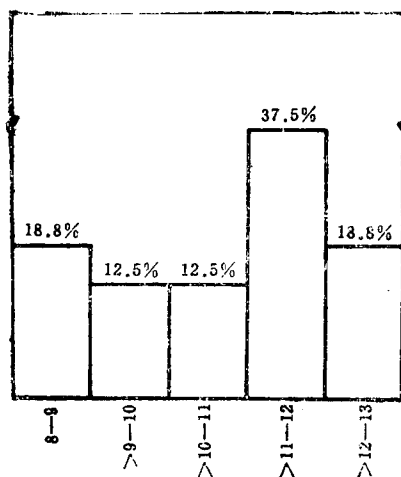


图1-2 东方红-75机组耕地生产率直方图

第三节 程序编制说明

本程序用Basic语言编制，可以在PC-1500袖珍型计算机，或略加修改后在APPLE-Ⅱ型微型计算机上运算。以后各程序也如此，不再一一说明。

该程序分两大部分，245句之前为计算部分，之后为绘图部分。

为了画出5个长方形和换成竖排打字，采用了两个子程序，分别从500句与600句开始。

为了使输出的结果醒目，采用了三种颜色和两种字体。

设计画图程序，重要的一点是解决好图形比例。本程序中处理图形比例的方法如下：

1. **画外框。**图形外框长宽比取1.5:1比较好看。因PC-1500计算机打印纸最大可用宽度为200绘图单位(五个绘图单位相当于一毫米)，故取高度为300绘图单位。程序中第280句为画外框。

2. **画小直方。**共五个小直方，直方的横向宽度为观测值范围，纵向高度为该范围内观测变量的个数占总观测变量数的百分比。

每一小直方宽40绘图单位，坐标位置L由320句给出： $L = L + 40$ 。

小直方纵向高度先由105—145句算出各直方中观测个数所占百分比 $Q(I)$ ，再由300句化为图形高： $H = 4 * Q(I)$ ，即每百分之一比例占4个绘图单位，小直方高度为H个绘图单位。

3. **小直方标注。**小直方的上标注是由105—145句算出的百分比 $Q(I)$ 。小直方下标注，即竖向打印的标注D1，由85句算出：

$$D1 = (D - M) / 5$$

其中：D为变量X的最大值，由65句求出：

M为变量X的最小值，由68句求出。

综 习

1. 试根据表1-3列出的东北某地39个大队的农机亩投资数据，计算其平均值、方差、标准差与变异系数，并画出分布直方图。

表1-3 东北某地30个大队的农机亩投资 (单位：元/亩)

33	48	32	52	34	55	20	43	34	23
32	23	51	35	37	17	32	47	17	24
30	25	19	14	35	18	39	20	34	33

2. 参考单一变量统计处理程序1-1，设计一个只计算出结果而不画直方图的程序，并计算上题。

3. 按照你认为理想的形式，变换程序1-1中的字体大小及字和图的颜色，并计算例题1-2。

第二章 两变量回归及其预报与控制

在农机化领域中，同时出现的几个变量，往往不是各自孤立地变化，而是相互联系的。变量之间相互联系的形式可分为两大类：一类是确定型关系，亦称函数关系，如发动机功率 N 与转速 n 和扭矩 M 的关系 $N=Mn/716.2$ ， N 与 M 和 n 成正比关系。另一类是非确定型关系，如拖拉机使用期与年维修费的关系，联合收获机价格与马力的关系，农机化程度与作物亩产量的关系等。拖拉机的年维修费与使用年限关系很大；但二者的关系却是非确定型的，不能说使用年限一定，维修费也确定了。事实上即使机型和使用时间都相同，但其年维修费也各不相同。这类关系亦称相关关系。

变量之间的相关关系只有在大量观测试验的基础上，对测试数据进行回归处理才能获得。所谓回归处理，实质上是寻求相应的函数关系来近似表述相关关系，亦即寻求回归函数。

回归处理大体可分为线性回归和非线性回归两类。对两个变量来说，就是直线回归与曲线回归。

第一节 直线回归及其预报与控制

一、直线回归方程

【例2-1】表2-1为某省1960—1980年农业机械化程度与粮食亩产量的统计数字。试求该省粮食亩产与农机化程度的关系。

表2-1 农业机械化程度与粮食亩产量

序 号	年 份	机械化程度(%)	亩产量(斤/亩)
1	1960—1961	4.70	117.75
2	1962—1963	6.05	132.2
3	1964—1965	8.30	162.1
4	1966—1967	8.25	162.35
5	1968—1969	9.10	166.65
6	1970—1971	11.40	203.8
7	1972—1973	15.15	207.95
8	1974—1975	19.70	266.7
9	1976—1977	23.95	261.75
10	1978—1979	27.15	275.95
11	1980—1981	27.31	272.0

为此,可设机械化程度为变量 X ,粮食亩产量为变量 Y ,将这批数据画在直角坐标上,如数据点近似成直线分布(图2-1),则预计两变量的相关关系可能是直线回归关系。此时,可在散点密集处画一条直线近似地代表两变量的相关关系。

直线方程为: $\hat{Y} = A + BX$

当然,直线上点的坐标与散点坐标不尽相同。设与变量 X_i 对应的变量实际观测值为 Y_i ,而直线点的对应值是 $\hat{Y}_i$,则二者之差 D_i (见图2-1)可写为:

$$D_i = Y_i - \hat{Y}_i = Y_i - (A + BX_i)$$

令偏差 D_i 的平方和为 Q ,则

$$Q = \sum_{i=1}^n D_i^2 = \sum_{i=1}^n (Y_i - A - BX_i)^2 \quad (2-1)$$

显然,我们为一组散点选择的直线应是可能选择的各条直线中 Q 值最小的那一条直线,即数学上所称“最小平方直线”。

根据数学极值原理,“最小平方直线”的 A , B 值,可通过对 Q 取偏导数并令其为零求得。

$$\frac{\partial Q}{\partial A} = -2 \sum_{i=1}^n (Y_i - A - BX_i) = 0 \quad (2-2)$$

$$\frac{\partial Q}{\partial B} = -2 \sum_{i=1}^n X_i (Y_i - A - BX_i) = 0 \quad (2-3)$$

解式(2-2), (2-3)得:

$$B = \frac{\sum_{i=1}^n X_i Y_i - \frac{1}{n} \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2} \quad (2-4)$$

$$A = \bar{Y} - B \bar{X} \quad (2-5)$$

式中: $\bar{Y}$ —— Y_i 值的算术平均值;

$\bar{X}$ —— X_i 值的算术平均值;

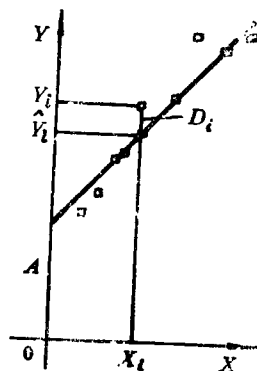


图2-1 农业机械化程度与粮食亩产量散点图

$$Y = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$X = \frac{1}{n} \sum_{i=1}^n X_i$$

A——回归常数；

B——回归系数。

由A、B构造的直线方程 $\hat{Y} = A + BX$ 称为Y对X的直线回归方程。

用式(2-4)，(2-5)对例1计算，结果得：

$B = 6.675$ 斤/亩 (1%农机化程度，即农机化程度提高1%，粮食亩产增加6.675斤)

$A = 104.949$ 斤/亩

该省粮食亩产与农机化程度的直线回归方程为：

$$\hat{Y} = 104.949 + 6.675X$$

二、相关检验

回归所得直线方程与实际观测数据散点的密切程度如何，或者说两变量之间的相关关系是不是直线关系，还需进一步检验。目前一般用相关系数R来完成检验工作。

$$R = \frac{\sum_{i=1}^n X_i Y_i - \frac{1}{n} \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sqrt{\left(\sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2 \right) \left(\sum_{i=1}^n Y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n Y_i \right)^2 \right)}} \quad (2-6)$$

R值的大小在 ± 1 之间，即 $-1 \leq R \leq 1$ 。按其大小不同，又分三种情况：

1. $R=0$ ，对照式(2-3)与式(2-4)，不难得出 $B=0$ ，回归直线平行于X轴，X的变化与Y无关，即X与Y之间没有相关关系。

2. $R=1$ ，所有的散点都在回归直线上，变量Y与X完全直线相关，或具有确定型直线函数关系。

3. $0 < |R| < 1$ ，这是绝大多数的回归情形，X与Y之间存在着一一定的直线相关关系。 $|R|$ 愈接近1，相关愈密切。反之， $|R|$ 愈接近0，相关程度愈低。

R的绝对值究竟要多大，才认为两变量间符合线性的相关关系，或叫线性相关显著呢？这与我们对置信度的要求有关：不同的置信度，要求显著相关的R值不一样。同时，测点数的多少，对R的显著值也有影响。附表1给出了不同自由度f ($f=n-2$ ，n为测点数)，不同的置信度水平 α ($\alpha=1-\text{置信度}$)时的显著相关最小R值，此值亦称显著值 R_r 。当回归方程的R值等于或大于 R_r 时，相关显著。反之，若R小于 R_r ，两变量相关不显著。

现仍计算例1，按式(2-6)算得 $R=0.966$ 。例中两个变量有11组观测数据，试检验置信度为99%时，两变量的直线相关显著性如何？

当 $f=11-2=9$ ， $\alpha=1-0.99=0.01$ 时，查附表1得：相关系数显著值 $R_r=0.7348$ 。

$\therefore R=0.966 > R_r=0.7348$

$\therefore$ 在99%置信度时，两变量直线相关显著。

三、直线回归程序

程序2-1是一个计算回归常数A、回归系数B和相关系数R，并列出行回归方程的计算机程序。

程序2-1 直线回归 “SREG”

```

10: REM STRAIGHT
    REGRESSION
20: CLEAR
30: CSIZE 1: COLOR 0
40: INPUT "NUMBER? "; N
50: DIM X(N), Y(N)
60: FOR I=1 TO N
70: INPUT "X(I)= "; X(I),
    "Y(I)= "; Y(I)
90: R=R+X(I)*Y(I)
100: S=S+X(I)*X(I)
110: V=V+Y(I)*Y(I)
120: T=T+X(I)
130: U=U+Y(I)
140: NEXT I
150: B= (N*R-T*U) / (N*S-T*T)
170: A= (U-B*T) / N
190: R= (N*R-T*U) / √ ((N*S-
    T*T) * (N*V-U*U))
200: R=INT (R*1000+0.5) /1000
215: A=INT (A*1000+0.5) /1000
225: B=INT (B*1000+0.5) /1000
230: FOR I=1 TO N
240: LPRINT TAB 1; "X (" ; I; ")
    =" ; X (I) ; TAB 18; "Y(" ;
    I; ")=" ; Y (I)
250: NEXT I
255: CSIZE 2: COLOR 3
260: LPRINT
270: LPRINT "A=" ; A
280: LPRINT "B=" ; B
290: LPRINT "R=" ; R
295: LPRINT
300: IFB<0 LPRINT "Y=" ; A; B;

```

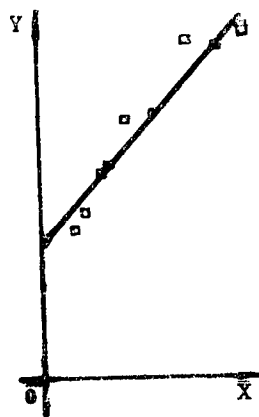
```

"X" : GOTO 310
302: LPRINT "Y=" ; A; "+" ; B;
    "X"
310: END

```

如果想更直观地看到回归密切程度，可采用程序2-2 “SREG-1” 画出变量数据散点图与回归直线。

对例1-1采用程序2-2处理的输出结果如图2-2所示。



$A = 104.949$
 $B = 6.675$
 $R = 0.966$
 $Y = 104.949 + 6.675X$

图2-2 直线回归加画散点图与回归直线

程序2-2 直线回归带画图 “SREG-1”

```

10: REM STRAIGHT
    REGRESSION AND PLOT
20: CLEAR
30: CSIZE 1: COLOR 0
40: INPUT "NUMBER? "; N

```