高等学校教学用书

计算机地质制图及应用

JISUANJI DIZHI ZHITU JI YINGYONG

吕国祥 张 瀛 编著

电子科技大学出版社

高等学校教学用书

计算机地质制图及应用

JISUANJI DIZHI ZHITU JI YINGYONG

吕国祥 张 瀛 编著

电子科技大学出版社

图书在版编目（CIP）数据

计算机地质制图及应用/吕国祥等编著. —成都：电子科技大学出版社，2009.10

ISBN 978-7-5647-0204-5

Ⅰ. 计… Ⅱ. 吕… Ⅲ. 地质图—计算机制图—应用软件 Ⅳ. P285.1-39

中国版本图书馆 CIP 数据核字（2009）第 183608 号

高等学校教学用书

计算机地质制图及应用

吕国祥　张　瀛　编　著

出　　版：	电子科技大学出版社（成都市一环路东一段 159 号电子信息产业大厦　邮编：610051）
策划编辑：	万晓桐
责任编辑：	万晓桐
主　　页：	www.uestcp.com.cn
电子邮箱：	uestcp@uestcp.com.cn
发　　行：	新华书店经销
印　　刷：	中国核动力研究设计院印刷厂
成品尺寸：	185mm×260mm　　印张 14.5　　字数 353 千字
版　　次：	2009 年 10 月第一版
印　　次：	2009 年 10 月第一次印刷
书　　号：	ISBN 978-7-5647-0204-5
定　　价：	32.00 元

◆ 本社发行部电话：028-83202463；邮购部电话：028-83208003。

◆ 本书如有缺页、破损、装订错误，请寄回印刷厂调换。

前　　言

当前，石油地质学研究正不断地走向定量化、精细化、精准化，需要更清晰而准确的图件去展示、去描绘我们所研究的地质现象，地质绘图工作在整个研究过程中是极其重要的，所以石油高校的地质专业都开设了计算机制图课程。

在石油地质研究工作中，绘制地质图件的工作往往烦琐而工作量巨大，如果软件选择的不恰当或者操作不熟练，往往使绘制工作苦不堪言。所以熟练的掌握各类地质绘图软件，对于石油地质工作者有巨大的帮助。

本书中，笔者优选了当前石油地质研究工作中常用和流行的几种绘图软件：Surfer 8.0，GeoMap 3.2，Gxplorer，ResForm，通过深入浅出地讲解这些软件的用法，让大家掌握地质绘图的基本方法，提高工作效率。第 1 章和第 2 章则是介绍与其相关的一些基础知识。

本书既可作为石油高校相关专业的教科书，也可作为石油地质工作者的参考书。

本书共有 6 个章节组成，包括：

第 1 章　相关分析

第 2 章　回归分析

第 3 章　Surfer 8.0 绘图软件的使用

第 4 章　Geomap 绘图软件的使用

第 5 章　Gxplorer 软件的使用

第 6 章　ResForm 软件的使用

限于作者的工作经验和学术水平，本书的缺点和错误在所难免，望读者斧正。

编著者

2009 年 8 月

目　　录

第1章 相关分析

在生产实际和科学研究中，我们经常要进行统计分析工作，常常会遇到各种不同的变量，但归纳起来这些变量分为数学变量和随机变量两类，而变量之间的关系分为确定性关系和相关关系。

1．确定性关系

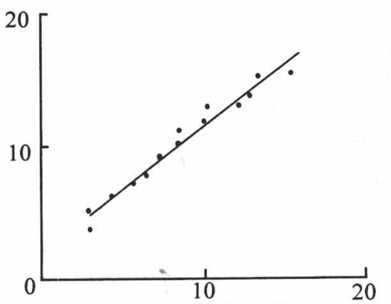
确定性关系也称为函数关系，是指变量之间存在着一种严格的对应关系，当一种现象确定时，相联系的另一种现象会随之确定，把这种关系用函数 $y=f(x)$ 表示，其中 x 称为自变量，y 称为因变量。如圆的面积与半径之间的关系：$s=\pi r^2$ 以及匀速直线运动的距离同速度和时间的关系：$L=vt$ 等，具有这种关系的变量称为数学变量。

2．相关关系

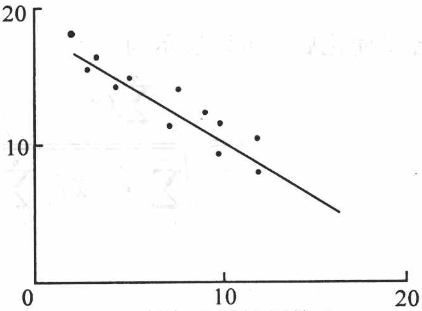
相关关系是指变量之间确实存在着关系，但不是严格对应的依存关系，而是一种不确定的依存关系，当一种现象发生变化时，会引起另一种现象的变化，当一种现象确定时，另一种现象不会随之完全确定。如居民的收入和储蓄之间的关系，岩石的孔隙度和渗透率之间的关系等，这些变量是相互联系和相互制约的，它们之间存在着一定的关系，称为相关关系。具有相关关系的变量称为随机变量。

1.1 相关关系的确定

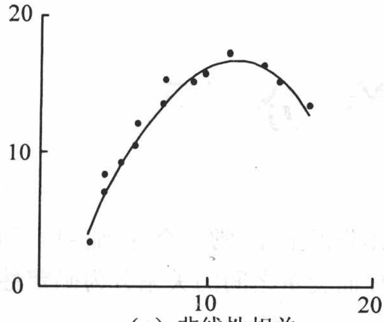
变量之间是否存在着相关关系？关系是否密切？是线性相关还是非线性相关？要回答这个问题，最简单的办法是绘制变量散点分布图。设 $x_i, y_i (i=1,2,\cdots,n)$ 是从总体中抽取的一个样本，以 x 为横坐标，y 为纵坐标，将 $x_i, y_i (i=1,2,\cdots,n)$ 绘制在直角坐标系中，由散点分布图大致可以看出相关关系的形式、密切程度和方向。图 1.1 是几种常见的散点图的形式。



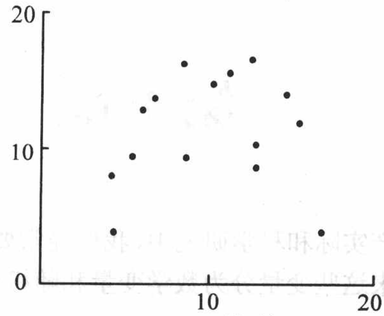
（a）正线性相关



（b）负线性相关



（c）非线性相关



（d）不相关

图 1-1

当 y 随着 x 的增大而增大称为正相关，当 y 随着 x 的增大而减小称为负相关，点越是分布在直线附近关系越密切。

1.2 相关系数的计算

由散点分布图大致可以看出相关关系的形式、密切程度和方向，当不能定量表示变量之间的相关关系，相关系数是确切表示变量之间相关关系密切程度的指标。最常用的相关系数是英国统计学家卡尔·皮尔逊（Karl Pearson）提出的相关系数，该相关系数是在线性相关条件下衡量两个变量之间相关关系密切程度的指标。公式如下

$$r=\frac{s_{xy}}{s_x s_y}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2\cdot\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2}} \tag{1.1}$$

式中 s_{xy} 表示变量 x 和 y 的样本协方差；s_x 和 s_y 表示变量 x 和变量 y 样本标准差，$\overline{x}$ 和 $\overline{y}$ 表示变量 x 和变量 y 的样本平均值。其中

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\quad,\quad\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

上式经过简化后还可以表示为

$$r=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\cdot\sum_{i=1}^{n}(y_i-\overline{y})^2}}=\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}=\frac{L_{xy}}{L_x L_y} \tag{1.2}$$

其中

$$L_{xy}=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}),\quad L_x=\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\quad,\quad L_y=\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}$$

（1.2）式进一步可以简化为

$$r = \frac{L_{xy}}{L_x L_y} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sqrt{(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2)\cdot(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2)}} \tag{1.3}$$

相关系数有如下特点：

①相关系数的取值范围在–1～+1之间，即 $-1 \leqslant r \leqslant 1$。

②当 $r>0$ 时，表明变量之间呈正相关，当 $r<0$ 时，表明变量之间呈负相关。

③相关系数的绝对值越接近于1，说明两个变量之间的相关关系越强，越接近于0说明相关关系越弱，当 $|r|=1$ 时，说明两个变量之间的关系属于确定性关系，当 $|r|=0$ 时说明两个变量之间完全没有线性相关关系，但并不说明两个变量之间不存在其他非线性相关关系。

1.3 相关系数的显著性检验

相关系数虽然可以反映两个变量之间关系的密切程度，但由于相关系数是由样本资料出发计算得出的，同一总体的不同样本可以算出不同的相关系数，到底哪一个能代表总体的相关程度呢？因此有必要对相关系数的显著性进行检验，常用的有以下两种检验法：

1．相关系数检验法

设总体的相关系数为 ρ，检验相关系数是否显著实际上是检验假设 $H_0:\rho=0$ 是否成立。给定信度（检验水平）α，根据自由度为 $f=n-2$ 查相关系数表得 $r_\alpha(n-2)$，当 $|r|>r_\alpha(n-2)$ 时拒绝原假设，认为相关系数显著，否则接受原假设，认为相关系数不显著。

2．t 检验法

设总体的相关系数为 ρ，检验假设 $H_0:\rho=0$，可以证明，当假设 $H_0:\rho=0$ 成立时，统计量

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t(n-2)$$

即服从自由度为 $n-2$ 的 t 分布。所以给定信度 α，根据自由度为 $f=n-2$ 查 t 分布表得 $t_{\alpha/2}(n-2)$，当 $|t|>t_{\alpha/2}(n-2)$ 时拒绝原假设，认为相关系数显著，否则认为相关系数不显著。

例1. 从某碳酸盐储层中取得20个样品，测得孔隙度（Φ%）和渗透率（10^{-3}MD）数据如表1-1所示，试计算渗透率和孔隙度之间的相关系数并进行显著性检验。

表 1-1

Φ	2.55	3.07	2.92	10.06	11.21	11.93	12	12.93	12.68	10.42
k	0.012	0.002	0.001 5	0.107	0.247	0.317	0.241	0.312	0.443	0.138
Φ	14.96	13.9	10.91	9.1	7.23	6.88	9.16	4.86	3.73	4.21
k	0.62	0.51	0.143	0.145	0.169	0.247	0.307	0.0346	0.0007	0.0008

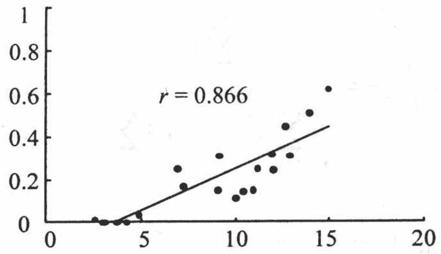


图 1-1

解：从散点图可以看出渗透率随孔隙度的增大而增大，呈正相关分布，根据计算相关系数的公式

$$r=\frac{L_{xy}}{L_xL_y}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\sqrt{(\sum_{i=1}^{n}x_i^2-n\overline{x}^2)\cdot(\sum_{i=1}^{n}y_i^2-n\overline{y}^2)}}$$

1．首先算出两个变量的平均值

$$\overline{\varphi}=\frac{1}{20}\sum_{i=1}^{20}\varphi_i=8.735\ 5 \qquad \overline{k}=\frac{1}{20}\sum_{i=1}^{20}k_i=0.1998\ 8$$

然后算出

$$L_{\varphi k}=\sum_{i=1}^{20}\varphi_ik_i-20\overline{\varphi}\overline{k}=11.879\ 86 \qquad L_{\varphi}=\sqrt{\sum_{i=1}^{20}\varphi_i^2-20\overline{\varphi}^2}=17.482\ 28$$

$$L_k=\sqrt{\sum_{i=1}^{20}k_i^2-20\overline{k}^2}=0.784\ 7$$

相关系数 r 为

$$r=\frac{L_{\varphi k}}{L_{\varphi}L_k}=\frac{11.879\ 86}{17.482\ 28\times0.784\ 7}=0.866$$

2．对相关系数进行显著性检验

给定信度 $\alpha=0.05$，对于自由度 $f=n-2=18$ 查相关系数检验表得

$$r_{0.05}(18)=0.443\ 8，$$

而算出的相关系数 $r=0.866>r_{0.05}(18)=0.443\ 8$，所以可以认为孔隙度与渗透率之间线性相关是显著的。

1.4 利用 Excel 进行相关分析

在 Microsoft Excel 中，可以利用数据分析宏的相关功能进行相关分析并算出相关系数，具体步骤如下：

（1）启动 Excel，录入数据。打开“工具”菜单，点击“数据分析”宏中的“相关系数”，如图 1-2 所示。

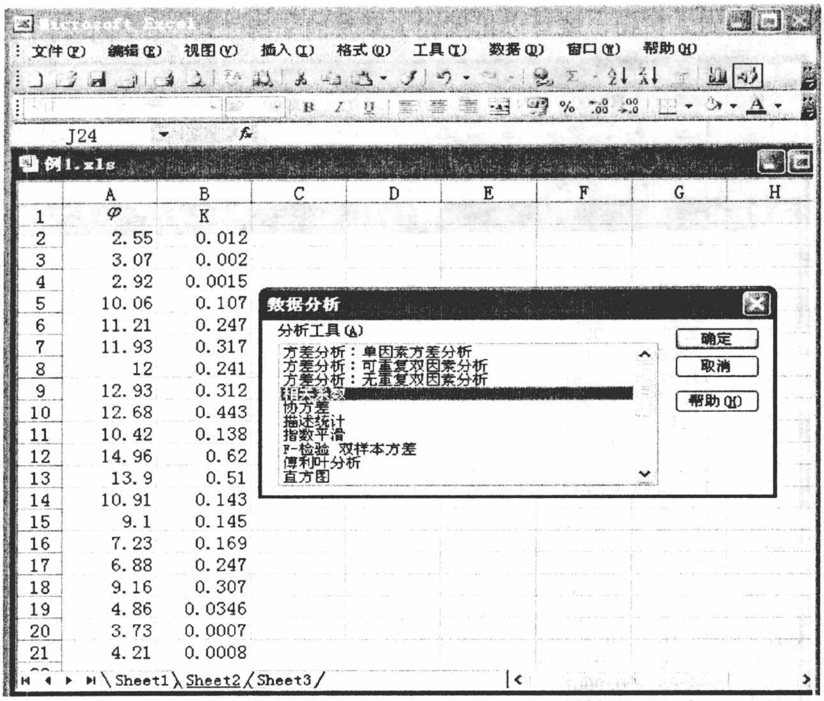


图 1-2

（2）单击“确定”按钮后，在“输入区域”中输入数据所在单元格区域“A1:B21”，选择“标志位于第一行”，输入“输出区域”为 D1 单元格，如图 1-3 所示。

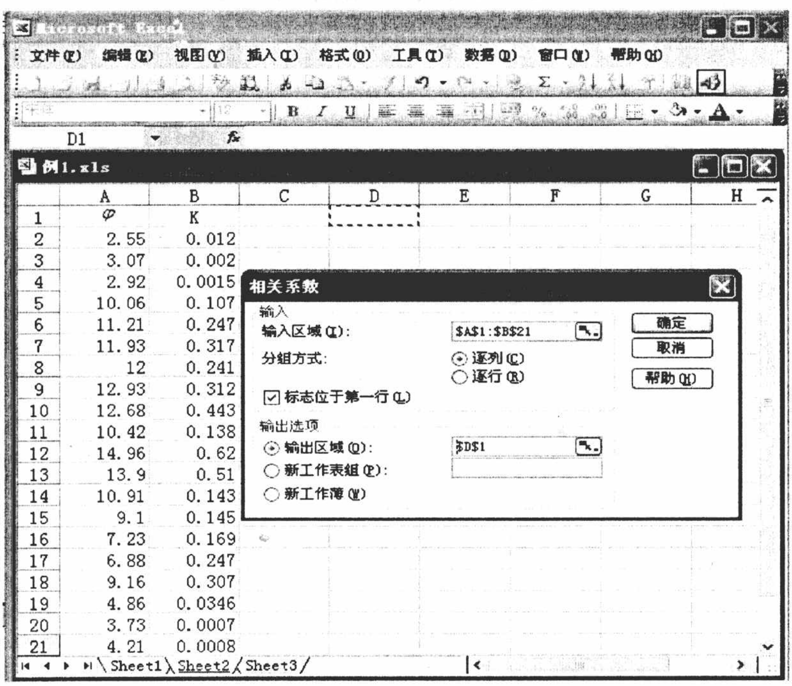
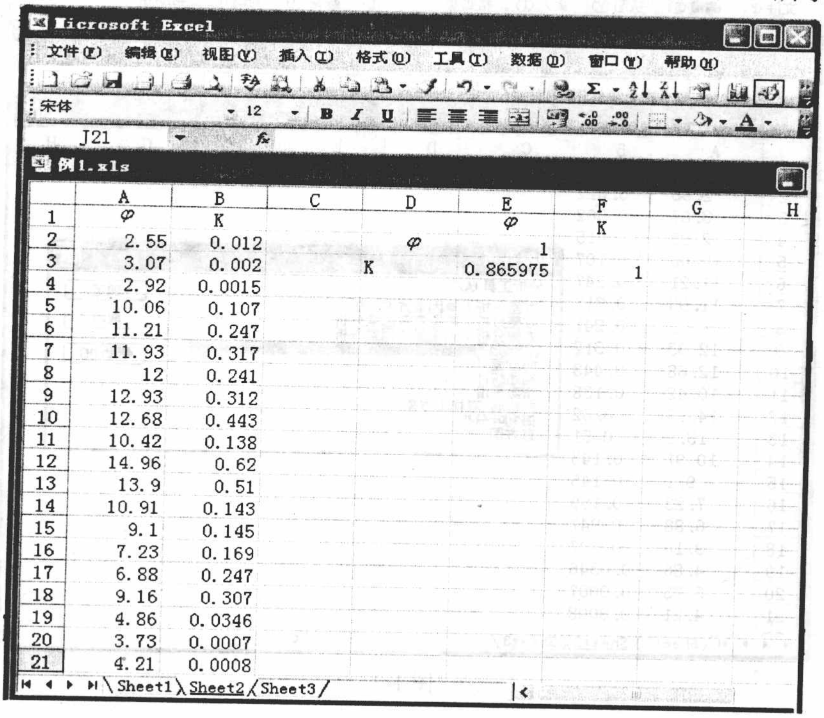


图 1-3

（3）单击“确定”按钮，就可以得到相关系数的分析结果，如图 1-4 所示。



	A	B	C	D	E	F	G	H
1	φ	K			φ	K		
2	2.55	0.012		φ	1			
3	3.07	0.002		K	0.865975	1		
4	2.92	0.0015						
5	10.06	0.107						
6	11.21	0.247						
7	11.93	0.317						
8	12	0.241						
9	12.93	0.312						
10	12.68	0.443						
11	10.42	0.138						
12	14.96	0.62						
13	13.9	0.51						
14	10.91	0.143						
15	9.1	0.145						
16	7.23	0.169						
17	6.88	0.247						
18	9.16	0.307						
19	4.86	0.0346						
20	3.73	0.0007						
21	4.21	0.0008						

图 1-4

所得相关系数为 $r = 0.866$，这与用公式算出的结果完全一致。

第2章 回归分析

相关分析是研究两个变量之间关系密切程度及方向的一种方法，其结果用相关系数r表示，回归分析也是研究变量之间的相关关系的一种统计分析方法。

回归分析的主要任务是在大量试验与观察数据的基础上，进行分析研究，以找出它们之间的内在规律。以后，我们把这种变量之间的相关关系称为回归关系。有关回归关系的计算方法和理论称为回归分析。回归分析的内容很多，主要解决以下几个方面的问题：

1．对于具有相关关系的变量，找出它们之间的数学表达式。

2．根据一个或几个相对而言较易测定或控制的变量值，来预测或控制另一个变量的取值，并确定这种预测的精度。

3．在共同影响某个特定变量的许多变量（因素）之间找出哪些是主要因素，哪些是次要因素，以及这些因素之间有什么关系，从而提供解决问题的方法。

2.1 一元线性回归

2.1.1 回归系数的确定

设$x_i, y_i (i=1,2,\cdots,n)$是从总体中抽取的一个样本，称为观测值，若它们之间存在着线性关系，则线性表达式为

$$\hat{y} = a + bx \tag{1.1}$$

式中a、b是待确定的参数，称为待定系数。在方程（1.1）中，任给一组数a、b，便可得到平面上的一条直线，当a、b取各种可能的值时，便可得到许许多多的直线，但究竟用哪一条直线来表示它们之间的关系最好呢？这就需要确定一个标准，一个常用的标准就是“最小二乘”原理。

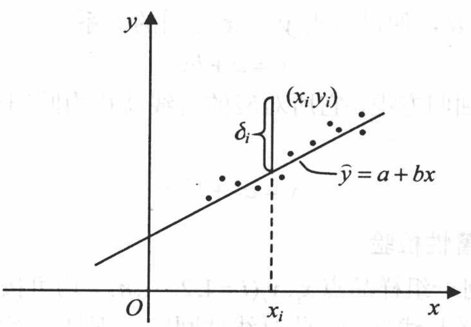


图 2-1

用$x_i, y_i (i=1,2,\cdots,n)$表示第$i$个样品点，如果用（1.1）式表示这$n$个点之间的线性关系，则将$x_i$代入（1.1）式得

$$\hat{y}_i = a + bx_i \qquad (i=1,2,\cdots,n)$$

$\hat{y}_i$称为回归值（或计算值），它与观测值y_i之间的误差用$\delta_i = y_i - \hat{y}_i$表示。所谓“最小二乘”原理就是要使误差平方和

$$Q = \sum_{i=1}^{n}\delta_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - a - bx_i)^2$$

达到最小，根据极值原理可知，只需将Q分别对a,b求偏导数并令其为 0，即令

$$\frac{\partial Q}{\partial a} = -2\sum_{i=1}^{n}(y_i - a - bx_i) = 0 \tag{1.2}$$

$$\frac{\partial Q}{\partial b} = -2\sum_{i=1}^{n}(y_i - a - bx_i)x_i = 0 \tag{1.3}$$

由式（1.2）可以解出

$$a = \overline{y} - b\overline{x}$$

其中
$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i \qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n}y_i$$

将式（1.2）两边同乘以（$-\overline{x}$）再与式（1.3）相加得

$$\sum_{i=1}^{n}(y_i - a - bx_i)(x_i - \overline{x}) = 0 \tag{1.4}$$

再将$a = \overline{y} - b\overline{x}$代入上式整理后可得

$$b = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{L_{xy}}{L_{xx}} \tag{1.5}$$

式中

$$L_{xy} = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n}x_iy_i - n\overline{xy}$$

$$L_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}x_i^2 - n\overline{x}^2$$

再由$a = \overline{y} - b\overline{x}$可求出$a$，便可得出$y$与$x$之间的关系

$$\hat{y} = a + bx$$

上式就称为y对x的回归方程，它所对应的直线就称为回归直线，同理，我们也可求出x对y的回归方程

$$\hat{x} = a^* + b^*y$$

2.1.2 回归方程的显著性检验

由上面可看出，对任何一组样品点$x_i, y_i (i = 1,2,\cdots,n)$，均可按最小二乘原理配一条直线。如果变量$y$与$x$之间的关系是线性的，此时线性回归方程是显著的；反之，如果变量y与x之间不是线性关系，我们却按线性关系予以处理，此时线性回归方程就无多大意义。因此，必须对回归方程的显著性进行检验。下面介绍回归方程的显著性检验方法。

1．相关系数检验法

首先计算总的离差平方和

$$L_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}[(y_i-\hat{y}_i)+(\hat{y}_i-\overline{y})]^2$$

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

而

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\overline{y})=\sum_{i=1}^{n}(y_i-a-bx_i)(a+bx_i-a-b\overline{x})$$

$$=b\sum_{i=1}^{n}(y_i-a-bx_i)(x_i-\overline{x})$$

由式（1.4）知

$$\sum_{i=1}^{n}(y_i-a-bx_i)(x_i-\overline{x})=0$$

故

$$L_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2 \tag{1.6}$$

称 $L_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}y_i^2-n\overline{y}^2$ 为总的离差平方和。$Q=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 为偏差平方和。$U=\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2$ 为回归平方和。

于是（1.6）式便可写成

$$L_{yy}=U+Q \tag{1.7}$$

在上式中，L_{yy} 反映了观测值总的波动情况；U 是回归值与平均值之差的平方和，它反映了变量 x 与变量 y 关系的密切程度。Q 是观测值与回归值之差的平方和，它反映了观测值与回归值之间的偏差程度，我们也称它为剩余平方和。

设 U 与 L_{yy} 的比值为 $r^2=\dfrac{U}{L_{yy}}$，将其开平方后得

$$r=\sqrt{\frac{U}{L_{yy}}} \tag{1.8}$$

我们称 r 为回归方程的复相关系数，给定信度 α，查相关系数表得 $r_\alpha(n-2)=r_\alpha$，若 $|r|>r_\alpha$，认为回归方程显著，否则就认为回归方程不显著。

复相关系数 r 也可以用下式计算

$$r=\sqrt{\frac{U}{L_{yy}}}=\sqrt{\frac{L_{yy}-Q}{L_{yy}}}=\sqrt{1-\frac{Q}{L_{yy}}} \tag{1.9}$$

2. F 检验法

检验回归方程是否显著实际上是检验假设 $H_0:b=0$ 是否成立。可以证明，当假设 $H_0:b=0$ 成立时，统计量

$$F=\frac{U}{Q/(n-2)} \tag{1.10}$$

服从第一自由度为 1、第二自由度为 $n-2$ 的 F 分布。这实际上就是方差分析。对于给定的信度 α，查 F 分布表可得相应的临界值 $F_\alpha(1,n-2)=F_\alpha$，如果 $F>F_\alpha$，则拒绝原假设，即认为回归方程显著，如果 $F<F_\alpha$，则接受原假设，即认为回归方程不显著。由此列出方差分析表如 2-1 所示：

表 2-1

方差来源	平　方　和	自由度	平均平方和	F 值	显著性
回 归	$U=b^2L_{xx}$	1	$U/1$	$F=\frac{U}{Q/n-2}$	
剩 余	$Q=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$	$n-2$	$Q/(n-2)$		
总　和	$L_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2$	$n-1$			

从（1.10）我们容易得出统计量 F 与复相关系数 r 之间的关系：

$$F=\frac{(n-2)U}{Q}=\frac{(n-2)U}{L_{yy}-U}=\frac{(n-2)U/L_{yy}}{1-U/L_{yy}}=\frac{(n-2)r^2}{1-r^2}$$

2.1.3　用回归方程进行预测

我们前面已解决了如何检验一个回归方程显著性的问题。如果回归方程是显著的，这就在一定程度上反映了两个相关变量之间的内在规律，于是我们就可以根据变量 x 的取值来预测或控制 y 的取值。我们自然要问：用回归方程来预测 y 其精度如何？为了研究预测的可靠程度，我们采用类似于区间估计的方法，假定随机变量 y 服从正态分布 $N(\hat{y},\sigma^2)$，由正态分布的性质可知，对于任一固定的 x_i、y_i 以 95%的概率落在区间 $(\hat{y}_i-1.96\sigma,\hat{y}_i+1.96\sigma)$ 之内。

σ 的估计值为 $\hat{\sigma}=\sqrt{\frac{Q}{n-2}}$，可以在回归直线 $\hat{y}=a+bx$ 的上下作二平行线：

$$L_1:\quad y'=a+bx-1.96\hat{\sigma}$$
$$L_2:\quad y''=a+bx+1.96\hat{\sigma}$$

可以认为，在全部可能出现的观测值 y_i（1, 2, ..., n）中，大约有 95%的点落在这两条直线所夹的范围内（如图 2-2 所示）。

因此，为了预测 x 在范围（x_1,x_2）内时相应的 y 值在什么范围内，可如图 2-3 作出平行线来找到 y 的相应范围（y_1,y_2）；反之，要控制 y 在范围（y_1,y_2）内，我们可事先将影响 y 的因素 x 控制在范围（x_1,x_2）内。

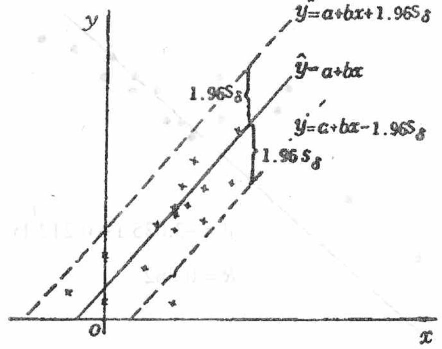


图 2-2

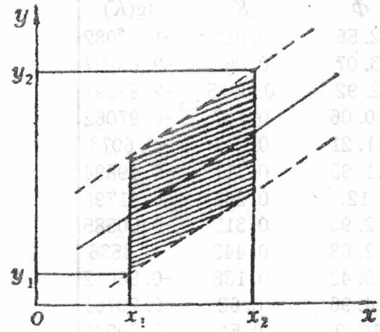


图 2-3

2.1.4　化非线性为线性的问题

前面介绍的回归分析的计算方法是基于变量之间的关系是线性关系，但有时在实际中所遇到的变量它们之间的关系不一定是线性关系，为了计算的简便，我们可以将数据先进行变换，然后用变换后的数据进行回归，再代回到原来的变量，从而求得原变量之间的关系，下面仅考虑几种情况。

1．指数关系

$$\hat{y}=ax^b$$

对 ax^b 取对数得

$$\hat{y}=\ln a+b\ln x$$

令 $a'=\ln a, x'=\ln x$，则化为线性的

$$\hat{y}=a'+bx'$$

2．对数关系

$$\ln\hat{y}=a+bx$$

令 $\hat{y}'=\ln\hat{y}$，则化为线性的

$$\hat{y}'=a+bx$$

3．双曲线关系

$$\frac{1}{\hat{y}}=a+b\frac{1}{x}$$

令 $\hat{y}'=\dfrac{1}{\hat{y}}, x'=\dfrac{1}{x}$，则化为线性的

$$\hat{y}'=a+bx'$$

对其他的线性或非线性的关系，也可进行类似的变换，利用线性回归分析求得回归系数后再换算成原变量之间的关系。

例 1．利用第 1 章例题中的数据，试建立渗透率与孔隙度之间的回归方程并进行显著性检验。