第九章　多元线性回归和相关分析

扬州大学 YANGZHOU UNIVERSITY

- 9.1　多元线性回归分析
- 9.2　多元线性相关分析

9.1 多元线性回归分析

多元线性回归（multiple linear regression）是依变量依两个或两个以上自变量且各自变量均为一次项（线性）的回归。

多元线性回归分析的主要内容包括：

① 建立多元线性回归方程；

② 进行多元线性回归方程及偏回归系数显著性测验，选择对依变量有显著效应的自变量，建立最优多元线性回归方程；

③ 评价各个自变量对依变量的相对重要性。

9.1 多元线性回归分析

1. 多元线性回归模型

若依变量Y同时受到m个**自变量**X_1、X_2、…、X_m的影响，且这m个自变量皆与Y成线性关系，则Y依m个变量的回归称为m元线性回归。

9.1 多元线性回归分析

1. 多元线性回归模型

表9-1 多元线性回归样本的数据结构

变量 试验号	x_1	x_2	…	x_m	y
1	x_{11}	x_{12}	…	x_{1m}	y_1
2	x_{21}	x_{22}	…	x_{2m}	y_2
			…		
n	x_{n1}	x_{n2}	…	x_{nm}	y_n

9.1 多元线性回归分析

1. 多元线性回归模型

***m*元线性回归样本的观察值组成为：**

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ \ldots \ + b_m x_{im} + e_i$$

其中b_0、b_1、…、b_m、e_i分别是β_0、β_1、…、β_m、ε_i的估计值。

*m*元线性回归总体的数学模型为：

$$Y_i = \beta_0 + \beta_i X_{11} + \beta_2 X_{i2} + \mathrm{L} \ + \beta_m X_{im} + \varepsilon_i$$

9.1 多元线性回归分析

揚州大学 YANGZHOU UNIVERSITY

1. 多元线性回归模型

一个m元线性回归方程可定义为：

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \ \ldots \ + b_mx_m$$

b_0是x_1、x_2、…、x_m 都为0时y 的点估计值；

b_j（j = ，2，…，m ）**称为自变量x_j对依变量y的偏回归系数**（**partial regression coefficient**）

偏回归系数：表示除自变量x_j以外的其余$m-1$个自变量皆保持一定时（取常量），自变量x_j每变化一个单位，依变量y平均变化的单位数值。

9.1 多元线性回归分析

2. 多元线性回归方程的求解和离回归标准误的计算

（1）多元线性回归方程的求解

$$Q=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(y_i-b_0-b_1x_{i1}-b_2x_{i2}-\cdots-b_mx_{im})^2=\text{最小}$$

$$\begin{cases}\dfrac{\partial Q}{\partial b_0}=-2\Sigma(y-b_0-b_1x_1-b_2x_2-\cdots-b_mx_m)=0\\ \dfrac{\partial Q}{\partial b_1}=-2\Sigma(y-b_0-b_1x_1-b_2x_2-\cdots-b_mx_m)x_1=0\\ \qquad\qquad\vdots\\ \dfrac{\partial Q}{\partial b_m}=-2\Sigma(y-b_0-b_1x_1-b_2x_2-\cdots-b_mx_m)x_m=0\end{cases}$$

$$\begin{cases}nb_0+b_1\Sigma x_1+b_2\Sigma x_2+\cdots b_m\Sigma x_m=\Sigma y\\ b_0\Sigma x_1+b_1\Sigma x_1^2+b_2\Sigma x_1x_2+\cdots b_m\Sigma x_1x_m=\Sigma x_1y\\ b_0\Sigma x_2+b_1\Sigma x_2x_1+b_2\Sigma x_2^2+\cdots b_m\Sigma x_2x_m=\Sigma x_2y\\ \qquad\qquad\vdots\\ b_0\Sigma x_m+b_1\Sigma x_mx_1+b_2\Sigma x_mx_2+\cdots b_m\Sigma x_m^2=\Sigma x_my\end{cases}$$

2. 多元线性回归方程的求解和离回归标准误的计算

(1) 多元线性回归方程的求解

若用矩阵形式表示正规方程组

$$
\begin{cases}
nb_0 + b_1\Sigma x_1 + b_2\Sigma x_2 + \cdots b_m\Sigma x_m = \Sigma y \\
b_0\Sigma x_1 + b_1\Sigma x_1^2 + b_2\Sigma x_1 x_2 + \cdots b_m\Sigma x_1 x_m = \Sigma x_1 y \\
b_0\Sigma x_2 + b_1\Sigma x_2 x_1 + b_2\Sigma x_2^2 + \cdots b_m\Sigma x_2 x_m = \Sigma x_2 y \\
\qquad\qquad \vdots \\
b_0\Sigma x_m + b_1\Sigma x_m x_1 + b_2\Sigma x_m x_2 + \cdots b_m\Sigma x_m^2 = \Sigma x_m y
\end{cases}
\Rightarrow
\begin{pmatrix}
n & \Sigma x_1 & \Sigma x_2 & \cdots & \Sigma x_m \\
\Sigma x_1 & \Sigma x_1^2 & \Sigma x_1 x_2 & \cdots & \Sigma x_1 x_m \\
\Sigma x_2 & \Sigma x_2 x_1 & \Sigma x_2^2 & \cdots & \Sigma x_2 x_m \\
\vdots & \vdots & \vdots & \cdots & \vdots \\
\Sigma x_m & \Sigma x_m x_1 & \Sigma x_m x_2 & \cdots & \Sigma x_m^2
\end{pmatrix}
\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}
=
\begin{pmatrix} \Sigma y \\ \Sigma x_1 y \\ \Sigma x_2 y \\ \vdots \\ \Sigma x_m y \end{pmatrix}
$$

$$令\mathbf{A}=\begin{pmatrix} n & \Sigma x_1 & \Sigma x_2 & \cdots & \Sigma x_m \\ \Sigma x_1 & \Sigma x_1^2 & \Sigma x_1 x_2 & \cdots & \Sigma x_1 x_m \\ \Sigma x_2 & \Sigma x_2 x_1 & \Sigma x_2^2 & \cdots & \Sigma x_2 x_m \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \Sigma x_m & \Sigma x_m x_1 & \Sigma x_m x_2 & \cdots & \Sigma x_m^2 \end{pmatrix},\quad \mathbf{b}=\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},\quad \mathbf{B}=\begin{pmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \\ \vdots \\ \sum x_m y \end{pmatrix}$$

则：$\mathbf{Ab}=\mathbf{B}$

$$若令\mathbf{X}=\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}$$

$$则:\ \mathbf{A}=\begin{pmatrix} n & \Sigma x_1 & \Sigma x_2 & \cdots & \Sigma x_m \\ \Sigma x_1 & \Sigma x_1^2 & \Sigma x_1 x_2 & \cdots & \Sigma x_1 x_m \\ \Sigma x_2 & \Sigma x_2 x_1 & \Sigma x_2^2 & \cdots & \Sigma x_2 x_m \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \Sigma x_m & \Sigma x_m x_1 & \Sigma x_m x_2 & \cdots & \Sigma x_m^2 \end{pmatrix}=\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ x_{11} & x_{21} & x_{31} & \cdots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{1m} & x_{2m} & x_{3m} & \cdots & x_{nm} \end{pmatrix}\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}=\mathbf{X'X}$$

常数项矩阵**B**亦可用矩阵**X′**和**Y**表示：

$$\mathbf{B}=\begin{pmatrix}\Sigma y\\ \Sigma x_1 y\\ \Sigma x_2 y\\ \vdots\\ \Sigma x_m y\end{pmatrix}=\begin{pmatrix}1 & 1 & 1 & 1 & 1\\ x_{11} & x_{21} & x_{31} & \cdots & x_{n1}\\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2}\\ \vdots & \vdots & \vdots & \cdots & \vdots\\ x_{1m} & x_{2m} & x_{3m} & \cdots & x_{nm}\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ y_3\\ \vdots\\ y_n\end{pmatrix}=\mathbf{X}'\mathbf{Y}$$

因此，正规方程组还可表示为：

$$(\mathbf{X}'\mathbf{X})\mathbf{b}=\mathbf{X}'\mathbf{Y}$$

因而有 $$\mathbf{b} = \mathbf{A}^{-1}\mathbf{B} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

假如设

$$\mathbf{C} = \mathbf{A}^{-1} = (c_{ij}) = \begin{pmatrix} c_{00} & c_{01} & c_{02} & \cdots & c_{0m} \\ c_{10} & c_{11} & c_{12} & \cdots & c_{1m} \\ c_{20} & c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_{m0} & c_{m1} & c_{m2} & \cdots & c_{mm} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_m \end{pmatrix}$$

则

$$\mathbf{b} = \mathbf{A}^{-1}\mathbf{B} = \mathbf{C}\mathbf{B} = \begin{pmatrix} c_{00} & c_{01} & c_{02} & \cdots & c_{0m} \\ c_{10} & c_{11} & c_{12} & \cdots & c_{1m} \\ c_{20} & c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_{m0} & c_{m1} & c_{m2} & \cdots & c_{mm} \end{pmatrix} \begin{pmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_m \end{pmatrix}$$

那么由正规方程组求出的偏回归系数最小二乘估计可表示为：

$$b_j = c_{j0}B_0 + c_{j1}B_1 + c_{j2}B_2 + \ldots + c_{jm}B_m$$

【例9-1】通过15个田块的超级稻沈农606测定数据（见表9-2），研究每公顷穗数（x_1，万）、每穗实粒数（x_2）与每公顷水稻籽粒产量（y，kg）的关系。试建立每公顷穗数每穗实粒数对每公顷水稻产量的多元线性回归方程。

表9-2 粳稻“沈农606”的每公顷穗数（x_1）、每穗实粒数（x_2）与每公顷产量（y）

田块	x_1	x_2	y
1	397.7	127.7	12585
2	434.4	111.3	12420
3	388.3	125.6	12315
4	404.1	129.9	12960
5	411.5	118.1	12390
6	414.7	109.8	11460
7	382.8	122.4	11910
8	413.1	113.4	11895
9	423.4	110.9	11685
10	402.6	119.3	12225
11	419.8	115.2	12360
12	411.8	112.5	11790
13	431.3	108.8	11985
14	412.1	119.8	12600
15	396.3	129.6	12315

① 写出结构矩阵（数据矩阵X）及依变量列向量Y

$$\mathbf{X}=\begin{pmatrix}1 & 397.7 & 127.7\\ 1 & 434.4 & 111.3\\ \vdots & \vdots & \vdots\\ 1 & 396.3 & 129.6\end{pmatrix},\quad \mathbf{Y}=\begin{pmatrix}12585\\ 12420\\ \vdots\\ 12315\end{pmatrix}$$

②求出系数矩阵A和常数项矩阵B

$$\mathbf{A}=\mathbf{X'X}=\begin{pmatrix}1 & 1 & \cdots & 1\\ 397.7 & 434.4 & \cdots & 396.3\\ 127.7 & 111.3 & \cdots & 129.6\end{pmatrix}\begin{pmatrix}1 & 397.7 & 127.7\\ 1 & 434.4 & 111.3\\ \vdots & \vdots & \vdots\\ 1 & 396.3 & 129.6\end{pmatrix}=\begin{pmatrix}15 & 6143.90 & 1774.30\\ 6143.90 & 2519503.53 & 725559.75\\ 1174.30 & 725559.75 & 210633.55\end{pmatrix}$$

$$\mathbf{B}=\mathbf{X'Y}=\begin{pmatrix}1 & 1 & \cdots & 1\\ 397.7 & 434.4 & \cdots & 396.3\\ 127.7 & 111.3 & \cdots & 129.6\end{pmatrix}\begin{pmatrix}12585\\ 12420\\ \vdots\\ 12315\end{pmatrix}=\begin{pmatrix}182895.0\\ 74898361.5\\ 21661990.5\end{pmatrix}$$

③求系数矩阵A的逆矩阵C

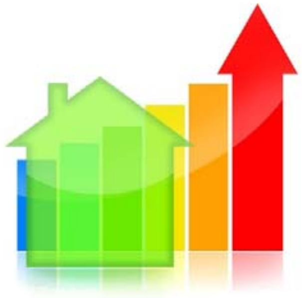
$$\mathbf{C}=\begin{pmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{pmatrix}=\mathbf{A}^{-1}=(\mathbf{X}'\mathbf{X})^{-1}=\begin{pmatrix} 322.9211738498 & -0.5123304638 & -0.9553685797 \\ -0.5123304638 & 0.0008623400 & 0.0013452212 \\ -0.9553685797 & 0.0013452212 & 0.0034186012 \end{pmatrix}$$

④求解偏回归系数矩阵

$$\mathbf{b}=\mathbf{CB}=\begin{pmatrix} 322.9211738498 & -0.5123304638 & -0.9553685797 \\ -0.5123304638 & 0.0008623400 & 0.0013452212 \\ -0.9553685797 & 0.0013452212 & 0.0034186012 \end{pmatrix}\begin{pmatrix} 182895.0 \\ 74898361.5 \\ 21661990.5 \end{pmatrix}=\begin{pmatrix} -7229.29433745 \\ 25.34485290 \\ 76.43474798 \end{pmatrix}$$

⑤ 写出线性回归方程

$$\hat{y}=-7229.29433745+25.34485290x_1+76.43474798x_2$$



生物统计与试验设计

Biostatistics and Experimental design

主讲：徐辰武

扬州大学农学院

9.1 多元线性回归分析

2. 多元线性回归方程的求解和离回归标准误的计算

（2）多元线性回归方程的离回归标准误的计算

$$SS_y = U_{y/12\cdots m} + Q_{y/12\cdots m} \qquad \nu_y = \nu_U + \nu_Q$$

$$\begin{cases} SS_y = \mathbf{Y'Y} - (\mathbf{1'Y})^2 / n \\ Q_{y/12\cdots m} = \mathbf{Y'Y} - \mathbf{b'}(\mathbf{X'Y}) \\ U_{y/12\ldots m} = \mathbf{b'}(\mathbf{X'Y}) - (\mathbf{1'Y})^2 / n = SS_y - Q_{y/12\cdots m} \end{cases}$$

总变异平方和SS_y为依变量y的离均差平方总和，反映的是依变量y的总变异，其自由度$\nu_y = n - 1$。

9.1 多元线性回归分析

2. 多元线性回归方程的求解和离回归标准误的计算

（2）多元线性回归方程的离回归标准误的计算

$Q_{y/12\ldots m}$ 称为多元离回归平方和，它反映了回归估计值和实测值y之间的差异。

$$Q = \sum (y - \hat{y})^2 = \text{最小}$$

$$S_{y/12\ldots m} = \sqrt{\frac{Q_{y/12\ldots m}}{n-(m+1)}}$$

9.1 多元线性回归分析

2. 多元线性回归方程的求解和离回归标准误的计算

（2）多元线性回归方程的离回归标准误的计算

离回归平方和（或剩余平方和）$Q_{y/12\ldots m}$与自变量x无关，仅反映除依变量y与m个自变量间存在线性关系以外的其他因素包括试验误差所引起的变异。由于在建立多元回归方程时，正规方程组包含m+1个方程，即有m+1个限制条件，故其自由度$\nu_Q = n - (m+1) = n - m - 1$。

$n-1$

回归平方和$U_{y/12\ldots m}$是由m个自变量x_j的不同引起的，即是依变量y受m个自变量综合线性影响所引起的变异。其自由度$\nu_U = m$。

扬州大学 YANGZHOU UNIVERSITY

定义多元回归方程的离回归标准误为：

$$s_{y/12\cdots m}=\sqrt{\frac{Q_{y/12\cdots m}}{n-(m+1)}}$$

$s_{y/12\ldots m}$的大小反映了回归平面与实测点的偏离程度，即回归估计值$\hat{y}$与实测值y偏离的程度。离回归标准误$s_{y/12\cdots m}$愈小，表明各个观察点愈靠近回归平面，则由回归方程估计y的精确度愈高；反之，离回归标准误$s_{y/12\cdots m}$愈大，各个观察点愈偏离回归平面，由回归方程估计y的精确度愈低，可见，$s_{y/12\cdots m}$是回归精确度的量度。

9.1 多元线性回归分析

揚州大學 YANGZHOU UNIVERSITY

3. 多元线性回归的假设测验

(1) 多元回归方程的假设测验

$H_0: \beta_1=\beta_2=\cdots=\beta_m=0$

$$F=\frac{U_{y/12\Lambda m}/\nu_U}{Q_{y/12\Lambda m}/\nu_Q}=\frac{U_{y/12\Lambda m}/m}{Q_{y/12\Lambda m}/[n-(m+1)]}$$

(2) 偏回归系数的假设测验

$H_0: \beta_j=0$

（A） *F*测验法

$$U_j=\frac{b_j^2}{c_{jj}}$$

$$F=\frac{U_j/\nu_j}{Q_{y/12L\ m}/\nu_Q}=\frac{U_j/1}{Q_{y/12L\ m}/[n-(m+1)]}$$

9.1 多元线性回归分析

3. 多元线性回归的假设测验

（1）多元回归方程的假设测验

$$F = \frac{U_{y/12\Lambda\ m}/\nu_U}{Q_{y/12\Lambda\ m}/\nu_Q} = \frac{U_{y/12\Lambda\ m}/m}{Q_{y/12\Lambda\ m}/[n-(m+1)]}$$

（2）偏回归系数的假设测验

（A） *F*测验法

$$U_j = \frac{b_j^2}{c_{jj}} \qquad F = \frac{U_j/\nu_j}{Q_{y/12\mathrm{L}\ m}/\nu_Q} = \frac{U_j/1}{Q_{y/12\mathrm{L}\ m}/[n-(m+1)]}$$

（B）t测验法

$$t = \frac{b_j - \beta_j}{s_{b_j}} \sim t_{\nu = n-m-1}$$

$$s_{b_j} = s_{y/12\Lambda m}\sqrt{c_{jj}}$$

$F = t^2$

【例9-3】试测验用表9-2资料建立的二元线性回归方程的显著性，并对二元回归方程中的两个偏回归系数做显著性测验。

（1）二元线性回归方程的假设测验

①假设H_0：$\beta_1=\beta_2=0$，H_A：β_1、β_2至少有一个不为0。

②确定显著水平，$\alpha=0.01$

③计算总变异平方和SS_y，回归平方和$U_{y/12}$及相应自由度

$$SS_y=\mathbf{Y'Y}-(\mathbf{1'Y})^2/n=2232265275-182895^2/15=2226540$$

$$U_{y/12}=SS_y-Q_{y/12}=2226540-450323.5667=1776216.4333$$

④计算F值

$$F=\frac{U_{y/12}/m}{Q_{y/12}/(n-m-1)}=\frac{1776216.4333/2}{450323.5667/(15-2-1)}=23.67$$

⑤推断

$F>F_{0.01\ (2,\ 12)}=6.93$，说明$P$（$H_0$）$<0.01$（实际$6.84\times10^{-5}$），所以应否定$H_0$，接受$H_A$，二元线性回归方程成立，表9-2的每公顷穗数$x_1$、每穗实粒数$x_2$与每公顷水稻产量$y$有真实的二元线性回归关系。

（2）偏回归系数的假设测验

①无效假设H_0：$\beta_j = 0$；对应假设H_A：$\beta_j \neq 0$。

②确定显著水平，$\alpha = 0.01$

③计算各偏回归平方和U_j及自由度

y依x_1的偏回归平方和及自由度：

$$U_1 = \frac{b_1^2}{c_{11}} = \frac{25.34485290^2}{0.00086234} = 744905.2213 \quad \nu_1 = 1$$

y依x_2的偏回归平方和及自由度：

$$U_2 = \frac{b_2^2}{c_{22}} = \frac{76.43474798^2}{0.00341860} = 1708965.8629 \quad \nu_2 = 1$$

④计算F值

$$F_1=\frac{U_1/1}{Q_{y/12}/\nu_Q}=\frac{744905.2213}{450323.5667/12}=19.85$$

$$F_2=\frac{U_2/1}{Q_{y/12}/\nu_Q}=\frac{1708965.8629}{450323.5667/12}=45.54$$

⑤推断　F_1、F_2均 $> F_{0.01(1,12)}=9.33$；说明H_0：$\beta_1=0$、H_0：$\beta_2=0$应被否定，即每公顷总穗数x_1、每穗实粒数x_2对每公顷水稻产量y的偏回归也都是极显著的。

若采用t测验的方法测验偏回归系数的显著性，

$$s_{b_1} = s_{y/12}\sqrt{c_{11}} = 193.7188 \times \sqrt{0.00086234} = 5.68867314$$

$$s_{b_2} = s_{y/12}\sqrt{c_{22}} = 193.7188 \times \sqrt{0.00341860} = 11.32650338$$

$$t_1 = \frac{b_1 - \beta_1}{s_{b_1}} = \frac{25.34485290}{5.68867314} = 4.46$$

$$t_2 = \frac{b_2 - \beta_2}{s_{b_2}} = \frac{76.43474798}{11.32650338} = 6.75$$

查附表，$t_{0.01\ (12)} = 3.055$，现实得t_1、t_2均$> t_{0.01\ (12)}$，所以否定H_0接受H_A，这与F测验法的结论完全一致，所以，两法可任择其一应用。

第九章　多元线性回归和相关分析

扬州大学 YANGZHOU UNIVERSITY

- 9.1　多元线性回归分析
- **9.2　多元线性相关分析**

9.2 多元线性相关分析

1. 多元相关分析

(1) 多元相关的意义及多元相关系数的计算

在 $M=m+1$个变量中，$\boldsymbol{m}$ 个变量的综合和一个变量的相关，叫做多元相关或复相关（**multiple correlation**）。虽然从相关分析角度来说，多元相关中的变量并没有依变量与自变量之分，但是在实际应用中，多元相关分析往往与多元线性回归分析联系在一起，因此，多元相关一般指依变量$\boldsymbol{y}$与$\boldsymbol{m}$个自变量$\boldsymbol{x}_1$、$\boldsymbol{x}_2$、…、$\boldsymbol{x}_m$的线性相关。

$$R^2_{y\cdot 12\cdots m}=\frac{U_{y/12\cdots m}}{SS_y}=1-\frac{Q_{y/12\cdots m}}{SS_y}$$

为y与x_1、x_2、…、x_m的多元决定系数（**multiple determination coefficient** ）。它表示回归方程的拟合程度。

$$R_{y\cdot 12\cdots m}=\sqrt{\frac{U_{y/12\cdots m}}{SS_y}}=\sqrt{1-\frac{Q_{y/12\cdots m}}{SS_y}}$$

为依变量y与 个自变量x_1、 x_2、…、x_m的多元相关系数或复相关系数（**multiple correlation coefficient**）。

【例9-4】由表9-2资料计算依变量y（每公顷水稻产量）与自变量x_1（每公顷穗数）和x_2（每穗实粒数）的二元相关系数，并与各自变量x_j与依变量y的简单相关系数作比较。

$$SS_y = 2226540 \quad U_{y/12} = 1776216.4333$$

$$R_{y\cdot 12} = \sqrt{\frac{U_{y/12}}{SS_y}} = \sqrt{\frac{1776216.4333}{2226540.0000}} = 0.8932$$

y分别与x_1、x_2的简单相关系数：

$$r_{1y} = \frac{SP_{1y}}{\sqrt{SS_{x_1} \cdot SS_y}} = \frac{-14211.20000000}{\sqrt{3003.04933333 \times 2226540.00000001}} = -0.17379380$$

$$r_{2y} = \frac{SP_{2y}}{\sqrt{SS_{x_2} \cdot SS_y}} = \frac{27950.60000001}{\sqrt{757.51733333 \times 2226540.00000001}} = 0.68058065$$

可见二元相关系数$R_{y\cdot 12}$比简单相关系数r_{1y}、r_{2y}的绝对值都大。

9.2 多元线性相关分析

1. 多元相关分析

（2）多元相关系数的假设测验

①F测验法

设ρ为y与x_1、x_2、…、x_m的总体多元相关系数，假设H_0：$\rho_{y\cdot 12\cdots m}=0$；$H_A$：$\rho_{y\cdot 12\cdots m}\neq 0$

$$F_R=\frac{R^2_{y\cdot 12\cdots m}/\nu_1}{(1-R^2_{y\cdot 12\cdots m})/\nu_2}\sim F_{\nu_1=m,\nu_2=n-m-1}$$

② 查R_{α}值法

$$R_{\alpha}=\sqrt{\frac{\nu_1 F_{\alpha}}{\nu_1 F_{\alpha}+\nu_2}}$$

并将其列成表。因此多元相关系数假设测验，亦可采用查R_{α}值法。

查表所需参数：变量的总个数$M=m+1$；自由度为$\nu_2=n-m-1$。若$R_{y\cdot 12\cdots m}\geq R_{\alpha}$，$P\leq\alpha$，则否定$H_0$，否则，$R_{y\cdot 12\cdots m}<R_{\alpha}$，$P>\alpha$，接受$H_0$。

9.2 多元线性相关分析

扬州大学
YANGZHOU UNIVERSITY

2. 偏相关分析

（1）偏相关的意义

在M个变量中，其余$M-2$个变量保持固定不变，指定的两个变量间的相关，叫做偏相关（partial correlation）。用来表示两个相关变量偏相关的性质与程度的统计数叫偏相关系数（partial correlation coefficient）。

（2）偏相关系数的计算

①由 M 个变量的简单相关系数r_{ij}，构建相关系数矩阵R

$$\mathrm{R}=(r_{ij})_{M\times M}=\begin{pmatrix} r_{11} & r_{12} & \mathrm{L} & r_{1M} \\ r_{21} & r_{22} & \mathrm{L} & r_{2M} \\ \mathrm{M} & \mathrm{M} & \mathrm{L} & \mathrm{M} \\ r_{M1} & r_{M2} & \mathrm{L} & r_{MM} \end{pmatrix}$$

②求相关系数矩阵R的逆矩阵C’

$$\mathbf{C'}=\mathbf{R}^{-1}=(c'_{ij})_{M\times M}=\begin{pmatrix} c'_{11} & c'_{12} & \mathrm{L} & c'_{1M} \\ c'_{21} & c'_{22} & \mathrm{L} & c'_{2M} \\ \mathrm{M} & \mathrm{M} & \mathrm{L} & \mathrm{M} \\ c'_{M1} & c'_{M2} & \mathrm{L} & c'_{MM} \end{pmatrix}$$

扬州大学 YANGZHOU UNIVERSITY

③由下式计算偏相关系数$r_{ij\cdot}$

$$r_{ij\cdot}=\frac{-c'_{ij}}{\sqrt{c'_{ii}c'_{jj}}}$$

纯净