第二节　多元线性回归预测

- 多元线性回归模型的基本假设
- 回归模型的参数估计
- 回归模型的假设检验
- 预测、控制和风险分析
- 可线性化的非线性回归

一、多元线性回归模型的基本假设

- 模型

设因变量 y 与自变量 x_1，$x_2,\cdots$，x_p 有统计的线性关系：

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p + \varepsilon$$

对 $(x_1$，$x_2,\cdots$，$x_p, y)$ 进行 n 次观测，得到观测值

$$(x_{1i}，x_{2i},\cdots，x_{pi}, y_i) \quad i = 1, 2, \cdots, n.$$

求出 b_0，$b_1,\cdots$，b_p 的估计值 $\hat{b}_0$，$\hat{b}_1,\cdots$，$\hat{b}_p$，建立预测方程

$$\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \cdots + \hat{b}_p x_p$$

通过预测方程进行预测，并给出预测的置信区间。

矩阵形式:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad X = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \cdots & \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix}$$

$$B = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \qquad Y = XB + \varepsilon$$

预测方程 $\hat{Y} = X\hat{B}, \quad e = Y - \hat{Y} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$

- 回归模型的基本假定

1. 零期望值假定 $E(\varepsilon_i)=0$

$$\Rightarrow E(\varepsilon)=E\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix}=\begin{pmatrix}E\varepsilon_1\\ E\varepsilon_2\\ \vdots\\ E\varepsilon_n\end{pmatrix}=0$$

2. 同方差假定 $D(\varepsilon_i)=\sigma^2$

3. 无自相关假定 $Cov(\varepsilon_i,\varepsilon_j)=E\varepsilon_i\varepsilon_j=0$

4. 随机误差项与解释变量 不相关： $Cov(x_{ij},\varepsilon_j)=0$

5. $\varepsilon_i \sim N(0,\sigma^2)$

$$E(\varepsilon\varepsilon') = E\left(\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}\begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{pmatrix}\right) = \begin{pmatrix} E(\varepsilon_1\varepsilon_1) & E(\varepsilon_1\varepsilon_2) & \cdots & E(\varepsilon_1\varepsilon_n) \\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2\varepsilon_2) & \cdots & E(\varepsilon_2\varepsilon_n) \\ \vdots & \vdots & \cdots & \vdots \\ E(\varepsilon_n\varepsilon_1) & E(\varepsilon_n\varepsilon_2) & \cdots & E(\varepsilon_n\varepsilon_n) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}$$

$$\Rightarrow \varepsilon \sim N(0, \sigma^2 I_n)$$

6. 解释变量之间不存在多重共线性

各解释变量的观测值之间线性无关

$$rank(X) = P + 1 \quad rank(X'X) = P + 1$$

二、回归模型的参数估计

- 普通最小二乘法（OLS）

残差平方和：$Q=\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}(y_i-\hat{b}_0-\hat{b}_1x_{1i}-\hat{b}_2x_{2i}-\cdots-\hat{b}_px_{pi})^2$

$$=Y'Y-2\hat{B}'X'Y+\hat{B}'X'X\hat{B}$$

$$\frac{\partial Q}{\partial \hat{B}}=-2X'Y+2X'X\hat{B}=0$$

$$\hat{B}=(X'X)^{-1}X'Y$$

■ 最小二乘估计量的性质

1.线性：$\hat{B}$分别为Y和ε的线性函数或线性组合

$$\hat{B}=(X'X)^{-1}X'Y$$

$$=(X'X)^{-1}X'(XB+\varepsilon)=(X'X)^{-1}X'XB+(X'X)^{-1}X'\varepsilon$$

$$=B+(X'X)^{-1}X'\varepsilon$$

2.无偏性：$E(\hat{B})=B$

$$E(\hat{B})=E\left(B+(X'X)^{-1}X'\varepsilon\right)=B$$

3.最小方差性：在所有线性、无偏估计量中$\hat{B}$的方差最小。

证明：设B^*是B的任意线性无偏估计，如果协方差矩阵之差

$$E\left[\left(B^*-B\right)\left(B^*-B\right)'\right]-E\left[\left(\hat{B}-B\right)\left(\hat{B}-B\right)'\right]$$

为半正定矩阵，则称最小二乘估计量$\hat{B}$是B的最小方差线性无偏估计。

由于B^*是B的线性无偏估计，记$B^*=AY$,由无偏性，应有

$$E\left(B^*\right)=E\left(AY\right)=E\left[A\left(XB+\varepsilon\right)\right]=AXB=B$$

从而 $AX=I$。有

$$B^*-B=AY-B=A\left(XB+\varepsilon\right)-B=A\varepsilon.$$

$$E\left[\left(B^*-B\right)\left(B^*-B\right)'\right]=E\left[A\varepsilon\varepsilon'A'\right]=\sigma^2AA'.$$

然而 $E\left[\left(\hat{B}-B\right)\left(\hat{B}-B\right)'\right]=\sigma^2(X'X)^{-1},$

因此 $E\left[\left(B^*-B\right)\left(B^*-B\right)'\right]-E\left[\left(\hat{B}-B\right)\left(\hat{B}-B\right)'\right]=\sigma^2\left[AA'-(X'X)^{-1}\right]$

考虑到 $\left[A-(X'X)^{-1}X'\right]\left[A-(X'X)^{-1}X'\right]'$

$$=AA'-(X'X)^{-1}X'A'-AX(X'X)^{-1}+(X'X)^{-1}X'X(X'X)^{-1}$$

$$=AA'-(X'X)^{-1}$$

由线性代数可知，对于任意一个非奇异矩阵，CC'为半正定矩阵。得证。

回归参数的区间估计

1. $\widehat{B}$的分布：$\widehat{B} \sim N(B, \sigma^2 (X'X)^{-1})$

$$D(\widehat{B}) = E\left(\left(\widehat{B} - E\widehat{B}\right)\left(\widehat{B} - E\widehat{B}\right)'\right) = E\left((X'X)^{-1} X'\varepsilon\varepsilon' X (X'X)^{-1}\right)$$

$$= (X'X)^{-1} X'\sigma^2 X (X'X)^{-1} = \sigma^2 (X'X)^{-1}$$

2.随机误差项方差的估计

σ^2的无偏估计量 $\quad \widehat{\sigma}^2 = \dfrac{\sum_{i=1}^{n} e_i^2}{n-p-1} = \dfrac{Y'Y - \widehat{B}'X'Y}{n-p-1} \quad E(\widehat{\sigma}^2) = \sigma^2$

证明：$e = Y - \widehat{Y} = Y - X\widehat{B} = Y - X(X'X)^{-1}X'Y = \left[I_n - X(X'X)^{-1}X'\right]Y = PY$

又 $e = \left[I_n - X(X'X)^{-1}X'\right](XB+\varepsilon) = \left[I_n - X(X'X)^{-1}X'\right]\varepsilon = P\varepsilon$

其中 $P = I_n - X(X'X)^{-1}X'$为幂等矩阵

$D(e) = Eee' = P\sigma^2$

又由矩阵迹的性质 $tr(ee') = e'e = \sum_{i=1}^{n} e_i^2$

$$E(\sum_{i=1}^{n} e_i^2) = E(e'e) = E[tr(ee')] = tr[E(ee')] = tr(P\sigma^2)$$

$$= \sigma^2 tr[I_n - X(X'X)^{-1}X'] = \sigma^2[tr(I_n) - tr(X(X'X)^{-1}X')]$$

$$= \sigma^2[n - tr(I_{p+1})] = (n-p-1)\sigma^2$$

三、回归模型的假设检验

1.拟和优度检验：衡量样本回归直线与观测值之间的拟和程度

多重可决系数 $r^2 = \dfrac{U}{S_{总}}$

其中： $U = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ $S_{总} = \sum_{i=1}^{n}(y_i - \bar{y})^2$ $Q = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

修正的可决系数 $\bar{r}^2 = 1 - \dfrac{Q/(n-p-1)}{S_{总}/(n-1)}$ $\bar{r}^2 \le r^2$

$\bar{r}^2$可能小于 0 ，此时作$\bar{r}^2$=0处理

2. F检验：检验y与 x_1，x_2,⋯，x_p之间是否确有线性关系

$F = \dfrac{U/p}{Q/(n-p-1)}$，对于给定的显著性水平$\alpha$，查$F$分布表得$F_\alpha(p, n-p-1)$， 若$F \ge F_\alpha(p, n-p-1)$，则回归模型可以用来预测。否则不能。

F与可决系数之间的关系： $F = \dfrac{n-p-1}{p} \cdot \dfrac{r^2}{1-r^2}$

3. t检验：检验 x_i 对 y 的作用是否显著。若显著，则保留 x_i 在回归方程中，否则去掉。

剩余标准差 $$S_y = \sqrt{\frac{Q}{n-p-1}} = \sqrt{\frac{Y'Y - \hat{B}'X'Y}{n-p-1}} = \hat{\sigma}$$

$$t_i = \frac{\hat{b}_i}{S_y \cdot \sqrt{C_{ii}}}$$

其中：C_{ii}为$(X'X)^{-1}$的第$i+1$个对角元。

对于给定的显著性水平α，查t分布表　得$t_\alpha(n-p-1)$，

若$|t_i| \geq t_\alpha(n-p-1)$，则$x_i$ 的作用显著，必须保留x_i 在回归方程中。

4. ***Chow*检验**：模型结构稳定性检验，检验在整个样本的各子样本中， 模型的系数是否相等。如果模型在不同的子样本中模型的系数不同，则说明该模型中存在着转折点。假设一个样本中有 n_1 个观测值，另一个样本有 n_2 个观测值。

(1) 合并两个样本，构成 n_1+n_2 的样本，进行回归，得到

$$\hat{y}_i=\hat{b}_0+\hat{b}_1x_{1i}+\hat{b}_2x_{2i}+\cdots+\hat{b}_px_{pi}$$

求得残差平方和 $\sum_{i=1}^{n_1+n_2}e_i^2$，自由度为（$n_1+n_2$−p−1）

(2)对两个小样本分别回归，得到

$$\hat{y}_{1i}=\hat{\alpha}_0+\hat{\alpha}_1x_{1i}+\hat{\alpha}_2x_{2i}+\cdots+\hat{\alpha}_px_{pi} \qquad \hat{y}_{2i}=\hat{\beta}_0+\hat{\beta}_1x_{1i}+\hat{\beta}_2x_{2i}+\cdots+\hat{\beta}_px_{pi}$$

求得残差平方和 $\sum_{i=1}^{n_1}e_{1i}^2,\sum_{i=1}^{n_2}e_{2i}^2$，自由度分别为（$n_1$−p−1），（$n_2$−p−1）。

(3)计算：

$$F=\frac{[\sum_{i=1}^{n_1+n_2}e_i^2-\sum_{i=1}^{n_1}e_{1i}^2-\sum_{i=1}^{n_2}e_{2i}^2]/(p+1)}{[\sum_{i=1}^{n_1}e_{1i}^2+\sum_{i=1}^{n_2}e_{2i}^2]/(n_1+n_2-2p-2)}$$

对于给定的显著性水平α，查F分布表得$F_\alpha(p+1,\ n_1+n_2-2p-2)$，若$F\ge F_\alpha$，则两个样本反映得两个经济关系显著不同，经济结构发生了变化。否则经济结构关系比较稳定。

例：某企业固定资产 x_1 职工人数 x_2 和创造利润 y 数据如下，试建立预测模型，并检验。

年份	1978	1979	1980	……	1986	1987	Σ
y_i	233	238	261	……	304	315	2741
x_{1i}	250	257	271	……	325	338	2958
x_{2i}	161	163	167	……	185	187	1740

解：$$X=\begin{pmatrix} 1 & 250 & 161 \\ 1 & 257 & 163 \\ \vdots & \vdots & \vdots \\ 1 & 325 & 185 \\ 1 & 338 & 187 \end{pmatrix} \quad Y=\begin{pmatrix} 233 \\ 238 \\ \vdots \\ 304 \\ 315 \end{pmatrix}$$

$$(X'X)^{-1}=\begin{pmatrix} 185.0231 & 0.5873 & -2.0611 \\ 0.5873 & 0.0014 & -0.0074 \\ -2.0611 & -0.0074 & 0.0245 \end{pmatrix}$$

$$\hat{B}=(X'X)^{-1}X'Y=\begin{pmatrix}-106.7218\\0.498921\\1.34047\end{pmatrix}$$

$$\hat{y}=-106.7218+0.498921x_1+1.34047x_2$$

F检验：F=20103，$\alpha=0.05$，查表得 $F_\alpha(2,7)=4.74$，回归效果显著。

t检验：剩余标准差 $S_y=\sqrt{\dfrac{Q}{n-p-1}}=\sqrt{\dfrac{Y'Y-\hat{B}'X'Y}{n-p-1}}=4.3411$

$$t_1=\frac{\hat{b}_1}{S_y\cdot\sqrt{C_{11}}}=\frac{0.498921}{4.3411\times\sqrt{0.0014}}=3.072$$

$$t_2=\frac{\hat{b}_2}{S_y\cdot\sqrt{C_{22}}}=\frac{1.34047}{4.3411\times\sqrt{0.0245}}=1.973$$

$\alpha=0.05$，查表得 $t_\alpha(7)=2.37$。

因为 $t_1>t_\alpha$，$t_2<t_\alpha$，所以 x_1 对 y 有显著影响，x_2 对 y 的影响不显著。

表 I　　置信限 t_α 的数值表（n—自由度；α—信度）

n \ α	0.10	0.05	0.02	0.01	0.001
1	6.31	12.71	31.82	63.66	636.62
2	2.92	4.30	6.97	9.93	31.60
3	2.35	3.18	4.54	5.84	12.94
4	2.13	2.78	3.75	4.60	8.61
5	2.02	2.57	3.37	4.03	6.86
6	1.94	2.45	3.14	3.71	5.96
7	1.90	2.37	3.00	3.50	5.41
8	1.86	2.31	2.90	3.36	5.04

四、预测、控制和风险分析

1.预测： $\hat{y}^{(0)} = \hat{b}_0 + \hat{b}_1 x_1^{(0)} + \hat{b}_2 x_2^{(0)} + \cdots + \hat{b}_p x_p^{(0)}$

2.风险分析：

$$\hat{y}^{(0)} - t_\alpha(n-p-1)\cdot\hat{\sigma}\cdot\sqrt{1+x^{(0)}(X'X)^{-1}x^{(0)'}} \le y^{(0)}$$

$$\le \hat{y}^{(0)} + t_\alpha(n-p-1)\cdot\hat{\sigma}\cdot\sqrt{1+x^{(0)}(X'X)^{-1}x^{(0)'}}$$

简化：$P(\hat{y}^{(0)} - 3S_y < y^{(0)} < \hat{y}^{(0)} + 3S_y) = 99.7\%$

$P(\hat{y}^{(0)} - 2S_y < y^{(0)} < \hat{y}^{(0)} + 2S_y) = 95.4\%$

$P(\hat{y}^{(0)} - S_y < y^{(0)} < \hat{y}^{(0)} + S_y) = 68.3\%$

$\hat{\sigma} = S_y$

证明：$E(y^{(0)} - \widehat{y}^{(0)})$

$$= E\left(b_0 + b_1 x_1^{(0)} + b_2 x_2^{(0)} + \cdots + b_p x_p^{(0)} + \varepsilon^{(0)} - \left(\hat{b}_0 + \hat{b}_1 x_1^{(0)} + \hat{b}_2 x_2^{(0)} + \cdots + \hat{b}_p x_p^{(0)}\right)\right)$$

$$= b_0 + b_1 x_1^{(0)} + b_2 x_2^{(0)} + \cdots + b_p x_p^{(0)} + E\varepsilon^{(0)} - E\left(\hat{b}_0 + \hat{b}_1 x_1^{(0)} + \hat{b}_2 x_2^{(0)} + \cdots + \hat{b}_p x_p^{(0)}\right)$$

$$\because E\widehat{b}_i = b_i \qquad \therefore E(y^{(0)} - \widehat{y}^{(0)}) = 0$$

$$D[y^{(0)} - \widehat{y}^{(0)}] = D[y^{(0)}] + D[\widehat{y}^{(0)}]$$

$$\because \quad D[y^{(0)}] = D\left[b_0 + b_1 x_1^{(0)} + b_2 x_2^{(0)} + \cdots + b_p x_p^{(0)} + \varepsilon^{(0)}\right] = \sigma^2$$

$$\therefore \quad D[y^{(0)} - \widehat{y}^{(0)}] = \sigma^2 + D[\widehat{y}^{(0)}]$$

又 $\because D[\hat{y}^{(0)}] = E[\hat{y}^{(0)} - E(\hat{y}^{(0)})]^2 = E[x^{(0)}\hat{B} - E(x^{(0)}\hat{B})]^2$

$$\text{上式} = E[x^{(0)}\hat{B} - E(x^{(0)}B + x^{(0)}(X'X)^{-1}X'\varepsilon)]^2$$

$$= E[x^{(0)}\hat{B} - x^{(0)}B]^2 = E[x^{(0)}(\hat{B}-B)x^{(0)}(\hat{B}-B)]$$

$$= E[x^{(0)}(\hat{B}-B)(\hat{B}-B)'x^{(0)}{}'] = x^{(0)}E[(\hat{B}-B)(\hat{B}-B)']x^{(0)}{}'$$

$$= x^{(0)}D(\hat{B})x^{(0)}{}' = \sigma^2 x^{(0)}(X'X)^{-1}x^{(0)}{}'$$

$$\therefore \quad D[y^{(0)} - \hat{y}^{(0)}] = \sigma^2\left[1 + x^{(0)}(X'X)^{-1}x^{(0)}{}'\right]$$

$$y^{(0)} - \hat{y}^{(0)} \sim N(0, \sigma^2\left[1 + x^{(0)}(X'X)^{-1}x^{(0)}{}'\right])$$

$$t = \frac{y^{(0)} - \hat{y}^{(0)}}{\hat{\sigma}\cdot\sqrt{1 + x^{(0)}(X'X)^{-1}x^{(0)}{}'}} \sim t(n-p-1)$$

简化：$D[y^{(0)} - \hat{y}^{(0)}] = \sigma^2\left[1 + x^{(0)}(X'X)^{-1}x^{(0)}{}'\right] \approx \hat{\sigma}^2$

$y^{(0)} - \hat{y}^{(0)} \sim N(0, \hat{\sigma}^2)$

3.控制问题：给定α，要使 y 以$1-\alpha$的概率落在$\left[y_1, y_2\right]$中，

求 x_1，x_2,…，x_p的控制范围。

解法：$\begin{cases} \hat{b}_0+\hat{b}_1 x_1^{(1)}+\hat{b}_2 x_2^{(1)}+\cdots+\hat{b}_p x_p^{(1)}-S_y \geq y_1 \\ \hat{b}_0+\hat{b}_1 x_1^{(2)}+\hat{b}_2 x_2^{(2)}+\cdots+\hat{b}_p x_p^{(2)}+S_y \leq y_2 \end{cases}$ **(68.3%)**

或者：$\begin{cases} \hat{b}_0+\hat{b}_1 x_1^{(1)}+\hat{b}_2 x_2^{(1)}+\cdots+\hat{b}_p x_p^{(1)}-2S_y \geq y_1 \\ \hat{b}_0+\hat{b}_1 x_1^{(2)}+\hat{b}_2 x_2^{(2)}+\cdots+\hat{b}_p x_p^{(2)}+2S_y \leq y_2 \end{cases}$ **(95.4%)**

或者：$\begin{cases} \hat{b}_0+\hat{b}_1 x_1^{(1)}+\hat{b}_2 x_2^{(1)}+\cdots+\hat{b}_p x_p^{(1)}-3S_y \geq y_1 \\ \hat{b}_0+\hat{b}_1 x_1^{(2)}+\hat{b}_2 x_2^{(2)}+\cdots+\hat{b}_p x_p^{(2)}+3S_y \leq y_2 \end{cases}$ **(99.7%)**

例、某家具厂对家具的总成本进行预测，书72页表6.1列出总成本 y 与直接劳动量 x_1，耗用木材量 x_2 的7组观测值

(1) 试建立预测方程；

(2) 若该厂9月份用了4.1百个劳动小时，耗用木材量2.5千立方，试预测该厂9月份的总成本为多少？（95%的概率）

(3) 若欲控制10月份的成本在3千元至3.8千元之间，那么直接劳动量 x_1 和耗用木材量 x_2 应满足什么条件？（95%的概率）

解：（1）$\hat{y} = -1.3956 + 0.7461\, x_1 + 0.6769\, x_2$

(2) $\hat{y} = -1.3956 + 0.7461 \times 4.1 + 0.6769 \times 2.5 = 3.356$ 千元

预测总成本的上、下限（95%）： $S_y = 0.173$

$$3.356 - 2 \times 0.173 = 3.010$$

$$3.356 + 2 \times 0.173 = 3.702$$

$\therefore$ 总成本为（3.010， 3.702）

$$(3)\begin{cases} -1.3956 + 0.7461\, x_1 + 0.6769\, x_2 - 2 \times 0.173 \geq 3 \\ -1.3956 + 0.7461\, x_1 + 0.6769\, x_2 + 2 \times 0.173 \leq 3.8 \end{cases}$$

$$\Rightarrow 4.7416 \leq 0.7461\, x_1 + 0.6769\, x_2 \leq 4.8496$$

五、 可线性化的非线性回归

- 几种可线性化的曲线

非线性式	变换	线性式（估计式）
（1） $y=\alpha x^{\beta}$	两边取对数 $\log y=\log\alpha+\beta\log x$ $Y=\log y \quad X=\log x \quad A=\log\alpha$	$\boldsymbol{Y=A+\beta X}$
（2） $y=\alpha\cdot\beta^{x}$	两边取对数 $\log y=\log\alpha+x\log\beta$ $Y=\log y \quad A=\log\alpha \quad B=\log\beta$	$\boldsymbol{Y=A+Bx}$
（3） $y=\alpha+\dfrac{\beta}{x}$	设 $X=\dfrac{1}{x}$	$\boldsymbol{y=\alpha+\beta X}$
（4） $y=\dfrac{1}{\alpha+\beta x}$	设 $Y=\dfrac{1}{y}$	$Y=\alpha+\beta x$
（5） $y=\alpha+\beta\log x$	设 $X=\log x$	$y=\alpha+\beta X$

非线性式	变换	线性式
（6）2次函数 $y=\alpha+\beta x+\gamma x^2$	设 $z=x^2$	$y=\alpha+\beta x+\gamma z$
（7）科布-道格拉斯函数 $z=\alpha\cdot x^{\beta}\cdot y^{\gamma}$	两边取对数 $\log z=\log\alpha+\beta\log x+\gamma\log y$ 设 $Z=\log z \quad A=\log\alpha$ $X=\log x \quad Y=\log y$	$Z=A+\beta X+\gamma Y$
（8）逻辑函数 I $y=\dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$	设 $\dfrac{1}{y}=\dfrac{1}{e^{\alpha+\beta x}}+1\Rightarrow e^{\alpha+\beta x}=\dfrac{y}{1-y}$ $Y=\ln\dfrac{y}{1-y}$	$Y=\alpha+\beta x$
（9）逻辑函数 II $y=\dfrac{\gamma}{1+e^{\alpha+\beta x}}$	设 $1+e^{\alpha+\beta x}=\dfrac{\gamma}{y}\Rightarrow$ $Y=\ln(\dfrac{\gamma}{y}-1)$	$Y=\alpha+\beta x$

- 例1 书86页 第6题

6. 某批发公司1975~1983年的销售额如下表所示，试预测1999年的销售额。

年份	1990	1991	1992	1993	1994	1995	1996	1997	1998
销售额（万元）y	15 810	17 618	20 824	22 342	25 014	35 721	44 068	63 920	75 763

（提示：拟合指数曲线 $\hat{y}=\alpha\beta^{x}$）

解： $\ln y=\ln\alpha+\ln\beta\cdot x$，令 $y'=\ln y\ \ a'=\ln\alpha\quad b'=\ln\beta\Rightarrow y'=a'+b'x$

x_i	1	2	$\cdots$	8	9	$\sum$
年份	1990	1991	$\cdots$	1997	1998	
y_i	15810	17618	$\cdots$	63920	75763	
$y_i'=\ln y_i$	9.6684	9.7767	$\cdots$	11.0654	11.2354	93.0081
x_i^2	1	4	$\cdots$	64	81	285
$y_i'^2$	93.4779	95.5834	$\cdots$	122.4428	126.2334	

$$\hat{b}'=\frac{\sum_{i=1}^{n}x_iy_i'-\bar{x}\sum_{i=1}^{n}y_i'}{\sum_{i=1}^{n}x_i^2-\bar{x}\sum_{i=1}^{n}x_i}=\frac{477.1430-\frac{45}{9}\times 93.0081}{285-\frac{45}{9}\times 45}=0.2017$$

$$\hat{a}'=\frac{93.0081}{9}-\hat{b}\cdot\frac{45}{9}=9.3257$$

$$\hat{\alpha}=e^{9.3257}=11223$$

$$\hat{\beta}=e^{0.2017}=1.2235$$

$$\hat{y}=11223\times 1.2235^{x}$$

1999年为： $\hat{y}=11223\times 1.2235^{10}=84362$

作业：86 页 3