山西大学学报(自然科学版)16(1): 9－13, 1993
Journal of Shanxi University (Nat. Sci. Ed.)

马尔可夫链二项模型的一个注记

张信东　刘维奇

(山西大学数学系)

摘要　Rudolfer S. M[1]提出马尔可夫链二项模型,以拟合某些n次相依试验中成功的次数,本文指出它的不足之处,给出正确的方差表达式和参数的矩估计量,并对n=3情况作了马尔可夫链二项模型和其它模型的比较.

关键词　马尔可夫链二项模型,方差函数,矩估计

中图分类号: O211. 62

0　引言

二项模型是用于拟合或描述n重Bernolli试验(见[2]),但是实际当中许多数据不能满足n重Bernolli试验中独立性的假定,这必然影响拟合的精度。为此曾有许多学者引入β－二项模型[3],相关二项模型[4],可加二项模型和可乘二项模型[5]以拟合这类数据。最近Rudolfer又讨论了马尔可夫链二项模型。

设$Z_1, Z_2, \cdots$是0－1状态马尔可夫链,具有转移概率矩阵

$$P=\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

其中$0<\alpha<1, 0<\beta<1$。令

$$X_n=\sum_{i=1}^{n} Z_i \quad n>1$$

可将X_n考虑为n次试验中成功的次数,成功和失败的概率分别为

$$p=\alpha/(\alpha+\beta) \qquad q=\beta/(\alpha+\beta)$$

这里不同于n重Bernolli试验,$Z_1, Z_2, \cdots$不具有独立性。假定Z_i服从参数为p的Bernolli分布,即

$$Z_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

易见这样的马尔可夫链$Z_1, Z_2, \cdots$是平稳序列,且有

引理　满足上述假定的马尔可夫链$\{Z_i\}$具有如下均值函数和相关函数:

$$EZ_i=\frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}(Z_i)=\frac{\alpha\beta}{(\alpha+\beta)^2} \quad i=1, 2, \cdots \tag{0.1}$$

收稿日期: 1992. 04. 28

$$r_{ij}=\frac{Cov(Z_i,Z_j)}{\sqrt{Var(Z_i)Var(Z_j)}}=(1-\alpha-\beta)^{|j-i|}\qquad i,j=1,2,\cdots$$

证明: (0.1)是显然的,下面证明(0.2)式,由于

$$Cov(Z_i,Z_j)=EZ_iZ_j-(EZ_i)(EZ_j)=P(Z_i=1,Z_j=1)-p^2=p\cdot p_{11}(|j-i|)-p^2 \tag{0.3}$$

其中 $p_{11}(|j-i|)$ 为由状态 1 到状态 1 的 $|j-i|$ 步转移概率。而

$$P=\begin{pmatrix}1-\alpha & \alpha\\ \beta & 1-\beta\end{pmatrix}=I+\begin{pmatrix}-\alpha & \alpha\\ \beta & -\beta\end{pmatrix}=I+\begin{pmatrix}\alpha\\ -\beta\end{pmatrix}(-1\ \ 1)\triangleq I+A$$

注意到 $A^2=-(\alpha+\beta)A\triangleq(\delta-1)A$, $\delta=1-\alpha-\beta$

$$P^K=(I+A)^K=I\cdot\sum_{i=0}^{K}\binom{K}{i}A^i=I+\sum_{i=1}^{K}\binom{K}{i}(\delta-1)^{i-1}A=I+\frac{(\delta^K-1)}{\delta-1}A$$

于是

$$p_{11}(|j-i|)=1+\frac{\delta^{|j-i|}-1}{\alpha+\beta}\beta \tag{0.4}$$

代入(0.3)式得

$$Cov(Z_i,Z_j)=\frac{\alpha}{\alpha+\beta}\left(1+\frac{\delta^{|j-i|}-1}{\alpha+\beta}\beta\right)-\frac{\alpha^2}{(\alpha+\beta)^2}=\frac{\alpha\beta}{(\alpha+\beta)^2}\delta^{|j-i|} \tag{0.5}$$

联合(0.1)即得(0.2)式,引理得证。

文献[1]中给出了 X_n 的概率函数

$$P(X_n=k)=\frac{\alpha^k(1-\alpha)^{n-1-2k}\beta^{k+1}}{\alpha+\beta}S(n,k,\alpha,\beta).\quad k=1,2\cdots n \tag{0.6}$$

其中

$$S(n,k,\alpha,\beta)=\begin{cases}1 & \text{当 } k=0\\ \sum_{r=0}^{k-r}[\psi_{00}(n,k,r)+2\left(\frac{1-\alpha}{\beta}\right)\psi_{01}(n,k,r)+\left(\frac{1-\alpha}{\beta}\right)^2\psi_{11}(n,k,r)]\cdot[\frac{(1-\alpha)(1-\beta)}{\alpha\beta}]^r & \text{当 } 0<k<n\\ (1-\alpha)^{n+1}(1-\beta)^{n-1}/(\alpha^{n-1}\beta^{n+1}) & \text{当 } k=n\end{cases}$$

而 $\psi_{ij}(n,k,r)$ 表示从 i 开始,j 结束长为 n,含有 k 个 1 且其中相连 r 对 1 的 0-1 序列数,

$$\psi_{ij}(n,k,r)=\begin{cases}1 & \text{当 } k=r=i=j=0 \text{ 或 } k=n=r+1,i=j=1\\ \binom{k-1}{r}\binom{n-k-1}{k-r-i-j} & \text{其它}\end{cases}$$

若取 $\alpha+\beta=1$

$$P(X_n=k)=\binom{n}{k}\alpha^k(1-\alpha)^{n-k}\qquad k=1,2\cdots$$

这说明二项模型是马尔可夫链二项模型的特例.

文献[1]中引理 4 给出错误的方差表达式,从而得出错误的结论: 文献[1]中引理 5、定理 2 及矩估计表达式。下面给出正确的结论。

1　马尔可夫链二项模型的方差计算及参数的矩估计

定理 1　X_n 的均值和方差分别为

$$EX_n = n\alpha/(\alpha+\beta) \tag{1.1}$$

$$Var(X_n) = \frac{n\alpha\beta}{(\alpha+\beta)^2} + \frac{2\alpha\beta}{(\alpha+\beta)^4}[\delta^{n+1} - n\delta^2 + (n-1)\delta] \tag{1.2}$$

且注意到 $|\delta|<1$。

证明　根据引理及 X_n 的定义

$$EX_n = E\sum_{i=1}^{n} Z_i = n\alpha/(\alpha+\beta)$$

$$Var(X_n) = \sum_{i=1}^{n} Var(Z_i) + 2\sum_{i<j} Cov(Z_i, Z_j) = \frac{n\alpha\beta}{(\alpha+\beta)^2} + 2\sum_{i<j} \frac{\alpha\beta}{(\alpha+\beta)^2}\,\delta^{j-i}$$

注意到

$$\sum_{i<j}\delta^{j-i} = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\delta^{j-i} = \sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\delta^{k} = \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}\delta^{k} = \sum_{k=1}^{n-1}\delta^{k} - \sum_{k=1}^{n-1} k\,\delta^{k} = \frac{\delta^{n+1} - n\delta^2 + (n-1)\delta}{(1-\delta)^2}$$

故

$$Var(X_n) = \frac{n\alpha\beta}{(\alpha+\beta)^2} + \frac{2\alpha\beta}{(\alpha+\beta)^4}\delta(\delta^n - n\delta + n - 1)$$

定理证毕。

二项分布 B(n,p) 的方差 $\sigma_B^2 = npq = n\alpha\beta/(\alpha+\beta)^2$，马尔可夫链二项模型方差与 σ_B^2 比较决定于下式的符号

$$h_n(\delta) = \delta(\delta^n - n\delta + n - 1) \tag{1.3}$$

由于

$$h_n(\delta) = \delta(\delta^n - 1 - n\delta + n) = \delta(\delta-1)(\delta^{n-1} + \delta^{n-2} + \cdots + \delta + 1 - n) \tag{1.4}$$

根据 $|\delta|<1$ 有 $\delta - 1 < 0$，$\delta^{n-1} + \delta^{n-2} + \cdots + \delta + 1 - n < 0$，于是有如下结论

定理 2　当 $-1<\delta<0$，　$Var(X_n) < \sigma_B^2$ (1.5)

当 $\delta = 0$，　$Var(X_n) = \sigma_B^2$ (1.6)

当 $0<\delta<1$，　$Var(X_n) > \sigma_B^2$ (1.7)

并且当 $-1<\delta<0$ 时 $Var(X_n)$ 随 δ 单调递增。

事实上 (1.5) ～ (1.7) 由 (1.2) 和 (1.4) 得到，当 $-1<\delta<0$ 时考虑

$$h_n'(\delta) = (n+1)\delta^n - 2n\delta + n - 1 = \delta[(n+1)\delta^{n-1} - 2n] + n - 1$$

由于 $(n+1)\delta^{n-1} - 2n < 0$，从而 $h_n'(\delta) > 0$，故当 $-1<\delta<0$ 时 $Var(X_n)$ 随 δ 单调递增。

文献 [1]指出马尔可夫链二项模型的方差和二项模型的方差的关系与试验次数 n的奇偶性有关，从刚才定理 2 可见实际上它们方差的关系与 n 的奇偶性无关。

为了拟合模型，估计参数，设样本 $(X_1, X_2, \cdots, X_m)$，样本容量为 m，其可能取值为 $(K_1, K_2, \cdots, K_m)$，记样本均值 $\bar{X} = (1/m)\sum_{i=1}^{m} X_i$，样本方差 $S^2 = \frac{1}{m}\sum_{i=1}^{m}(X_i - \bar{X})^2$，则利用 (0.6) 式[1]给出参

数(α,β)的对数似然函数

$$L(\alpha,\beta)=\sum_{k=1}^{n} f_k\{k\log\alpha+(n-1-2k)\log(1-\alpha)+(k+1)\log\beta-\log(\alpha+\beta)+\log[S(n,k,\alpha,\beta)]\} \quad (1.8)$$

其中f_k表示样本中k的频数。但要求出参数的最大似然估计量比较困难，而需借助于数值计算方法。利用定理1不难得到参数的矩估计量

定理3 马尔可夫链二项模型参数的矩估计量为

$$\hat{\alpha}=\frac{\bar{X}}{n}(1-\hat{\delta}),\ \hat{\beta}=(1-\frac{\bar{X}}{n})(1-\hat{\delta}) \quad (1.9)$$

其中$\hat{\delta}$为如下方程在$(-1,1)$内的解

$$\delta^{n+1}-(D+n)\delta^2+(n-1+2D)\delta-D=0 \quad (1.10)$$

而$D=\frac{n^2S^2}{2\bar{X}(n-\bar{X})}-\frac{n}{2}$。

关于方程(1.10)在$(-1,1)$解的存在唯一性以及矩估计的渐近性质我们另有文章讨论。

文献[1]中引用了Skellam的337个Brassica数据(见文献[3])对马尔可夫链二项模型同其它模型做了比较。文献[1]的结果指出矩估计效果不好，其原因是主要结果错误所致。我们给出正确结论，说明矩估计效果很好，并不比最大似然估计差，这也给我们实际计算提供了方便。

当$n=3$时，X_n的概率函数为

$$P(X_n=k)=\begin{cases}(1-\alpha)^2\beta/(\alpha+\beta) & k=0\\ \alpha\beta[2[1-\alpha]+\beta]/(\alpha+\beta) & k=1\\ \alpha\beta[\alpha+2(1-\beta)]/(\alpha+\beta) & k=2\\ \alpha(1-\beta)^2/(\alpha+\beta) & k=3\end{cases} \quad (1.11)$$

根据下表数据可计算得：$\bar{X}=1.7418$, $S^2=0.8562$, $D=0.2580$代入(1.10)解之得$\hat{\delta}=0.1216$. 再由(1.9)得出$\hat{\alpha}=0.5100$, $\hat{\beta}=0.3684$. 转移概率矩阵

$$\hat{P}=\begin{pmatrix}0.4900 & 0.5100\\ 0.3684 & 0.6316\end{pmatrix}$$

$Z_1\sim\begin{pmatrix}0 & 1\\ 0.4194 & 0.5806\end{pmatrix}$. 其拟合效果见下表。

附表：对Skellam Brassica数据各种模型拟合结果

k	0	1	2	3	E(X)	Var(X)
f_k	32	103	122	80	1.742	0.859
可加二项模型	33.89	97.63	126.25	97.18	1.745	0.860
可乘二项模型	33.43	97.06	128.63	77.88	1.745	0.850
$\beta-$二项模型	33.96	97.18	127.69	78.17	1.742	0.857
马尔可夫二项模型(矩估计)	33.94	97.20	127.81	78.05	1.742	0.793
马尔可夫二项模型(最大似然估计)	33.98	97.13	127.73	78.16	1.742	0.857
二项模型	24.86	103.24	142.93	65.96	1.742	0.731

参 考 文 献

1 Rudolfer S M. A Markov chain model of extrabinomial Variation, Biometrika, 1990, 77(2): 255～264

2 复旦大学。概率论(第一册)。北京:人民教育出版社 1979, 75～78

3 Skellam J G. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. J. R. Statist soc B, 1948, 10, 257～267

4 Kupper L L, Haseman J K. The use of a correlated binomial model for the analysis of certain toxicological experiments. Biometrics, 1978, 34: 69～76

5 Altham P M E. Two generalizations of the binomial distribution. Appl. Statist 1978, 27: 162～167

A NOTE ON THE MARKOV BINOMIAL MODEL

Zhang Xingdong　Liu Weiqi

(Department of Mathematics, Shanxi University)

Abstract

Rudolfer S, M.[1] proposed the Markov chain binomial model to fit the number of successes in n, not necessarily independent, trials. In this paper, the faults of Rudolfer's model are indicated and the correct expression for the variance function and correct moment estimate of the parameters are given. Also, for n=3, the Markov chain binomial model is compared with the other models.

Key words Markov chain binomial model, variance function, moment estimate