情報・システム工学概論

統計モデルの数理

— 第２回：強さをはかる統計モデル —

駒木　文保

工学部 計数工学科

2018 年 11 月 5 日

Bradley–Terry モデル (BT モデル)

野球，サッカー，相撲，囲碁・将棋などの勝敗データ

プレイヤーの “強さ” を数値化して将来の結果を予測できる.

スポーツの統計学

詳しくは，竹内・藤野 (1988) などを参照.

BT モデルを基本とした，さまざまな拡張が提案され続けている.

２人のプレイヤーの場合

プレイヤー A, B

A, B **の強さ**

$$\pi_{\mathsf{A}}, \pi_{\mathsf{B}} \ \in (0, \infty)$$

A **と** B **２人の対戦で** A **が勝つ確率**

$$p_{\mathsf{AB}} = \frac{\pi_{\mathsf{A}}}{\pi_{\mathsf{A}} + \pi_{\mathsf{B}}}.$$

A **が勝つ確率と** B **が勝つ確率の和**

$$p_{\mathsf{AB}} + p_{\mathsf{BA}} = \frac{\pi_{\mathsf{A}}}{\pi_{\mathsf{A}} + \pi_{\mathsf{B}}} + \frac{\pi_{\mathsf{B}}}{\pi_{\mathsf{A}} + \pi_{\mathsf{B}}} = 1$$

例. A と B が同じ強さ

$$\pi_{\mathrm{A}} = \pi_{\mathrm{B}} = 1$$

A が勝つ確率

$$p_{\mathrm{AB}} = \frac{\pi_{\mathrm{A}}}{\pi_{\mathrm{A}} + \pi_{\mathrm{B}}} = \frac{1}{1+1} = \frac{1}{2}.$$

例. A が強く，B が弱い

$$\pi_A = 10, \quad \pi_B = 0.1$$

A が勝つ確率

$$p_{AB} = \frac{\pi_A}{\pi_A + \pi_B} = \frac{10}{10 + 0.1} = \frac{100}{101} \simeq 0.990099.$$

N 人のプレイヤーの場合

プレイヤー $1, 2, \ldots, N$ **の強さ**

$$\pi_1, \pi_2, \ldots, \pi_N.$$

プレイヤー i **とプレイヤー** j **が対戦して** i **が勝つ確率**

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

N 人全員のプレイヤーの強さを $c > 0$ 倍しても意味は変わらない.

強さ

$$c\pi_1, c\pi_2, \ldots, c\pi_N.$$

プレイヤー i とプレイヤー j が対戦して i が勝つ確率

$$p_{ij} = \frac{c\pi_i}{c\pi_i + c\pi_j} = \frac{\pi_i}{\pi_i + \pi_j}.$$

強さのパラメータには定数倍の不定性がある.

BT モデルの限界

"苦手" は表現できない.

じゃんけんのような関係は表せない.

しかし，近似モデルとしては有効なことが多い.
全てのモデルは近似.

３人のプレイヤーがじゃんけんの石 G，はさみ C，紙 P だったら，

$$p_{\mathrm{GC}} p_{\mathrm{CP}} p_{\mathrm{PG}} = 1,$$
$$p_{\mathrm{CG}} p_{\mathrm{GP}} p_{\mathrm{PC}} = 0.$$

したがって，

$$p_{\mathrm{GC}} p_{\mathrm{CP}} p_{\mathrm{PG}} \neq p_{\mathrm{CG}} p_{\mathrm{GP}} p_{\mathrm{PC}}.$$

BT モデルの場合

勝敗の結果として３すくみがおこる確率

$$p_{ik}p_{kj}p_{ji} = \frac{\pi_i}{\pi_i + \pi_k}\frac{\pi_k}{\pi_k + \pi_j}\frac{\pi_j}{\pi_j + \pi_i} = \frac{\pi_i\pi_j\pi_k}{(\pi_i + \pi_j)(\pi_j + \pi_k)(\pi_k + \pi_i)}.$$

逆の向きの３すくみがおこる確率

$$p_{ki}p_{ij}p_{jk} = \frac{\pi_k}{\pi_k + \pi_i}\frac{\pi_i}{\pi_i + \pi_j}\frac{\pi_j}{\pi_j + \pi_k} = \frac{\pi_i\pi_j\pi_k}{(\pi_i + \pi_j)(\pi_j + \pi_k)(\pi_k + \pi_i)}.$$

したがって

$$p_{ik}p_{kj}p_{ji} = p_{ki}p_{ij}p_{jk}.$$

３すくみがないことを,

$$p_{ij}p_{jk}p_{ki} = p_{ji}p_{ik}p_{kj}$$

が常に成立することと定義する.

“３すくみ” が無いことから BT モデルが導ける.

３すくみが無いとき，すべての k に対し，

$$\frac{p_{ji}}{p_{ij}} = \frac{p_{jk}p_{ki}}{p_{ik}p_{kj}}$$

が成立．ここで，

$$\pi_1 := 1, \quad \pi_j := \frac{p_{j1}}{p_{1j}} \quad (j \neq 1)$$

とおく．すると，

$$\frac{p_{j1}}{p_{1j}} = \pi_j = \frac{\pi_j}{\pi_1}.$$

また，$i \neq 1$, $j \neq 1$ のとき，

$$\frac{p_{ji}}{p_{ij}} = \frac{p_{j1}p_{1i}}{p_{i1}p_{1j}} = \frac{\pi_j}{\pi_i}.$$

したがって，すべての i, j に対し，

$$\frac{p_{ji}}{p_{ij}} = \frac{p_{j1}p_{1i}}{p_{i1}p_{1j}} = \frac{\pi_j}{\pi_i}.$$

が成立する．
また，

$$\frac{p_{ji}}{p_{ij}} = \frac{1 - p_{ij}}{p_{ij}} = \frac{1}{p_{ij}} - 1$$

であるから，

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

これは BT モデルに他ならない．

パラメータの推定法

前回触れた最尤推定を２項分布の例で説明.
本質的な考え方は他のモデルでも同様.

例. ２項分布の確率関数 (θ の値を与えたもとで x の関数と見る)

$$P(x;\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}.$$

$P(x;\theta)$ を x の値を与えたもとで θ の関数と見るとき,
尤度（ゆうど）関数とよぶ.

尤度関数の対数

$$\log P(x;\theta) = \log\binom{n}{x} + x\log\theta + (n-x)\log(1-\theta).$$

を対数尤度関数と呼ぶ.

対数尤度をパラメータで偏微分して 0 とおいて得られる方程式

$$\frac{\partial}{\partial\theta}\log P(x;\theta) = \frac{x}{\theta} - \frac{n-x}{1-\theta} = 0$$

を尤度方程式と呼ぶ.

最尤推定値

$$\hat{\theta}(x) = \frac{x}{n}.$$

は尤度方程式を解いて得られる.

推定量は ^ (ハット) をつけて表すことが多い.

プレイヤーが２人（A と B）の場合.

π_{A}: A **の強さ**,　π_{B}: B **の強さ**

$$\theta := \frac{\pi_{\mathsf{A}}}{\pi_{\mathsf{A}} + \pi_{\mathsf{B}}}$$

とおけば２項分布モデルの推定と本質的に同じ.

２人のプレイヤーの強さが $\pi_{\mathrm{A}}, \pi_{\mathrm{B}}$ であるのと $c\pi_{\mathrm{A}}, c\pi_{\mathrm{B}}, (c > 0)$ であるのは同じことなので,

$$\pi_{\mathrm{A}} + \pi_{\mathrm{B}} = 2$$

という制約をつける.

制約をつけることにより, $\pi_{\mathrm{A}}, \pi_{B}$ の値が一意に決まる.

強さ π_{A} が 1 なら A と B は同じ強さ,

π_{A} が 1 より大きければ, A は B より強い.

プレイヤーが２人の時のパラメータ推定は，θ の最尤推定値 $\hat{\theta}$ を求めてから

$$\frac{\hat{\pi}_{\mathsf{A}}}{\hat{\pi}_{\mathsf{A}} + \hat{\pi}_{\mathsf{B}}} = \hat{\theta},$$

$$\hat{\pi}_{\mathsf{A}} + \hat{\pi}_{\mathsf{B}} = 2$$

を満たすように $\hat{\pi}_1$, $\hat{\pi}_2$ を求めればよい．

すると，

$$\hat{\pi}_{\mathsf{A}} = 2\hat{\theta}, \quad \hat{\pi}_{\mathsf{B}} = 2(1-\hat{\theta})$$

となる．

一般の場合のパラメータ推定

N 人のプレイヤー

n_{ij}: i と j の勝負の数 $(n_{ij} = n_{ji})$

x_{ij}: i が j に勝った数 $(x_{ij} = n_{ji} - x_{ji})$

π_i: プレイヤー i の強さ

i と j が対戦した時 i が勝つ確率

$$\frac{\pi_i}{\pi_i + \pi_j}$$

i と j が n_{ij} 回対戦した時 i が x_{ij} 回勝つ確率

$$\binom{n_{ij}}{x_{ij}} \left(\frac{\pi_i}{\pi_i + \pi_j} \right)^{x_{ij}} \left(\frac{\pi_j}{\pi_i + \pi_j} \right)^{n_{ij} - x_{ij}}$$

全体の対戦の結果の確率

$$\prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} \left(\frac{\pi_i}{\pi_i + \pi_j} \right)^{x_{ij}} \left(\frac{\pi_j}{\pi_i + \pi_j} \right)^{n_{ij} - x_{ij}}$$

確率をパラメータ $\pi_1, \ldots, \pi_N$ の関数と見たものが尤度関数.

対数尤度 (パラメータ $\pi_1, \ldots, \pi_N$ の関数)

$x_{ji} = n_{ij} - x_{ij}$ **だから**

$$
\begin{aligned}
&\log \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} \left(\frac{\pi_i}{\pi_i + \pi_j}\right)^{x_{ij}} \left(\frac{\pi_j}{\pi_i + \pi_j}\right)^{n_{ij} - x_{ij}} \\
&= \log \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} (\pi_i + \pi_j)^{-n_{ij}} \pi_i^{x_{ij}} \pi_j^{x_{ji}} \\
&= \log \left[\left\{ \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} (\pi_i + \pi_j)^{-n_{ij}} \right\} \left(\prod_{i=1}^{N} \prod_{j:j \neq i} \pi_i^{x_{ij}} \right) \right] \\
&= \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \log \binom{n_{ij}}{x_{ij}} - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij} \log(\pi_i + \pi_j) + \sum_{i=1}^{N} \sum_{j:j \neq i} x_{ij} \log \pi_i .
\end{aligned}
$$

第 1 項はパラメータに関係がないので C とおき，
$T_i := \sum_{j:j\neq i} x_{ij}$ **とおくと，対数尤度関数は，**

$$C - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij} \log(\pi_i + \pi_j) + \sum_{i=1}^{N} T_i \log \pi_i .$$

$T_i \ (i = 1, \ldots, N)$ **は十分統計量．**

制約

$$\sum_{i=1}^{N} \pi_i = N.$$

ラグランジュの未定乗数法

ラグランジュ関数

$$C - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij} \log(\pi_i + \pi_j) + \sum_{i=1}^{N} T_i \log \pi_i - \lambda(\sum_{i=1}^{N} \pi_i - N)$$

を $\pi_1, \ldots, \pi_N, \lambda$ **で偏微分して** 0 **とおいて得られる式を解く.**

π_i **で偏微分して得られる式**

$$\frac{T_i}{\pi_i} - \sum_{j:j\neq i} \frac{n_{ij}}{\pi_i + \pi_j} - \lambda = 0 \tag{1}$$

λ **で偏微分して得られる式**

$$\sum_{i=1}^{N} \pi_i = N. \tag{2}$$

(1), (2) **を解けば良い.**

(1) より

$$T_i = \sum_{j:j\neq i} n_{ij} \frac{\pi_i}{\pi_i + \pi_j} + \lambda \pi_i.$$

左辺を i について和をとる．チーム i の勝数の i に関する和はゲームの総数に等しいから，

$$\sum_i T_i = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij}.$$

右辺を i について和をとる．

$$\sum_{i=1}^{N-1} \sum_{j:j\neq i} n_{ij} \frac{\pi_i}{\pi_i + \pi_j} + \lambda \sum_{i=1}^{N} \pi_i = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij} + \lambda \sum_{i=1}^{N} \pi_i.$$

したがって $\lambda = 0$.

したがって，(1), (2) の代わりに

$$\sum_{j:j\neq i} n_{ij}\frac{\pi_i}{\pi_i+\pi_j} = T_i$$

$$\sum_{i=1}^{N} \pi_i = N$$

を解けば良い.

この式は直観的にわかりやすい.
しかし，陽には解けないので数値的に解く必要がある.

ここでは簡便な反復法を用いる.

書き換え

$$\pi_i = \frac{T_i}{\displaystyle\sum_{j:j\neq i} n_{ij} \frac{1}{\pi_i + \pi_j}},$$

$$\sum_{i=1}^{N} \pi_i = N.$$

初期値 $\hat{\pi}_1^{(0)}, \ldots, \hat{\pi}_N^{(0)}$ を適当に設定.

$$
\begin{aligned}
\tilde{\pi}_i^{(1)} &= \frac{T_i}{\displaystyle\sum_{j:j\neq i} n_{ij} \frac{1}{\hat{\pi}_i^{(0)} + \hat{\pi}_j^{(0)}}}, \\
\hat{\pi}_i^{(1)} &= N \frac{\tilde{\pi}_i^{(1)}}{\sum_{i=1}^{N} \tilde{\pi}_i^{(1)}}.
\end{aligned}
$$

これを繰り返して $\hat{\pi}_i^{(1)}, \hat{\pi}_i^{(2)}, \ldots$ と更新して行くと

$$
\lim_{l\to\infty} \hat{\pi}_i^{(l)} = \hat{\pi}_i
$$

が成立することが知られている.

数値例

表．３人のプレイヤーの対戦結果．
プレイヤー i が プレイヤー j に勝利した回数 x_{ij}．
例えば， $x_{12} = 7$.

i \ j	1	2	3
1		7	8
2	3		5
3	2	5	

$$n_{12} = n_{13} = n_{23} = 10.$$

$$T_1 = x_{12} + x_{13} = 7 + 8 = 15,$$
$$T_2 = x_{21} + x_{23} = 3 + 5 = 8,$$
$$T_3 = x_{31} + x_{32} = 2 + 5 = 7.$$

以下を解けば最尤推定値が求まる.

$$\pi_1 = \frac{3}{2}\frac{1}{\frac{1}{\pi_1+\pi_2}+\frac{1}{\pi_1+\pi_3}},$$

$$\pi_2 = \frac{4}{5}\frac{1}{\frac{1}{\pi_2+\pi_1}+\frac{1}{\pi_2+\pi_3}},$$

$$\pi_3 = \frac{7}{10}\frac{1}{\frac{1}{\pi_3+\pi_1}+\frac{1}{\pi_3+\pi_2}},$$

$$\sum_{i=1}^{3}\pi_i = 3.$$

初期値 $\hat{\pi}_1^{(0)} = 1$, $\hat{\pi}_2^{(0)} = 1$, $\hat{\pi}_3^{(0)} = 1$.

$$
\begin{aligned}
\tilde{\pi}_1^{(1)} =& \frac{3}{2} \frac{1}{\frac{1}{1+1} + \frac{1}{1+1}} = \frac{3}{2}, \\
\tilde{\pi}_2^{(1)} =& \frac{4}{5} \frac{1}{\frac{1}{1+1} + \frac{1}{1+1}} = \frac{4}{5}, \\
\tilde{\pi}_3^{(1)} =& \frac{7}{10} \frac{1}{\frac{1}{1+1} + \frac{1}{1+1}} = \frac{7}{10}.
\end{aligned}
$$

和が 3 になるように正規化

$$\hat{\pi}_1^{(1)} = \frac{3\tilde{\pi}_1^{(1)}}{\tilde{\pi}_1^{(1)} + \tilde{\pi}_2^{(1)} + \tilde{\pi}_3^{(1)}} = \frac{3\frac{3}{2}}{\frac{3}{2} + \frac{4}{5} + \frac{7}{10}} = 30\frac{\frac{3}{2}}{\frac{30}{10}} = \frac{3}{2}.$$

同様に

$$\hat{\pi}_2^{(1)} = \frac{4}{5}, \quad \hat{\pi}_3^{(1)} = \frac{7}{10}.$$

1 回目の更新が終了.

以下収束するまで繰り返す.

$$\tilde{\pi}_1^{(2)} = \frac{3}{2}\frac{1}{\frac{1}{3/2+4/5} + \frac{1}{3/2+7/10}} = \frac{3}{2}\frac{1}{\frac{15+8}{10} + \frac{15+7}{10}} = \frac{3}{2}\frac{1}{\frac{45}{10}} = \frac{1}{3},$$

$$\vdots$$

収束の様子

l	$\hat{\pi}_1^{(l)}$	$\hat{\pi}_2^{(l)}$	$\hat{\pi}_3^{(l)}$				
0	1.000000	1.000000	1.000000				
1	1.500000	0.800000	0.700000	16	1.799039	0.644140	0.556821
2	1.665950	0.717395	0.616656	17	1.799043	0.644138	0.556819
3	1.735875	0.679141	0.584984	18	1.799045	0.644137	0.556818
4	1.768215	0.661194	0.570591	19	1.799046	0.644136	0.556818
5	1.783801	0.652558	0.563642	20	1.799046	0.644136	0.556818
6	1.791460	0.648323	0.560217	21	1.799047	0.644136	0.556818
7	1.795259	0.646225	0.558516	22	1.799047	0.644136	0.556818
8	1.797153	0.645180	0.557667	23	1.799047	0.644136	0.556818
9	1.798099	0.644658	0.557242	24	1.799047	0.644136	0.556817
10	1.798572	0.644397	0.557030	25	1.799047	0.644136	0.556817
11	1.798809	0.644267	0.556924	26	1.799047	0.644136	0.556817
12	1.798928	0.644201	0.556871	27	1.799047	0.644136	0.556817
13	1.798987	0.644169	0.556844	28	1.799047	0.644136	0.556817
14	1.799017	0.644152	0.556831	29	1.799047	0.644136	0.556817
15	1.799032	0.644144	0.556824	30	1.799047	0.644136	0.556817

実装

アルゴリズムの実装は難しくない.

簡単な例であれば電卓でも計算可能.

フリーの統計解析用プログラミング言語 R などを使うと容易.

R については多くの情報が RjpWiki などインターネットで入手できる．また，竹村 (2007) も参考になる.

Bradley–Terry モデルの赤池情報量規準（AIC）

定義

$$\mathrm{AIC} := -2 \times \textbf{最大対数尤度} + 2 \times \textbf{パラメータ数}.$$

Bradley–Terry **モデルのパラメータ数**: $N-1$

AIC:

$$\begin{aligned}
&-2 \log \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} \left(\frac{\hat{\pi}_i}{\hat{\pi}_i + \hat{\pi}_j} \right)^{x_{ij}} v \left(\frac{\hat{\pi}_j}{\hat{\pi}_i + \hat{\pi}_j} \right)^{n_{ij} - x_{ij}} + 2 \times (N-1) \\
&= -2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \log \binom{n_{ij}}{x_{ij}} + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} n_{ij} \log(\hat{\pi}_i + \hat{\pi}_j) \\
&\quad - 2 \sum_{i=1}^{N} \sum_{j:j \neq i} x_{ij} \log \hat{\pi}_i + 2(N-1).
\end{aligned}$$

その他のモデル 1 すべてのプレイヤーの強さが等しい

すべての対戦組合せ (i,j) $(i<j)$ **でプレイヤー** i **が勝つ確率が** $\frac{1}{2}$.

パラメータ数は 0.

AIC:

$$
\begin{aligned}
&-2\log\prod_{i=1}^{N-1}\prod_{j=i+1}^{N}\binom{n_{ij}}{x_{ij}}\left(\frac{1}{2}\right)^{x_{ij}}\left(\frac{1}{2}\right)^{n_{ij}-x_{ij}}+2\times 0\\
&=-2\log\prod_{i=1}^{N-1}\prod_{j=i+1}^{N}\binom{n_{ij}}{x_{ij}}\left(\frac{1}{2}\right)^{n_{ij}}\\
&=-2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\log\binom{n_{ij}}{x_{ij}}+(2\log 2)\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}n_{ij}.
\end{aligned}
$$

その他のモデル 2 フルモデル

すべての対戦組合せ (i,j) $(i<j)$ **についてプレイヤー** i **が** j **に勝つ確率** p_{ij} **をパラメータとするモデル.**

パラメータ数は $N(N-1)/2$**.**

パラメータ p_{ij} **の最尤推定値は** $\hat{p}_{ij} = \dfrac{x_{ij}}{n_{ij}}$.

AIC:

$$
\begin{aligned}
&-2\log \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \binom{n_{ij}}{x_{ij}} \hat{p}_{ij}^{x_{ij}} (1-\hat{p}_{ij})^{n_{ij}-x_{ij}} + 2 \times \frac{N(N-1)}{2} \\
&= -2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left\{ \log \binom{n_{ij}}{x_{ij}} + x_{ij} \log \frac{x_{ij}}{n_{ij}} + (n_{ij}-x_{ij}) \log \left(1 - \frac{x_{ij}}{n_{ij}}\right) \right\} \\
&\quad + N(N-1).
\end{aligned}
$$

参考文献

竹内啓，藤野和建 (1988) スポーツの数理科学，共立出版.

竹村彰通 (2007) 統計 第２版, 共立講座 21 世紀の数学 14, 共立出版.