

# Excel表計算による血液型時系列分析

日大生産工(学部) 茂木 渉  
日大生産工 篠原 正明

## 1. はじめに

2001年3月10日現在、日本人の血液型分布は田中克己氏らによる4,464,349名の集計結果によると、A型：38.7%、B型：22.2%、O型：29.2%、AB型：9.9%となっている。この比率は今後の時変動に対して一定なのか、あるいは血液型分布の変動は未だ過渡期にあり、これからも変動するのか。変動するのならば、いつ収束するのであるだろうか。それとも収束は起こらないのか。本論文では、これらの疑問を解消すべく、血液型分布変動の様子をMarkov Chain Modelとして定式化し、Microsoft Excelを用いて解析を試みる。

なお、一般的に血液型は上記の例のように表現型と呼ばれるA型、B型、O型、AB型の4つで表されるが、本研究における血液型とはAA型、AO型、BB型、BO型、OO型、AB型の6つで表される遺伝子型の血液型のことを指す。

## 2. マルコフ連鎖概説

「時点 $n$ で状態 $i$ にいたとき、時点 $n+1$ で状態 $j$ に推移する確率は、時点 $n-1$ 以前にどの状態にいたかは無関係である」という仮定をマルコフ性と呼び、マルコフ性を持った確率過程をマルコフ過程と呼ぶ。そして、マルコフ過程は連鎖的に $t$ 時間後の確率分布が求まることから、マルコフ連鎖とも呼ばれる。

また、推移図における推移の確率が時点 $n$ でも時点 $n-1$ でも全く同じであるような性質を推移確率の定常性と呼び、定常な推移確率を持つマルコフ連鎖のことを斉時的マルコフ連鎖と呼ぶ。本論文におけるマルコフ連鎖の推移確率は状態に依存し、時間的一様でない非斉時的マルコフ連鎖である。

## 3. 世代間血液型分布変動

2節で示したように、「未来の挙動が現在の値だけで決定され、過去の挙動と無関係で

ある状況」は全てマルコフ連鎖として表現できる。これより、ある夫婦間から生まれる子供の血液型はその親の血液型からのみ決定され、それ以前の世代の血液型分布には影響されないため、世代間血液型分布もマルコフ連鎖として取り扱うことができる。

まず、遺伝子型の血液型、AA型、AO型、BB型、BO型、OO型、AB型を6つの状態ととらえ、状態推移確率行列を作成する。そして、第 $i$ 世代における血液型分布ベクトルを状態確率ベクトル $x_i$ として計算を行う。

また、本研究においては1節に示されている血液型分布に近い、次の初期値を使用した。

表1. 実験に使用した初期値

AA	0.23
AO	0.15
BB	0.06
BO	0.17
OO	0.29
AB	0.10

## 4. ランダム婚

ランダム婚とは、ある血液型に偏ることなく、無作為に相手を選ぶことである。

ランダム婚の場合、例えば片方の親の血液型(これを親1と表現する)をAA型に固定して考えた時、生まれてくる子供の血液型はもう片方の親(これを親2と表現する)の血液型に依存して決定される。

その状態遷移図は図1のように表される。図中の矢印の横の、式中の血液型は親2の血液型である。

このようにして親1の全ての血液型に対して子供の血液型出現確率を求めると、表2のような遷移確率行列 $P$ が作成できる。表の縦軸は親1の血液型、横軸は子供の血液型、表中の血液型は親2の血液型を示している。

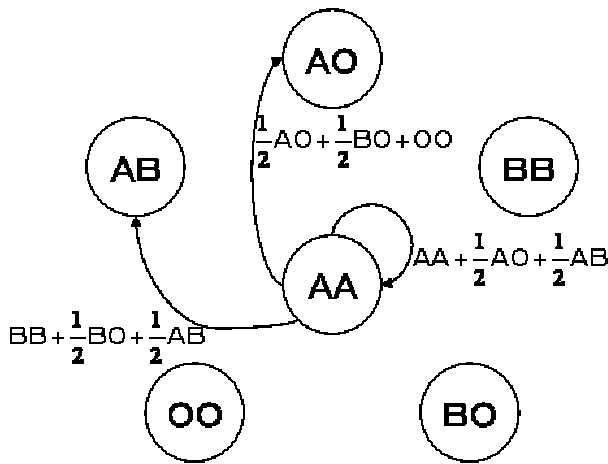


図1 . ランダム婚の状態遷移図

表2 . ランダム婚の遷移確率行列  $P$

	AA	AO
AA	$AA+(AB+AO)/2$	$OO+(AO+BO)/2$
AO	$AA/2+(AB+AO)/4$	$(AA+AO+OO)/2+(AB+BO)/4$
BB	0	0
BO	0	$AA/2+(AO+AB)/4$
OO	0	$AA+(AO+AB)/2$
AB	$AA/2+(AB+AO)/4$	$OO/2+(AO+BO)/4$

	BB	BO
AA	0	0
AO	0	$BB/2+(BO+AB)/4$
BB	$BB+(BO+AB)/2$	$OO+(AO+BO)/2$
BO	$BB/2+(BO+AB)/4$	$(BB+BO+OO)/2+(AO+AB)/4$
OO	0	$BB+(BO+AB)/2$
AB	$BB/2+(BO+AB)/4$	$OO/2+(AO+BO)/4$

	OO	AB
AA	0	$BB+(BO+AB)/2$
AO	$OO/2+(AO+BO)/4$	$BB/2+(BO+AB)/4$
BB	0	$AA+(AO+AB)/2$
BO	$OO/2+(AO+BO)/4$	$AA/2+(AO+AB)/4$
OO	$OO+(AO+BO)/2$	0
AB	0	$(AA+BB+AB)/2+(AO+BO)/4$

第  $i$  世代の各血液型比率ベクトル  $x_i$  に表 2 で得られた遷移行列  $P$  の転置行列  $P^T$  を左から掛けたものが、第  $(i+1)$  世代の各血液型を持つ人口の比率ベクトルに相当する。ただし、遷移行列  $P$  は状態確率ベクトル  $x_i$  に依存して決定されるので、 $P = P^T$  である。以上より、 $(i+1)$  世代の血液型出現確率  $x_{i+1}$  は次式で表される。

$$x_{i+1} = P_i^T x_i \quad (1)$$

シミュレーションの結果は表 3 のようになった。

表 3 . ランダム婚シミュレーション結果

	AA	AO	BB	BO	OO	AB
第0世代	0.23	0.15	0.06	0.17	0.29	0.1
第1世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第2世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第3世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第4世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第5世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第6世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1385
第7世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第8世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第9世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第10世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第11世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第12世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第13世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第14世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第15世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第16世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第17世代	0.126025	0.1755	0.038025	0.1755	0.2025	0.1384
第18世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第19世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384
第20世代	0.126025	0.3195	0.038025	0.1755	0.2025	0.1384

結果は第 1 世代で全て収束する、つまり血液型分布の変動は第 1 世代で収束するということである。これは (1) 式が、 $i+1$  世代が生まれた時点で  $i$  世代が消滅することになっていることが原因である。即ちこの結果は、「時点  $i+1$  に誕生する新生児の血液型分布は、 $i$  によらず一定である」ことを示している。

しかし、現実世界では子供が生まれた時点で親が死滅するわけではない。そこで、この点を改良したモデルを 5 節に示す。

## 5 . 出生率付きランダム婚

一般に出生率とは合計特殊出生数を指すことが多い。合計特殊出生数とは 15 歳から 49 歳までの女性の年齢別出生率を合計したもので、女性 1 人が仮にその年次の年齢別出生率のまま一生の間子供を産むとしたときの子供の数に相当するが、本来の出生率とは人口 1000 人に対するその年の 1 月 ~ 12 月までに生まれた子供の数を示した値である。本論文では、人口 1 人当たりが 1 年間に出生する割合、すなわち出生率を 1000 で割った値を  $\lambda$  とする。

4 節のランダム婚に、この出生率という概念を付加することにより「出産を行った個体はその瞬間に死滅する」ことを取り除き、世代間血液型分布の時変動モデルについて考察する。

$i+1$  世代に誕生する新生児の血液型分布ベクトル  $P_i^T x_i$  に  $\lambda$  を掛けることにより、新生児の血液型分布ベクトルが全人口に対して影響する割合がわかる。これと  $i$  世代の血液型分

布ベクトル  $x_i$  を足し、(2)式を得る。

$$y_i = x_i + P_i^T x_i \quad (2)$$

そして(2)式のベクトルの要素を合計した値が1になるように正規化したベクトルを、 $i+1$ 世代の血液型分布ベクトル  $x_{i+1}$  とする。

$$x_{i+1} = \frac{y_i}{\sum_{i=1}^6 y_i} \quad (3)$$

また、(2)(3)式を変形すると次のマルコフ連鎖と同値となる。

$$x_{i+1} = A_i^T x_i \quad (4)$$

$$A_i = \frac{I + P_i}{1 +} \quad (5)$$

ここで、 $1$  は1年に生まれる子供の数の割合なので、このモデルにおける1世代と1年は等価である。また  $1$  は2001年日本国内の平均出生率9.3を1000で割った0.0093とする。

シミュレーションの結果は以下のようになった。

表4 . 出生率付きランダム婚 結果表

	AA	AO	BB	BO	OO	AB
第0世代	0.23	0.15	0.06	0.17	0.29	0.1
第100世代	0.167225	0.252335	0.046733	0.173321	0.237172	0.123214
第201世代	0.1422	0.293131	0.041444	0.174644	0.216112	0.132468
第300世代	0.132494	0.308954	0.039392	0.175158	0.207944	0.136058
第400世代	0.128588	0.315321	0.038567	0.175364	0.204657	0.137502
第500世代	0.127041	0.317844	0.03824	0.175446	0.203355	0.138074
第593世代	0.126454	0.3188	0.038116	0.175477	0.202861	0.138291
第594世代	0.12645	0.318806	0.038115	0.175477	0.202858	0.138293
第595世代	0.126447	0.318813	0.038114	0.175478	0.202855	0.138294
第596世代	0.126443	0.318819	0.038113	0.175478	0.202851	0.138296
第597世代	0.126439	0.318825	0.038112	0.175478	0.202848	0.138297
第598世代	0.126435	0.318832	0.038112	0.175478	0.202845	0.138298
第599世代	0.126431	0.318838	0.038111	0.175479	0.202842	0.1383
第600世代	0.126427	0.318844	0.03811	0.175479	0.202839	0.138301

表4を見ると、どの血液型も600世代付近で小数点第3桁まで一致している。図2のグラフでも、収束していることが見て取れる。

つまり、表1で与えられた血液型分布の初期値から、600年後には血液型分布は収束をむかえらる。

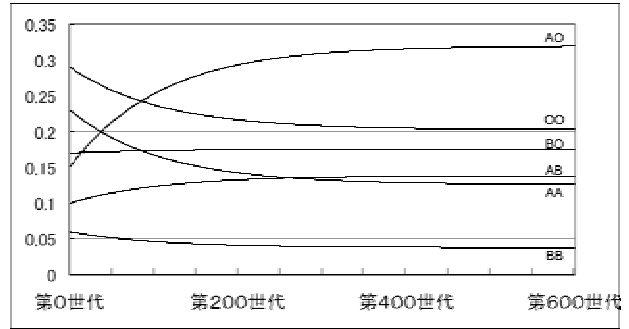


図2 . 出生率付きランダム婚 結果グラフ

## 6 . 現代婚

同じ血液型同士でしか結婚しないことを完全同型婚と呼び、完全同型婚と4節のランダム婚を組み合わせたものを現代婚と呼ぶ。

まず同型完全婚の場合、親1の血液型をAA型に固定した時の状態遷移図は、図3のように表される。このようにして子供の血液型出現確率を求めると、表5のような遷移確率行列  $Q$  が作成できる。

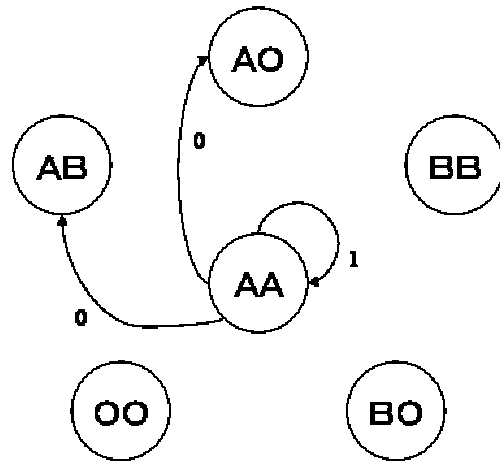


図3 . 同型完全婚の状態遷移図

表5 . 同型完全婚の遷移確率行列  $Q$

	AA	AO	BB	BO	OO	AB
AA	1	0	0	0	0	0
AO	1/4	1/2	0	0	1/4	0
BB	0	0	1	0	0	0
BO	0	0	1/4	1/2	1/4	0
OO	0	0	0	0	1	0
AB	1/4	0	1/4	0	0	1/4

これにより、現代婚の計算式は(6)(7)式、状態遷移図は図4、遷移確率行列は表6で表せる。 $R$ は現代婚、 $P$ はランダム婚、 $Q$ は同型完全婚である。また、 $m=1$ はランダム婚、 $m=0$ は完全同型婚と同値である。

$$x_{i+1} = R_i^T x_i \quad (6)$$

$$R_i = mP_i + (1-m)Q_i \quad (7)$$

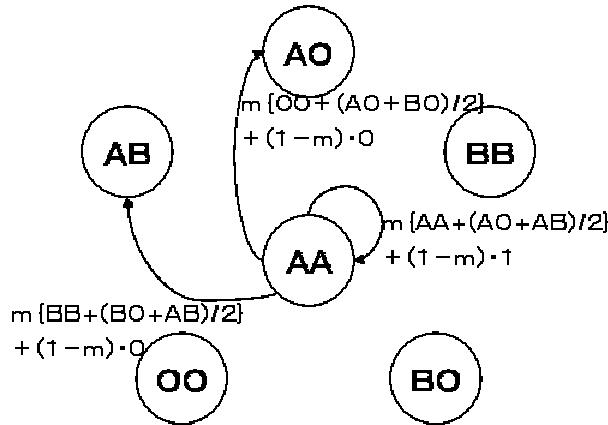


図4．現代婚の状態遷移図

表6．現代婚の遷移確率行列 R

	AA	AO
AA	$m[AA+(AB+AO)/2]+(1-m)·1$	$m[OO+(AO+BO)/2]+(1-m)·0$
AO	$m[AA/2+(AB+AO)/4]+(1-m)·1/4$	$m[(AA+AO+OO)/2+(AB+BO)/4]+(1-m)·1/2$
BB	$m·0+(1-m)·0$	$m·0+(1-m)·0$
BO	$m·0+(1-m)·0$	$m[AA/2+(AO+AB)/4]+(1-m)·0$
OO	$m·0+(1-m)·0$	$m[AA+(AO+AB)/2]+(1-m)·0$
AB	$m[AA/2+(AB+AO)/4]+(1-m)·1/4$	$m[OO/2+(AO+BO)/4]+(1-m)·0$

	BB	BO
AA	$m·0+(1-m)·0$	$m·0+(1-m)·0$
AO	$m·0+(1-m)·0$	$m[BB/2+(BO+AB)/4]+(1-m)·0$
BB	$m[BB+(BO+AB)/2]+(1-m)·1$	$m[OO+(AO+BO)/2]+(1-m)·0$
BO	$m[BB/2+(BO+AB)/4]+(1-m)·1/4$	$m[(BB+BO+OO)/2+(AO+AB)/4]+(1-m)·1/2$
OO	$m·0+(1-m)·0$	$m[BB+(BO+AB)/2]+(1-m)·0$
AB	$m[BB/2+(BO+AB)/4]+(1-m)·1/4$	$m[OO/2+(AO+BO)/4]+(1-m)·0$

	OO	AB
AA	$m·0+(1-m)·0$	$m[BB+(BO+AB)/2]+(1-m)·0$
AO	$m[OO/2+(AO+BO)/4]+(1-m)·1/4$	$m[BB/2+(BO+AB)/4]+(1-m)·0$
BB	$m·0+(1-m)·0$	$m[AA+(AO+AB)/2]+(1-m)·0$
BO	$m[OO/2+(AO+BO)/4]+(1-m)·1/4$	$m[AA/2+(AO+AB)/4]+(1-m)·0$
OO	$m[OO+(AO+BO)/2]+(1-m)·1$	$m·0+(1-m)·0$
AB	$m·0+(1-m)·0$	$m[(AA+BB+AB)/2+(AO+BO)/4]+(1-m)·1/2$

現代婚のシミュレーション結果は以下のようになった。

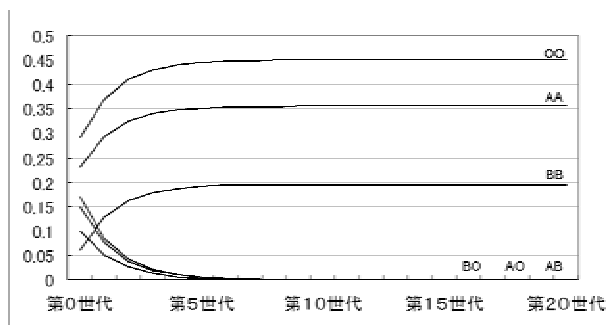


図5．現代婚(m=0) 結果グラフ

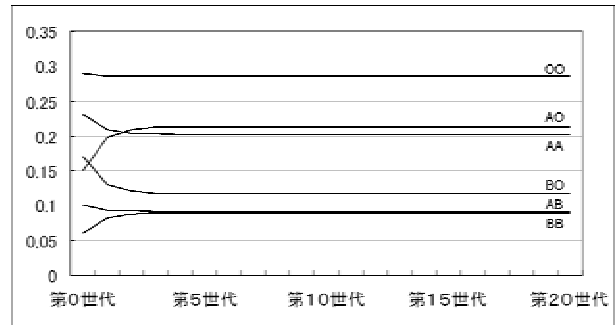


図6．現代婚(m=0.5) 結果グラフ

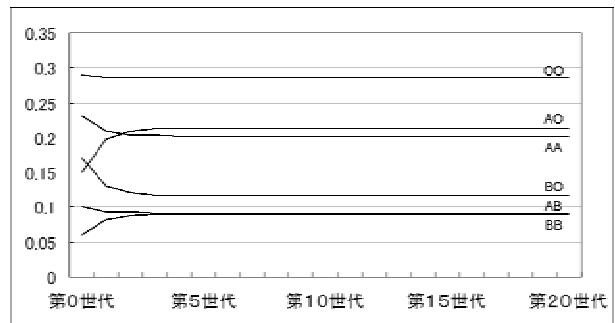


図7．現代婚(m=0.9) 結果グラフ

## 7. 終わりに

ランダム婚では出生率を付加しても最終結果は変わらないことから、血液型分布は時点  $i+1$  に誕生する新生児の血液型分布に収束すると考えられる。現代婚では  $m$  の値を小さくし、同型完全婚に近づけるとAO型、BO型、AB型が死滅するという結果が示された。

今後の課題としては、出生率の変動を予測したモデルなどの他のシミュレーションモデルの作成や血液型以外の解析への応用などである。

## 8. 参考文献

- [1] 岩楯健寛、鈴木洋臣、日高桂、播磨砂登美、窪川亮平、篠原正明：「状態確率依存マルコフ連鎖による血液型分布シミュレーションとHardy-Weinberg則」『第35回日本大学生産工学部学術講演会数理情報部会』(2002)、pp75-78
- [2] 内山貴夫、岩楯健寛、中野隼人、藤本聡、蜂須和則、篠原正明：「状態確率依存マルコフ連鎖による血液型分布時系列シミュレーション」『第35回日本大学生産工学部学術講演会数理情報部会』(2002)、pp79-82
- [3] 上野剛：『血液型時系列分析』日本大学生産工学部数理情報工学科 平成16年度 卒業研究論文