論　　文

ロト 6 分析 *

ガンマ・ポアソン分布の適用例

松尾　精彦

概要

2000 年 10 月 5 日の第一回抽選から始まった番号選択式の宝くじ LOTO6 も，去年 2015 年 9 月 7 日には第 1000 回目の抽選を迎えた．この宝くじは，購入者が 1 ～ 43 の数字の中から 6 個の数字の組合せを選び，主催者が用意した抽選器が抽出する 6 つの番号 (本選番号) のうち最低 3 個同じ番号が共有されれば賞金がもらえる．この抽選器を製作した業者や抽選器を操作する主催者が未熟であったり，不正を犯したりしない限り，この宝くじの本選番号を当てることは不可能である．ましてや，当選確率の高い組合せを見つけることも同様に不可能である．

それにもかかわらず，いくつかの雑誌や Web サイトで本選番号の予測法が述べられている．これらの予測法は，本選番号となりうるいずれの組合せも互いに独立で等しい確率で選ばれることを無視し，何らかの傾向があると主張するものである．大抵は，ある種の性質を有する組合せを選ぶと当選する確率が高くなると主張する．しかしどの方法も，その事象の確率が高いだけのことで，その事象が数多く生起しても何ら不思議はないものである．これらの予測法は多岐に渡り，事象の起こる確率を計算することが困難であるものがほとんどであるので，このような俗説についてここでは述べない．あえて言っておくと，3 回の異常な抽選を除いた 997 回の抽選で期待される 1 等の口数は 2606 口であるのに対し，実際の当選口数は 2521 口数でしかない．つまり，予想は役に立たないということだ．ここではあくまでも，抽選においては，どの本選番号も独立で同確率で選ばれることを仮定して計算を行う．

この論文で議論するのは，ロト 6 購入者の番号の選び方に傾向があるのかを調べることである．雑誌の情報などではなく，購買者集団が自らの意思で，どのような組合せを選ぶのかを調べることである．機械 (抽選器) には傾向はないが，人には傾向がある．その傾向を明らかにすれば，他の人々が選びそうな組合せを避けることができ，仮に 1 等が当たった場合には賞金を独り占めにする確率が高まる．

キーワード：ロト 6，一般化線形モデル，ガンマ・ポアソン分布，　過分散モデル
経済学文献季報分類番号：16-10

1　はじめに

ロト方式の宝くじで日本最初のものは 1999 年 4 月 13 日抽選のミニロトであり，約一年半後の 2000 年 10 月 5 日にロト 6 の抽選が初めて行われた．さらに，高額の賞金が期待できるロト 7 (それだけ当たる確率も低くなるが) も 2013 年 4 月 5 日に最初の抽選が行われている．これら 3 つの

* 本研究は，平成 28 年度関西大学研修員研究費によって行った．

ロトを区別しない場合は単にロトと書くことにする．ロトの売り上げ総額の 45% が賞金に回されることは他の宝くじと同じだが，購入者が自分自身で番号を選択することができる．そのため当選口数が変動し，賞金を当選口数で等分するため，賞金金額が抽選の度に変動することが醍醐味なのだろう．

これらのロトで話題になるのが，どうやって本選番号を当てるかということだ．しかし，ロトの抽選器は業者が，どの番号も同じ確率で独立に選ばれるように作っている．抽選の際決められた数の玉を攪拌し, ひとつずつ番号を抽出する抽選器は東京と大阪にひとつずつあり，また，1 ～ 43 までの番号のついた玉のセットは 10 揃いあり抽選毎にランダムに選ばれる．もしこの抽選器が，特定の番号を出しやすい，あるいは，ある番号は他の番号と一緒に出る傾向がある，といった事象が生ずるならば大きな問題である．もちろん，抽選器が特定の番号を出しやすいという傾向があっても，主催者側は損もしないし得もしない．賞金に回す金額は売上金額の 45% であることに変わりはない．だからといって，手を抜いているとは思えない．

それよりむしろ，購入者の間でどのような組合せに人気があるかということが問題なのだ．ロトでは各賞に割り当てる金額は毎回同じ割合であるが，その金額を何口で分配するかを心配すべきである．例えば 1 等賞金総額が 2 億円で当選口数が 1 口ならば賞金は 2 億円，しかし 10 口ならば 2 千万円になる．ここでは詳しく説明しないが，競馬で儲けを出すには人気馬ではなく，人気は無くても実力のある馬に賭けるべきである．ロトの場合も同様で，本選番号なぞ前もって分かりはしないのだから，どうせなら配分の多い，地味で人気のない組合せを探すべきなのである．

第 2 節では，まず，ロト 6 の仕組みについて説明する．第 3 節では, ロトでの 1 等口数の分布が, ガンマ・ポアソン分布に従っていないとは言えないことを示す．第 4 節では, 数字の組合せと当選口数との関係について述べ，どのような組合せをすればその組合せの入札口数が少なくなるかについて議論する．第 5 節では, データを一般化線形モデルの枠組みの中で分析し，その示唆するところを解説する．そしてどのような数字の組合せをしたら賞金金額が高くなるかについて述べる．

第 6 節では，抽選器がランダムな数字を選び出していることを示す．つまり，どの数字も等しい確率で選ばれ，互いに独立に選ばれることを示す．

2　ロト 6 とは

1 から 43 までの 43 個の数字のなかから 6 個の数字を選び，専用の申し込み用紙にマークして，宝くじ売り場等[*1] で購入する．このくじは他の宝くじ同様に 売上総額の 55% を差し引いた 45% が賞金に割り振られる．まず数字が 3 個一致する 5 等の組合せに固定額 1000 円を分配する．そして残りの額が 1 等から 4 等賞金に一定の割合で分配され，さらに当選口数で等分される．

抽選は同じ大きさ，同じ重さ，同じ摩擦係数の球を 43 個を抽選器に入れて回転させ，そこから互いに独立に等しい確率で 6 個抽出し（これを本選番号と呼ぶ），さらにもう 1 個ボーナス番号を抽出する．その後主催者はすべての番号が抽出器の中に入っていたことを，抽選に立ち会った人々

[*1] 最近は ATM やインターネットでも購入できる

を証人として明らかにする．結果として得られる本数字 6 個の組合せは全部で，$\binom{43}{6} = 6{,}096{,}454$ 通りあり，1 等は本数字 6 個がすべて一致した場合であり確率は $1/\binom{43}{6} = 1/6{,}096{,}454$, 2 等は本数字のうち 5 個一致し，残りの 1 個がボーナス数字に一致する場合であり，組合せの数は本数字を一つ取り去った組合せの数だから，$6/\binom{43}{6} = 6/6{,}096{,}454$, である．同様に 3 等は，本数字のうち 5 つが一致するもので，残りの 1 個が本数字でもボーナス数字でもない組合せの数だけ当たりがあるので，確率は，$\binom{6}{5}\binom{36}{1} \,/\, \binom{43}{6} = 216/6{,}096{,}454$ となる．4 等は，本選番号のうち 4 個が一致し，残りの 2 個は本数字以外の組合せだから，確率は，$\binom{6}{4}\binom{37}{2}/\binom{43}{6} = 9{,}990/6{,}096{,}454$. 最後の 5 等は，本数字が 3 個一致し，あとの 3 個は本選番号ではない組合せだから，確率は，$\binom{6}{3}\binom{37}{3}/\binom{43}{6} = 155{,}400/6{,}096{,}454$ となる．

そこで, 表 1 ([5] の中の表を用いた) の当選金額を用いて，一口 200 円を買うときの期待値を計算すると，

$$100{,}000{,}000 \times 1/6096454 + 15000000 \times 6/6096454 + 500000 \times 216/6096454 + 9500 \times 9990/6096454 + 1000 \times 155400/6096454 = 89.938 \quad (1)$$

となる．これは購入価格 200 円のほぼ 45% が賞金に割りふられることを示している．

普通ならば 200 円を出して 90 円しか返ってこない賭けなどしない．しかし，宝くじの場合，期待値で購入を考えるには大数の法則が働くまでに時間がかかりすぎる．大富豪ならば大金を使えるので，大数の法則が比較的早く働き，仮にロトに参加する大富豪が居たとしてもすぐに手を引くことが容易に推察される．資金さえあれば，人々は勝つ確率ではなく期待値を重視すべきである．

3　1 等当選口数の分布

分析をする前に，1000 回の抽選の中で明らかなアウトライアー 3 ケース (第 230 回，第 518 回，第６９１回) を除外する．まず，第 230 回だが 1 等当選者口数が 167 口もの応募があった．原因は，[7] に解説があったが，説明は表 2 を用いると分かりやすい．表の左上から斜め下に数字をたどってゆけば，第 230 回抽選の本選番号と一致する．この論文における分析では，雑誌等で推奨された等々の，他人が推奨する番号が 1 等（あるいは 2 等）となった抽選はとり除くことにする．つぎに，第 691 回だが，ここでは 1 等口数ではなく 2 等口数が異常に多い 3670 口であり，海外ドラマの『LOST』に現れる謎の数字の組合せらしい．最後は，第 518 回だが，ここでは 1 等当選口数が 0 なのにかかわらず，2 等当選口数が 330 にものぼっている．なぜこのような現象が起こったのかはっきりしないが，かなり異常なケースであってアウトライアーとして除外して分析を行う．

項 3.2 でみるように，ロト購入者は全ての組合せを等しい確率で購入するのではなく，組合せにより異なる確率で購入することが分かる．ある組合せは他の組合せよりも数多く選らばれる．もしもロト 6 の購入者すべてがクイックピックを使うならば，すべての数字の組合せが等しい確率で購入されることになる．すると，第 i 回抽選における 1 等当選者の数は，N_i を総購入口数とすれば，平均が $N_i/\binom{43}{6}$ のポアソン分布で近似できる．

この節では，まず，当選口数をポアソン分布では説明できないことを示したうえで，ポアソン分

表１　本数字 [06][11][23][31][39][43]　ボーナス数字 【17】のとき

等級	当せんパターン例	当せん金額（理論値）
1等	申込数字が本選番号に6個すべて一致［この1通りのみ］ [06][11][23][31][39][43]	約1億円
2等	申込数字6個のうち5個が本選番号に一致し、 さらに申込数字の残り1個がボーナス数字に一致 ［この6通りのみ］ [06][11][23][31][39]【17】, [06][11][23][31][43]【17】 [06][11][23][39][43]【17】, [06][11][31][39][43]【17】 [06][23][31][39][43]【17】, [11][23][31][39][43]【17】	約1,500万円
3等	申込数字6個のうち5個が本選番号に一致 ［この3通りを含め216通り］ [06][11][23][31][39] ●、 [06][11][23][31][43] ●、 [06][11][23][39][43] ●......	約50万円
4等	申込数字6個のうち4個が本選番号に一致 ［この3通りを含め9,990通り］ [06][11][23][31] ●●、 [06][11][23][39] ●●、 [06][11][31][39] ●●......	約9,500円
5等	申込数字6個のうち3個が本選番号に一致 ［この3通りを含め155,400通り］ [06][11][23] ●●●、 [06][11][31] ●●●、 [06][23][31] ●●●......	1,000円 (原則固定)

表２　第230回抽選以前の本選番号

回	抽選日	1	2	3	4	5	6	ボーナス	1等口数
222	2005-01-20	[4]	14	15	22	29	31	20	4
223	2005-01-27	7	[9]	19	26	30	32	3	4
224	2005-02-03	12	17	[18]	19	30	35	13	1
225	2005-02-10	9	11	21	[31]	35	38	13	4
226	2005-02-17	8	26	36	40	[42]	43	21	1
227	2005-02-24	6	13	17	29	32	[36]	35	9
228	2005-03-03	5	8	17	26	36	39	34	4
229	2005-03-10	8	16	21	24	27	36	20	3
230	2005-03-17	4	9	18	31	36	42	34	167

布をその構成分布として含む分布族で当てはめることをかんがえる．ここでは，その唯一の候補としてガンマ・ポアソン分布を考える．

3.1 ポアソン分布と負の二項分布

期待値 μ のポアソン分布の確率関数は，

$$f(y;\mu) = \frac{\mu^y}{y!}e^{-\mu}$$

であり，期待値も分散もともに μ である．同時密度関数は，N_i を第 i 回目の総売上数としたとき $\mu_i = N_i/\binom{43}{6}$ として，

$$\prod f(y_i;\mu_i) = \prod \frac{\mu_i^{y_i}}{y_i!}e^{-\mu_i}$$

となり尤度関数は，

$$l(\boldsymbol{\mu};\mathbf{y}) = \sum_i (y_i \log \mu_i - \mu_i - \log y_i!) \tag{2}$$

となる．

一方各 μ_i が独立で同一なガンマ分布，

$$g(\mu) = \frac{\beta^\alpha\ \mu^{\alpha-1}}{\Gamma(\alpha)}\exp(-\mu\beta),\ \mu = E[M] = \frac{\alpha}{\beta},\ V(M) = \frac{\alpha}{\beta^2} = \frac{\mu}{\beta}$$

に従うとすれば，

$$g(y;\alpha,\beta) = \frac{\Gamma(y+\alpha)}{\Gamma(\alpha)y!}\left(\frac{\beta}{1+\beta}\right)^\alpha \left(\frac{1}{1+\beta}\right)^y$$

ただし，$\alpha = \mu\beta$ であり, $\mu = E[Y] = \alpha/\beta,\ V[Y] = \mu(1+\frac{1}{\beta})$ である．α が正の整数ならば, 成功確率が $\dfrac{\beta}{1+\beta}$ のベルヌイ試行において α 回成功するまでの失敗の数 Y の分布であり，負の二項分布になる．ここでは，α が正の整数とは限らないので, 一般性を考慮してガンマ・ポアソン分布と呼ぶ．また，$\phi = (1+1/\beta)$ を過分散パラメータとも呼ぶこともある．尤度関数はこの場合，

$$l(\alpha,\beta;\mathbf{y}) = \sum_{i=1}\left(-y_i\log(1+\beta) + \alpha\log\left(\frac{\beta}{1+\beta}\right) + \log\Gamma(\alpha+y_i) - \log\Gamma(\alpha) - \log y_i!\right)$$

となるが，次のようにパラメトライズすることもできる．

$$l(\boldsymbol{\mu},\beta;\mathbf{y}) = \sum_{i=1}\left(-y_i\log(1+\beta) + \mu_i\beta\log\left(\frac{\beta}{1+\beta}\right) + \log\Gamma(\mu_i\beta+y_i) - \log\Gamma(\mu_i\beta) - \log y_i!\right) \tag{3}$$

3.2　ポアソン分布，ガンマ・ポアソン分布のあてはめ

まず，(2) には未知数がないので，尤度が計算でき，$l(\boldsymbol{\mu};\mathbf{y}) = -2159.33$ となる．この適合度を調べるために 1000 回のシミュレーションを行ったところ，最小が -1865.23 であり最大が -1739.64 であった．言い換えれば，1 等当選口数の分布は (2) のポアソン分布ではないことが容易に分かる．このことからロト購入者は数字の組合せを等しい確率で買ってはいない，つまり，組合せ毎に異なる確率で選んでいることがわかる．この現象を説明するのに真っ先に挙げられるのが，ガンマ・ポアソン分布である．

そこで，ガンマ・ポアソン分布をあてはめてみる．この場合 β のみが未知数なので β について (3) の尤度を最大化したところ，$\beta = 0.977816$ のときの -1995.68 であった．この β 値をもちいて，β と (3) の値を 1000 回シミュレートしてみた．すると，(3) の値が -1995.68 よりも小さくなる割合は，$\dfrac{517}{1000}$ となり，宝くじの当選口数はガンマ・ポアソン分布であるという帰無仮説は棄却されない．このシミュレーションの弱点は，β として推定値を用いた点にある．先ほどのシミュレーションで求められた 1000 個の β 値では，最小値が 0.739823, 最大値が 1.39336 であった．念のために β 値を，0.7 から 1.4 まで，間隔 0.1 ずつ変化させて，それぞれ 1000 回のシミュレーションを行ったときの p 値は表 3 のようになる．

表 3　さまざまな β 値に対する p 値の推定値

β	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4
p 値	$\frac{867}{1000}$	$\frac{764}{1000}$	$\frac{619}{1000}$	$\frac{512}{1000}$	$\frac{352}{1000}$	$\frac{246}{1000}$	$\frac{153}{1000}$	$\frac{80}{1000}$

表 3 から β 値が小さいほど有意確率が大きくなっていることがわかる．これはガンマ・ポアソン分布の平均と分散が，μ と $\mu(1+1/\beta)$ であるため，β が大きくなるほど分散が小さくなるからである．

さて，1 等口数が，$\beta = 0.977816$ のガンマ・ポアソン分布のとき，どれだけの確率で一等口数が 0 になるかを見てみよう．μ は 1 等口数の期待値であり, 総売上枚数を $\binom{43}{6} = 6096454$ で割ったものである. この μ を 1 から 8 まで変化したときの確率を表したのが表 4 である.

表 4　さまざまな μ 値に対する 1 等当選者が居ない確率の推定値

μ	1	2	3	4	5	6	7	8
確率	0.50494	0.25496	0.12874	0.065006	0.032824	0.01657	0.008369	0.004226

表 4 はあくまで概算であることを，心に留めておいて実際の結果をながめよう．997 回のデータをグループ分けして，期待一等口数と実際の一等口数の関係を調べてみよう．まず，期待一等口数

μ により 997 回の抽選を 4 つのグループに分ける; $0.5 \sim 1.5$, $1.5 \sim 2.5$, $2.5 \sim 3.5$, $3.5 \sim 4.5$, 代表値をそれぞれ，1,2,3,4 とする．これらのグループのにおける，一等口数が 0 の割合を調べてみた．すると，次のような表が得られる．

表 5　実際の抽選における一等口数 0 の割合

抽選のグループ分け	1	2	3	4
割合	0.337	0.217	0.103	0.074

期待一等口数が $1(0.5 \sim 1.5)$ の場合の一等口数が 0 の割合は 0.337, $2(1.5 \sim 2.5)$ の場合は 0.217, 以下同様である．平均すれば 1 枚当たる売上枚数 ($\binom{43}{6}$=6096454) があるとき，実に 1/3 から半分の組合せが購入されていないことが分かる．また，期待一等口数が 3 枚当たる抽選があっても，10% を超える 確率で 1 等当選者は現れない．10% と確率は低いが，約 60 万ほどの組合せがあり，それゆえ，1 等賞金を独り占めする組合せを探ることには意味がある．

4　ロト 6 の購入方法

この論文を通して，本選番号はランダムに選ばれるという仮定に立っているので，$\{1,2,3,4,5,6\}$ のような極端に単純な組合せであろうが，$\{13,22,34,37,39,42\}$ のように特徴のない組合せであろうが等しい確からしさで実現すると仮定する．そうすると，1 等から 4 等までの当選者は，各賞に割り当てられた金額を各賞に当選した口数で割った額を手にすることになる．となると，なるだけ他の購入者が選びそうにない組合せを選べば良いのである．前にも述べたが，第 230 回（2005/3/17 抽選）の 1 等組合せは $\{4,9,18,31,46,42\}$ であり，167 口の応募があった．そのため平均 1 億円といわれる 1 等賞金がなんと 1,814,700 円で，2 等当選口数 29 口, 9,405,200 円の賞金を大きく下回ることになった．このような，大勢の応募者が選びそうな組合せを避けるとよい．そうすれば，1 等に振り分けられる金額を独り占めにできるかもしれない．

どうせ本選番号が前もって分かるわけないのだから，他の購入者が選ばないような数字の組合せを探せばよい．そのためにまず，ロトの購入者がどのような方法で番号を選んでいるか推察すると，次のような方法が考えられる．

- 家族の誕生日や記念日を選ぶ.
- 機器も用いず自分なりにランダムに適当に選ぶ.
- ラッキーナンバーを選ぶ.
- パターンのある数列を選ぶ (例：{7,14,21,28,35,42}).
- 申し込み用紙の上で視覚的パターンができるように選ぶ.
- 過去の本選番号を分析して様々な法則を見つけて選ぶ.

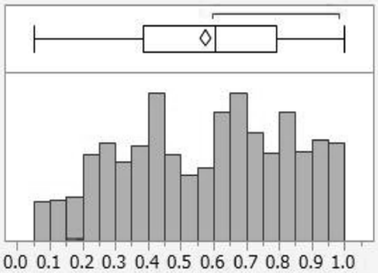
これらの方法は単独ではなく，いくつかの組合せで選ばれているようだ．つまり，「好きな番号を 3 つ選び，残りの番号はパターンで選ぶ」，「しばらく現れていない番号を外してえらぶ」，といった

ようなことが行われているようだ．これら以外にクイックピックを選択することができる．これはコンピュータが発生する数の組合せを購入するもので，購入者の意思は反映しない．それゆえ自分と同じ考えを持つ人との競合を避けることができる．人がランダムと思って購入する組合せは，実はランダムではなく，何らかの傾向がある．本選番号に傾向はないが，人の選ぶ番号には傾向がある．これは他のギャンブルでも言えることである．まず最初に 1 等当選口数の多い本選番号の組合せについて検討してゆこう．

この節において 1 等口数の多寡を，次の代替値をもちいて判断する．いま x_i を i 回目の 1 等当選口数，n_i を応募数とするとき，X_i が平均 $n_i/\binom{43}{6}$, 過分散パラメータが $\phi=1+1/\beta\,(\beta=0.977816)$ を用いたのガンマ・ポアソン分布にしたがっているとして，累積分布 $p_i=P(X_i\le x_i)$ が大きいときを言うことにしよう．この p_i が 1 に近いときは，1 等口数が多くて，それゆえ賞金金額が少なく，逆に 0 に近いときは 1 等当選者が居ないかあるいは非常に多くの口数の応募があるのに 1 等口数が極端に少ない時を言う．この量は数値分析には用いないが，実際の一等口数の分布に影響しそうな要因の効果を調べるために用いる．997 回を重ねた抽選における p_i の分布は次の図 1 のように表される

図 1　ガンマ・ポアソン分布関数

GammaPoissonDist



分位点

100.0%	最大値	0.9999983723
99.5%		0.9998841893
97.5%		0.9778112798
90.0%		0.9177095146
75.0%	四分位点	0.7897870888
50.0%	中央値	0.6059132204
25.0%	四分位点	0.3832117619
10.0%		0.2395666983
2.5%		0.1029575788
0.5%		0.0617010185
0.0%	最小値	0.0509852831

要約統計量

平均	0.5768482
標準偏差	0.2501027
平均の標準誤差	0.0079208
平均の上側95%	0.5923916
平均の下側95%	0.5613047
N	997

4.1　1 等本選番号が極端に多い組合せ

この項では，1 等当選口数の多い抽選から順に検討してみよう．第 2 位は第 301 回（2006/7/27 抽選) であり本選番号は {4,5,6,9,10,11} ボーナス番号は 28，1 等は 28 口，2 等は 24 口あり，キャリーオーバーがなければ，どちらもほぼ同額の賞金を得ていることになっただろう．

Haigh[2] にもあるように，本選番号が {1,2,3,4,5,6} だったら，何千人もの人々が 1 等を獲得するのだろう．くじを購入する人々のうち，誰も選ばないと思われる組合せの中で最も安易な組合せなのだから，数多くの人々がこの組合せに群がると言ってよい．今だ全世界でこの組合せが本選番号として実現していない．LOTO について何の情報が与えられない者は推測でものを言うしかない．第 2 位の 301 回の本選番号は二組の 3 連続数ということで，応募する人が多かったと考え

表6　1等口数の多い回 2-7 位

301 回 1 等 28，2 等 24

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

146 回 1 等 21，2 等 31

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

259 回 1 等 21，2 等 75

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

331 回 1 等 21，2 等 42

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

84 回 1 等 17，2 等 8

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

199 回 1 等 16，2 等 30

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

られる．このことは，2等当選口数が1等口数より小さいことからも容易に推察できる．表6は1等の出る確率の大きい方から，2位から7位までの結果を視覚的に表したものである．ここで，○で囲った数字が本選番号で，アンダーラインのついた数字がボーナス番号である．

いずれの場面においても4ないし6つの数字にパターンがあることが分かる．第146回においては7,17,27,37が，第259回においては1,12,23,34が，第331位においては2,8,32,38が，第84回では35,26,7と37,28,19が，最後の第199回においては，3,14,25,36がパターンを見せている．しかし，8位以降の本選番号では，視覚的なパターンがあまり見られなくなってくる．同じ観察を1等確率 p_i の低い抽選（1等口数はほとんど0）についても行ったところ，ボーナス番号をどの本数字と入れ替えてもパターンは見られなかった．

繰り返しになるが，さまざまな雑誌やWebサイトで推奨されている，過去の本選番号のパターンを利用した予想法等をすべてフォローするわけにはゆかない．不幸にして，このような予想と同

じ組合せが万が一本選番号と一致したら，少ない賞金額に我慢しなければならないだろう．くれぐれも述べるがギャンブルは勝率ではなく期待値で考えなければならない．

4.2　単独で人気のある番号

次に，単独で人気のある番号について調べてみよう．方法はいくつかあると思うが，ここでは 1 から 43 までの番号，例えば 5 が本選番号に含まれるならば 1 等の当選口数を 5 に割り当て，ボーナス番号である場合には，2 等の当選口数を割り当てる．そうすれば，5 を必ず含む組合せがランダムに選ばれることになる．この作業をすべての数字について，すべての抽選について足し合わせる．すると，すべての番号が公平にカウントされ，順序付けが可能となる．この数が大きいほど人気がある数であるという解釈が可能である．ただし，抽選のランダム性より本選番号やボーナス番号として選ばれる比率が異なるので，参考程度に見てほしい．なお，この情報はデータから得られたものであるので，モデル作成の際の要因をして用いることはできない．

結果は，人気のあるほうから 10 個をあげると，12,11,17,8,3,5,27,37,15,23 次に 11 位から 20 位は， 7,10,6,14,35,38,36,35,28,39, 21 位から 30 位は， 31,2,19,22,18,9,26,33,4,24, 31 位から 40 位は， 30,13,21,42,16,32,34,29,20,40, 残るワースト 3 は， 1,41,43 となる．

繰り返すが，この順序はあくまでデータから推測したものであって主催者側が公表したものではないことに注意されたい．もしもこの順番が正しくて，著者以外の誰もこの順序を知らなければ，$\{1,20,29,40,41,43\}$ という選択も魅力的であるが，この順序を知っている人が多ければ多いほど，1 等口数は増え賞金額は少なくなるだろう．このような情報を主催者が公表しないのは，こういった混乱が生じることをおそれているのだろう．実際みずほ銀行では，ロトの情報については一切公開しないと述べている．

4.3　本選番号が占める行数と列数との関係

6 つの本選番号が占める行の数と列の数別に，997 回の抽選を分類してみよう．例えば本選番号の組合せが，$\{3,5,6,16,38,41\}$ ならば，次の表 7 のように $\{1,3,5,6,8\}$ は 5 つの行と，$\{1,2,4,5\}$ の 4 つの列に位置する．このように本選番号の行数と列数に着目し，ガンマ・ポアソン分布を仮定して求めた各確率　$p_i = P(X_i \leq n_i)$ の平均を求めてみる．以前と同様，確率 p_i の平均が大きいほど一等当選口数が多い傾向にあり，小さいほど一等当選口数が少ない傾向にある．まず，行の数，列の数についてそれぞれ見てゆくと，次の表にまとめられる．

表 8 から分かることは，行数が 3 と 6 の場合に確率が高くなっていて，行数が 4 か 5 のときやや確率が小さいことが分かる．このことから行数の 2 乗の項をモデルに加えることが示唆される．

次に表 9 で列数と p_i の平均値との関係を見てみよう．表 9 では，列数が増えるほど p_i の平均値は高くなる．言い換えれば，列数が多いほど，1 等口数が増え賞金が少なくなることが分かる．列数の場合は直線的な関係である一方，行数の方はそういった関係は見られない，

次に表 10 で行数と列数を同時に見てみよう．これ以降，交差表の中の数値は該当抽選における

表 7　選ばれた組合せの数字の位置

1	11	21	31	41
2	12	22	32	42
3	13	23	33	43
4	14	24	34	
5	15	25	35	
6	16	26	36	
7	17	27	37	
8	18	28	38	
9	19	29	39	
10	20	30	40	

表 8　行数と確率との関係

行数	該当数	p_i の平均
3	28	0.603
4	257	0.575
5	510	0.570
6	202	0.593

表 9　列数と確率との関係

列数	該当数	p_i の平均
2	44	0.547
3	367	0.558
4	496	0.579
5	90	0.656

p_i の平均とし，その後にある () 内の数値は該当抽選の数を表す．すると，該当数は 6 と少ないが行数 3 列数 3 の組合せで p_i の平均は最小の 0.353 となっている．該当数が 70 以上ある組合せに限ると，行数 3 列数 5 の場合が最も p_i の平均の値が低い．各列とも行数が増加するほど，p_i の平均値は大きくなっている．

表 10　行と列を同時に見たときの確率平均

	列数			
行数	3	4	5	6
2		0.494(6)	0.563(18)	0.548(20)
3	0.353(6)	0.564(74)	0.555(201)	0.0.575(86)
4	0.657(18)	0.576(145)	0.566(252)	0.606(81)
5	0.733(4)	0.616(32)	0.672(39)	0.682(15)

4.4　本選番号の平均と標準偏差と p_i の平均との関係

この項では，本選番号の平均と標準偏差が p_i の平均にどう影響しているかを調べる．まず，平均と標準偏差を 4 分位偏差により 4 つのグループに分け，それぞれのグループの p_i の平均値を求める．ここで，平均 (mean) は 19 以下の抽選を XS,　20 以上 22 以下を S, 23 以上 25 以下を L, 26 を超えるものを XL と割り振っている．また，標準偏差 (std) の場合は，10.25 以下を XS, 10.25 より大きく 12.24 以下を S, 12.24 より大きく 14.17 以下を L, 14.17 を超えるものを XL に割り振っている．表 11 の　本選番号の平均と p_i の平均との関係をみると，平均が大きくなるほど確率 p_i の平均が小さくなることが分かる．平均だけをみると，本選番号の平均が大きくなるほど，高額な 1 等賞金が期待できる．

表 11　本選番号の平均と p_i の平均との関係

平均による分類	該当数	p_i の平均
XS	250	0.603
S	257	0.602
L	288	0.585
XL	202	0.506

次に，本選番号の標準偏差と p_i の平均 との関係をみると，表 12 のように標準偏差による分類では，XS と XL のときに，p_i の平均が小さくなっている．もしもこの要素がモデルに含まれているなら，2 次の項を入れる必要がありそうだ．

最後に，表 13 の本選番号の平均と標準偏差と，p_i の平均との同時関係を眺めてみる．まず平均が L 以下 で標準偏差が L 以下の場合に高い値が集まっており，いずれかが XL のとき低い値が集まっている．それゆえ，この場合，平均と標準偏差の交差項を入れなければならないことが分かる．

表 12　本選番号の標準偏差と p_i の平均との関係

標準偏差による分類	該当数	p_i の平均
XS	248	0.560
S	249	0.623
L	250	0.592
XL	250	0.533

表 13　本選番号の平均と標準偏差の同時関係

	標準偏差			
平均	XS	S	L	XL
XS	0.635(83)	0.676(65)	0.620(50)	0.492(52)
S	0.589(35)	0.610(57)	0.637(80)	0.568(85)
L	0.563(41)	0.639(63)	0.586(68)	0.549(75)
XL	0.476(89)	0.565(64)	0.502(52)	0.481(38)

4.5　本選番号の最大値と最小値

6 つの本選番号の最大値と最小値が，p_i の平均値に影響を与えているかを調べる．この要因は先の本選番号の平均と標準偏差と似てはいるものの，考慮に値する．前項の分析と同様に，四分偏差をもとに，ほぼ同じサイズのグルーピングを行い，最大値と最小値の p_i への影響をしらべることにする．本選番号の最小値は，2 以下を XS, 3 以上 5 以下を S, 6 以上 9 以下を L, 10 以上を XL に振り分け．最大値は，35 以下を XS, 36 以上 39 以下を S, 40 以上 41 以下を L, 42 以上を XL に割り振っている．

表 14 を眺めると，最小値が XS かあるいは XL のとき p_i の平均 が小さくなっていて，S あるいは L という中途半端な値では，p_i の平均値が大きくなっていることが分かる．このことはモデルに，最小値の 2 乗の項を付け加える必要性を示している．

表 14　本選番号の最小数と p_i の平均との関係

最小値	該当数	p_i の平均
XS	253	0.559
S	274	0.622
L	233	0.590
XL	237	0.531

続いて表 15 で最大数と p_i の平均との関係をみてみると，最大値が大きくなるほど，p_i の平均が単調に小さくなっていることが分かる．

表 15　本選番号の最大数と p_i の平均との関係

最大値	該当数	p_i の平均
XS	249	0.624
S	297	0.593
L	203	0.563
XL	248	0.521

最後に表 16 を見てみよう．この表をみると，上で述べた表 14, 表 15 の性質を受け継いでいるようであり，交互作用は有意になるかどうか怪しい．

表 16　行数と列数を同時に考えたときの p_i 値の平均

	最大数			
最小数	XS	S	L	XL
XS	0.619(76)	0.555(63)	0.501(46)	0.536(68)
S	0.670(83)	0.621(89)	0.613(47)	0.559(55)
L	0.612(58)	0.622(76)	0.582(51)	0.520(48)
XL	0.539(32)	0.562(69)	0.556(59)	0.480(77)

4.6　本選番号間隔の最小値と最大値

この項では，本選番号間の間隔に着目する．例えば本選番号が，$\{3,12,18,41,42,43\}$ とすれば間隔値は $\{9,6,23,1,1\}$ この間隔値の最小値と最大値と，p_i の平均値の関係についてしらべよう．ここで，最小間隔は，1 のときを XS, 2 のときを S, 3 のときを L, 4 以上を XL に振り分けている．また最大間隔は, 10 以下を XS, 11 以上 13 以下を S, 14 以上 16 以下を L, 17 れ以上を XL に割り振っている．

まず表 17 を見てみよう．これを見ると間隔が小さいほど p_i 値の平均が低く，間隔を十分に取っているほうが p_i の平均が大きい方がよいことが分かる．つまり，間隔の最小が小さいときが，大きいときよりも良好で，数字の組合せの中に連番あるいは接近したペアを入れておくとよいことが分かる．

次に表 18 をみてみると今度は間隔の最大値が一番大きなグループ XL で p_i の平均が一番小さいことが分かる．

次に表 19 をみてみると，いつも通り XL-XL のペアで最小となっている．注意すべきは，最小間隔が XL のときの，最大間隔が XS,S のときである．これに属する組み合わせに p_i 値は非常に大き

表 17 本選番号間隔の最小数と p_i の平均との関係

最小数	該当数	p_i の平均
XS	452	0.563
S	271	0.544
L	155	0.597
XL	119	0.679

表 18 本選番号間隔の最大数と p_i の平均との関係

最大数	該当数	p_i の平均
XS	280	0.596
S	264	0.636
L	217	0.571
XL	248	0.493

く, 選ばないほうがよいことが分かる．また，p_i 値の傾向は行毎，列ごとに異なるので，これらの要因の交差項が必要であることが分かる．

表 19 本選番号の最小間隔と最大間隔と，p_i 値の平均

	最大値			
最小値	XS	S	L	XL
XS	0.574(113)	0.594(119)	0.591(96)	0.501(124)
S	0.567(78)	0.603(54)	0.533(65)	0.488(74)
L	0.579(49)	0.676(49)	0586(35)	0.493(26)
XL	0.726(44)	0.750(42)	0.573(21)	0.447(12)

5 ポアソン分布，ガンマ・ポアソン分布のあてはめ

項 3.2 で述べたように，一等当選口数の分布はポアソン分布には従ってない．しかし，ガンマ・ポアソン分布に従っている確証はない．こういった分析を一般化線形モデルで行うには，[過分散モデルのあてはめ] を用いる．一等口数の分布を一般化線形モデルの枠組みで分析するとき，2 通りの可能性がある．前節で紹介した説明変数を用いてやれば，ポアソン回帰モデルとして解釈できるか，あるいは，すべての説明変数を用いても過分散回帰分布であると解釈できる場合の 2 つの可能性である．

このデータで問題になるのが，期待一等口数のほとんどが 5 以下であるということだ．それゆ

え，ポアソン変量の正規近似が期待できず，形式的な統計的推測の信頼性が低くなる．このような状況における統計的推測の詳細については，別の機会にゆだねるものとして，ここでは議論しない．

まず，分布はポアソン，リンク関数は対数の一般化線形モデルをあてはめよう．まず図 2 はモデル全体の検定である．ここでモデルに加えた変数は，本選番号が占める行数と列数とその交差項 (Ncol, Nrow, Ncol*Ncol, Nrow*Nrow, Nrow*Ncol)，平均と標準偏差 (Mean, Std, Mean*Mean, Std*Std, Mean*Std)，本選番号の平均値と標準偏差とその交差項 (Min, Max, Min*Min, Max*Max, Min*Max)，番号間の間隔の最小値・最大値とその交差項 (MinInterval,MaxInterval,MinInterval*MinInterval, MaxInterval*MaxInterval,MiniIntreval*MaxInterval) の合計 20 要素である．これらすべての要因をすべてモデルに入れて，推測を行うと, 自由度 20 のモデルは高度に有意となるが，適合度部分を見ると，Pearson 統計量をみてもデビアンスにおいても，分布がポアソン分布であることは否定される．

図 2

モデル全体の検定

モデル	(-1)*対数尤度	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
差分	145.642466	291.2849	20	<.0001*
完全	2169.83537			
縮小	2315.47783			

適合度統計量	カイ2乗	自由度	p値(Prob>ChiSq)
Pearson	2229.076	976	<.0001*
デビアンス	2048.327	976	<.0001*

AICc
4382.6184

次に図 3 の各要因の効果の検定を見てみよう．すると，行数と列数の交差項と行数，そして最大値と最小値の交差項が有意ではない．そこで，変数減少法を用いて最も AICc[*2] の高いモデルを探索する．図 3 において有意ではなかった 3 つの要因をひとつずつ取り去り AICc も最小で，しかもすべての要因が有意であるモデルが得られた．これは先のもでるから，行の数 (nrow)，行と列の交差項 (nrow*ncol) そして，最大数と最小数の交差項 (max*min) を取り除いたものである．しかしながら，ここでも次の 図 4 のようにポアソン回帰モデルであるとは言えない．しかし，このデータでは観測値（ひいては推定値）の正規近似が望めないため，もう少し厳密な議論は別の機会で詳細を述べる予定である．

ともあれ，ポアソン回帰分布では適合が十分でないことが分かるので，最後に過分散を導入する．過分散モデルでは，平均にかかる母数をポアソン分布と同じ方法で推測し，過分散パラメータを（カイ 2 乗統計量）/(残差自由度) で推定し，このパラメータを正規線形モデルの分散と同様の扱いをして統計的推測を行う．

まず図 5 から見て行こう．9 個すべてのパラメータが高度に有意である．次の図 6 は過分散パラ

[*2] この量は通常の AIC の改良版で，AICc =-2 loglikelihood + $2k = \frac{2k(k+1)}{n-k-1}$ により求められる

図 3　各要因の効果

効果の検定

要因	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
Ncol	1	19.739423	<.0001*
Ncol*Ncol	1	3.4022514	0.0651
Nrow	1	0.1697508	0.6803
Nrow*Nrow	1	0.697795	0.4035
Ncol*Nrow	1	0.5887959	0.4429
Mean	1	13.307143	0.0003*
Mean*Mean	1	12.249984	0.0005*
Std	1	24.947053	<.0001*
Std*Std	1	11.716968	0.0006*
Mean*Std	1	1.4494306	0.2286
Min	1	25.470292	<.0001*
Min*Min	1	0.5753225	0.4482
Max	1	26.518513	<.0001*
Max*Max	1	2.8494823	0.0914
Min*Max	1	14.991663	0.0001*
MinInterval	1	11.568625	0.0007*
MinInterval*MinInterval	1	1.5264612	0.2166
MaxInterval	1	12.106789	0.0005*
MaxInterval*MaxInterval	1	1.4304205	0.2317
MinInterval*MaxInterval	1	14.744135	0.0001*

図 4　モデルのあてはめの結果

モデル全体の検定

モデル	(-1)*対数尤度	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
差分	127.331789	254.6636	9	<.0001*
完全	2188.14604			
縮小	2315.47783			

適合度統計量	カイ2乗	自由度	p値(Prob>ChiSq)
Pearson	2310.565	987	<.0001*
デビアンス	2084.949	987	<.0001*

AICc
4396.5152

メータで差分，完全の値を，過分散パラメータ $\phi = 2.341$ で割ったものである．モデルとしての有意性が得られるが，適合度は図 4 の場合と同様，データがポアソン回帰モデルに従うという帰無仮説を非常に強く否定している．

図 5

効果の検定

要因	自由度	尤度比カイ2乗	p値(Prob>ChiSq)
Ncol	1	26.502831	<.0001*
Mean	1	14.680413	0.0001*
Std	1	13.0957	0.0003*
Mean*Std	1	10.995066	0.0009*
Min	1	15.267638	<.0001*
Max	1	15.732403	<.0001*
MinInterval	1	14.67161	0.0001*
MaxInterval	1	10.910921	0.0010*
MinInterval*MaxInterval	1	12.884211	0.0003*

次に，図 6 を見てみよう．自由度 9 のモデルはデータをよく説明している．過分散パラメータは

2.3410 と大きく，ポアソン回帰モデルとは言えないことがよく分かる．

図 6

モデル全体の検定

モデル	(-1)*対数尤度	尤度比カイ2乗	自由度	p値(Prob>ChiSq)
差分	54.3920922	108.7842	9	<.0001*
完全	934.706427			
縮小	989.098519			

適合度統計量	カイ2乗	自由度	p値(Prob>ChiSq)	過分散
Pearson	2310.565	987	<.0001*	2.3410
デビアンス	2084.949	987	<.0001*	

AICc
1891.6809

図 7 はパラメータの推定結果を表しているが，要因がすべて有意であるので解釈が楽になる．列の数 (Ncol) は推定値がプラスなので，大きいほど確率が小さくなる．平均は大きい高い賞金が望める．同様に，最小値 (Min) は小さいほど確率が小さくなり，最大値は大きいほど確率が小さくなる．交互作用を有する，平均 (Mean) と標準偏差 (Std) は表 12 を見て判断することになり，最小間隔 (MinInterval) と最大間隔 (MaxInterval) も同様に，表 18 を見て判断することになる．Haigh [2] では，データを一般化線形モデルにあてはめてはいない．ただ，6 個の数字の組合せをランダムに発生させ，それが 4 節で述べたような，申し込み口数の小さくなるような組合せを選ぶよう勧めている．この論文では，Haigh [2] と同様に，コンピュータにより組合せをランダムに発生させ，その組合せのモデルにより計算される p_i の値が小さいものを選ぶ方法を提案する．

図 7

パラメータ推定値

項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	下側信頼限界	上側信頼限界
切片	1.5061787	0.2785151	28.700531	<.0001*	0.9585713	2.0505093
Ncol	0.209883	0.0625309	11.321167	0.0008*	0.0875037	0.3326333
Mean	-0.030622	0.0122526	6.2710059	0.0123*	-0.054694	-0.006647
Std	0.07846	0.0332698	5.5940668	0.0180*	0.013413	0.1438612
(Mean-22.0664)*(Std-12.1918)	0.005654	0.0025729	4.6967427	0.0302*	0.0005473	0.0106312
Min	0.0441032	0.0173065	6.5218497	0.0107*	0.0102401	0.0780917
Max	-0.046753	0.0180536	6.7203822	0.0095*	-0.082177	-0.011398
MinInterval	0.0769466	0.0303888	6.2672455	0.0123*	0.0168568	0.1360095
MaxInterval	-0.026554	0.0123506	4.6607986	0.0309*	-0.050857	-0.002438
(MinInterval-2.01404)*(MaxInterval-13.5095)	-0.016085	0.0068886	5.5037253	0.0190*	-0.029643	-0.002637

6　抽選器はランダムに番号を抽出しているか

これまで，抽選番号は等しい確率で独立に選ばれるという仮定のもとで議論を進めてきた．しかし，誰も選んでいない組合せを探るための前提として，この仮定が妥当なものであることを述べておかなければならない．そこで，この節では，抽選器による番号の出方がランダムと言えるかについて考える．いま $\mathbf{x}_i$ を，第 i 回目の抽選の本選番号とボーナス番号から作られる 43 次元縦ベクトルで，本選番号あるいはボーナス番号が位置する箇所の値を 1，その

他の値を 0 としたものとし，$\mathbf{X}$ は $\mathbf{x}_i^T$ を縦方向に並べた 997×43 の行列とする．次に示す上三角行列は, 対称行列 $\mathbf{X}^T\,\mathbf{X}$ の上三角行列のみを示したものである．まず，各数値が同じ割合で出現するかをみるため，カイ二乗検定を行う．各数値の観測数は，$\mathbf{X}^T\,\mathbf{X}$ の対角線上に現れ, $\{139, 165, 146, 160, 161, 179, 148, 164, 142, 182, 158, 175, 169, 164, 159, 162, 153, 170, 160, 163, 169, 166, 172, 143, 163, 165, 186, 175, 150, 168, 165, 151, 157, 153, 168, 160, 180, 175, 169, 155, 157, 151, 162\}$

となり，共通の推定値 $997 \times 7/43 = 162.302$ を用いて検定統計量,

$$\frac{(139-162.302)^2}{162.302} + \frac{(165-162.302)^2}{162.302} + \cdots + \frac{(146-162.302)^2}{162.302} = 30.530$$

が求められる．これは自由度 43 のカイ二乗分布に漸近的に従う．このときの p 値は，0.995　となり，数字毎に出現回数が異なるとは全く言えない.

同様にして，全 $\binom{43}{2} = 903$ 通りのペアが同じ割合で現れているかを検定する．共通の期待回数が，$948 \times \binom{6}{2} / \binom{43}{2} = \dfrac{948 \times 15}{903} = 14.6977$　であることから，自由度 903 のカイ二乗分布に従う統計量の値 914.953 が得られ，p 値は，0.923 となり，ペアによっては異なる頻度で現れるとは言えない.

抽選器の中の球には記憶も意思もない．だから，過去の当選数字の組合せが，これからの本選番号の数字に影響することはない．また，各抽選において各数字を持つ玉の出現確率は同等で，しかもある数字が出れば他の数字が出やすい傾向もない．ごく普通に考えれば検証する必要もないことだが，念のために行った適合度検定は我々の常識と一致する.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	139	17	21	24	22	18	14	15	17	15	20	17	21	18
2		165	22	28	25	29	15	23	22	18	21	32	21	33
3			146	22	19	18	21	19	16	21	26	25	23	23
4				160	25	29	26	20	23	23	25	23	22	22
5					161	24	23	22	25	22	15	28	25	22
6						179	23	28	18	29	22	27	20	27
7							148	18	22	21	24	21	23	20
8								164	21	26	22	24	31	18
9									142	25	20	23	16	23
10										182	23	40	29	23
11											158	25	25	26
12												175	24	21
13													169	24
14														164

7　終りに

ブラックジャックの世界では，情報を収集し，適切な選択をすればディーラーを負かすことができると言われている．しかしながら広い意味で同じギャンブルである LOTO で，他の購入者が買わないだろう組合せを選んだらどうなるだろう．儲けることは出来なくても賞金の期待値は高ま

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	21	26	16	21	22	16	24	23	26	18	21	20	21	18
2	19	26	22	22	17	26	22	31	22	22	30	29	23	22
3	20	24	24	24	19	21	17	20	23	21	24	28	24	23
4	18	21	23	31	20	21	16	28	20	15	22	22	22	26
5	19	20	22	26	25	29	26	22	24	18	17	26	28	28
6	23	25	27	28	25	27	27	30	28	28	22	20	29	30
7	21	22	20	30	31	16	15	12	23	16	19	21	25	14
8	26	25	24	17	20	29	23	20	26	25	28	34	32	21
9	18	22	16	27	19	22	18	19	21	16	18	24	22	18
10	29	21	21	27	33	30	29	29	31	21	28	22	27	37
11	28	16	24	23	23	29	20	18	26	18	24	23	29	24
12	23	31	20	24	26	23	20	28	19	17	24	29	22	35
13	31	22	22	28	20	26	22	35	20	16	30	15	27	27
14	17	16	20	27	27	15	25	30	26	24	29	28	27	32
15	159	15	22	21	19	18	28	23	26	21	22	22	25	13
16		162	22	30	21	32	20	19	17	23	21	26	32	24
17			153	25	24	20	20	18	22	23	18	21	20	26
18				170	24	22	21	24	24	22	24	23	29	25
19					160	23	28	23	19	20	29	28	25	22
20						163	26	19	31	25	20	17	28	30
21							169	26	32	26	24	31	30	26
22								166	27	10	33	23	35	27
23									172	23	21	25	25	26
24										143	19	26	23	27
25											163	31	29	25
26												165	25	25
27													186	29
28														175

る．しかし，当選確率が非常に低い宝くじでは期待値に近づくまでに時間と資金が不足してしまう．私の専門はポアソン分布データの回帰分析であるため，一等当選口数がポアソン分布に従うとかんがえられる LOTO データを分析することに興味を持ち作成したものである．この論文におけるシミュレーションは Wolfram 社の Mathematica ® ver.10 を用い，統計処理では SAS 社の JMP ® ver.12 を用いて行った．

参考文献

[1] Annete J. Dobson(2002) *An Introduction to Generalized Linear Models, second ed.*, Chapman and Hall (和訳：一般化線形モデル入門，第 2 版，田中豊，森川敏彦ら，共立出版)

[2] Haigh, John(2002) *Taking Chances, 2nd ed.*, Oxford University Press.

[3] JMP® Pro 12.0.1（2015） SAS institute.

[4] Mathematica® ver.10.4 (2016) Wolfram Research Inc.

[5] みずほ銀行　宝くじコーナー , http://www.mizuhobank.co.jp/takarakuji/

[6] 宝くじネット (http://www.takarakuji-loto.jp/)

	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
1	18	18	19	16	19	22	20	20	26	22	23	20	21	17	21
2	28	25	25	20	19	24	21	23	21	33	14	18	22	29	29
3	21	19	20	20	18	16	22	22	20	14	19	22	20	16	19
4	13	24	28	22	23	22	25	19	25	32	19	21	27	17	26
5	14	22	16	20	22	29	23	24	24	28	36	19	28	16	20
6	24	30	36	27	24	28	20	28	23	31	25	24	31	19	23
7	16	19	22	17	21	23	25	24	18	26	26	19	28	24	24
8	13	27	21	21	26	27	21	25	21	32	19	22	23	24	25
9	21	22	18	17	25	18	17	15	27	23	19	22	23	18	16
10	20	27	28	21	24	26	24	28	28	28	34	30	20	28	26
11	13	30	25	25	22	17	18	25	30	22	29	13	15	16	29
12	23	24	25	25	21	27	29	26	27	26	26	26	23	22	29
13	28	34	26	27	24	18	27	22	25	30	21	17	25	23	22
14	19	19	27	24	30	20	24	18	28	25	15	24	22	25	23
15	22	34	25	17	18	23	18	30	28	17	23	28	24	28	31
16	27	18	14	20	20	31	19	27	23	24	29	19	21	30	31
17	22	28	20	18	20	20	28	18	21	24	25	24	21	24	23
18	22	29	20	19	22	17	33	18	25	25	24	27	31	16	23
19	19	27	24	19	24	13	26	22	21	28	24	19	24	19	19
20	22	22	24	23	31	18	22	18	26	25	18	27	16	24	21
21	29	27	30	22	32	15	26	25	31	28	21	21	25	21	19
22	27	19	28	19	15	23	27	24	27	18	22	23	27	17	28
23	23	29	27	22	30	17	26	23	27	34	25	24	18	26	28
24	19	14	16	26	19	26	21	20	23	21	19	15	20	15	21
25	24	21	18	18	17	23	25	17	27	21	24	24	20	26	21
26	22	22	18	18	21	27	17	18	23	28	17	23	14	20	25
27	25	25	19	24	31	32	24	30	26	28	33	28	28	22	28
28	27	17	23	20	19	21	28	21	34	30	24	28	26	24	28
29	150	23	20	19	26	16	26	19	28	22	28	21	18	21	12
30		168	34	22	24	18	26	22	30	26	23	25	20	21	12
31			165	28	17	25	26	24	30	28	24	27	24	21	20
32				151	19	16	29	19	23	25	29	24	20	22	24
33					157	24	30	24	23	20	16	24	26	16	16
34						153	20	21	21	27	31	24	14	23	25
35							168	21	30	28	30	19	30	19	17
36								162	30	24	31	19	25	23	23
37									180	33	38	29	18	17	20
38										175	21	17	18	25	24
39											169	20	17	25	22
40												155	18	14	21
41													157	25	20
42														151	23
43															162

[7] All About (http://allabout.co.jp/)
[8] AMEBLO CANTALE, http://ameblo.jp/bz-125-norinori/entry-11347781417.html