# 極値分布とその応用に関する研究

2002MM042 川田 将之
指導教員　尾崎 俊治

## 1　はじめに

私たちの生活の周りには数多くの数値が存在する. ある一つの項目について数値を集めると, それはデータになる. 例えば気温・降水量などの数値は, 毎日観測されているので一年間のデータを集めると膨大な量になる. これらのような日々変化しているデータについては折れ線グラフにして表わされることが多く, 簡潔で誰が見ても分かりやすい数値である. 一方, 日本国内で過去に起こった地震の震度の数値を集めて, 大きい数値から順番に並べたデータもある. スポーツの世界でも, 試合の記録をランキング形式で並べて誰が 1 番であるかを決めるものがたくさんある. このようなランキング形式のデータは, 過去の記録を上回る記録が現れた場合, 前のものと入れ替わるという仕組みになっている. 他にも様々な種類のデータが存在するが, 数値のなかでも" 極値的な数値 "について考える. ここで考える極値とは, 普通には考えられないような数値のことである. 極値を知ることによって, 実際に極値的な現象が起こった場合にどう対応すればよいかを考える基準を得ることができる.

さて, 極値を調べるにあたって統計的分布である極値分布がある. 極値分布は集めたデータから極値がどの程度の確率で起こりえるかを示す分布である. あるデータを用いて考えるとき, データの性質を示すパラメータが存在するのだが, そのパラメータを推定して極値分布を導く. その導いた分布から極値が求まる. 本論文ではパラメータの求め方, 極値分布と極値の導き方について研究する. また, 極値分布の応用として数値解析ソフト *Mathematica* を用いて実際のデータから極値を求める.

## 2　極値分布

分布関数 $F(x)$ のパラメータを $\theta, \xi > 0$ とし, その定義域を $-\infty < x < \infty$ とする. 極値分布の分布関数は 3 つのタイプが存在することが知られているが, 最も一般的で標準形とされているのがガンベル型の分布である（蓑谷 [1] , Johnson et al. [2] 参照）.

$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\xi}{\theta}\right)\right]\right\}. \qquad (1)$$

この分布関数は $x$ の値が増加するにつれて $F(x)$ の値も増加し 1 に収束する. $F(x)$ の値が 0.99 のとき 100 回に 1 回くらいの確率で $x$ の値の極値が現れると考えられる.

## 3　パラメータ推定

最尤推定法を用いると 2 つのパラメータ $\theta, \xi$ の最尤推定量 $\hat{\theta}, \hat{\xi}$ は,

$$\hat{\theta} = \bar{X} - \frac{\sum_{i=1}^{n} X_i \exp\left(-\frac{X_i}{\hat{\theta}}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{X_i}{\hat{\theta}}\right)}, \qquad (2)$$

$$\hat{\xi} = -\hat{\theta}\log\left\{\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\frac{X_i}{\hat{\theta}}\right)\right\}, \qquad (3)$$

で求められる. (2) 式の非線形方程式を解いて $\hat{\theta}$ を求め, (3) 式に代入して $\hat{\xi}$ が求まる. 極値分布の場合はパラメータの推定量が求まってもそれ単体ではデータの性質を知ることはできない.

## 4　日本中央競馬においての三連単の高額払戻金

競馬は日本のスポーツの中でもギャンブル度が高く人気がある. 競馬は十数頭の馬に騎手が乗って競争をするスポーツであるが, 観客はその順位を予想して一口 100 円の馬券を買うというものである. もちろん, 強い馬には人気が集中して多くの人がその馬が一着でゴールすると予想する. しかし強い馬の馬券は倍率が小さく, 払戻金も小額になる. 逆に弱い馬の馬券は倍率が大きく, 払戻金も高額になる.

今回の研究では順位予想が一番難しいとされる“ 三連単 ”について考える. 三連単とは, 一着から三着の順位をすべて予想するもので, 競争する馬の数が 16 頭の場合は, 16 頭のうち上位 3 頭の順列になるので, $16 \times 15 \times 14$ を計算して 3360 通りの組み合わせが考えられる. 平成 17 年 10 月 22 日, 東京競馬場で人気順にして 3360 組の中で 3344 番目の馬券が正解となった. このときの払戻金は一口 100 円に対してなんと 18,469,120 円であった. では, 参考文献 [3]（1986 年以降の記録で平成 17 年 12 月 19 日現在, 表 1）から 15 個のデータを用いて極値を求める.

表 1: 三連単高額払戻金ランキング（ [3] 参照）

| 日付 | 場所 | 払戻金額（円） | 人気順 |
|---|---|---|---|
| 05/10/22 | 東京 | 18,469,120 | 3344/3360 |
| 05/4/9 | 福島 | 10,149,930 | 2136/2184 |
| 05/1/23 | 小倉 | 4,050,530 | 2316/3360 |
| 05/3/12 | 阪神 | 4,049,110 | 1247/1320 |
| 05/8/7 | 函館 | 3,029,920 | 1451/1716 |
| 04/11/20 | 東京 | 2,857,890 | 2680/3360 |
| 05/8/21 | 札幌 | 2,759,500 | 1775/2184 |
| 05/12/18 | 中山 | 2,614,840 | 2172/3360 |
| 05/4/17 | 阪神 | 2,505,770 | 2086/2730 |
| 05/6/4 | 中京 | 2,417,070 | 2704/4688 |
| 05/10/2 | 札幌 | 2,416,680 | 1661/2184 |
| 05/5/29 | 東京 | 2,306,910 | 1827/3360 |
| 05/4/3 | 阪神 | 2,303,230 | 2176/4688 |
| 05/2/5 | 東京 | 2,270,060 | 1784/3360 |
| 05/6/5 | 中京 | 2,214,210 | 3584/4688 |

### 考察

結果は表 2 のようになった. 払戻金 18,000,000 円という数値は極値分布で考えた場合, 分布が 0.9998 となりおよそ 1 万分の 2 の確率で起こる極値中の極値であると言える. 12,000,000 円でも 分布が 0.99 を越える結果になったのは, もともとのデータのランキング 1 位と 2 位の数値と 3 位以下の数値に大きな開きがあったからだと考えられる. 今回の例は極値現象が現実に起こったものであると考えられ, 今後 18,000,000 円という数値が現れる確率というのは極めて低いと言える.

表 2: 三連単の高額払戻金額

| 払戻金 $x$ | 分布 $F(x)$ |
|---|---|
| 18,000,000 | 0.999862 |
| 14,000,000 | 0.998545 |
| 12,000,000 | 0.995284 |
| 10,000,000 | 0.984771 |
| 8,000,000 | 0.951407 |

## 5 おわりに

本論文では, 極値現象について扱い統計的に解析してきた. 研究を進めるにあたって *Mathematica* を利用することで, 極値を具体的な数値で導くことができた. 反省・課題としては, 最小値としての極値に関することや, 他の方法による極値の導き方についてなど研究したいことはまだまだ残ってしまったことは心残りである.

今回用いたデータの他にも, 世の中には数え切れないほどのデータが存在する. それらのデータについても, 今回と同じ方法で極値を得ることができる. しかし, 得られた極値をどう理解するかによって, 極値の価値も大きく違ってくる. 私は極値という概念を知って, 世の中がさらに広く見えるようになった気がする. 今ある現実を見ることも大切であるが, 少し余裕を持って極値を考えることの面白さが理解できた. 極値は普通には考えられない“ まさか ”という数値であるが, 良い方にも悪い方にも起こりえる. 今後, 極値に対する関心が広まってどのような分野で応用されるかは分からないが, 私自身の今後の人生で極値に出会うことを楽しみにしている.

## 参考文献


[1] 蓑谷千凰彦：統計分布ハンドブック, pp. 290-309, 朝倉書店 (2003).

[2] N.L. Johnson, S. Kots and N. Balakrishnan：Continuous Univariate Distributions, Volume 2, Second Edition, pp. 1-112, Wiley, New York (1995).

[3] `http://keiba.yahoo.co.jp/`（Yahoo!スポーツ競馬）.