ベンフォード則の統計的分析

57 番　SS 科学部　2 年 7 組　佐直　駿太

1　はじめに

(1)　ベンフォード則とは

数字の最高桁が一様な分布ではなく、特定の値が多くなっているという法則である。この法則によると、数値の分布の割合が、$\log_{10}(N+1)-\log_{10}(N)$で求められる。

(2)研究の目的

(ア)身近な事柄について、ベンフォード則が成立するのかどうかについて調査する。今回、選挙の各候補者の得票数や、世界の国の人口・面積においてベンフォード則が成立するのかを調べた。また、χ^2 検定を用いて測定値と理論値との比較を行った。

(イ)ベンフォード則の成立する理由を理解する。

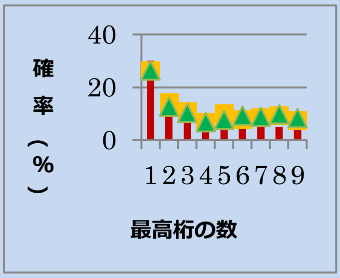
(ウ)10 進数以外では規則性があるのか、また、どのような分布を示しているのかを世界の国の人口・面積において調査する。

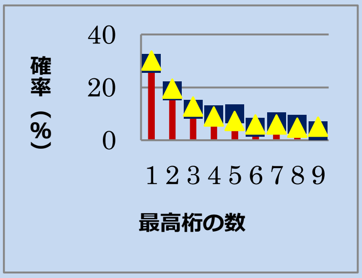
2　研究の方法・結果

(1)世界の国の人口、面積で最高桁の 1 から 9 の数の分布を調査し、χ^2 検定を行った。今回、対立仮説はベンフォード則に従う、帰無仮説は従わないとした。χ^2 の値が 4.46,1.12 となり棄却域に入ったので、ベンフォード則に従うとした。(左下図、四角が人口、三角が面積)。

(2)　ア)第 47 回の衆議院選挙の小選挙区において各候補者数の得票数とベンフォード則の理論値と比較を行った。そして、χ^2 検定を用いて検証した。値は 122 となり、棄却域に入らなかったので、ベンフォード則に従わないとした。(右下図、四角)。

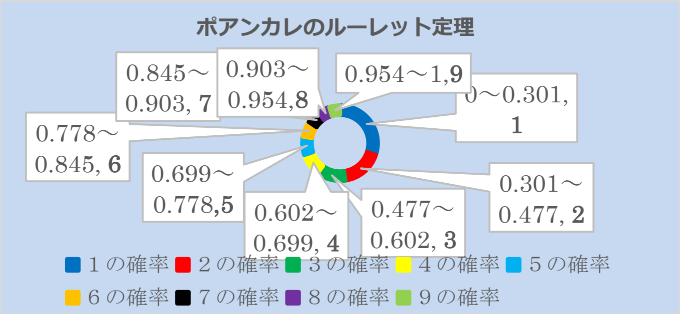
イ)ア)で χ^2 検定の値が棄却域に入らなかったので、さらにもう一年分のデータで検証を行った。値は 105 となり、棄却域に入らなかったので、ベンフォード則に従わないとした（右下図、三角)。





(3)ベンフォード則の説明

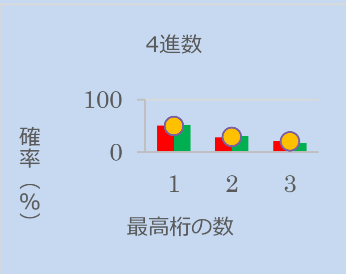
ベンフォード則で求められる数値の分布になる理由は、ポアンカレのルーレット定理（下図）で説明することができる。十進数を十倍するごとに一桁上がる構造を円に表し、それを 1 周とする。さらに十倍すると 2 周すると考える。1 や 10、100 は対数関数$\log_{10} x$で円周の周回数と対応できる。1 から 9 までの 1 桁の数は、 0 周から 1 周の間にあることになり、それらが何周に相当するかは、0.301 (周) $=\log_{10} 2$　0.477 (周) $=\log_{10} 3$　0.602 (周) $=\log_{10} 4$　0.699 (周) $=\log_{10} 5$　0.778 (周) $=\log_{10} 6$　0.845 (周) $=\log_{10} 7$　0.903 (周) $=\log_{10} 8$　0.954 (周) $=\log_{10} 9$となるので 1 は 0 から 0.301 周、2 は 0.301 周から 0.477 周とこれ以降も同様に続いていき、9 は 0.954 周から 1 周となる。また、ベンフォード則では最高桁の数だけに着目していてその数値が何桁であるかは無視することができる。

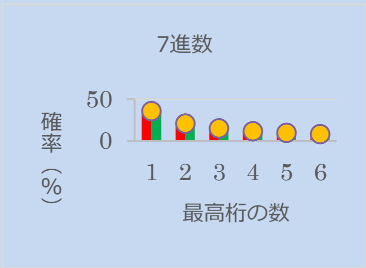


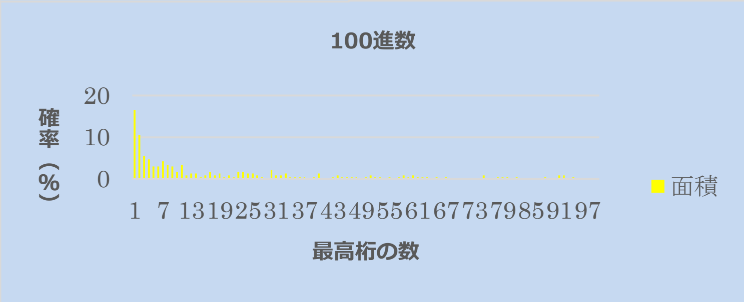
(4)世界の国の人口・面積のＮ進数(Ｎ＝3,4,…9)における調査

次に、世界の国の人口、面積において 3 進数から 9 進数の最高桁の数の分布について調査したそれぞれの進法での最高桁の分布は

$\log_m(N+1)-\log_m(N)$（mは進数の数､Ｎは最高桁の数に対応している）と予測した。予測値が正しいことを統計的に示すために χ^2 検定を行った。その結果、仮説が正しいことが証明された。100 進数でも数の分布を調べた。その結果、他の進数と同様に最高桁の出現確率は 1 が一番多く、100 進数の場合は 1 から 9 までに 50%以上が存在している。







3　まとめ・今後の展望

(1)人口、面積の最高桁の分布は、χ^2 検定からベンフォード則に従っていることが言えるが、選挙では、棄却域に入らなく、二回とも同じような分布を示したので、特有の分布が存在していると考察した。

(2) 他の進法について最高桁の数の分布を調査したところ､他の進法でも最高桁の分布が一様なものでなく特定のものになっているので、ベンフォード則を応用できると考察した。

(3) 今回､調査したデータ数が少なかったので､データ数を増やし､人口・面積以外についても調査したい。

4　参考文献

1）数学で身に着ける柔らかい思考力　ダイヤモンド社
　ロブ・イースタウェイ　ジェレミー・ウィンダム　著　水谷淳　訳
2）第 6 回マス・フェスタ資料
3）例題で学ぶ統計的方法　創成社　井上洋　野澤昌弘　著