

研究班番号【87】
ベンフォードの法則が従う規則性

数学班:上田晋太郎、笠原柊生

Abstract

We watched TV and the Internet and became curious about Benford's Law, a method for exposing data fraud. We researched the regularity with which Benford's law is obeyed. The study concluded that one of the regularities is that the range of data is not infinite.

要約

本研究では、ベンフォードの法則が適用されるデータの規則性を調べた。調査によってデータの範囲が無限ではないということが規則性の一つということが結論付けられた。

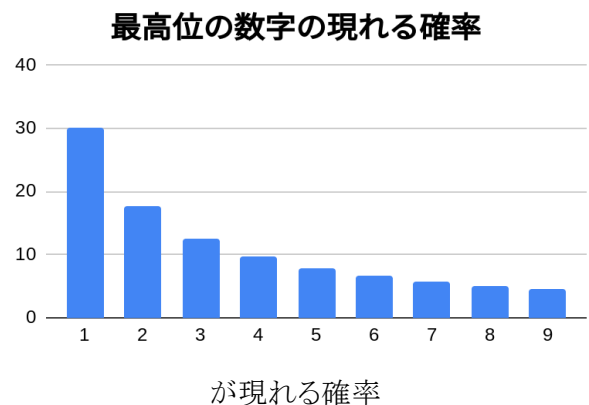
1. はじめに

テレビやネットで国や企業の不正のニュースをみて、不正を暴く方法について興味を持った。そこでその方法の1つであるベンフォードの法則について調べることにした。本研究ではベンフォードの法則が適用できるデータにどのような規則性があるか研究する。

2. 研究手法

総務省統計局より公開されているデータを用いて、そのデータの先頭の数を取り出し、1~9までの個数を調べた。Googleスプレッドシートを用いて棒グラフを作成し、右の理論値のグラフと比較することでベンフォードの法則に従うかどうかを調べた。

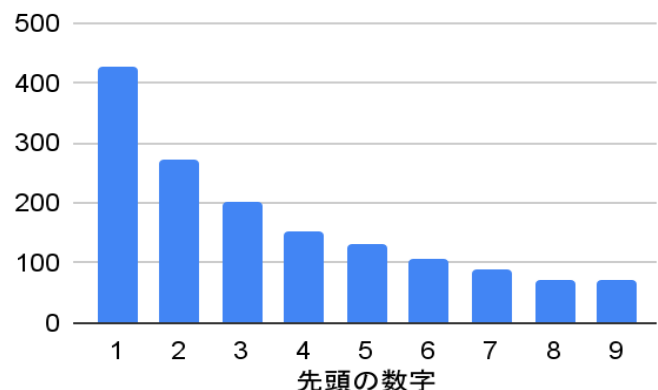
↑横軸は先頭の数字、縦軸は先頭の数字



《実験1-1》

- ①令和3年住民基本台帳年齢階級別市区町村別人口(データ数1528個)のデータの先頭の数の1~9までの個数を調べた。
- ②それを基に棒グラフを作成した。
- ③理論値のグラフと比較した。

先頭の数字	人口の合計
1	427
2	273
3	203
4	152
5	133
6	108
7	89
8	73
9	70



↑(左)先頭の数字(右)各市町村の人口の先頭の数字の合計

↑左のグラフを棒グラフにしたもの。縦軸は各市町村の人口の先頭の数字の合計

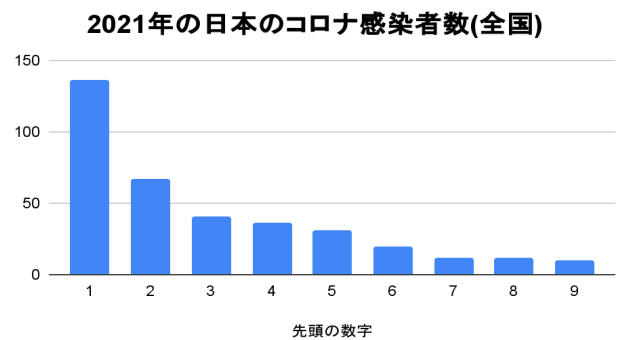
《実験1-2》

- ①2021年の日本全国のコロナウイルス感染者数(データ数365個)のデータの先頭の数の1~9ま

での個数を調べた。

- ②それを基に棒グラフを作成した。
- ③理論値のグラフと比較した。

1	136
2	67
3	41
4	36
5	31
6	20
7	12
8	12
9	10



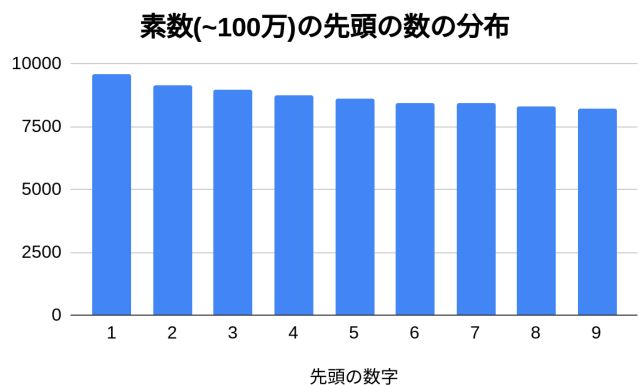
↑(左)先頭の数字(右)2021年の感染者数の先頭の数字の合計

↑左のグラフを棒グラフにしたもの。縦軸は2021年の感染者数の先頭の数字の合計

《実験2》

- ①1~100万までの素数(データ数78498個)のデータの先頭の数の1~9までの個数を調べた。
- ②それを基に棒グラフを作成した。
- ③理論値のグラフと比較した。

1	9585
2	9142
3	8960
4	8747
5	8615
6	8458
7	8435
8	8326
9	8230



↑(左)先頭の数字(右)1から100万までの素数の先頭の数字

↑左のグラフを棒グラフにしたもの。縦軸は素数の先頭の数字

3. 結果

《実験1》

市区町村の人口と2021年の感染者数はベンフォードの法則に従った。

《実験2》

1から100万までの素数はベンフォードの法則に従わなかった。

4. 考察

ベンフォードの法則が適用できるデータは数え上げられるものであり、データの範囲がきっちり決まっているわけではないものの、素数のように無限にあるようなものは適用がしにくいまたはデータをとる範囲を考える必要がある。したがって人口の他に選挙の投票数やSNS等の各アカウント別のフォロワー数にも適用できると考える。

5. 結論

ベンフォードの法則に従うデータの条件が明確に定まったわけではないが、無限であるものには注意しなければならない。今後の展望としては乱数のようなランダムで存在するデータについて考えるようになるか調べたい。

6. 参考文献ならびに参考Webページ

織田哲平.“データの不正を一瞬で暴く！？「ベンフォードの法則」について”.データの時間.
2021-4-21.<https://data.wingarc.com/benfords-law-30399>

永野裕之.“Googleを世界一にした経済学者も応用した「ベンフォードの法則」で、帳簿や選挙などの不正を見抜”.DIAMONDonline.2020-6-17.<https://diamond.jp/articles/-/240539>.