Loi de Benford et implémentation python

Franck Jeannot

Montréal, Canada, Avril 2023, AA799, v1.1

Abstract

A reminder on Benford's law simplifications and a simple python implementation.

Keywords: Benford's law, log simplification, python

1. Introduction

La loi de Frank Benford [1], également connue sous le nom de *loi des nombres anormaux*, est une loi empirique qui décrit la distribution des chiffres significatifs dans de nombreux ensembles de données du monde réel. Cette loi a été introduite dans l'article intitulé *"The law of anomalous numbers"* présenté en 1937 et publié en 1938 dans les *Proceedings of the American Philosophical Society* par Frank Benford, un physicien américain.

L'article commence par noter que dans de nombreux ensembles de données, tels que les populations de villes, les longueurs de rivières, les prix des produits et les résultats financiers, le premier chiffre significatif des nombres ne semble pas être uniformément distribué, comme on pourrait s'y attendre à première vue. Au lieu de cela, il existe une tendance distincte où les chiffres plus petits (1, 2, 3) apparaissent plus fréquemment que les chiffres plus grands (7, 8, 9). Benford appelle cette tendance les « nombres anormaux » et cherche à expliquer cette observation.

Benford propose ensuite une formule mathématique pour décrire cette loi des nombres anormaux, appelée maintenant la loi de Benford. La formule stipule que la probabilité que le **premier** chiffre significatif d'un nombre soit d'un certain chiffre d est donnée par la formule :

Loi de Benford : équivalences et simplifications

$$\begin{aligned} P(d) &= \log_{10}(d+1) - \log_{10}(d) \\ &= \log_{10}((d+1)/d) \\ P(d) &= \log_{10}(1+\frac{1}{d}) \end{aligned}$$

$$\log(a) - \log(b) = \log(a/b) \qquad (\text{ propriété des logarithmes})$$

Ainsi, la formule $P(d) = \log_{10}(d+1) - \log_{10}(d)$ est équivalente à la formule $P(d) = \log_{10}(1 + \frac{1}{d})$, et les deux formules peuvent être utilisées indifféremment pour décrire la loi de Benford.

Exemple simplifié d'implémentation en python :

Python Function :

```
import math
def benfordlaw(d):
    return math.log10(1 + 1/d)
```

```
print("Value for d=2:", benfordlaw(2));
Value for d=2: 0.17609125905568124;
```

Cette formule montre que la probabilité d'observer un chiffre plus petit est plus élevée que celle d'observer un chiffre plus grand, conformément à l'observation empirique de Benford.

L'article de Benford souligne également que cette loi des nombres anormaux a des implications importantes dans de nombreux domaines, tels que la détection de fraudes, l'analyse de données financières et la vérification de l'authenticité de données scientifiques. Il note également que cette loi a été vérifiée empiriquement dans de nombreux ensembles de données du monde réel, ce qui confirme sa validité en tant que modèle statistique.

En conclusion, l'article de Benford introduit la loi des nombres anormaux, également connue sous le nom de loi de Benford, qui décrit la distribution des chiffres significatifs dans de nombreux ensembles de données du monde réel. Il propose une formule mathématique pour décrire cette loi et souligne son importance dans divers domaines de recherche et d'application.

Références

[1] Franck Jeannot, Expérimentations sur la loi de Benford (08 2017).
URL https://franckybox.com/wp-content/uploads/Benfords_law.pdf