Benford’s Project

Figs 5,6,8,9,10: Summary Charts

Edward Stansfeld

January 2022

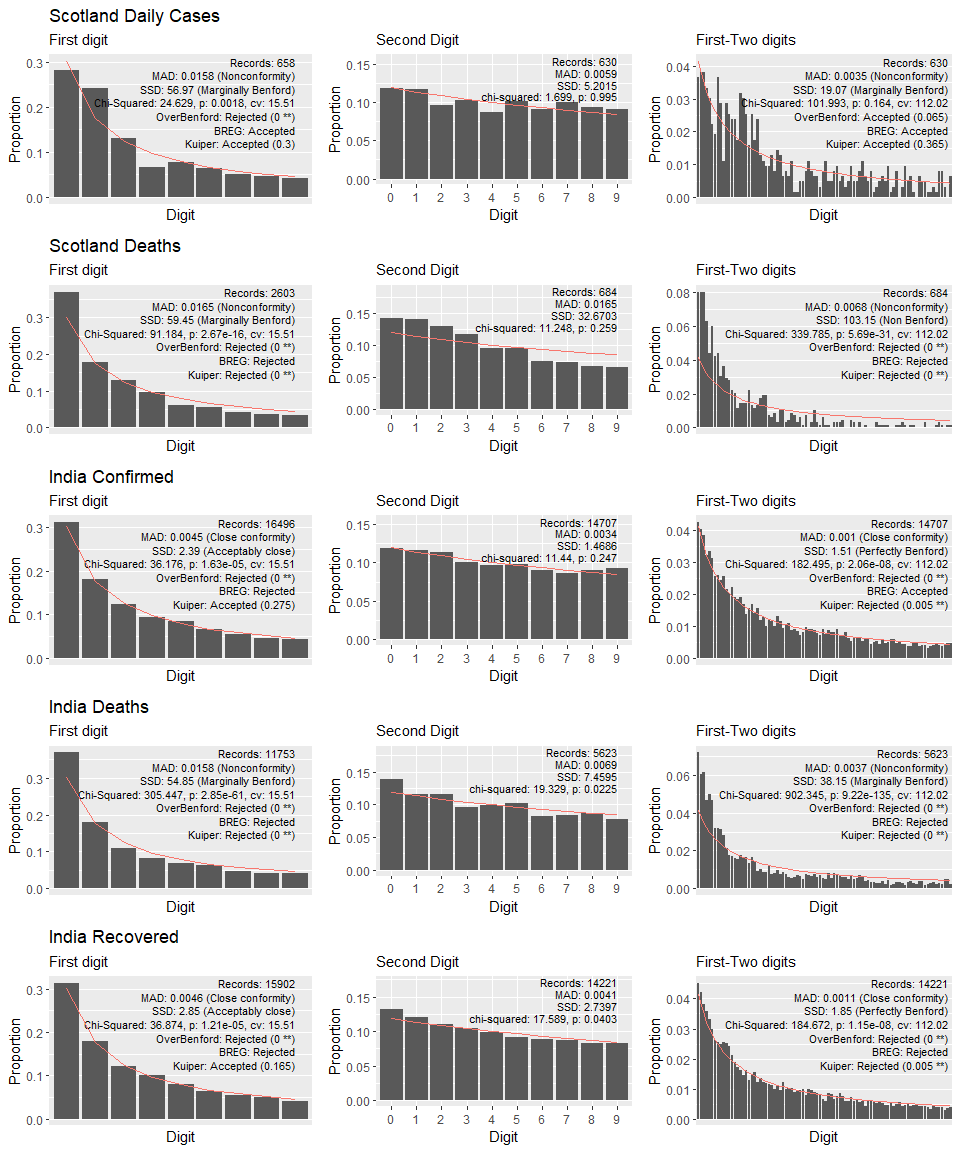
## Introduction

After the suitability of the data for Benford’s tests is established, Nigrini suggests that first and second digit tests should be run to provide an overview, and if there are more than about 300 records, the first two digits test should be used for the actual analysis (while allowing for more deviation from the expected values). The R benfords.analysis package (Cinelli 2019) contains a suite of fast Benfords tests and some statistical analysis, and is based on the approach of Nigrini 2012. This algorithm was applied to the daily case data and the daily death data and its results are shown in Appendix 1. The algorithm was supplemented with additional functions to extract and plot the second digit and calculate statistics.

## Results

### Does the data follow Benford’s Law?

The Benford’s plots for the first, second and first-two digits are shown below. Removing records below 10 affected the mantissa curve and it is expected there will be a similar effect here. Similarly, the deaths data are not expected to behave in the classic manner due to its restricted range.



The plots show marked Benford’s-like behaviour, with steep frequency curves descending from the lowest digit to the highest. The deaths curves have steeper descents with a higher ratio of 1’s than the classical Benford’s distribution. This is known as “Benford’s extreme” and is associated with a limited range of orders of magnitude.

## Summary statistics

## Warning: package 'knitr' was built under R version 4.1.2

STATISTICAL SUMMARY FOR SCOTLAND AND INDIA AGGREGATE DATA

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dataset | Test | Records | MAD | SSD | Chi.squared | overBenford | BREG | Kuiper |
| Scotland confirmed | First digit | 658 | 0.016 | 56.966 | 24.629 | 0 \*\* | Accepted | 0.3 |
|  | First-Two digits | 630 | 0.003 | 19.066 | 101.993 | 0.065 | Accepted | 0.365 |
| Scotland deaths | First digit | 2,603 | 0.017 | 59.453 | 91.184 | 0 \*\* | Rejected | 0 \*\* |
|  | First-Two digits | 684 | 0.007 | 103.146 | 339.785 | 0 \*\* | Rejected | 0 \*\* |
| India confirmed | First digit | 16,496 | 0.005 | 2.394 | 36.176 | 0 \*\* | Rejected | 0.275 |
|  | First-Two digits | 14,707 | 0.001 | 1.510 | 182.495 | 0 \*\* | Accepted | 0.005 \*\* |
| India deaths | First digit | 11,753 | 0.016 | 54.852 | 305.447 | 0 \*\* | Rejected | 0 \*\* |
|  | First-Two digits | 5,623 | 0.004 | 38.146 | 902.345 | 0 \*\* | Rejected | 0 \*\* |
| India recovered | First digit | 15,902 | 0.005 | 2.852 | 36.874 | 0 \*\* | Rejected | 0.165 |
|  | First-Two digits | 14,221 | 0.001 | 1.854 | 184.672 | 0 \*\* | Rejected | 0.005 \*\* |

## Conclusion

By eye, these datasets are conformant with Benford’s Law, despite some of the calculated measures rejecting conformance.

## References

Cinelli, C. (2019) benford.analysis: Tools that make it easier to use Benford’s law for data validation and forensic analytics., GitHub. Available at: <https://github.com/carloscinelli/benford.analysis> (Accessed: 23 October 2021).

Nigrini, M. (2012) Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud … - Mark J. Nigrini - Google Books. Available at: <https://books.google.co.uk/books?hl=en&lr=&id=Bh5Vr_I1NZoC&oi=fnd&pg=PP11&ots=qcigiizrOi&sig=Zyv0EV3snhVKH3yOC9z117ieSlI&redir_esc=y#v=onepage&q&f=false> (Accessed: 11 August 2021).