A Personal Paper on Statistics and its Real World Applications

Matthew Balogh

Table of contents

Concepts

Logarithmic scale

Logarithmic scale or simply **log scale** is used to display data that would span a broad range of values on a *linear scale*, especially when there are significant magnitude differences between the individual data points. [1]

While on a *linear scale*, each unit corresponds to the same increment on the scale, on a *logarithmic scale* each unit corresponds to a multiple of the base value and each subsequent unit is the multiplication of the previous one using the base value. [1]

Semi-log plot

A **semi-log plot** or graph has one axis on a logarithmic and one axis on a linear scale. [2]

Log-log plot

A **log-log plot** or graph has both its abscissa and ordinate in a logarithmic scale.

> **Note**
>
> A semi-log scaled plot can help in the following:
>
> - shrink the data points into a smaller area
> - draw the best-fit line if the original data follows an exponential trend

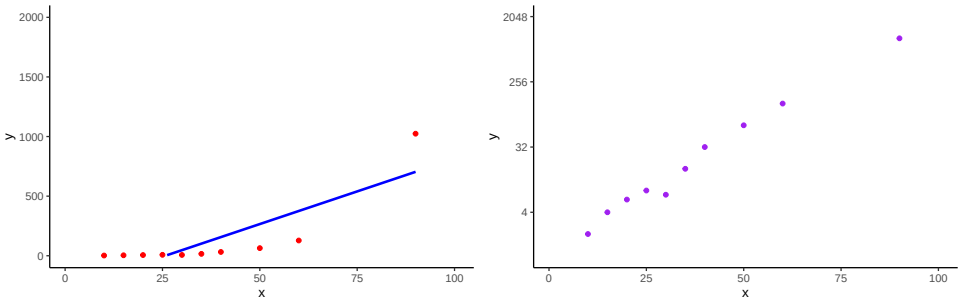
Tip

See Example 0.1 in the Appendix section for a practical example on this topic.

Appendix

Example 0.1. Finding the best-fit line (might be a curve?) - utilizing logarithmic scale.

Given a dataset which shows a raising but strange tendency depicted in Figure 1a. If one tries to fit a straight line between these data points they observe that those do not hug the linear so well. It seems that the differences between values on the ordinate get bigger as the values on the abscissa increase. In such case, switching to a logarithmic scale on the ordinate might be a good decision as presented in Figure 1b.

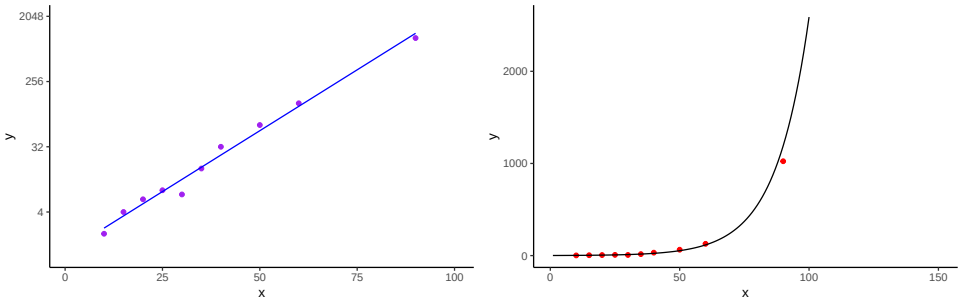


(a) Linear scales, the data shows non-linear uptrend.

(b) Semi-log scales, the data shows an uptrend close to linear.

Figure 1: The same dataset plotted on two graphs with linear and semi-log scales, respectively.

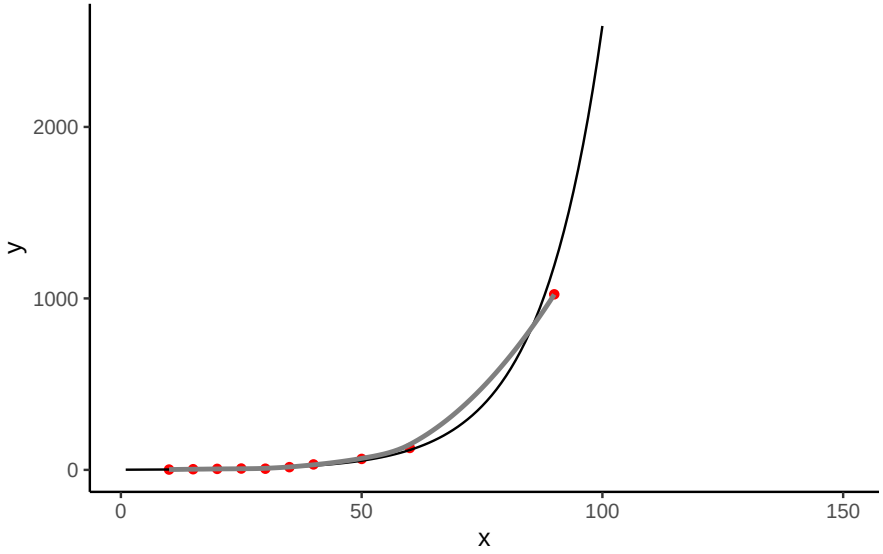
The data on the semi-log (`log2`) plot shows a tendency close to linear which means that the data on the original scale follows an exponential trend. One could find the best-fit line on the semi-log plot (Figure 2a) and then transform it back to the original scales to arrive at a non-linear, exponential best-fit curve (Figure 2b) that describes the original data on the original scales.



(a) (b)

Warning

Applying `geom_smooth` on the original data indicates a different curve from the one acquired above. This might be because data points on the semi-log plot do not perfectly fit the straight line, nor the exponential curve on the original plot.



References

[1] *Wikipedia*. Available: https://en.wikipedia.org/wiki/Logarithmic_scale

[2] *Wikipedia*. Available: https://en.wikipedia.org/wiki/Semi-log_plot