

Real Versus Spurious Correlation

Don't rush to a chocolate shop after looking at the left panel of this picture (*Figure 1*)! A strong correlation between two variables as measured by a high correlation coefficient may not necessarily mean that a causal relationship exists between the two. In the graph (or plot of points) on the left panel, the ability to win Nobel Prizes (as measured by Nobel Laureates per 10 million) by a country is highly correlated with the per capita consumption of chocolate of that country. Does that mean eating lots of chocolate increases mental ability or creativity? Most probably not. An explanation for this 'spurious correlation' may be that both variables are influenced by other factors such as human development index, per capita income and lifestyle of the country.

On the other hand, real correlations as seen in the scatter plot on the right are useful in establishing relationships between variables and for predicting future outcomes. This plot displays data on water flow in Kootenai River at two places, Libby (Montana, USA) and Newgate (British Columbia, Canada) in January of each year during 1931–1943. A dam was being planned on the river at Newgate, BC, where it crossed the Canadian border. The question was how the flow at Newgate could be estimated from that at Libby. We could have very well displayed data from River Kaveri instead, but why court controversy unnecessarily?

We note further that spurious correlations simply disappear when one calculates these correlations conditional on the influencing factors.

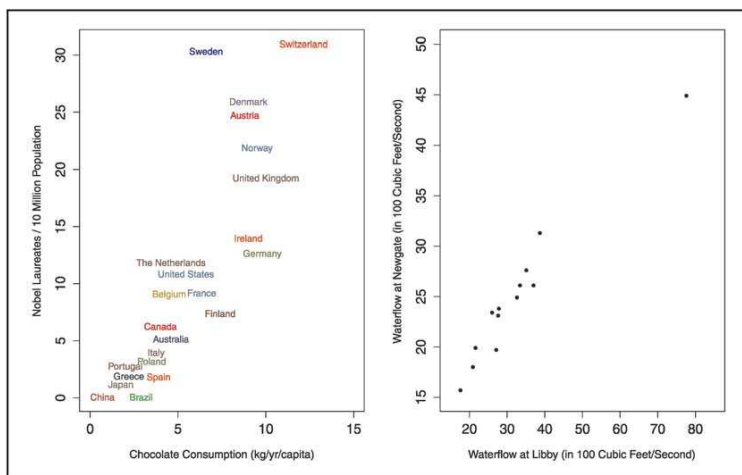


Figure 1. Graphs were drawn using the R package (R: a language and environment for statistical computing).

R Development Core Team (2008). R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Sources

- [1] <http://en.wikipedia.org>
- [2] M Ezekiel and F A Fox, *Methods of Correlation and Regression Analysis*, Wiley, New York, 1959.
- [3] F H Messerli, Chocolate Consumption, Cognitive Function, and Nobel Laureates, *New England Journal of Medicine*, Vol.367, pp.1562–1564, 2012.

Mohan Delampady