

# How NOT to do with Statistics

*Larry Shepp*

*Wharton School, University of Pennsylvania*

*Rutgers University (emeritus)*

*Department of Statistics*

`shepp@stat.rutgers.edu`

## Abstract

The first part of my talk is not too controversial: if a person tries to show you that two time series are correlated by showing that the empirical correlation of the partial sums of the time series is non-zero, then this person is either dishonest or incompetent. A colleague of mine, Abraham Wyner, pointed out that this abuse of statistics is not as rare as it ought to be; it has been used, for example, to argue for the hypothesis that carbon burning causes global warming. I will show why this is wrong and then pose an interesting if somewhat unnecessary mathematical problem on it. The second part is a critique of the current trend in statistics based on what I would call the blind use of regression methodology to train an algorithm for data-mining purposes. This will be more controversial which explains why I refer to the axe-swinging fellow in old China.



班门弄斧