

Some Hypotheses about ARORA, the Financial Turing Test

Patrick Burns

8th November 2010

Abstract

Abstract: We suggest some hypotheses about why people are successful in ARORA. We offer ways of generating the random data to eliminate the use of those tricks. We also argue that success in ARORA is not necessarily an indication of useful knowledge.

1 Introduction

The ARORA (A Random Or Real Array) website at <http://arora.ccs.neu.edu> allows you to test yourself in deciding which of two series is real data and which is randomly generated. It is great fun.

If you have not yet played the game, I encourage you to play it before continuing. Assuming you have success, what do you think you are seeing that allows you to distinguish real from random?

2 Some Details

[Hasanhodzic et al., 2010] describes the experiment that resulted in the ARORA website.

Table 1 shows the eight datasets that were used and the p-values of the tests on each.

Table 1: Datasets used.

Data	Frequency	p-value
Nasdaq Composite	tick	<1e-5
Russell 2000 Index	tick	<1e-5
US Dollar Index	tick	<1e-5
Gold spot	tick	<1e-5
Dow Jones Corporate Bond	daily	<1e-5
Dow Jones Industrial Average	daily	<1e-5
Canada/US exchange rate	daily	.0004
S&P GSCI Commodity Index	daily	.00503

3 My Hypotheses

I formed my hypotheses before knowing the details of the data that were used.

3.1 Momentum

The main clue that I seem to use is the fine grained texture. Often the real data have a smoother look than the random data. This indicates positive autocorrelation in the real data — that is, momentum. I had hypothesized that the data were intraday, and this turned out to be partially true.

[Hasanhodzic et al., 2010] quotes one of the subjects who partially confirms my experience. But that person says that for some sets of data, the reverse is true. That is, the real data is rougher than the random data. This would suggest mean reversion (negative autocorrelation).

It is possible to test my hypothesis. Instead of permuting the returns, simulate the returns as an AR(1) or possibly higher order autoregressive model with innovations that have a Student's t distribution with degrees of freedom that are fit from the original data. My expectation is that there would be only a small ability to discriminate real from random. Such randomly generated series would be predicable, but I would think the predictability of the prices would be minimal.

3.2 Change

A secondary characteristic that I use is that sometimes the texture of the real data changes. In market terms, the amount of momentum that is exhibited changes over time. This is not going to happen with the permuted data which has constant zero autocorrelation.

The use of this feature could be eliminated also by simulating an autoregressive model with changing parameters. I hypothesize that people would not be able to distinguish real data against this data generating mechanism.

4 Other Opinions

The experience of at least one person is that they distinguished real from random via support and resistance lines. That was not my experience because I don't believe in them. It is highly likely that at least one of us is fooling ourselves about how we succeed at selecting the real data.

[Hasanhodzic et al., 2010] seems to lend some support to the technical analysis line of thinking:

This gap between human and algorithmic pattern recognition may explain the gulf separating technical analysis (a largely human endeavor) and quantitative financial analysis (a more analytical and algorithmic approach), and why the former practice persists despite the lack of support from the latter.

The persistence of a practice doesn't mean it is actually useful. Astrology persists despite massive evidence that it doesn't work. This comparison is too harsh to technical analysis. In my opinion astrology should be rejected while technical analysis should merely be held up to severe skepticism.

If a test using the random generation mechanism suggested in Section 3 showed that the real data could be successfully identified, then the existence of support and resistance lines might be an explanation.

5 What about Prediction?

Distinguishing real from random is only useful if it allows you to predict. You can not predict a random series. But just because you know which is the real data, that doesn't mean you possess the useful knowledge of where it is going.

The Technical Analysis Challenge was a test in this direction. You can take the test yourself by going to <http://www.burns-stat.com/pages/techanchal.html>. This does not have a nice interface like ARORA. It also requires honesty since the answers are available.

I think a test of this sort comes closer to answering the question that we really want answered. A complaint from technical analysts about the Technical Analysis Challenge was that there was no volume data. Hence another study of this sort should probably include an indication of volume. However, that could make it much easier to identify the data.

References

[Hasanhodzic et al., 2010] Hasanhodzic, J., Lo, A., and Viola, E. (2010). Is it real, or is it randomized?: A financial Turing test. *arXiv.org*.