Introduction to Financial Technology - CA3

Finn Formica

1 Introduction

In 1992 Rust *et al.* published a paper on the topic of Double Auction Tournaments [1]. The goal was to gain new insights into effective trading strategies and to compare the performance of automated and human traders. The top ranked program was submitted by Todd Kaplan and was incredibly simple. The strategy involved waiting in the background for the other traders to reach a deal and as the market drew to a close, start shaving pennies off the price and steal the deal [1].

20 years later there are many newer types of automated trading agents that are widely used in research and testing. This report will concern itself with the comparison of these newer trading strategies with Todd Kaplan's automated trading strategy, dubbed SNPR. Although SNPR performed excellently in past tournaments, developments in the field of automated trading may have left the strategy in the dust. The hypothesis of this report is that SNPR will perform the worst out of the available trading agents during the experiment.

2 Methodology

The Bristol Stock Exchange (BSE) is a simulation written by Dave Cliff for educational and research purposes. It simplifies a real financial stock exchange to a Continuous Double Auction (CDA) with a Limit Order Book (LOB) that tracks bids and asks - if there is an overlap in the spread BSE will complete the transaction. This simulation can then be populated by several different automated trading agents, dubbed traders henceforth, which can buy or sell stock on BSE within their own market limits. BSE tracks the performance of the traders and logs all data to several files for later analysis.

The market conditions were kept simple and static to provide a benchmark of SNPR's performance against the rest of the traders populating BSE. Each simulation trial was run for 3 minutes, with traders being fed orders at 10 second fixed intervals with slight randomness (drip-jitter). Both the buyer and seller trading groups contained three of each agent type (ZIP, ZIC, GVWY, SHVR, SNPR) for a total of 30 traders in the market. The supply and demand ranges for the buyers and sellers were equal, with a minimum of 100 and maximum of 200 and a step-mode of jittered to provide some randomness. The simulation was run for a total of 50 trials and the average profits for each trader type were stored for later analysis.

3 Results

Before performing any tests, an understanding of the sample statistics and distribution of the data was gathered. The type of hypothesis test utilised is dependent on the distribution of the data, ie parametric tests assume a Gaussian distribution, and the range of the means and variances of the sample.

A quick and easy way to visualise this was a box-plot, shown in Fig. 1. It is difficult to get a good sense of the distribution of the data, however, it is clear that SNPR has the lowest mean overall. The rest of the traders seem to have similar means. This is likely due to SNPR only trading as the market is about to close, whereas the rest of the traders could generate profit throughout the market period. The highest mean was SHVR, which implements a similar strategy to SNPR but is not constrained to placing orders as the market is about to close.

To get a better sense for the distribution of the data a density plot and QQ-plot were generated. The density plot, shown in Fig. 2, shows that a Gaussian model would be reasonable for all samples - the data looks unimodal and symmetric about the peaks. The QQ-plot, shown in Fig. 3, verifies this since the points of the sample lie close to a straight line. The Shapiro-Wilk normality test was used to statistically verify the claim of a Gaussian fit. Table 1 shows that, for a significance level of $\alpha = 0.05$,

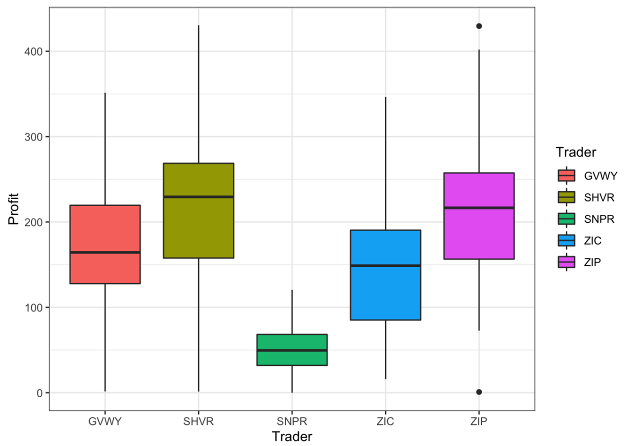

Figure 1: Box-plot of average profits

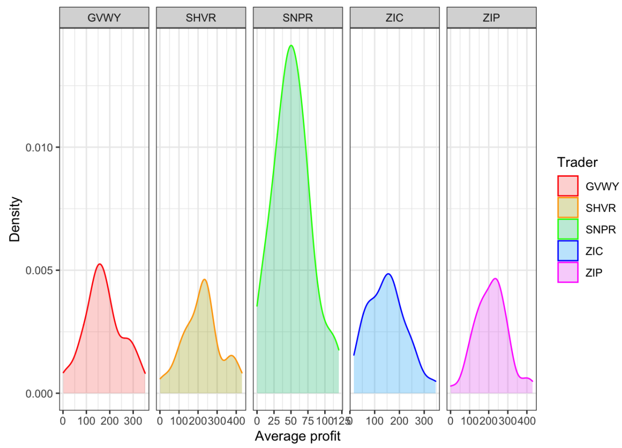

Figure 2: Density plot of sample distributions

there was no statistical significance to say that the samples did not follow a Gaussian distribution. The result from the Shapiro-Wilk test allowed the assumption that a parametric hypothesis test would be a good fit, since the underlying assumptions for those tests are a Gaussian distribution.

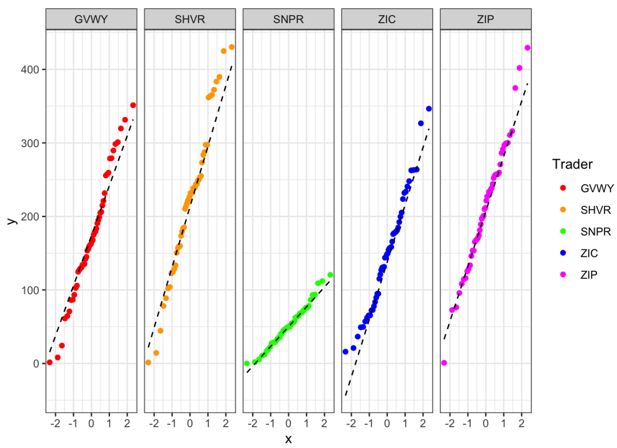

Figure 3: QQ-plot

Trader	p-value
ZIP	0.9181
ZIC	0.3503
GVWY	0.6813
SHVR	0.4587
SNPR	0.6113

Table 1: Shapiro-Wilk p-value statistics

4 Analysis

To test the hypothesis, the statistical test must be able to compare the differences in means and spreads of the distributions across the different samples. This can be accomplished by a Student's T-test, however, completing these tests pairwise causes an increase in the family-wise error rate and there would be a high chance of incorrectly rejecting the null hypothesis. Instead the ANOVA test was used, which measures all of the means and spreads of the samples against each other concurrently.

4.1 ANOVA

Since there was only one independent variable in the data, the traders, a one-way ANOVA test was utilised. The hypotheses for the ANOVA test were as follows:

- Null hypothesis, H_0: all population means are equal.
- Alternate hypothesis, H_1: at least one population mean is different from the others.

The test summary, shown in Table 2, showed a very high significance within the data corresponding to a significance level of 0.001. The null hypothesis was rejected and there was at least one population mean that was different from the others. However, the ANOVA test is unable to determine which mean, and a post-hoc test is required for that discrimination.

	Trader	**Residuals**
Df	4	245
Sum Sq	928325	1521150
Mean Sq	232081	6209
F-value	37.38	
p-value	$< 2 \times 10^{-16}$	

Table 2: Summary of ANOVA test

Trader pair	GVWY	SHVR	SNPR	ZIC	ZIP
GVWY	-	0.02	0.00	0.57	0.08
SHVR	0.02	-	0.00	0.00	0.98
SNPR	0.00	0.00	-	0.00	0.00
ZIC	0.57	0.00	0.00	-	0.00
ZIP	0.08	0.98	0.00	0.00	-

Table 3: p-value statistics from Tukey's HSD test

4.2 Post-hoc Tukey's HSD

A common follow up to ANOVA is Tukey's HSD, which performs multiple pairwise comparisons between groups, whilst accounting for the issues described earlier with the pairwise T-test.

Table 3 displays the p-value statistics from the Tukey's HSD test. SNPR is shown to have p-value < 0.05 across all other traders which indicates that the low mean shown in Fig. 1 is statistically significant, and hence, SNPR is the least profitable out of all the trading agents on BSE within static market conditions. The result is in-line with the hypothesis that the newer traders available would perform better due to advances in research.

4.3 Kruskal-Wallis

The ANOVA and Tukey's HSD tests showed that there was a clear and significant difference between the means of the samples. Although the data was not skewed there were a large range of variances between the samples. Since the Kruskal-Wallis is less sensitive to distributions than ANOVA this test was conducted as a sanity check. The hypotheses for the Kruskal-Wallis test are equivalent to that of ANOVA but look at mean rank instead.

Chi-squared	**Df**	**p-value**
103.6	4	$<2.2 \times 10^{-16}$

Table 4: Summary of Kruskal-Wallis test

The Kruskal-Wallis test also produced a very small p-value, shown in Table 4. This verified the analysis with the ANOVA test, as the null hypothesis was rejected and there was at least one population mean rank that was different to the others.

5 Conclusion

This report was concerned with the analysis of the profits of the SNPR trading agent taken from the experiment by Rust *et al.*. Through the ANOVA and Tukey's HSD tests SNPR was shown to have a statistically lower mean than the other traders populating BSE in static market conditions.

However, financial markets are often not static, so further experimentation in dynamic markets, and introducing supply shocks is still required. It would also be beneficial to vary the types and proportions of traders on the market, and other historically significant scenarios. To make the simulation more realistic, introducing fees depending on the market time and equilibrium price would also be significant, and would allow further research into trader efficiency.

References

[1] J Rust, JH Miller, and R Palmer. Behavior of trading automata in a computerized double auction market. in the double auction market: institutions, theories, and evidence. *edited by Daniel Friedman and John Rust. Santa Fe Institute Studies of the Science of Complexity, New York: Addison-Wesley*, 1992.