An application of the Foster - Vohra method of calibration to stock market games.

Vladimir V'yugin

Institute for Information Transmission Problems
Russian Academy of Sciences

International Workshop on Statistical Learning, June 26-28, Moscow

logo

Assymptotic calibration

- Forecasting a sequence of outcomes from an unknown source.
- Nonstandard measure of performance: calibration tests (Dawid(1984)).
- Known applications: (1) convergence to correlated equilibrium (Foster and Vohra (1997)); (2) Vovk's defensive forecasting (Vovk (2006); (3) constructing aggregating strategies (Vovk (2010)).
- **We present applications to algoritmic trading.**

Online forecasting task

FOR $n = 1, 2, \ldots$
Forecaster announces a forecast $p_n \in [0, 1]$.
Nature announces an outcome $S_n \in [0, 1]$.
ENDFOR

Example for binary case $S_n \in \{0, 1\}$:
p_n is a probability that it will rain ($S_n = 1$ means rain).

In general, p_n can be interpreted as the expected value of S_n.

How to evaluate performance of the forecaster?

Dawid's (1984) method of calibration: informal setting

Informally: It should rain about 80% of the days for which $p_n = 0.8$, and so on.

A sequence of forecasts $p_1, p_2, \ldots$ is "calibrated" for an infinite binary sequence $S_1, S_2 \ldots$ if for each $p^* \in [0,1]$

$$\frac{\sum_{p_i \approx p^*} S_i}{\sum_{p_i \approx p^*} 1} \approx p^*$$

as the denominator of this relation tends to infinity when $n \to \infty$.

logo

Formal definition:

A sequence of forecasts $p_1, p_2, \ldots$ is well-calibrated for an infinite sequence $S_1, S_2 \ldots$ if for the indicator function $I(p)$ of any subinterval $I \subseteq [0,1]$ the calibration error tends to zero:

$$\frac{\sum_{i=1}^{n} I(p_i)(S_i - p_i)}{\sum_{i=1}^{n} I(p_i)} \longrightarrow 0$$

as the denominator of this relation tends to infinity when $n \to \infty$.

Commonly used the weaker condition: for each checking rule $I(p)$,

$$\frac{1}{n}\sum_{i=1}^{n} I(p_i)(S_i - p_i) \to 0 \text{ as } n \to \infty$$

Adversarial Nature (Oakes (1985))

Any total deterministic forecasting algorithm f

$$p_n = f(S_1, S_2, \ldots, S_{n-1})$$

is not calibrated for the sequence $S_1, S_2 \ldots$, where

$$S_i = \begin{cases} 1 \text{ if } p_i < 0.5 \\ 0 \text{ otherwise.} \end{cases}$$

The condition of calibration fails for $I = [0, 0.5)$ or $I = [0.5, 1]$.

logo

Probability forecasting game

Foster and Vohra (1994) – first positive result

FOR $n = 1, 2, \dots$
Forecaster announces a probability distribution P_n on $[0, 1]$.
Nature announces an outcome $S_n \in [0, 1]$.
Forecaster draws $\tilde{p}_n$ from P_n and observes outcome S_n.
ENDFOR

The forecast $\tilde{p}_n$ is hidden from **Nature** at step n.
Distributions P_n, $n = 1, 2, \dots$ defines the overall probability distribution on infinite trajectories $\tilde{p}_1, \tilde{p}_2, \dots$ of forecasts.

logo

Kakade and Foster's (2004) calibration theorem:

P_n can be taken in form of random rounding of a real number.

For any $\Delta > 0$, given binary $S_1, \ldots, S_{i-1}$ some algorithm computes deterministic forecast p_i and randomly rounds it up to Δ to $\tilde{p}_i$ such that:

For the characteristic function $I(p)$ of any subinterval of $[0,1]$

$$\limsup_{n\to\infty} \left| \frac{1}{n} \sum_{i=1}^{n} I(\tilde{p}_i)(S_i - \tilde{p}_i) \right| \le \Delta$$

almost surely.

This result is also valid for real outcomes $S_i \in [0,1]$.

logo

Probability forecasting game with side information

FOR $n = 1, 2, \ldots$
Nature announces a signal $x_n \in [0, 1]$.
Forecaster announces a probability distribution P_n on $[0, 1]$.
Nature announces an outcome $S_n \in [0, 1]$.
Forecaster draws $\tilde{p}_n$ from P_n and observes outcome S_n.
ENDFOR

logo

Side information and history depended checking rules

History at step n: $c_n = (\tilde{p}_1, S_1, \dots, \tilde{p}_{n-1}, S_{n-1})$.

Extracton of information: $\bar{z}_n = f(x_n, c_n) \in [0,1]^k$, where x_n – signal, c_n – history, f – extraction function.

Let $R \subseteq [0,1]^{k+1}$ and its indicator function:

$$I_R(p, \bar{z}) = \begin{cases} 1 \text{ if } (p, \bar{z}) \in R, \\ 0 \text{ otherwise,} \end{cases}$$

where $p \in [0,1]$ – forecast and
$\bar{z} \in [0,1]^k$ – extracted information vector.

logo

General calibration theorem:

Theorem

Given $\varepsilon > 0$ and extraction function f with range in $\mathscr{R}^k$ we can compute forecasts $p_1, p_2, \ldots$ and randomize them and extracted vectors such that: for any subset $R \subseteq [0.1]^{k+1}$, n, and for any $\delta > 0$, with probability at least $1 - \delta$,

$$\left| \sum_{i=1}^{n} I_R(\tilde{p}_i, \tilde{z}_i)(S_i - \tilde{p}_i) \right| \leq$$

$$\leq 22 \left(\frac{k+1}{4} \right)^{\frac{2}{k+3}} n^{1 - \frac{1}{k+3} + \varepsilon} + \sqrt{\frac{n}{2} \ln \frac{2}{\delta}},$$

where $\tilde{p}_i$ and $\tilde{z}_i$ are randomizations of p_i and information vector $\bar{z}_i = f(x_i, c_i) \in [0,1]^k$.

Aplications to Stock Market games

Trading game with side information

FOR $n = 1, 2 \ldots$
Stock Market announces side information $x_n \in [0, 1]$.
Trader buys C_n shares of the stock by S_{n-1} each.
Stock Market announces the price S_n of the stock.
Trader sells all shares by S_n and updates his capital:
$\mathscr{K}_n = \mathscr{K}_{n-1} + C_n(S_n - S_{n-1})$, where $\mathscr{K}_0 = 0$.
ENDFOR

Trading strategy is a rule:
$x_n \Rightarrow C_n$ or
c_n, $x_1, \ldots x_n \Rightarrow C_n$, where c_n – history

logo

Trading strategies

We compete trading strategies of two types:

- **Universal trading strategy:**
  $C_n = \tilde{M}_n$ – output of randomized algorithm which uses well calibrated forecasts $\tilde{p}_n$ of future price S_n (will be defined below).
- **Benchmark class of stationary trading strategies:**
  $C_n = D(x_n)$ – stationary trading strategy, where
  $D : [0,1] \to \mathscr{R}$ is a continuous function.
  We approximate continuous functions by functions from RKHS (Reproducing Kernel Hilbert Space).

Benchmark class of stationary trading strategies:

- RKHS is a Hilbert space $\mathscr{F}$ of real-valued functions on a compact metric space X such that the evaluation functional $f \to f(x)$ is continuous for each $x \in X$.
- $f(x) = (f \cdot \Phi(x))$.
- $K_2(x,y) = (\Phi(x) \cdot \Phi(y))$ – kernel.
- $\|\cdot\|_{\mathscr{F}}$ be a norm in $\mathscr{F}$.
- We consider RKHS $\mathscr{F}$ with finite embedding constant $c_{\mathscr{F}} = \sup_x \sup_{\|f\|_{\mathscr{F}} \le 1} |f(x)| = \|\Phi(\bar{x})\|_{\mathscr{F}} < \infty$.
- Universal RKHS: for any continouus f and $\varepsilon > 0$ and $g \in \mathscr{F}$ exists such that $\sup_{0 \le x \le 1} |f(x) - g(x)| < \varepsilon$.
- Sobolev space of absolutely continuous functions $f : [0,1] \to \mathscr{R}$ with $\|f\|_{\mathscr{F}} \le 1$, where $\|f\|_{\mathscr{F}} = \sqrt{\int_0^1 (f(t))^2 dt + \int_0^1 (f'(t))^2 dt}$ is universal.

logo

Calibration theorem:

Theorem

Given an RKHS $\mathcal{F}$ and $\varepsilon > 0$ we can compute forecasts $p_1, p_2, \ldots$ and randomize them and past prices such that for any $\delta > 0$ and for any n, with probability $1 - \delta$,

$$\left| \sum_{i=1}^{n} I(\tilde{p}_i > \tilde{S}_{i-1})(S_i - \tilde{p}_i) \right| \leq$$

$$\leq 18(c_{\mathcal{F}}^2 + 1)^{\frac{1}{4}} n^{3/4+\varepsilon} + \sqrt{\frac{n}{2} \ln \frac{2}{\delta}},$$

$$\left| \sum_{i=1}^{n} D(x_i)(S_i - p_i) \right| \leq \|D\|_{\mathcal{F}} \sqrt{(c_{\mathcal{F}}^2 + 1)n}$$

for all $D \in \mathcal{F}$, where $I(p > x)$ is the indicator of the condition $p > x$.

Universal trading strategy

At each round n, **Trader** buys (or sells)

$$\tilde{M}_n = \begin{cases} 1 \text{ if } \tilde{p}_n > \tilde{S}_{n-1}, \\ -1 \text{ otherwise.} \end{cases}$$

shares of the stock by S_{n-1} each, where
$\tilde{p}_n$ – randomized well calibrated forecast
S_{n-1} – randomized past price

logo

Capitals of traders

$\mathscr{K}_n^M = \sum_{i=1}^{n} \tilde{M}_i \Delta S_i$ – capital of universal trading strategy (random quantity)

$\mathscr{K}_n^D = \sum_{i=1}^{n} D(x_i) \Delta S_i$ – capital of stationary strategy $D(x)$,
where $\Delta S_i = S_i - S_{i-1}$

logo

Main result: universality

Theorem

There exists a universal tradimg strategy $\tilde{M}_n$ such that for any continuous function $D(x)$,

$$\liminf_{n\to\infty} \frac{1}{n}\left(\mathcal{K}_n^M - \|D\|_+^{-1}\mathcal{K}_n^D\right) \geq 0$$

almost surely.

$\|D\|_+ = \max\{1, \sup_{0\leq x\leq 1} |D(x)|\}$ – *normalization factor.*

We can take $D(x) = 0$ for all x.

Corollary

The universal trading strategy is asymptotically non-risk:

$$\liminf_{n\to\infty} \frac{1}{n} \mathscr{K}_n^M \geq 0$$

almost surely.

Rate of convergence for D in RKHS:

Theorem

Given $\varepsilon > 0$, for any n and $\delta > 0$, with probability $1 - \delta$, for all $D \in \mathscr{F}$ (RKHS)

$$\mathscr{K}_n^M \geq \|D\|_+^{-1} \mathscr{K}_n^D -$$
$$-38(c_{\mathscr{F}}^2 + 1)^{\frac{1}{4}} n^{\frac{3}{4}+\varepsilon} - \|D\|_+^{-1} \|D\|_{\mathscr{F}} \sqrt{(c_{\mathscr{F}}^2 + 1)n} - \sqrt{\frac{n}{2} \ln \frac{2}{\delta}}$$

logo

Competing with discontinuous trading strategies

Deterministic signals x_i: counterexample

Theorem

Let $\tilde{M}_i$ be a sequence of independent random variables (randomized trading strategy) such that $|\tilde{M}_i| \le 1$ for all i. Consider the protocol of trading game with signals $x_i = P\{\tilde{M}_i > 0\}$. Then a binary decision rule $D(x)$ and a sequence $S_1, S_2, \dots$ of prices can be defined such that with probability one

$$\limsup_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{M}_i\Delta S_i - \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n}D(x_i)\Delta S_i\right) \le 0. \qquad (1)$$

logo

Competing with discontinuous trading strategies

Randomized signals x_i: positive result

Theorem

An algorithm for computing forecasts and a sequential method of randomization of forecasts $\tilde{p}_i$ and past prices $\tilde{S}_{i-1}$ can be constructed such that for any nontrivial decision rule D for any $\delta > 0$, with probability at least $1 - \delta$,

$$\sum_{i=1}^{n} \tilde{M}_i \Delta S_i \geq \|D\|_{\infty}^{-1} \sum_{i=1}^{n} D(\tilde{\mathbf{x}}_i) \Delta S_i - O\left(n^{\frac{4}{5}+\varepsilon} + \sqrt{\frac{n}{2} \ln \frac{2m}{\delta}}\right),$$

where $\tilde{x}_i$ – randomized side information.

logo

Numerical experiments (with V.G.Trunov)

Data has been downloaded from FINAM site: www.finam.ru
Number of trading points in each game is 88000–116000 min.
(From March 26 2010 to March 25 2011)
We buy or sell 5 shares at each round
Transaction costs are ignored
It was found that $\mathscr{K}_n > 0$, i.e., we never incur debt in our experiments.

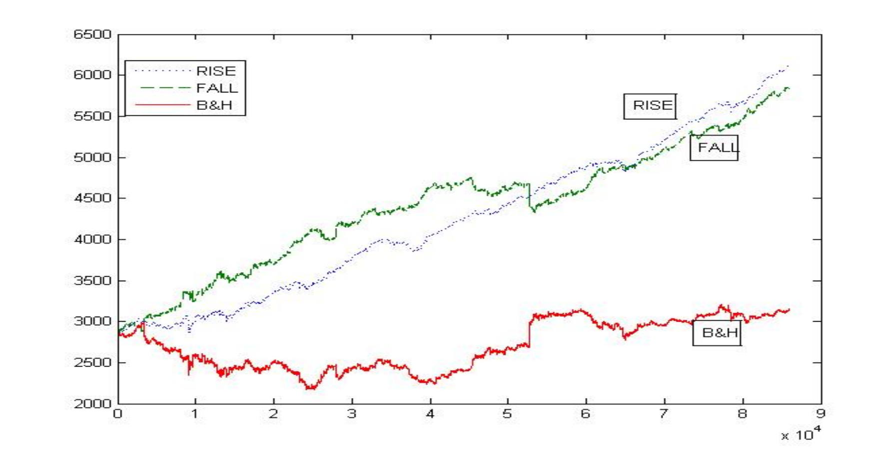
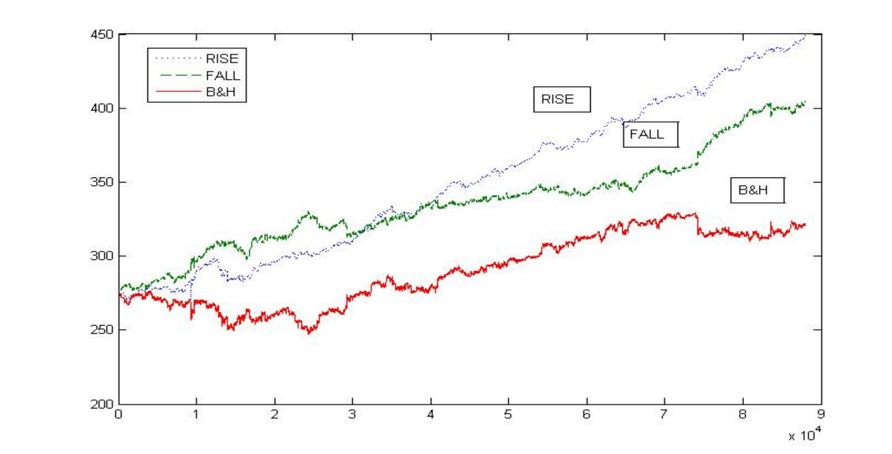

RISE
FALL
B&H
RISE
FALL
B&H
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
0
1
2
3
4
5
6
7
8
9
x 10^4

Figure: GOOG



450
400
350
300
250
200
0
1
2
3
4
5
6
7
8
9
x 10^4
RISE
FALL
B&H
RISE
FALL
B&H

Figure: KOCO

TICKER	BUY& HOLD PROFIT %	UN FOR A RISE PROFIT %	UN FOR A FALL PROFIT %	ARMA FOR A RISE PROFIT %	ARM FOR PRO
TEST	6.85	-1.39	-8.19	9.88	3.08
AT-T	7.71	137.40	129.70	30.73	23.0
KOCO	16.55	62.66	46.15	2.90	-13.
GOOG	10.25	114.85	104.62	12.85	2.62
INBM	24.28	85.38	61.09	29.31	5.02
INTL	4.29	111.70	107.50	25.86	21.6
MSD	10.71	58.32	47.60	18.66	7.95
US1.AMT	22.01	22.74	0.77	28.46	6.49
US1.IP	2.40	19.83	17.47	9.36	7.00
US2.BRCM	25.30	53.62	28.28	20.06	-5.2
US2.FSLR	40.15	143.92	103.61	-9.86	-50.
SIBN	-6.54	732.87	739.33	357.74	364
GAZP	22.75	101.20	78.45	31.75	9.00
LKOH	19.39	261.84	242.45	87.08	67.6
MTSI	-1.61	669.16	670.68	326.12	327

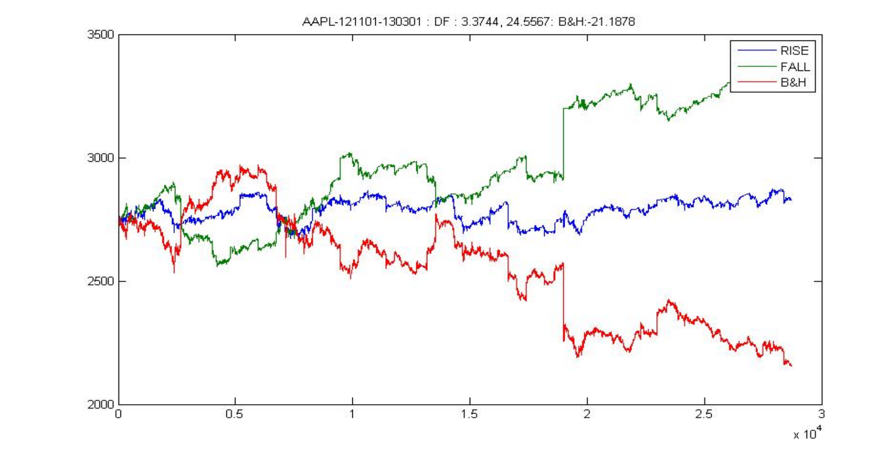

AAPL-121101-130301 : DF : 3.3744, 24.5567: B&H:-21.1878
RISE
FALL
B&H
3500
3000
2500
2000
0
0.5
1
1.5
2
2.5
3
x 10^4

Figure: AAPL

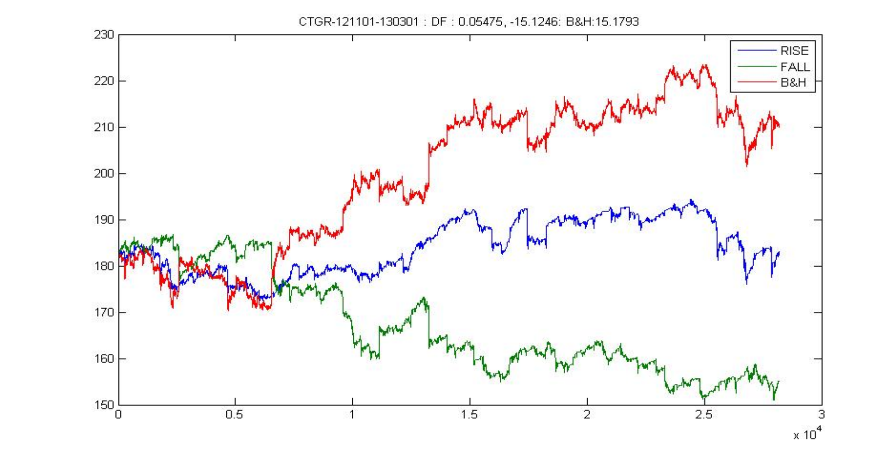

CTGR-121101-130301 : DF : 0.05475, -15.1246: B&H:15.1793
RISE
FALL
B&H
230
220
210
200
190
180
170
160
150
0
0.5
1
1.5
2
2.5
3
x 10^4

Figure: CTGR

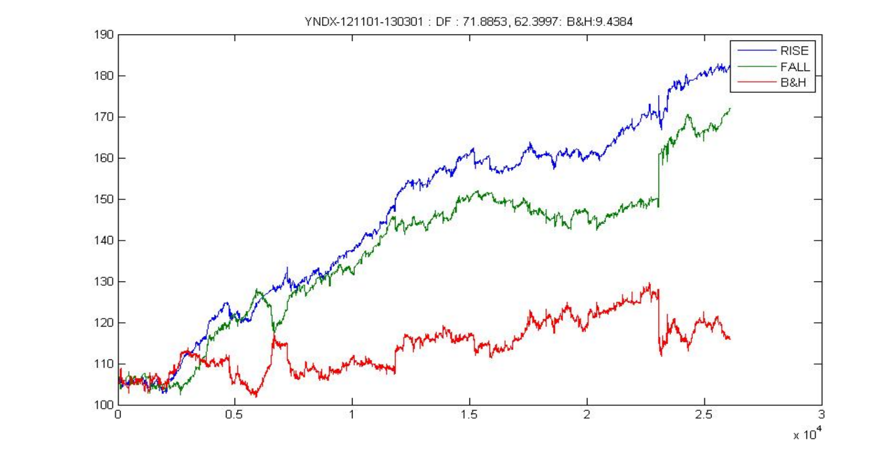

YNDX-121101-130301 : DF : 71.8853, 62.3997: B&H:9.4384
RISE
FALL
B&H
190
180
170
160
150
140
130
120
110
100
0
0.5
1
1.5
2
2.5
3
x 10^4

Figure: CTGR

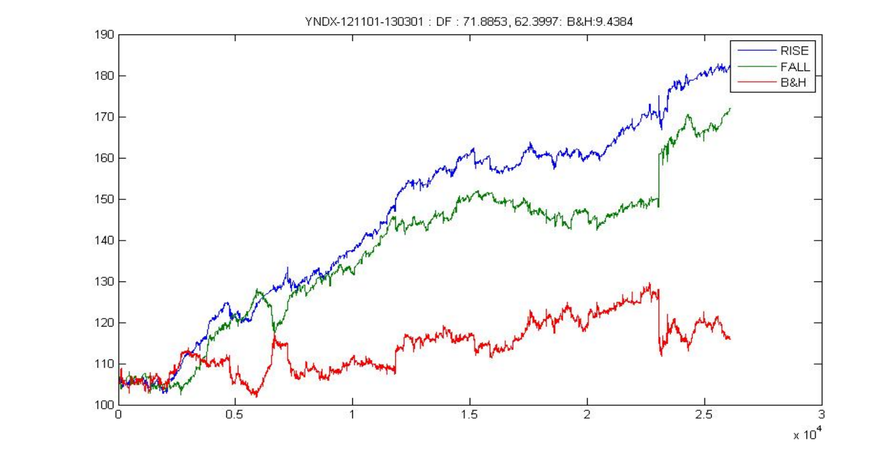

YNDX-121101-130301 : DF : 71.8853, 62.3997: B&H:9.4384
RISE
FALL
B&H
190
180
170
160
150
140
130
120
110
100
0
0.5
1
1.5
2
2.5
3
x 10^4

Figure: CTGR