

# Calibration with Changing Checking Rules and Its Application to Short-Term Trading

Vladimir Trunov and Vladimir V'yugin

Institute for Information Transmission Problems, Russian Academy of Sciences,  
Bol'shoi Karetnyi per. 19, Moscow GSP-4, 127994, Russia  
email: vyugin@iitp.ru

**Abstract.** We provide a natural learning process in which a financial trader without a risk receives a gain in case when Stock Market is inefficient. In this process, the trader rationally choose his gambles using a prediction made by a randomized calibrated algorithm. Our strategy is based on Dawid's notion of calibration with more general changing checking rules and on some modification of Kakade and Foster's randomized algorithm for computing calibrated forecasts.

## 1 Introduction

Predicting sequences is the key problem of machine learning and statistics. The learning process proceeds as follows: observing a finite-state sequence given on-line a forecaster assigns an subjective estimate to future states. The method of evaluation of these forecasts depends on an underlying learning approach.

A minimal requirement for testing any prediction algorithm is that it should be calibrated (see Dawid [3]). Dawid gave an informal explanation of calibration for binary outcomes as follows. Let a binary sequence  $\omega_1, \omega_2, \dots, \omega_{n-1}$  of outcomes be observed by a forecaster whose task is to give a probability  $p_n$  of a future event  $\omega_n = 1$ . In a typical example,  $p_n$  is interpreted as a probability that it will rain. Forecaster is said to be well-calibrated if it rains as often as he leads us to expect. It should rain about 80% of the days for which  $p_n = 0.8$ , and so on.

A more precise definition is as follows. Let  $I(p)$  denote the characteristic function of a subinterval  $I \subseteq [0, 1]$ , i.e.,  $I(p) = 1$  if  $p \in I$ , and  $I(p) = 0$ , otherwise. An infinite sequence of forecasts  $p_1, p_2, \dots$  is calibrated for an infinite binary sequence of outcomes  $\omega_1 \omega_2 \dots$  if for characteristic function  $I(p)$  of any subinterval of  $[0, 1]$  the calibration error tends to zero, i.e.,

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

as the denominator of the relation (1) tends to infinity.

The indicator function  $I(p_i)$  determines some "checking rule" which selects indices  $i$  where we compute the deviation between forecasts  $p_i$  and outcomes  $\omega_i$ .

If the weather acts adversatively, then Oakes [8] and Dawid [4] show that any deterministic forecasting algorithm will not always be calibrated.

Foster and Vohra [5] show that calibration is almost surely guaranteed with a randomizing forecasting rule, i.e., where the forecasts  $p_i$  are chosen using internal randomization and the forecasts are hidden from the weather until weather makes its decision whether to rain or not.

The origin of calibration algorithm is the Blackwell's [1] approachability theorem but, as its drawback, the forecaster has to use linear programming to compute the forecasts. We modify a more computationally efficient method from Kakade and Foster [7], where "an almost deterministic" randomized rounding universal forecasting algorithm is presented. For any sequence of outcomes and for any precision of rounding  $\Delta > 0$ , an observer can simply randomly round the deterministic forecast  $p_i$  up to  $\Delta$  in order to calibrate for this sequence with probability one :

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n I(\tilde{p}_i)(\omega_i - \tilde{p}_i) \leq \Delta, \quad (1)$$

where  $\tilde{p}_i$  is a random forecast. This algorithm can be easily extended such that the calibration error tends to zero as  $n \rightarrow \infty$ .

The goal of this paper is to extend Kakade and Foster's algorithm to arbitrary real valued outcomes and to a more general notion of calibration with changing parameterized checking rules. We present also convergence bounds for calibration depending on the number of parameters.

A closely related approach for weak calibration is presented in Vovk [10].

We apply this algorithm to technical analysis in finance. We consider real valued outcomes (for example, prices of a stock). In this case, predictions could be interpreted as mean values of future outcomes under some unknown for us probability distributions. We need not any form of such distribution – we should predict only future means.

We provide a natural learning process in which a financial trader (speculator for a rise or for a decline) without a risk of complete ruin receives a gain if the market is inefficient. In this process, the trader rationally choose his gambles using a prediction made by a randomized calibrated algorithm.

The learning process is the most traditional one. At each step *Forecaster* makes a prediction of future price of a stock and *Speculator* takes the best response to this prediction. He chooses a strategy: dealing for a rise or for a fall, or pass the step. *Forecaster* uses some randomized algorithm for computing calibrated forecasts.

Let us give a more precise formulation. Consider a game between *Speculator* and *Stock Market*. Let  $S_1, S_2, \dots$  – be a sequence of prices of a stock. We suppose that prices are bounded and rescaled such that  $0 \leq S_i \leq 1$  for all  $t$  and  $S_1 = S_0$ .

The protocol of a game is described as follows. Let  $k$  is a positive integer number. The initial capital of *Speculator* is  $\mathcal{K}_0 = 0$ .

**FOR**  $i = 1, 2 \dots$

At the beginning of the step  $i$  *Speculator* and *Forecaster* observe past prices

$S_1, \dots, S_{i-1}$  of a financial instrument (a stock) and some side information.  
*Forecaster* announces a random forecast of a stock future price – random variable  $\tilde{p}_i \in [0, 1]$ .  
*Speculator* bets by buying or selling a number  $M_i$  of shares of the stock by  $S_{i-1}$  each.<sup>1</sup>  
*Stock Market* announces a price  $S_i$  of a stock.  
*Speculator* receives his total gain (or suffer loss) at the end of step  $i$  :  
 $\mathcal{K}_i = \mathcal{K}_{i-1} + M_i(S_i - S_{i-1})$ .

#### ENDFOR

In that follows we consider only playing for a rise and will buy only one share of a stock, so  $M_i = 0$  or  $M_i = 1$ .

Let  $\epsilon > 0$  be a threshold for entering the game. A decision rule for entering will be the following: at step  $i$  enter the game (buy  $M_i = 1$  of shares) if  $\tilde{p}_i > \tilde{S}_{i-1} + \epsilon$ ; pass the step otherwise (get  $M_i = 0$ ), where  $\tilde{S}_{i-1}$  is randomized past price. Thereby, we need changing checking rules depending on past outcomes :

$$I(p_i > S_{i-1} + \epsilon) = \begin{cases} 1, & \text{if } p_i > S_{i-1} + \epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

It will follow from Theorem 1 (Section 2) that there exists a randomized algorithm computing forecasts calibrated almost surely in a modified sense :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)(S_i - \tilde{p}_i) = 0,$$

where  $\tilde{p}_i$  is a random forecast,  $\tilde{S}_{i-1}$  is a randomized past price of a stock, and  $\epsilon > 0$  is a threshold for entering a game.

In Section 3 we construct trading strategies based on calibrated forecasts.

## 2 Computing calibrated forecasts

Let  $y_1, y_2, \dots$  be an infinite sequence of real numbers. An infinite sequence of random variables  $\tilde{y}_1, \tilde{y}_2, \dots$  is called a *randomization* of  $y_1, y_2, \dots$  if  $E_n(\tilde{y}_n) = y_n$  for all  $n$ , where  $E_n$  is the symbol of mathematical expectation.

We specify a side information – we add to the protocol of the game signals  $\bar{x}_1, \bar{x}_2, \dots$  given online: for any  $n$ , a  $k$ -dimensional vector  $\bar{x}_n \in [0, 1]^k$  is given to *Forecaster* before he announces a forecast  $\tilde{p}_n$ . We consider checking rules of general type:

$$I(p, \bar{x}) = \begin{cases} 1, & (p, \bar{x}) \in \mathcal{S}, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\mathcal{S} \subseteq [0, 1]^{k+1}$  and  $\bar{x} \in [0, 1]^k$  is a signal. In Section 3 we use a set  $\mathcal{S} = \{(p, x) : p > x + \epsilon\}$ , where  $p, x \in [0, 1]$  and  $\epsilon > 0$ . At any step  $i$  we check  $(\tilde{p}_i, \tilde{x}_i) \in \mathcal{S}$ , where  $x_i = S_{i-1}$  and  $\tilde{p}_i, \tilde{x}_i$  are randomization of  $p_i, x_i$ .

<sup>1</sup> In case  $M_i > 0$  *Speculator* playing for a rise, in case  $M_i < 0$  *Speculator* playing for a fall, *Speculator* pass the step if  $M_i = 0$ . We suppose that *Speculator* can borrow money for buying  $M_i$  shares of a stock and return them after selling.

The following theorem on calibration is the main tool for technical analysis presented in Section 3.

**Theorem 1.** *Given  $k$  an algorithm  $f$  for computing forecasts and a method of randomization can be constructed such that for any sequence of real numbers  $S_1, S_2, \dots$  and for any sequence of  $k$ -dimensional signals  $\bar{x}_1, \bar{x}_2, \dots$  the event*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) = 0, \quad (2)$$

has  $Pr$ -probability 1, where  $Pr$  is a probability distribution generated by a sequence of tuples  $(\tilde{p}_i, \tilde{x}_i)$  of random variables,  $i = 1, 2, \dots$ , and  $I$  is the characteristic function of an arbitrary subset  $\mathcal{S} \subseteq [0, 1]^{k+1}$ . Here  $\tilde{p}_i$  is the randomization of a forecast  $p_i$  computed by the forecasting algorithm  $f$  and  $\tilde{x}_i$  is obtained by independent randomization of each coordinate  $x_{i,j}$  of the vector  $\bar{x}_i$ ,  $j = 1, \dots, k$ . Also  $\text{Var}_n(\tilde{p}_n) \rightarrow 0$  and  $\text{Var}_n(\tilde{x}_{i,j}) \rightarrow 0$  as to  $n \rightarrow \infty$  for all  $i$  and  $j$ .<sup>2</sup>

*Proof.* We modify a weak calibration algorithm of Kakade and Foster [7] using also ideas from Vovk [10].

At first, we construct an  $\Delta$ -calibrated forecasting algorithm, and after that we apply some double trick argument for it.

**Lemma 1.** *Given  $k$  an algorithm for computing forecasts and a method of randomization can be constructed such that for any sequence of real numbers  $S_1, S_2, \dots$  and for any sequence of signals  $\bar{x}_1, \bar{x}_2, \dots$  the event*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) \leq \Delta$$

has  $Pr$ -probability 1, where  $Pr$  and  $I$  as in Theorem 1. Also  $\text{Var}_n(\tilde{p}_n) \leq \Delta$  and  $\text{Var}_n(\tilde{x}_{i,j}) \leq \Delta$  for all  $n$ , for all  $i$  and  $j$ .

*Proof.* At first we define a deterministic forecast and after that we randomize it.

Divide the interval  $[0, 1]$  on subintervals of length  $\Delta = 1/K$  with rational endpoints  $v_i = i\Delta$ , where  $i = 0, 1, \dots, K$ . Let  $V$  denotes the set of these points.

Any number  $p \in [0, 1]$  can be represented as a linear combination of two neighboring endpoints of  $V$  defining subinterval containing  $p$  :  $p = \sum_{v \in V} w_v(p)v = w_{v_{i-1}}(p)v_{i-1} + w_{v_i}(p)v_i$ , where  $p \in [v_{i-1}, v_i]$ ,  $i = \lfloor p/\Delta + 1 \rfloor$ ,  $w_{v_{i-1}}(p) = 1 - (p - v_{i-1})/\Delta$ , and  $w_{v_i}(p) = (p - v_{i-1})/\Delta$ . Define  $w_v(p) = 0$  for all other  $v \in V$ .

In that follows we round some deterministic forecast  $p_n$  to  $v_{i-1}$  with probability  $w_{v_{i-1}}(p_n)$  and to  $v_i$  with probability  $w_{v_i}(p_n)$ . We also round the each coordinate  $x_{n,s}$ ,  $s = 1, \dots, k$ , of a signal  $\bar{x}_n$  to  $v_{j_s-1}$  with probability  $w_{v_{j_s-1}}(x_{n,s})$  and to  $v_{j_s}$  with probability  $w_{v_{j_s}}(x_{n,s})$ , where  $x_{n,s} \in [v_{j_s-1}, v_{j_s}]$ .

Let also  $W_v(Q_n) = w_{v^1}(p_n)w_{v^2}(\bar{x}_n)$ , where  $v = (v^1, v^2)$ ,  $v^1 \in V$ ,  $v^2 = (v_1^2, \dots, v_k^2) \in V^k$ ,  $w_{v^2}(\bar{x}_n) = \prod_{s=1}^k w_{v_s^2}(x_{n,s})$ , and  $Q_n = (p_n, \bar{x}_n)$ . For any  $Q_n$ ,  $W_v(Q_n)$  is a probability distribution in  $V^{k+1}$  :  $\sum_{v \in V^{k+1}} W_v(Q_n) = 1$ .

<sup>2</sup>  $\text{Var}_n(\tilde{p}_n) = E_n(\tilde{p}_n - p_n)^2$ .

In that follows we define a deterministic forecast  $p_n$ . Let the forecasts  $p_1, \dots, p_{n-1}$  already defined (put  $p_1 = 1/2$ ). Let us define for  $v = (v^1, v^2)$  and  $Q_i = (p_i, \bar{x}_i)$

$$\mu_{n-1}(v) = \sum_{i=1}^{n-1} W_v(Q_i)(S_i - p_i).$$

We have

$$\begin{aligned} (\mu_n(v))^2 &= (\mu_{n-1}(v))^2 + \\ &+ 2W_v(Q_n)\mu_{n-1}(v)(S_n - p_n) + (W_v(Q_n))^2(S_n - p_n)^2. \end{aligned} \quad (3)$$

Summing (3) by  $v \in V^{k+1}$ , we obtain:

$$\begin{aligned} \sum_{v \in V^{k+1}} (\mu_n(v))^2 &= \sum_{v \in V^{k+1}} (\mu_{n-1}(v))^2 + \\ &+ 2(S_n - p_n) \sum_{v \in V^{k+1}} W_v(Q_n)\mu_{n-1}(v) + \sum_{v \in V^{k+1}} (W_v(Q_n))^2(S_n - p_n)^2. \end{aligned} \quad (4)$$

Change the order of summation:

$$\begin{aligned} \sum_{v \in V^{k+1}} W_v(Q_n)\mu_{n-1}(v) &= \sum_{v \in V^{k+1}} W_v(Q_n) \sum_{i=1}^{n-1} W_v(Q_i)(S_i - p_i) = \\ &= \sum_{i=1}^{n-1} \left( \sum_{v \in V^{k+1}} W_v(Q_n)W_v(Q_i) \right) (S_i - p_i) = \\ &= \sum_{i=1}^{n-1} (\bar{W}(Q_n) \cdot \bar{W}(Q_i))(S_i - p_i) = \sum_{i=1}^{n-1} K(Q_n, Q_i)(S_i - p_i), \end{aligned}$$

where  $\bar{W}(Q_n) = (W_v(Q_n) : v \in V^{k+1})$ ,  $\bar{W}(Q_i) = (W_v(Q_i) : v \in V^{k+1})$  be vectors of probabilities of rounding. The dot product of corresponding vectors defines the kernel

$$K(Q_n, Q_i) = K(p_n, \bar{x}_n, p_i, \bar{x}_i) = (\bar{W}(Q_n) \cdot \bar{W}(Q_i)). \quad (5)$$

Let  $p_n$  be equal to the root of the equation

$$S_n(p_n) = \sum_{v \in V} W_v(p_n, \bar{x}_n)\mu_{n-1}(v) = \sum_{i=1}^{n-1} K(p_n, \bar{x}_n, p_i, \bar{x}_i)(S_i - p_i) = 0, \quad (6)$$

if a solution exists. Otherwise, if the left hand-side of the equation (6) (which is a continuous by  $p_n$  function) strictly positive for all  $p_n$  define  $p_n = 1$ , define  $p_n = 0$  if it is strictly negative. Announce  $p_n$  as a deterministic forecast.

The third term of (4) is upper bounded by 1. Indeed, since  $|S_i - p_i| \leq 1$  for all  $i$ ,

$$\sum_{v \in V^{k+1}} (W_v(Q_n))^2(S_i - p_n)^2 \leq \sum_{v \in V^{k+1}} W_v(Q_n) = 1.$$

Then by (4),  $\sum_{v \in V^{k+1}} (\mu_n(v))^2 \leq n$ . Recall that for any  $v \in V^{k+1}$

$$\mu_n(v) = \sum_{i=1}^n W_v(Q_i)(S_i - p_i). \quad (7)$$

Insert the term  $I(v)$  in the sum (7), where  $I$  is the characteristic function of an arbitrary set  $\mathcal{S} \subseteq [0, 1]^{k+1}$ , sum by  $v \in V^{k+1}$ , and exchange the order of summation. Using Cauchy–Schwartz inequality for vectors  $(I(v) : v \in V^{k+1})$ ,  $(\mu_n(v) : v \in V^{k+1})$  and Euclidian norm, we obtain

$$\begin{aligned} & \left| \sum_{i=1}^n \sum_{v \in V^{k+1}} W_v(Q_i) I(v) (S_i - p_i) \right| = \\ & = \left| \sum_{v \in V^{k+1}} I(v) \sum_{i=1}^n W_v(Q_i) (S_i - p_i) \right| \leq \\ & \leq \sqrt{\sum_{v \in V^{k+1}} I(v)} \sqrt{\sum_{v \in V^{k+1}} (\mu_n(v))^2} \leq \sqrt{|V^{k+1}|} n \end{aligned} \quad (8)$$

for all  $n$ , where  $|V^{k+1}| = 1/\Delta^{k+1}$  – is the cardinality of the partition.

Let  $\tilde{p}_i$  be a random variable taking values  $v \in V$  with probabilities  $w_v(p_i)$  (only two of them are nonzero). Recall that  $\tilde{x}_i$  is a random variable taking values  $v \in V^k$  with probabilities  $w_v(\tilde{x}_i)$ .

Let  $\mathcal{S} \subseteq [0, 1]^k$  and  $I$  be its indicator function. For any  $i$  the mathematical expectation of a random variable  $I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)$  is equal to

$$E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) = \sum_{v \in V^{k+1}} W_v(Q_i) I(v) (S_i - v^1), \quad (9)$$

where  $v = (v^1, v^2)$ .

By the strong law of large numbers, for some  $\mu_n = o(n)$  (as  $n \rightarrow \infty$ ), *Pr*-probability of the event

$$\left| \sum_{i=1}^n I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) - \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq \mu_n \quad (10)$$

tends to 1 as  $n \rightarrow \infty$ . A form of  $\mu_n$  will be specified later.

By definition of deterministic forecast

$$\left| \sum_{v \in V^{k+1}} W_v(Q_i) I(v) (S_i - p_i) - \sum_{v \in V^{k+1}} W_v(Q_i) I(v) (S_i - v^1) \right| < \Delta$$

for all  $i$ , where  $v = (v^1, v^2)$ . Summing (9) by  $i = 1, \dots, n$  and using the inequality (8), we obtain

$$\left| \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| =$$

$$= \left| \sum_{i=1}^n \sum_{v \in V^{k+1}} W_v(Q_i) I(v)(S_i - v^1) \right| < \Delta n + \sqrt{|V^{k+1}|n} \quad (11)$$

for all  $n$ , where  $|V^{k+1}| = 1/\Delta^{k+1}$  is the cardinality of the partition.

By (11) and (10) we obtain that  $Pr$ -probability of the event

$$\left| \sum_{i=1}^n I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) \right| \leq \Delta n + \mu_n + \sqrt{n/\Delta^{k+1}} \quad (12)$$

tends to 1 as  $n \rightarrow \infty$ . In particular,  $Pr$ -probability of the event

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

is equal to 1. Lemma is proved.

To prove that (2) holds almost surely choose a monotonic sequence of rational numbers  $\Delta_1 > \Delta_2 > \dots$  such that  $\Delta_s \rightarrow 0$  as  $s \rightarrow \infty$ . We also define an increasing sequence of natural numbers  $n_1 < n_2 < \dots$ . For any  $s$ , we use on steps  $n_s \leq n < n_{s+1}$  the partition of  $[0, 1]$  on subintervals of length  $\Delta_s$ .

We choose  $n_s$  such that  $n_s \geq \left(\frac{k+2}{2}\right)^2 \Delta_s^{-(k+3)}$  for all  $s$ .<sup>3</sup> Put  $n_0 = 0$  and  $\Delta_0 = 1$ . Also, define the numbers  $n_1, n_2, \dots$  such that the inequality

$$\left| \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq 4(s+1)\Delta_s n \quad (13)$$

holds for all  $n_s \leq n \leq n_{s+1}$  and for all  $s$ .

We define this sequence by mathematical induction on  $s$ . Suppose that  $n_s$  ( $s \geq 1$ ) is defined such that the inequality

$$\left| \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq 4s\Delta_{s-1}n \quad (14)$$

holds for all  $n_{s-1} \leq n \leq n_s$ , and the inequality

$$\left| \sum_{i=1}^{n_s} E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq 4s\Delta_s n_s \quad (15)$$

also holds. Let us define  $n_{s+1}$ . Consider all forecasts  $\tilde{p}_i$  defined by the algorithm given above for discretization  $\Delta = \Delta_{s+1}$ . We do not use first  $n_s$  of these forecasts (more correctly we will use them only in bounds (16) and (17); denote these forecasts  $\hat{\mathbf{p}}_1, \dots, \hat{\mathbf{p}}_{n_s}$ ). We add the forecasts  $\tilde{p}_i$  for  $i > n_s$  to the forecasts defined

<sup>3</sup> This is the minimum point of (11).

before this step of induction (for  $n_s$ ). Let  $n_{s+1}$  be such that the inequality

$$\begin{aligned} & \left| \sum_{i=1}^{n_{s+1}} E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq \left| \sum_{i=1}^{n_s} E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| + \\ & + \left| \sum_{i=n_s+1}^{n_{s+1}} E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) + \sum_{i=1}^{n_s} E(I(\hat{\mathbf{p}}_i, \tilde{x}_i)(S_i - \hat{\mathbf{p}}_i)) \right| + \\ & + \left| \sum_{i=1}^{n_s} E(I(\hat{\mathbf{p}}_i, \tilde{x}_i)(S_i - \hat{\mathbf{p}}_i)) \right| \leq 4(s+1)\Delta_{s+1}n_{s+1} \end{aligned} \quad (16)$$

holds. Here the first sum of the right-hand side of the inequality (16) is bounded by  $4s\Delta_s n_s$  – by the induction hypothesis (15). The second and third sums are bounded by  $2\Delta_{s+1}n_{s+1}$  and by  $2\Delta_{s+1}n_s$ , respectively. This follows from (11) and by choice of  $n_s$ . The induction hypothesis (15) is valid for

$$n_{s+1} \geq \frac{2s\Delta_s + \Delta_{s+1}}{\Delta_{s+1}(2s+1)} n_s.$$

Analogously,

$$\begin{aligned} & \left| \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq \left| \sum_{i=1}^{n_s} E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| + \\ & + \left| \sum_{i=n_s+1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) + \sum_{i=1}^{n_s} E(I(\hat{\mathbf{p}}_i, \tilde{x}_i)(S_i - \hat{\mathbf{p}}_i)) \right| + \\ & + \left| \sum_{i=1}^{n_s} E(I(\hat{\mathbf{p}}_i, \tilde{x}_i)(S_i - \hat{\mathbf{p}}_i)) \right| \leq 4(s+1)\Delta_s n \end{aligned} \quad (17)$$

for  $n_s < n \leq n_{s+1}$ . Here the first sum of the right-hand inequality (16) is also bounded by  $4s\Delta_s n_s \leq 4s\Delta_s n$  – by the induction hypothesis (15). The second and the third sums are bounded by  $2\Delta_{s+1}n \leq 2\Delta_s n$  and by  $2\Delta_{s+1}n_s \leq 2\Delta_s n$ , respectively. This follows from (11) and from choice of  $\Delta_s$ . The induction hypothesis (14) is valid.

By (13) for any  $s$

$$\left| \sum_{i=1}^n E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i)) \right| \leq 4(s+1)\Delta_s n \quad (18)$$

for all  $n \geq n_s$  if  $\Delta_s$  satisfies the condition  $\Delta_{s+1} \leq \Delta_s(1 - \frac{1}{s+2})$  for all  $s$ .

By the law of large numbers (31), the relation (10) can be specified:

$$Pr \left\{ \sup_{n \geq n_s} \left| \frac{1}{n} \sum_{i=1}^n V_i \right| > \Delta_s \right\} \leq (\Delta_s)^{-2} e^{-2n_s \Delta_s^2} \quad (19)$$

for all  $s$ , where  $V_i = I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) - E(I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i))$  is a sequence of martingale-differences.



Combining (18) with (19), we obtain

$$Pr \left\{ \sup_{n \geq n_s} \left| \frac{1}{n} \sum_{i=1}^n I(\tilde{p}_i, \tilde{x}_i)(S_i - \tilde{p}_i) \right| \geq (4s + 5)\Delta_s \right\} \leq (\Delta_s)^{-2} e^{-2n_s \Delta_s^2} \quad (20)$$

for all  $s$ . The series  $\sum_{s=1}^{\infty} (\Delta_s)^{-2} e^{-2n_s \Delta_s^2}$  is convergent if  $n_s$  satisfies

$$n_s \geq \frac{\ln s + 2 \ln \ln s - 2 \ln(\Delta_s)}{2 \Delta_s^2}$$

for all  $s$ . Let also  $\Delta_s = o(1/s)$  as  $s \rightarrow \infty$ . Then Borel–Cantelli Lemma implies convergence of (2) almost surely.

It is easy to verify that the sequences  $n_s$  and  $\Delta_s$  satisfying all the conditions above exist.

### 3 Applications to technical analysis

#### 3.1 Simple trading for a rise

Let  $S_1, S_2, \dots$  be a sequence of a stock prices. At any step  $i$  we use one-dimensional signals  $x_i = S_{i-1}$ ,  $i = 1, 2, \dots$ , and an indicator function  $I(p_i > x_i + \epsilon)$ , where  $\epsilon$  is a parameter.<sup>4</sup>

At the end of the trading period *Speculator* receives a gain or suffer loss  $\Delta S_I = S_i - S_{i-1}$  for one share of the stock. The total gain or loss is equal to

$$\mathcal{K}_n = \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) \Delta S_i,$$

where  $\tilde{p}_i$  and  $\tilde{S}_{i-1}$  are randomized forecast and past price of the stock respectively.

Let us specify details of rounding. The expression  $\Delta n + \sqrt{n/\Delta^{k+1}}$  from (12) takes its minimal value for  $\Delta = (\frac{k+1}{2})^{\frac{2}{k+3}} n^{-\frac{1}{k+3}}$ . In this case, the right-hand side of the inequality (11) is equal to  $\Delta n + \sqrt{n/\Delta^{k+1}} = 2\Delta n = 2(\frac{k+1}{2})^{\frac{2}{k+3}} n^{1-\frac{1}{k+3}}$ .

We have  $k = 1$ , and hence, we use at any step  $n$  the rounding  $\Delta_s = n_s^{-1/4}$ , where  $s$  is such that  $n_s < n \leq n_{s+1}$ .

We write  $A \sim B$  if positive constants  $c_1$  and  $c_2$  exist such that  $c_1 B \leq A \leq c_2 B$  for all values of parameters from the expressions  $A$  and  $B$ .

Define  $n_s = s^M$  and  $\Delta_s = s^{-M/4}$ , where  $M$  is a positive integer number. Then  $s \sim n_s^{1/M}$  (the constants  $c_1$  and  $c_2$  depend on  $M$ ).

Easy to verify that all requirements for  $n_s$  and  $\Delta_s$  given in Section 2 are valid.

By (20) we can define  $\mu_n = (4s+5)\Delta_s n$ , where  $s$  is such that  $n_s < n \leq n_{s+1}$ . For  $n_s < n \leq n_{s+1}$  it holds  $n \sim n_s$ , hence,  $\mu_n \sim n^{3/4+1/M}$ .

<sup>4</sup> In other approach we can consider a sequence of signals  $\epsilon_1, \epsilon_2, \dots$

We represent the total gain by  $n$  steps in a form

$$\begin{aligned} \mathcal{K}_n = \sum_{\tilde{p}_i > \tilde{S}_{i-1} + \epsilon} \Delta S_i &= \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)(S_i - \tilde{p}_i) + \\ &+ \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)(\tilde{S}_{i-1} - S_{i-1}) + \\ &+ \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)(\tilde{p}_i - \tilde{S}_{i-1}). \end{aligned} \quad (21)$$

By (12) the probability that the first addend of the sum (21) is more than  $-(\Delta_s n + \mu_n + 2\sqrt{n/\Delta_s^2})$  tends to 1 as  $n \rightarrow \infty$ , where  $s$  is such that  $n_s < n \leq n_{s+1}$ .

According to Section A.1 the probability that the second addend of the sum (21) is more than  $-\Delta_s n$  tends to 1 as  $n \rightarrow \infty$ . By definition the third addend of the sum (21) is more than  $\epsilon \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)$  for all  $n$ .

Then the probability that the average income per one gamble  $k_n$  satisfies

$$\begin{aligned} k_n = \frac{\mathcal{K}_n}{\sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)} &\geq \left( \epsilon - \frac{2\Delta_s n + \mu_n + 2\sqrt{n/\Delta_s^2}}{\sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)} \right) \sim \\ &\sim \left( \epsilon - \frac{3n^{3/4} + n^{3/4+1/M}}{\sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)} \right) \end{aligned} \quad (22)$$

tends to 1 as  $n \rightarrow \infty$ . Using inequalities of Section A.1 and definition of  $\Delta_s$  and  $\mu_n$ , one can check that the corresponding convergence rate is  $e^{-c\sqrt{n}}$ , where  $c > 0$ .

We summarize this result in the following proposition.

**Proposition 1.** *A randomized trading strategy exists such that given  $0 < \epsilon < 1$  and  $0 < \gamma < 1$  with (internal) probability  $1 - e^{-c'\sqrt{n}}$  if*

$$\sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) \geq \frac{cn^{3/4+\nu}}{\gamma\epsilon}, \quad (23)$$

*then  $k_n \geq (1 - \gamma)\epsilon$ , where  $\nu = 1/M$ ,  $c'$  and  $c$  are positive constants.*

### 3.2 Trading with a limited risk

The most important requirement for a trading strategy is guarantee conditions. We present a defensive strategy for *Speculator* in sense of Shafer and Vovk's book [9]. This means that starting with some initial capital *Speculator* never goes

to debt and receives a gain when a sufficiently long subsequence of forecasts like (23) exists.

We modify the strategy given in Section 3.1 to a defensive strategy.

Let  $\mathcal{K}_0 > 0$  be a starting capital of *Speculator*. Define  $M_i = \delta \mathcal{K}_{i-1}$ , where  $\mathcal{K}_{i-1}$  is the capital of *Speculator* at step  $i-1$  and  $\delta$  is a parameter such that  $0 \leq \delta \leq 1$ . As usual, we suppose that all prices are scaled such that  $0 \leq S_{i-1} \leq 1$   $S_1 = S_0$ .

*Speculator's* capital after  $i$ th step is equal to

$$\mathcal{K}_i = \mathcal{K}_{i-1} + \delta \mathcal{K}_{i-1} \Delta S_i. \quad (24)$$

At any step  $n$  the logarithm of the capital is equal to

$$\begin{aligned} \ln \mathcal{K}_n &= \ln \mathcal{K}_0 + \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) \ln(1 + \delta \Delta S_i) \geq \\ &\geq \ln \mathcal{K}_0 + \delta \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) \Delta S_i - \\ &\quad - \delta^2 \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) (\Delta S_i)^2. \end{aligned} \quad (25)$$

Here we have used the inequality  $\ln(1+x) \geq x - x^2$  for  $|x| \leq 1$ .

Let  $\tilde{L}_n = \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon)$ . Since  $|\Delta S_{i-1}| \leq 1$ , a bound

$$\sum_{i=1}^n I(p_i > \tilde{S}_{i-1} + \epsilon) (\Delta S_i)^2 \leq \tilde{L}_n \quad (26)$$

is valid for all  $n$ .

By (22) probability of the event

$$\sum_{i=1}^n I(p_i > \tilde{S}_{i-1} + \epsilon) \Delta S_i \geq \epsilon \tilde{L}_n - cn^{3/4+1/M} \quad (27)$$

tends to 1 as  $n \rightarrow \infty$ , where  $c$  is a positive constant.

Denote

$$\overline{\text{Var}}_n(S) = \frac{1}{\tilde{L}_n} \sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) (\Delta S_i)^2.$$

For practical applications, where  $\overline{\text{Var}}_n(S) \ll 1$ , we can replace in (26)  $\tilde{L}_n$  on  $\tilde{L}_n \overline{\text{Var}}_n(S)$ . Then by (25) and (27) probability of the event

$$\ln \mathcal{K}_n \geq \ln \mathcal{K}_0 + \delta(\epsilon \tilde{L}_n - cn^{3/4+1/M}) - \delta^2 \overline{\text{Var}}_n(S) \tilde{L}_n \quad (28)$$

tends to 1 as  $n \rightarrow \infty$ . Therefore, probability of the event

$$\mathcal{K}_n \geq \mathcal{K}_0 \exp \left( \delta \left( \tilde{L}_n (\epsilon - \delta \overline{\text{Var}}_n(S)) - cn^{3/4+1/M} \right) \right) \quad (29)$$

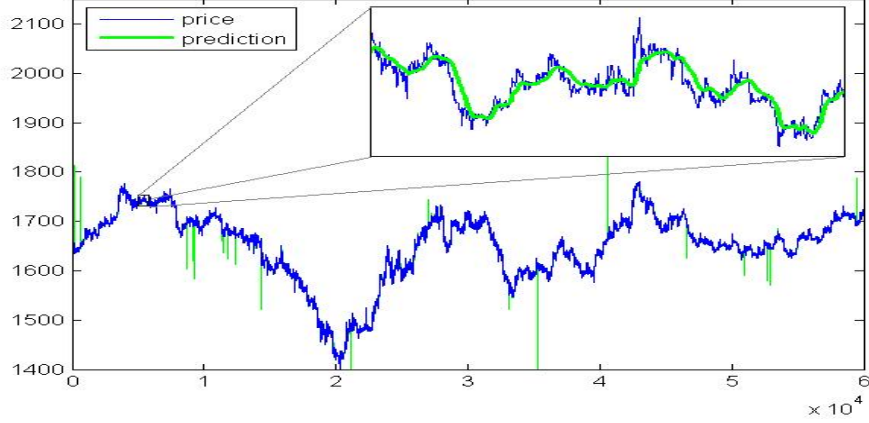


Fig.1 Calibrated forecasts of prices of LKOH stock

tends to 1 as  $n \rightarrow \infty$ .

We summarize the result of this section in the following proposition.

**Proposition 2.** *A randomized trading strategy exists such that given  $0 < \epsilon < 1$  and  $0 < \delta < 1$  with (internal) probability  $1 - e^{-c'\sqrt{n}}$  the capital  $\mathcal{K}_n$  of Speculator has a lower bound (29), where  $c'$  and  $c$  are positive constants. The capital increases if*

$$\sum_{i=1}^n I(\tilde{p}_i > \tilde{S}_{i-1} + \epsilon) \geq \frac{cn^{3/4+\nu}}{(\epsilon - \delta \overline{\text{Var}}_n(S))},$$

where  $\nu = 1/M$ . Also,  $\mathcal{K}_n > 0$  for all  $n$ .

## 4 Conclusion

Calibration is an intensively developing area of recent research where several algorithms for computing calibrated forecasts were developed. It is attractive to find some practical applications of these results. Using calibrated forecasts for constructing short-term trading strategies in Stock Market looks very natural.

In this paper, we construct such strategies and perform numerical experiments. These experiments show a positive return for six main Russian stocks, and for two stocks we receive a gain even when transaction costs are subtracted.

To construct trading strategies we develop a more general notion of calibration and prove convergence results for it. We also present some sufficient conditions under which our trading strategies receive a gain.

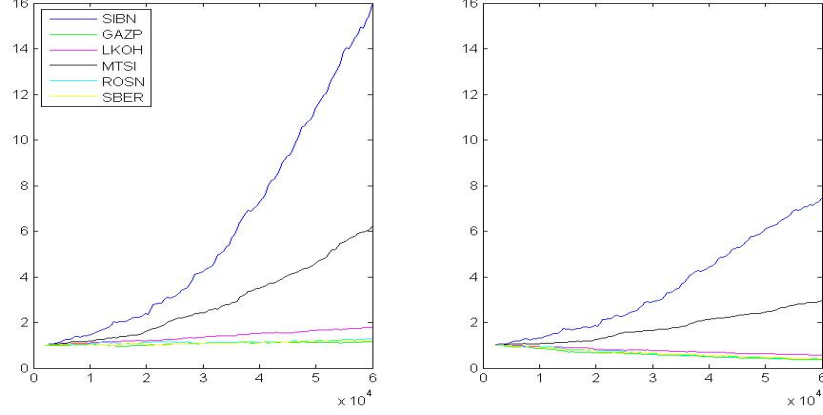


Fig.2 Capitals of speculators playing for a rise on six Russian stocks (with no transaction costs – on the left figure, with transaction costs 0.01% – on the right figure)

## A Appendix

### A.1 Large deviation inequality for martingales

A sequence  $V_1, V_2, \dots$  is called martingale-difference with respect to a sequence of random variables  $X_1, X_2, \dots$  if for any  $i > 1$  the random variable  $V_i$  is a function of  $X_1, \dots, X_i$  and  $E(V_{i+1}|X_1, \dots, X_i) = 0$  almost surely. The following inequalities are consequences of Hoffding-Azuma inequality [2]:

Let  $V_1, V_2, \dots$  be a martingale-difference with respect to  $X_1, X_2, \dots$ , and  $V_i \in [A_i, A_i + 1]$  for some random variable  $A_i$  measurable with respect to  $X_1, \dots, X_i$ . Let  $S_n = \sum_{i=1}^n V_i$ . Then for any  $t > 0$

$$P \left\{ \left| \frac{S_n}{n} \right| > t \right\} \leq 2e^{-2nt^2} \quad (30)$$

for all  $n$ . A strong law of large numbers is also holds: for any  $t$

$$P \left\{ \sup_{k \geq n} \left| \frac{S_k}{k} \right| > t \right\} \leq t^{-2} e^{-2nt^2} \quad (31)$$

for all  $n$ . Since the series of the exponents from the right-hand side of the inequality (30) convergent, by Borel-Cantelli Lemma we obtain the martingale strong law of large numbers

$$P \left\{ \lim_{n \rightarrow \infty} \frac{S_n}{n} = 0 \right\} = 1.$$

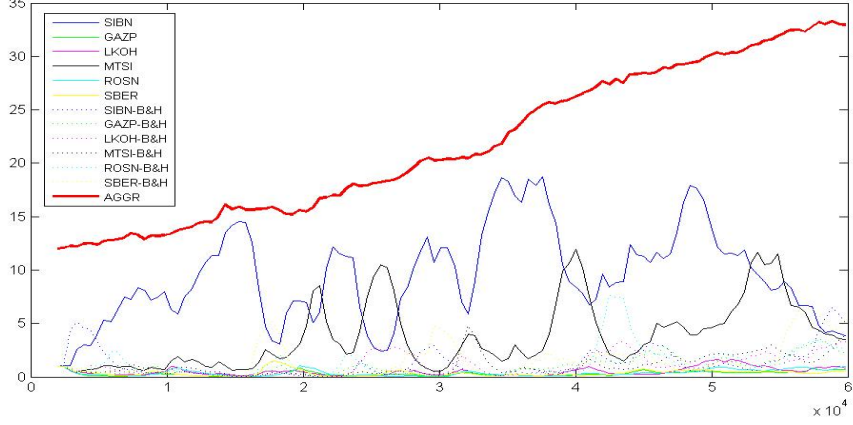


Fig.3 Capital of the weighted average of 12 strategies (two strategies for each stock – calibration strategy and “buy and hold” (B&H) strategy)

## A.2 Numerical experiments

In the numerical experiments, we have used historical data in form of per minute time series of prices of six main stocks of Russian Stock Market in 2010 (From 2010-03-26T10:31 to 2010-09-16T12:15). downloaded from FINAM site: *www.finam.ru*. Number of trading points in each game is  $6 \cdot 10^4$  min. In our experiments, we dealing only for a rise starting with the same initial capital  $\mathcal{K}_0$ .

We have used the threshold  $\epsilon = \epsilon' \sigma$ , where  $\sigma$  is the standard deviation of a price calculating using some sliding window,  $0 < \epsilon' < 1$ . A kernel  $K(p, p') = \cos(\pi(p - p'))$  was used as a smooth approximation of (5).

Results of numerical experiments are shown in Table 1. In the first column, ticker symbols of six stocks from Russian Stock Market are shown. The second column contains the frequencies of steps  $i$  where  $p_i > S_{i-1} + \epsilon$ . In the third column, the average duration of a gamble is shown. We sell all shares of a stock at step  $i$  in case  $\tilde{p}_i \leq S_{i-1} + \epsilon$  or  $S_i \leq S_{i-1}$ . In fourth and in fifth columns, a relative return (in percentage wise on initial capital) for six main stocks from Russian Stock Market in 2010. We have used a transaction cost at the rate 0.01%. In the sixth column, a return of “buy and hold” strategy is shown. By this strategy, we buy a holdings of shares for  $\mathcal{K}_0$  and sell them at the end of the trading period.

On Fig.1 the evolution of LKOH prices and their predictions are shown. On Fig. 2 the relative returns of calibration strategies for all six stocks are shown (without and with transaction costs). On Fig.3 a relative return of short-term trading for six stocks are shown including calibration and buy and hold strategy for each stock. The extra bold line represents a relative return of some averaging strategy AGGR. This strategy is similar to the Freund and Shapire [6] exponential weighting algorithm, where one-day steps are used (see also the last line of Table 1).

**Table 1.** Relative return in percentage wise on capital used in dealing for a rise for six main stocks of Russian Stock Market in 2010 and for the aggregating strategy AGGR.

Ticker symbol of a stock	frequency of entry points	average duration of a gamble	without transaction costs	with transaction costs	buy and hold
GAZP	0.100	1.88	15.73%	−63.67%	−6.30%
LKOH	0.099	1.85	78.87%	−43.53%	2.19%
MTSI	0.065	2.51	527.05%	196.15%	1.15%
ROSN	0.097	1.86	27.25%	−58.66%	−12.88%
SBER	0.092	1.94	19.72%	−58.86%	−2.61%
SIBN	0.066	2.86	1504.39%	646.01%	−21.94%
AGGR			761.17%	321.67%	

## References

1. D.Blackwell, An analog of the minimax theorem for vector payoffs. Pacific Journal of Mathematics. 6 (1956) 1-8.
2. N.Cesa-Bianchi, G.Lugosi, Prediction, Learning and Games. Cambridge University Press, 2006.
3. A.P.Dawid, The well-calibrated Bayesian [with discussion]. J. Am. Statist. Assoc. 77 (1982) 605-613.
4. A.P.Dawid. Calibration-based empirical probability [with discussion]. Ann. Statist. 13, 1251-1285, 1985.
5. D.P.Foster, R. Vohra. Asymptotic calibratio. Biometrika. 85, 379–390, 1998.
6. Freund Y, Schapire R.E. A decision-theoretic generalization of online learning and an application to boosting. Journal of Computer and System Sciences, 55 (1997), 119-139.
7. S.M.Kakade, D.P.Foster Deterministic calibration and Nash equilibrium. LNCS 3120 (John Shawe Taylor and Yoram Singer. ed). P.33–48, 2004.
8. D.Oakes. Self-Calibrating Priors Do not Exist [with discussion]. J. Am. Statist. Assoc. V.80. P.339–342, 1985.
9. Shafer, G., Vovk, V., Probability and Finance. It's Only a Game! New York: Wiley, 2001.
10. V.Vovk. Defensive Forecasting for Optimal Prediction with Expert Advice. arXiv:0708.1503v1. 2007.