On games of continuous and discrete randomized forecasting

Vladimir V. V'yugin*

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

Abstract. Using the game-theoretic framework for probability, Vovk and Shafer [7] have shown that it is always possible, using randomization, to make sequential probability forecasts that pass any well-behaved statistical test. We show that Vovk and Shafer's result is valid only when the forecasts are computed with unrestrictedly increasing degree of accuracy. We present a test failing any given method of randomized forecasting which uses a fixed level of discreteness.

Using the game-theoretic framework for probability [6], Vovk and Shafer have shown in [7] that it is always possible, using randomization, to make sequential probability forecasts that pass any well-behaved statistical test. This result generalizes work by other authors, among them are Foster and Vohra [2], Kakade and Foster [3], Lehrer [4], Sandrony et al. [5], who consider only tests of calibration.

We complement this result with a lower bound. We show that Vovk and Shafer's result is valid only when the forecasts are computed with unrestrictedly increasing degree of accuracy. We present a test failing any given method of discrete randomized forecasting. To formulate this example, we use the forecasting game presented by Vovk and Shafer [7], namely Binary Forecasting Game II.

Let $\mathcal{P}\{0,1\}$ be the set of all measures on the two-element set $\{0,1\}$. Any measure from $\mathcal{P}\{0,1\}$ is represented by a number $p \in [0,1]$ - the probability of $\{1\}$. Let $\mathcal{P}[0,1]$ be the set of all probability measures on the unit interval $[0,1]$ supplied with the standard Borel σ-field $\mathcal{F}$.

Randomizing forecasting is defined as follows. For each n, a forecaster given a binary sequence of past outcomes $\omega_1 \dots \omega_{n-1}$ (and a sequence of past forecasts $p_1, \dots, p_{n-1}$) outputs a probability distribution $P_n \in \mathcal{P}[0,1]$. The forecasts p_n of the the future event $\omega_n = 1$ are distributed according to this probability distribution.

Vovk and Shafer's [7] *Binary Forecasting Game II* between three players - Forecaster, Skeptic, Reality, Random Number Generator is described by the following *protocol*:

* This research was partially supported by Russian foundation for fundamental research: 09-07-00180a

Let $\mathcal{K}_0 = 1$ and $\mathcal{F}_0 = 1$.
FOR $n = 1, 2, \ldots$
Skeptic announces $S_n : [0,1] \to \mathcal{R}$.
Forecaster announces a probability distribution $P_n \in \mathcal{P}[0,1]$.
Reality announces $\omega_n \in \{0,1\}$.
Forecaster announces $f_n : [0,1] \to \mathcal{R}$ such that $\int f_n(p)P_n(dp) \leq 0$.
Random Number Generator announces $p_n \in [0,1]$.
Skeptic updates his capital $\mathcal{K}_n = \mathcal{K}_{n-1} + S_n(p_n)(\omega_n - p_n)$.
Forecaster updates his capital $\mathcal{F}_n = \mathcal{F}_{n-1} + f_n(p_n)$.
ENDFOR

Restriction on Skeptic: Skeptic must choose the S_n so that his capital $\mathcal{K}_n$ is nonnegative for all n no matter how the other players move.

Restriction on Forecaster: Forecaster must choose the P_n and f_n so that his capital $\mathcal{F}_n$ is nonnegative for all n no matter how the other players move.

Vovk and Shafer [7] showed that Forecaster has a winning strategy in the Forecasting Game II, where Forecaster wins if either (i) his capital $\mathcal{F}_n$ is unbounded or (ii) Skeptic's capital $\mathcal{K}_n$ stays bounded; otherwise the other players win.

Using some specific forms of $S_n(p)$, Shafer and Vovk [6] have shown that Forecaster has strategies forcing the strong law of large numbers and the law of iterated logarithm.

Theorem 1. *Forecaster has a winning strategy in Binary Forecasting Game II.*

For completeness of the presentation, we give a sketch of the proof from [7].

At first, at any round n of Binary Forecasting Game II, a simple auxiliary game between Realty and Forecaster is considered: Forecaster chooses $p_n \in [0,1]$, Realty chooses $\omega_n \in \{0,1\}$. Forecaster losses (and Realty gains) $S(p_n)(\omega_n - p_n)$.

For any mixed strategy of Realty $Q_n \in \mathcal{P}\{0,1\}$, let Forecaster's strategy be $p_n = Q\{1\}$. So, the Realty's expected gain is $S(p_n)(1 - Q\{1\})Q\{1\} + S(p_n)(0 - Q\{1\})Q\{0\} = 0$, where $Q\{0\} = 1 - Q\{1\}$.

In order to apply von Neumann's minimax theorem, which requires that move space be finite, we replace Forecaster move space $[0,1]$ with a finite subset of $[0,1]$ dense enough that the value of the game is smaller than some arbitrary small positive number Δ (depending on n). This is possible, since $|S_n(p)| \leq \mathcal{K}_{n-1} \leq 2^{n-1}$. [2] The minimax theorem asserts that Forecaster has a mixed strategy $P \in \mathcal{P}[0,1]$ such that

$$\int S_n(p)(\omega_n - p)P(dp) \leq \Delta \tag{1}$$

for both $\omega_n = 0$ and $\omega_n = 1$.

Let E_Δ be the subset of $\mathcal{P}[0,1]$ consisting all probability measures P satisfying (1) for $\omega_n = 0$ and $\omega_n = 1$. Endowed with the weak topology, $\mathcal{P}[0,1]$ is

[2] Skeptic must choose $S_n(p)$ such that $\mathcal{K}_n \geq 0$ for all n no matter the other players move.

compact. Since each E_Δ is closed, $\cap E_{\Delta_i} \neq \emptyset$, where Δ_i, $i = 1, 2, \ldots$, is some decreasing to 0 sequence of real numbers. So there exists $P_n \in \mathcal{P}[0,1]$ such that

$$\int S_n(p)(\omega_n - p)P_n(dp) \leq 0$$

for both $\omega_n = 0$ and $\omega_n = 1$.

In Binary Forecasting Game II, consider the strategy for Forecaster that uses at any round n the probability distribution P_n just defined and uses as his second move the function f_n defined $f_n(p) = S_n(p)(\omega_n - p)$. Then $\mathcal{F}_n = \mathcal{K}_n$ for all n. So either Skeptic's capital will stay bounded or Forecaster's capital will be unbounded. $\triangle$

In that follows we consider some modification of Binary Forecasting Game II in which Skeptic (but not Forecaster) announces $f_n : [0,1] \to \mathcal{R}$. This means that Skeptic defines the test of randomness he needs.

Also, at each step n, Skeptic divide his capital into two accounts: $\mathcal{K}_n = \mathcal{Q}_n + \mathcal{F}_n$; he uses the capital $\mathcal{F}_n$ to force Random Number Generator to generate random numbers which pass the test f_n.

Let $\mathcal{K}_0 = 2$.
FOR $n = 1, 2, \ldots$
Skeptic announces $S_n : [0,1] \to \mathcal{R}$.
Forecaster announces a probability distribution $P_n \in \mathcal{P}[0,1]$.
Reality announces $\omega_n \in \{0,1\}$.
Skeptic announces $f_n : [0,1] \to \mathcal{R}$ such that $\int f_n(p)P_n(dp) \leq 0$.
Random Number Generator announces $p_n \in [0,1]$.
Skeptic updates his capital $\mathcal{K}_n = \mathcal{K}_{n-1} + S_n(p_n)(\omega_n - p_n) + f_n(p_n)$.
ENDFOR

We divide the Skeptic's capital into two parts:
$\mathcal{K}_n = \mathcal{Q}_n + \mathcal{F}_n$ for all n, where
$\mathcal{Q}_0 = 1$ and $\mathcal{F}_0 = 1$.
$\mathcal{Q}_n = \mathcal{Q}_{n-1} + S_n(p_n)(\omega_n - p_n)$ and
$\mathcal{F}_n = \mathcal{F}_{n-1} + f_n(p_n)$ for all $n > 0$.

Restriction on Skeptic: Skeptic must choose the S_n and f_n so that his capital $\mathcal{K}_n$ is nonnegative for all n no matter how the other players move.

Actually, Skeptic will choose the S_n and f_n so that both of his capitals $\mathcal{Q}_n$ and $\mathcal{F}_n$ are nonnegative for all n no matter how the other players move.

Assume for each n, the probability distribution P_n is concentrated on a finite subset D_n of $[0,1]$, say, $D_n = \{p_{n,1}, \ldots, p_{n,m_n}\}$. The number $\Delta = \liminf_{n\to\infty} \Delta_n$, where

$$\Delta_n = \inf\{|p_{n,i} - p_{n,j}| : i \neq j\},$$

is called *the level of discreteness* of the corresponding forecasting scheme on the sequence $\omega_1\omega_2\ldots$. In general case D_n is measurable with respect to the σ-field $\mathcal{F}^{n-1}$, depending on $\omega_1 \ldots \omega_{n-1}$.

A typical example is the uniform rounding: for each n, rational points $p_{n,i}$ divide the unit interval into equal parts of size $0 < \Delta < 1$ and P_n is concentrated on these points. In this case the level of discreteness equals Δ for an arbitrary sequence $\omega_1\omega_2\ldots$.

We prove that when Forecaster uses finite subsets of $[0,1]$ for randomization Realty and Skeptic can defeat Forecaster (and Random Number Generator) in this forecasting game, where Realty and Skeptic win if Skeptic's capital $\mathcal{K}_n$ is unbounded; otherwise Forecaster and Random Number Generator win.

Theorem 2. *Assume Forecaster's uses a randomized strategy with a positive level of discreteness on each infinite sequence ω. Then Realty and Skeptic win in the modified Binary Forecasting Game II.*

Proof. Define a strategy for Realty: at any step n Realty announces an outcome

$$\omega_n = \begin{cases} 0 \text{ if } P_n((0.5,1]) > 0.5 \\ 1 \text{ otherwise.} \end{cases}$$

We follow Shafer and Vovk's [6] method of defining a defensive strategy for Skeptic.

Let $\epsilon_k = 2^{-k}$, $k = 1, 2, \ldots$. We define recursively by n: $\mathcal{Q}_0^{s,k} = 1$, $S_0^{s,k}(p) = 0$, $s = 1, 2$, and for $n \geq 1$

$$S_n^{1,k}(p) = -\epsilon_k \mathcal{Q}_{n-1}^{1,k} \xi(p > 0.5), \tag{2}$$

$$S_n^{2,k}(p) = \epsilon_k \mathcal{Q}_{n-1}^{2,k} \xi(p \leq 0.5), \tag{3}$$

where $\xi(true) = 1$, $\xi(false) = 0$, and for $n \geq 1$

$$\mathcal{Q}_n^{1,k} = \mathcal{Q}_{n-1}^{1,k} + S_n^{1,k}(p_n)(\omega_n - p_n)), \tag{4}$$

$$\mathcal{Q}_n^{2,k} = \mathcal{Q}_{n-1}^{2,k} + S_n^{2,k}(p_n)(\omega_n - p_n)). \tag{5}$$

We combine $S_n^{1,k}(p)$ and $S_n^{2,k}(p)$ in the Skeptic's strategy $S_n(p) = \frac{1}{2}(S_n^1(p) + S_n^2(p))$, where $S_n^1(p) = \sum\limits_{k=1}^{\infty} \epsilon_k S_n^{1,k}(p)$ and $S_n^2(p) = \sum\limits_{k=1}^{\infty} \epsilon_k S_n^{2,k}(p)$. It can be proved by the mathematical induction on n that $0 \leq Q_n^{i,k} \leq 2^n$ and $|S_n^{i,k}(p)| \leq 2^{n-1}$ for $i = 1, 2$ and for all k, p and n. Then these sums are finite for each n and p.

By (4)-(5) the Skeptic's capital $\mathcal{Q}_n$ at step n, when he follows the strategy $S_n(p)$, equals $\mathcal{Q}_n = \frac{1}{2}\sum\limits_{k=1}^{\infty} \epsilon_k (\mathcal{Q}_n^{1,k} + \mathcal{Q}_n^{2,k})$.

Define for each n the function $g_n(p) = (2\xi(p \leq 0.5) - 1)(\omega_n - p)$. Let $E_{P_n}(g_n) = \int g_n(p) P_n(dp)$.

Assume Forecaster uses some randomized strategy P_n, $n = 1, 2, \ldots$.

We define recursively by n: $\mathcal{F}_0^k = 1$, $g_0^k(p) = 0$, and for $n \geq 1$

$$g_n^k(p) = -\epsilon_k \mathcal{F}_{n-1}^k (g_n(p) - E_{P_n}(g_n)), \tag{6}$$

where $\epsilon_k = 2^{-k}$, and $\mathcal{F}_n^k = \mathcal{F}_{n-1}^k + g_n^k(p_n)$ for $n \geq 1$.

By definition for any k and n,

$$\mathcal{F}_n^k = \prod_{j=1}^{n} (1 - \epsilon_k (g_j(p_j) - E_{P_j}(g_j))). \tag{7}$$

By (7) $0 \le \mathcal{F}_n^k \le 2^n$ for all n and k. Finally, Skeptic defines at step n, $f_n(p) = \sum\limits_{k=1}^{\infty} \epsilon_k g_n^k(p)$. By definition $\int f_n(p)P_n(dp) \le 0$.

By (7) the Skeptic's capital $\mathcal{F}_n$ at step n, when he follows the strategy $f_n(p)$, equals $\mathcal{F}_n = \sum\limits_{k=1}^{\infty} \epsilon_k \mathcal{F}_n^k$. Also, $\mathcal{F}_n \ge 0$ for all n.

Suppose that $\sup\limits_n \mathcal{F}_n = C < \infty$, where $C > 0$. Then $\sup\limits_n \mathcal{F}_n^k < \frac{C}{\epsilon_k}$ for each k.

We have for each k,

$$\ln \mathcal{F}_n^k \ge -\epsilon_k \sum_{j=1}^{n} (g_j(p_j) - E_{P_j}(g_j)) - n\epsilon_k^2.$$

Here we use the inequality $\ln(1+r) \ge r - r^2$ for all $|r| \le \frac{1}{2}$.

Since $\mathcal{F}_n$ is bounded by $C > 0$, we have for any k

$$\frac{1}{n}\sum_{j=1}^{n}(g_j(p_j) - E_{P_j}(g_j)) \ge \frac{-\ln C + \ln(\epsilon_k)}{n\epsilon_k} - \epsilon_k \ge -2\epsilon_k \tag{8}$$

for all sufficiently large n.

Define two variables $\vartheta_{n,1} = \sum\limits_{j=1}^{n} \xi(p_j > 0.5)(\omega_j - p_j)$ and $\vartheta_{n,2} = \sum\limits_{j=1}^{n} \xi(p_j \le 0.5)(\omega_j - p_j)$. By definition of g_j, $\vartheta_{n,2} - \vartheta_{n,1} = \sum\limits_{j=1}^{n} g_j(p_j)$. Define $g_{1,j}(p) = \xi(p > 0.5)(\omega_j - p)$ and $g_{2,j}(p) = \xi(p \le 0.5)(\omega_j - p)$. Then $g_j(p) = g_{2,j}(p) - g_{1,j}(p)$.

Assume for any n the probability distribution P_n is concentrated on a finite set $\{p_{n,1}, \dots, p_{n,m_n}\}$.

For technical reason, if necessary, we add 0 and 1 to the support set of P_n and set their probabilities to be 0. Denote $p_n^- = \max\{p_{n,t} : p_{n,t} \le 0.5\}$ and $p_n^+ = \min\{p_{n,t} : p_{n,t} > 0.5\}$.

By definition ω_n, p_n^+ and p_n^- are predictable and $p_n^+ - p_n^- \ge \Delta$ for all n, where $\Delta > 0$. We have

$$\sum_{j=1}^{n} E_{P_j}(g_{1,j}) \le \sum_{\omega_j=0} P_j\{p > 0.5\}(-p_j^+) + \sum_{\omega_j=1} P_j\{p > 0.5\}(1 - p_j^+) \le$$
$$-0.5\sum_{j=1}^{n}\xi(\omega_j = 0)p_j^+ + 0.5\sum_{j=1}^{n}\xi(\omega_j = 1)(1 - p_j^+). \tag{9}$$

$$\sum_{j=1}^{n} E_{P_j}(g_{2,j}) \ge \sum_{\omega_j=0} P_j\{p \le 0.5\}(-p_j^-) + \sum_{\omega_j=1} P_j\{p \le 0.5\}(1 - p_j^-) \ge$$
$$-0.5\sum_{j=1}^{n}\xi(\omega_j = 0)p_j^- + 0.5\sum_{j=1}^{n}\xi(\omega_j = 1)(1 - p_j^-). \tag{10}$$

Subtracting (9) from (10), we obtain

$$\sum_{j=1}^{n} E_{P_j}(g_j) = \sum_{j=1}^{n} E_{P_j}(g_{2,j}) - \sum_{j=1}^{n} E_{P_j}(g_{1,j}) \ge 0.5\Delta n.$$

Using (8), we obtain for all sufficiently large n

$$\frac{1}{n}(\vartheta_{n,2} - \vartheta_{n,1}) = \frac{1}{n}\sum_{j=1}^{n} g_j(p_j) \ge \frac{1}{n}\sum_{j=1}^{n} E_{P_j}(g_j) - 2\epsilon_k \ge 0.5\Delta - 2\epsilon_k. \quad (11)$$

Now we compute a lower bound of Skeptic's capital. We have from the definition (2)-(3) $\mathcal{Q}_n^{1,k} = \prod_{j=1}^{n}(1 - \epsilon_k \xi(p_j > 0.5)(\omega_j - p_j))$, and $\mathcal{Q}_n^{2.k} = \prod_{j=1}^{n}(1 + \epsilon_k \xi(p_j \le 0.5)(\omega_j - p_j))$. By these inequalities, $0 \le \mathcal{Q}_n^{i,k} \le 2^n$ for all n and for $i = 1, 2$, no matter how the other players move. Also at step n, $\ln \mathcal{Q}_n^{1,k} \ge -\epsilon_k \vartheta_{n,1} - \epsilon_k^2 n$ and $\ln \mathcal{Q}_n^{2,k} \ge \epsilon_k \vartheta_{n,2} - \epsilon_k^2 n$. These inequalities and (11) imply

$$\limsup_{n\to\infty} \frac{\ln \mathcal{Q}_n^1 + \ln \mathcal{Q}_n^2}{n} \ge 0.5\epsilon_k \Delta - 2\epsilon_k^2 \ge 2\epsilon_k^2 \quad (12)$$

for all sufficiently large n, where $\epsilon_k \le \frac{1}{8}\Delta$. From this, we obtain

$$\limsup_{n\to\infty} \frac{\ln \mathcal{Q}_n^{i,k}}{n} \ge \epsilon_k^2$$

for $i = 1$ or for $i = 2$, and for all sufficiently large n.

Hence, we obtain for the total capital of Skeptic $\mathcal{K}_n = \mathcal{Q}_n + \mathcal{F}_n$

$$\limsup_{n\to\infty} \mathcal{K}_n = \infty$$

no matter how Forecaster moves if Realty uses her strategy defined above.

References

1. A.P.Dawid, The well-calibrated Bayesian [with discussion], J. Am. Statist. Assoc. 77 (1982) 605-613.
2. D.P. Foster, R. Vohra, Asymptotic calibration, Biometrika 85 (1998) 379-390.
3. S.M. Kakade, D.P. Foster, Deterministic calibration and Nash equilibrium, LNCS 3120 (John Shawe Taylor and Yoram Singer. ed) (2004) 33-48.
4. E. Lehrer, Any Inspection Rule is Manipulable, *Econometrica,* **69**-5, (2001) 1333-1347.
5. A. Sandroni, R. Smorodinsky, R. Vohra, Calibration with Many Checking Rules, Mathematics of Operations Research, 28-1, (2003) 141-153.
6. G. Shafer, V. Vovk, Probability and Finance. It's Only a Game! New York: Wiley, 2001.
7. Vladimir Vovk, Glenn Shafer, Good randomized sequential probability forecasting is always possible, J. Royal Stat. Soc. B, 67 (2005) 747-763.
8. Vladimir Vovk, Akimichi Takemura, Glenn Shafer, Defensive Forecasting, Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (2005) 365-372 (http://arxiv.org/abs/cs/0505083).