Central limit theorem without probability

V. G. Vovk

Technical Report

CLRC-TR-99-02

June 1995

Royal Holloway
University of London

Department of Computer Science
Egham, Surrey TW20 0EX, England

Probability does not exist
Bruno de Finetti

Abstract

Let $X_1, \dots, X_n$ be a sequence of bounded unprobabilized random variables which are disclosed in this order. Before X_i, $i = 1, \dots, n$, is disclosed, the Bookmaker announces the price E_i for a ticket which will pay X_i and the price V_{i-1} for a ticket which will pay $\sum_{j=i}^{n}(X_j - E_j)^2$. We prove that, under mild conditions,

$$\frac{1}{\sqrt{V_0}} \sum_{i=1}^{n} (X_i - E_i)$$

can be approximated by the standard normal distribution $\mathcal{N}_{0,1}$, in the sense that, for each bounded continuous function U and large n, $\int U(w)\,\mathcal{N}_{0,1}(dw)$ is close to being the fair price for the ticket which will pay

$$U\left(\frac{1}{\sqrt{V_0}} \sum_{i=1}^{n} (X_i - E_i)\right).$$

The final version of this paper was prepared for the seminar on the foundations of probability (Aalborg University, June 1995) organized by Steffen Lauritzen.

1 Theorem

We consider the following perfect-information game involving the Statistician, the Bookmaker, and the Nature. The game proceeds in trials. At each trial i, $i = 1, \dots, n$, the Nature produces a real number X_i. Before that, at the beginning of trial i, the Bookmaker announces E_i, his prediction for X_i. The quality of the prediction E_i is measured by $(X_i - E_i)^2$ (sometimes this is called the Brier loss function). At the end of each trial i the Bookmaker announces V_i, his prediction for his cumulative loss $\sum_{j=i+1}^{n}(X_j - E_j)^2$ over the remaining trials; the initial prediction V_0 is announced at the beginning of the game. We give the following operative interpretation to the Bookmaker's predictions: before X_i is disclosed, the Bookmaker allows the Statistician to buy any amount, positive or negative, of *variance tickets* for $\$V_{i-1}$ each and *X_i-tickets* for $\$E_i$ each. A variance ticket is a contract which obliges the Bookmaker to pay the Statistician $\$(X_i - E_i)^2$ every trial i and an X_i-ticket is a contract which obliges the Bookmaker to pay the Statistician $\$X_i$ after X_i is disclosed. After the Bookmaker has announced V_{i-1} and E_i but before X_i is disclosed, the Statistician buys as many variance tickets and X_i-tickets as he wishes; after that he will have, say, v_i variance tickets and e_i X_i-tickets. In this game the Statistician plays against the Bookmaker and the Nature who can coordinate their actions; we will use the word "Environment" to mean the Bookmaker and the Nature together.

Suppose the Statistician's initial capital is K_0. We can describe unfolding of the game, including the evolution of the Statistician's capital, as follows:

Environment chooses $V_0 \in]0, \infty[$
FOR $i = 1, \dots, n$:
  Statistician chooses $v_i \in \mathbb{R}$
  $K_i := K_{i-1} - v_i V_{i-1}$ (1)
  Environment chooses $E_i \in \mathbb{R}$
  Statistician chooses $e_i \in \mathbb{R}$
  $K_i := K_i - e_i E_i$ (2)
  Environment chooses $X_i \in \mathbb{R}$
  $K_i := K_i + e_i X_i + v_i (X_i - E_i)^2$ (3)
  Environment chooses $V_i \in]0, \infty[$
  $K_i := K_i + v_i V_i.$ (4)

We call the pair (v_i, e_i) the *portfolio* held by the Statistician at trial i. Equations (1) and (2) show how the Statistician's capital decreases when he buys the new portfolio; (3) describes the proceeds from holding the portfolio; (4) shows how the Statistician's capital increases when he sells the old portfolio (the X_i-tickets become useless and he throws them away). We can summarize (1) through (4) as follows:

$$K_i := K_{i-1} + v_i(V_i - V_{i-1}) + v_i(X_i - E_i)^2 + e_i(X_i - E_i). \tag{5}$$

The addend $v_i(V_i - V_{i-1})$ represents the capital gain from holding the portfolio, $v_i(X_i - E_i)^2$ represents the dividends paid by the variance tickets, and $e_i(X_i - E_i)$ represents the profit brought by the X_i-tickets.

We are going to prove a central limit theorem, and to do this we have to impose some restrictions on the Environment. We will require that $V_n = 0$ and, for some constant $C > 1$,

$$|X_i| \le C, \quad i = 1, \dots, n, \tag{6}$$

$$|V_{i+1} - V_i| \le C, \quad i = 0, \dots, n-1, \tag{7}$$

$$V_0 \ge n/C. \tag{8}$$

Inequalities (6) simply mean that X_i are known to be uniformly bounded; (7) means that each observation carries not too much information about the predictability of the future observations (note that the sequence V_i is not necessarily monotonic); and (8) means that the Bookmaker believes that on the average his error $|X_i - E_i|$ will not be excessively small. Therefore, the parameter C reflects the strength of the limitations imposed on the Environment.

To complete the description of the game between the Statistician and the Environment, we must specify a rule which says who won the game given its *record*

$$V_0 v_1 E_1 e_1 X_1 V_1 \dots v_{n-1} E_{n-1} e_{n-1} X_{n-1} V_{n-1} v_n E_n e_n X_n.$$

Let $f\colon \mathbb{R}^{3n} \to \mathbb{R}$ be a function of a *partial record*

$$V_0 E_1 X_1 V_1 \dots E_{n-1} X_{n-1} V_{n-1} E_n X_n.$$

By $\mathcal{G}(K_0, f \mid n, C)$ we denote the game which is played according to the above protocol (K_0 is the Statistician's initial capital, n is the duration of the game, and C is the constant from (6)–(8)) with the following winning condition: the Statistician wins if his final capital satisfies

$$K_n \ge f(V_0 E_1 X_1 V_1 \dots E_{n-1} X_{n-1} V_{n-1} E_n X_n);$$

otherwise, the Environment wins. It is well known that in a perfect-information game of finite duration such as $\mathcal{G}(K_0, f \mid n, C)$ one of the players has a winning strategy; we let $\mathcal{G}(K_0, f \mid n, C) \smile \mathrm{S}$ stand for "the Statistician has a winning strategy in $\mathcal{G}(K_0, f \mid n, C)$".

Now we can define our basic notions, the *upper* and *lower expectation*:

$$\mathbf{E}^+(f \mid n, C) := \inf\{K_0 \mid \mathcal{G}(K_0, f \mid n, C) \smile \mathrm{S}\};$$
$$\mathbf{E}^-(f \mid n, C) := -\,\mathbf{E}^+(-f \mid n, C).$$

We interpret $\mathbf{E}^+(f \mid n, C)$ as an upper bound on the maximal price that the Statistician can pay at the beginning of the game for the ticket which pays $\$f(V_0 \dots X_n)$ at the end of the game. Indeed, having $\$\,\mathbf{E}^+(f \mid n, C) + \epsilon$, whatever small $\epsilon > 0$ can be, is better for the Statistician than having the ticket: if the Statistician has $\$\,\mathbf{E}^+(f \mid n, C) + \epsilon$, he can consume $\$\frac{\epsilon}{2}$ and still make more than $\$f(V_0,\ldots,X_n)$ out of the remaining $\$\,\mathbf{E}^+(f \mid n, C) + \frac{\epsilon}{2}$. Analogously, $\mathbf{E}^-(f \mid n, C)$ is a lower bound on the minimal price that the Statistician can pay at the beginning of the game for that ticket. Indeed, having the ticket is better than having $\$\,\mathbf{E}^-(f \mid n, C) - \epsilon$: if the Statistician has the ticket, he can borrow $\$\,\mathbf{E}^-(f \mid n, C) - \frac{\epsilon}{2}$ and set aside $\$\,\mathbf{E}^-(f \mid n, C) - \epsilon$ consuming $\$\frac{\epsilon}{2}$; after that he can use his winning strategy in

$$\mathcal{G}\left(-\,\mathbf{E}^-(f \mid n, C) + \frac{\epsilon}{2}, -f \mid n, C\right)$$

to make his debt of $\$\,\mathbf{E}^-(f \mid n, C) - \frac{\epsilon}{2}$ into a debt of $\$f(V_0 \ldots X_n)$, which can be repaid with the ticket. (The Statistician can narrow the price span $[\mathbf{E}^-(f \mid n, C), \mathbf{E}^+(f \mid n, C)]$ if he has additional information about the Environment.)

Now we can state the main result of this paper; we let $\mathcal{N}_{\mu,\sigma^2}$ stand for the normal distribution in the real line $\mathbb{R}$ with mean μ and variance σ^2.

Theorem 1 *For arbitrarily large C and all bounded continuous $U\colon \mathbb{R} \to \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbf{E}^{\pm}\left(U\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^{n}(X_i - E_i)\right) \mid n, C\right) = \int U(w)\,\mathcal{N}_{0,1}(dw). \qquad (9)$$

2 Proof

First we prove two auxiliary results.

Lemma 2 $\mathbf{E}^+(f + g \mid n, C) \le \mathbf{E}^+(f \mid n, C) + \mathbf{E}^+(g \mid n, C)$.

Proof If the Statistician has winning strategies in $\mathcal{G}(a, f | n, C)$ and $\mathcal{G}(b, g | n, C)$, he can win $\mathcal{G}(a + b, f + g \mid n, C)$ acting as follows: if at trial i the winning strategies in $\mathcal{G}(a, f \mid n, C)$ and $\mathcal{G}(b, g \mid n, C)$ recommend portfolios (v_i', e_i') and (v_i'', e_i''), respectively, he chooses the portfolio $(v_i' + v_i'', e_i' + e_i'')$. $\square$

Lemma 3 (coherence) $\mathbf{E}^-(f \mid n, C) \le \mathbf{E}^+(f \mid n, C)$.

Proof We are required to prove

$$\mathbf{E}^+(f \mid n, C) + \mathbf{E}^+(-f \mid n, C) \ge 0.$$

By Lemma 2, it suffices to prove $\mathbf{E}^+(0 | n, C) \ge 0$. To disprove $\mathbf{E}^+(0 | n, C) < 0$, note that if the Statistician's initial capital is $K_0 < 0$, the Environment can ensure $0 > K_0 \ge K_1 \ge \ldots$ choosing $E_i := 0$, $V_i := n - i$, and

$$X_i := \begin{cases} 1, & \text{if } e_i < 0, \\ -1, & \text{otherwise.} \end{cases}$$

$\square$

We begin the proof of Theorem 1 by noting that it actually suffices to prove

$$\overline{\lim_{n\to\infty}}\,\mathbf{E}^{+}\left(U\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^{n}(X_i-E_i)\right)\mid n,C\right)\le\int U(w)\,\mathcal{N}_{0,1}(dw). \tag{10}$$

Indeed, substituting in (10) $-U$ in place of U, we will obtain

$$\underline{\lim}_{n\to\infty}\,\mathbf{E}^{-}\left(U\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^{n}(X_i-E_i)\right)\mid n,C\right)\ge\int U(w)\,\mathcal{N}_{0,1}(dw),$$

which together with (10) and Lemma 3 implies (9).

So we will prove (10). First we will assume that U is a smooth function whose 3rd and 4th derivatives are bounded: for some constant c and all $w\in\mathbb{R}$, $|U^{(3)}(w)|\le c$ and $|U^{(4)}(w)|\le c$. Without loss of generality we also assume that the Environment always chooses $E_i\in[-C,C]$.

Put

$$u(D,S):=\int U(w)\,\mathcal{N}_{S,D}(dw),$$

where $D\ge 0$ and $S\in\mathbb{R}$; we will construct a strategy for the Statistician which will enable him to have about $u(D_i,S_i)$, where

$$D_i:=\frac{V_i}{V_0},\quad S_i:=\frac{1}{\sqrt{V_0}}\sum_{j=1}^{i}(X_j-E_j),$$

at the end of each trial i provided he started off with $u(D_0,S_0)$. It is well known (and easy to check) that u satisfies Kolmogorov's backward equation:

$$D>0\Longrightarrow\frac{\partial u}{\partial D}=\frac{1}{2}\frac{\partial^2 u}{\partial S^2}. \tag{11}$$

Putting $\Delta A_i:=A_{i+1}-A_i$, $i=0,\ldots,n-1$, we find:

$$\begin{gathered}\Delta u(D_i,S_i)=\tfrac{\partial u}{\partial D}(D_i,S_i)\Delta D_i+\tfrac{\partial u}{\partial S}(D_i,S_i)\Delta S_i\\+\tfrac{1}{2}\tfrac{\partial^2 u}{\partial D^2}(D_i',S_i')(\Delta D_i)^2+\tfrac{1}{2}\tfrac{\partial^2 u}{\partial D\partial S}(D_i',S_i')\Delta D_i\Delta S_i+\tfrac{1}{2}\tfrac{\partial^2 u}{\partial S^2}(D_i',S_i')(\Delta S_i)^2,\end{gathered} \tag{12}$$

where (D_i',S_i') is a point strictly between (D_i,S_i) and (D_{i+1},S_{i+1}). Applying Taylor's formula to $\frac{\partial^2 u}{\partial S^2}$, we find

$$\tfrac{\partial^2 u}{\partial S^2}(D_i',S_i')=\tfrac{\partial^2 u}{\partial S^2}(D_i,S_i)+\tfrac{\partial^3 u}{\partial D\partial S^2}(D_i'',S_i'')\Delta D_i'+\tfrac{\partial^3 u}{\partial S^3}(D_i'',S_i'')\Delta S_i', \tag{13}$$

where (D_i'',S_i'') is a point strictly between (D_i,S_i) and (D_i',S_i'), and $\Delta D_i'$ and $\Delta S_i'$ satisfy $|\Delta D_i'|\le|\Delta D_i|$, $|\Delta S_i'|\le|\Delta S_i|$. Plugging (13) and (11) into (12), we find:

$$\begin{gathered}\Delta u(D_i,S_i)=\tfrac{\partial u}{\partial D}(D_i,S_i)\left(\Delta D_i+(\Delta S_i)^2\right)+\tfrac{\partial u}{\partial S}(D_i,S_i)\Delta S_i\\+\tfrac{1}{2}\tfrac{\partial^2 u}{\partial D^2}(D_i',S_i')(\Delta D_i)^2+\tfrac{1}{2}\tfrac{\partial^2 u}{\partial D\partial S}(D_i',S_i')\Delta D_i\Delta S_i\\+\tfrac{1}{2}\tfrac{\partial^3 u}{\partial D\partial S^2}(D_i'',S_i'')\Delta D_i'(\Delta S_i)^2+\tfrac{1}{2}\tfrac{\partial^3 u}{\partial S^3}(D_i'',S_i'')\Delta S_i'(\Delta S_i)^2.\end{gathered} \tag{14}$$

We can see that it is natural for the Statistician to choose the portfolio

$$\Bigl(\underbrace{\frac{1}{V_0}\frac{\partial u}{\partial D}(D_i,S_i)}_{\text{variance tickets}},\underbrace{\frac{1}{\sqrt{V_0}}\frac{\partial u}{\partial S}(D_i,S_i)}_{X_{i+1}\text{-tickets}}\Bigr)$$

at trial $i+1$. In this case, the first two addends of (14) give the increment of the Statistician's capital (cf. (5)), and we can rewrite (14) as

$$\begin{array}{c}\Delta u(D_i,S_i)=\Delta K_i\\ +\frac{1}{2}\frac{\partial^2 u}{\partial D^2}(D_i',S_i')(\Delta D_i)^2+\frac{1}{2}\frac{\partial^2 u}{\partial D\partial S}(D_i',S_i')\Delta D_i\Delta S_i\\ +\frac{1}{2}\frac{\partial^3 u}{\partial D\partial S^2}(D_i'',S_i'')\Delta D_i'(\Delta S_i)^2+\frac{1}{2}\frac{\partial^3 u}{\partial S^3}(D_i'',S_i'')\Delta S_i'(\Delta S_i)^2.\end{array}\tag{15}$$

¿From (11) we deduce

$$\begin{array}{c}\frac{\partial^2 u}{\partial D^2}=\frac{1}{2}\frac{\partial^3 u}{\partial D\partial S^2};\\ \frac{\partial^2 u}{\partial D\partial S}=\frac{\partial^2 u}{\partial S\partial D}=\frac{1}{2}\frac{\partial^3 u}{\partial S^3};\\ \frac{\partial^3 u}{\partial D\partial S^2}=\frac{\partial^3 u}{\partial S^2\partial D}=\frac{1}{2}\frac{\partial^4 u}{\partial S^4}.\end{array}$$

Since $\frac{\partial^3 u}{\partial S^3}$ and $\frac{\partial^4 u}{\partial S^4}$, being averages of $U^{(3)}$ and $U^{(4)}$, cannot exceed c, we obtain from (15) and (6)–(8):

$$\begin{array}{c}|\Delta u(D_i,S_i)-\Delta K_i|\\ \le c\left(|\Delta D_i|^2+|\Delta D_i||\Delta S_i|+|\Delta D_i||\Delta S_i|^2+|\Delta S_i|^3\right)\\ \le c\left(\frac{C^2}{V_0^2}+\frac{C}{V_0}\frac{2C}{\sqrt{V_0}}+\frac{C}{V_0}\frac{4C^2}{V_0}+\frac{(2C)^3}{V_0^{3/2}}\right)\\ \le \frac{C_1}{(n/C)^{3/2}}\le C_2 n^{-3/2}\end{array}$$

(C_1 and C_2 are some constants). Summing over $i=0,\dots,n-1$, we obtain

$$|(u(D_n,S_n)-u(D_0,S_0))-(K_n-K_0)|\le C_2 n^{-1/2},$$

i.e.,

$$\left|U\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^n(X_i-E_i)\right)-\int U(w)\,\mathcal{N}_{0,1}(dw)-(K_n-K_0)\right|\le C_2 n^{-1/2}.$$

Therefore, for an arbitrarily small $\delta>0$,

$$\mathcal{G}\left(\int U(w)\,\mathcal{N}_{0,1}(dw)+\delta, U\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^n(X_i-E_i)\right)\mid n,C\right)\smile \mathrm{S},\tag{16}$$

provided n is sufficiently large.

Now we drop the assumption that U is a smooth function with bounded $U^{(3)}$ and $U^{(4)}$. Let $\delta>0$. We are required to prove that, from some n on, (16) holds. Since U is bounded and continuous, we can find, for arbitrarily large c and arbitrarily small ϵ, a smooth $\overline{U}$ such that:

- $\diamond$ outside $[-c-1, c+1]$, $\overline{U}$ is constant and equals $\sup_w |U(w)|$;
- $\diamond$ inside $[-c, c]$, $U \le \overline{U} \le U + \epsilon$;
- $\diamond$ inside $[-c-1, -c]$ and inside $[c, c+1]$, $\overline{U}$ is monotonic and $\overline{U} \ge U$.

Taking c sufficiently large and ϵ sufficiently small, we will have

$$U \le \overline{U}, \int \left(\overline{U}(w) - U(w)\right) \mathcal{N}_{0,1}(dw) < \frac{\delta}{2}. \tag{17}$$

Since $U^{(3)}$ and $U^{(4)}$ are smooth and finite (i.e., zero outside a finite interval), they are bounded. Therefore, we already proved that, from some n on,

$$\mathcal{G}\left(\int \overline{U}(w)\,\mathcal{N}_{0,1}(dw) + \frac{\delta}{2}, \overline{U}\left(\frac{1}{\sqrt{V_0}}\sum_{i=1}^{n}(X_i - E_i)\right) \mid n, C\right) \smile \mathrm{S}. \tag{18}$$

It remains to note that (16) follows from (17) and (18).

Acknowledgments

The research described in this publication was made possible in part by Grant No. MRS000 from the International Science Foundation.