Variance Bounds
Some Old and Some New

Clément Deslandes

+

C.H.

Gaussian Poincaré Inequality

$X \sim N(0,1)$, $f : \mathbb{R} \to \mathbb{R}$
f "nice", then

$$\operatorname{Var} f(X) \le \mathbb{E}(f'(X))^2 . \quad (1)$$

Many proofs of (1).

Let H_n be nth order Hermite polynomial, then $(H_n)_{n\geq 0}$ is CONS for $L^2(e^{-x^2/2}/\sqrt{2\pi})$.

Hence

$$f(X) = \sum_{n=0}^{\infty} a_n H_n(X)$$

$$\mathbb{E} f(X) = a_0, \text{ so}$$

$$\operatorname{Var} f(X) = \sum_{n=1}^{\infty} a_n^2$$

$$f'(X) = \sum_{n=1}^{\infty} a_n H_n'(X)$$

But $H_n' = \sqrt{n}\, H_{n-1}$, so

$$\mathbb{E}(f'(X))^2 = \sum_{n=1}^{\infty} a_n^2 n$$

Clearly $\sum_{n=1}^{\infty} a_n^2 \leq \sum_{n=1}^{\infty} n a_n^2$.

(and cases of equality).

Another proof: Take

$$f(x) = \sum_{k=1}^{N} a_k e^{i t_k x}, \quad \text{then}$$

(1) is equivalent to

$$\sum_{k=1}^{N} \sum_{\ell=1}^{N} a_k \bar{a}_\ell e^{-t_k^2/2} e^{-t_\ell^2/2} \left\{ 1 - e^{t_k t_\ell} - t_k t_\ell e^{t_k t_\ell} \right\} \geqslant 0$$

i.e., the matrix

$$\left(1 - e^{t_k t_\ell} - t_k t_\ell e^{t_k t_\ell}\right)_{k,\ell} \text{ is p.s.d } \checkmark$$

I learned both methods from A. Kagan ($\simeq 92$), the first method is in a paper of Chernoff (AOP $\simeq 1981$)

Both methods lead to 2-sided bounds. For any $n \geq 1$

$$\sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k!} \mathbb{E}\left(f^{(k)}(X)\right)^2 \leq \operatorname{Var} f(X)$$

$$\leq \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} \mathbb{E}\left(f^{(k)}(X)\right)^2 \quad (2)$$

i.e.,

$$\mathbb{E}(f'(X))^2 - \frac{1}{2}\mathbb{E}(f''(X))^2$$

$$\leq \operatorname{Var} f(X) \leq \mathbb{E}(f'(X))^2.$$

Extensions to Wiener Space (Malliavin derivative, Poisson space (difference operator), Infinitely Divisible Laws, $X \sim ID(b, \sigma, \nu)$

i.e., $\varphi_X(t) = e^{itb + \int_{-\infty}^{+\infty} (e^{itu} - 1 - itu \mathbb{1}_{|u| \leq 1}) \nu(du)}$

where $b \in \mathbb{R}$ and $\nu(\{0\}) = 0$, $\int (1 \wedge u^2)\,\nu(du) < +\infty$

Then,

$$\operatorname{Var} f(X) \leq \mathbb{E} \int_{-\infty}^{+\infty} |f(X+u) - f(X)|^2\,\nu(du)$$

$$= \mathbb{E} \int_{-\infty}^{+\infty} |\Delta_u f(X)|^2\,\nu(du)$$

(Chen). Extension to two-sided with iterations of the difference operator.

Inconvenience, LHS of (2) is not necessarily ≥ 0, e.g.;

$\mathbb{E}(f'(X))^2 - \frac{1}{2}\mathbb{E}(f''(X))^2$ could be ≤ 0.

To bypass this problem, another lower bound:

$$\operatorname{Var} f(X) \geq \sum_{k=1}^{n} \left(\mathbb{E}\, \frac{f^{(k)}(X)}{k!}\right)^2$$

Efron-Stein Inequality.

$S : \mathbb{R}^n \to \mathbb{R}$, $X_1, X_2, \ldots, X_n$ iid

S symmetric in its arguments. Then,

$$\operatorname{Var} S(X_1, \ldots, X_n) \leq \mathbb{E}\, J_1,$$

where

$$J_1 = \sum_{k=1}^{n} \left(\frac{S(X_1, \ldots, X_n) - S(X_1, \ldots, X_{k-1}, \hat{X}_k, X_{k+1}, \ldots, X_n)}{2} \right)^2 \qquad (3)$$

$\hat{X}_k$ is an independent copy of X_k.

$$= \sum_{k=1}^{n} \left(\frac{S - S_k}{2} \right)^2,$$ where S_k change the kth coordinate.

"Statistical Gradient"

$$\nabla S = (S - S_1, S - S_2, - , S - S_n)$$

$$\mathbb{E} J_1 = \mathbb{E} \|\nabla S\|^2 , \text{ so}$$

$$\operatorname{Var} S \leq \frac{1}{2} \mathbb{E} J_1 .$$

Similarity with Gaussian Poincaré.

$$\sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k \leq \operatorname{Var} S$$

$$\leq \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k$$

C.H. ($\simeq$ 1997)

O. Bousquet + CH (HDP $\underline{\text{VIII}}$)

(2020)

No symetry assumption and just $\perp\!\!\!\perp$.

$S : \mathbb{R}^n \longrightarrow \mathbb{R}$,

$$\mathbb{E}^{(i)} S = \mathbb{E}(S \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$$

$$\mathbb{E}^{(0)} S = S$$

$$\mathbb{E}^{(i)} \mathbb{E}^{(j)} S = \mathbb{E}^{(j)} \mathbb{E}^{(i)} S$$

$$= \mathbb{E}(S \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n)$$

$$\mathbb{E}^{(i,j)} S = \mathbb{E}^{(j,i)} S$$

$$\mathrm{Var}^{(i)} S = \mathbb{E}^{(i)}\left(S - \mathbb{E}^{(i)} S\right)^2$$
$$= \mathbb{E}^{(i)} S^2 - \left(\mathbb{E}^{(i)} S\right)^2$$

$$\mathrm{Var}^{(i,j)} S = \mathbb{E}^{(i)} \mathrm{Var}^{(j)} S$$
$$- \mathrm{Var}^{(j)} \mathbb{E}^{(i)} S$$
$$= \mathrm{Var}^{(j,i)} S \geq 0$$

Then, iterate,

$$J_1 = \sum_{i=1}^{n} \mathrm{Var}^{(i)} S$$

$$J_k = \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} \mathrm{Var}^{(i_1, i_2, \ldots, i_k)} S$$
$$= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \mathrm{Var}^{(i_1, \ldots, i_k)} S$$

For $p = 1, 2, \ldots, [\frac{n}{2}]$

$$\sum_{k=1}^{2p} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k \leq \operatorname{Var} S \leq \sum_{k=1}^{2p-1} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k \qquad (4)$$

$$\operatorname{Var} S = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k$$

Again, LHS of (4) is not necessarily non-negative, i.e.,

$$\frac{1}{2} \mathbb{E} J_2 - \frac{1}{3!} \mathbb{E} J_3 \leq \mathbb{E} J_1 - \operatorname{Var} S \leq \frac{1}{2} \mathbb{E} J_2$$

How to get a non-negative lower bound on $\mathbb{E} J_1 - \operatorname{Var} S$?

Let

$$K_k = k! \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \mathrm{Var}^{(i_1,\dots,i_k)} \mathbb{E}^{\overline{(i_1,\dots,i_k)}} S$$

where $\overline{(i_1,\dots,i_k)}$ the complement of the indices $(i_1,\dots,i_k)$.

Fact (Jensen's + Convexity)

$$\mathbb{E} K_k \le \mathbb{E} J_k .$$

$$\frac{1}{2}\mathbb{E} K_k \le \mathbb{E} J_1 - \mathrm{Var} S \le \frac{1}{2}\mathbb{E} J_2$$

More generally:

For any $p = 1, 2, \dots, [n/2]$

$$\sum_{k=1}^{2p} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k + \frac{1}{(2p+1)!} \mathbb{E} K_{2p+1}$$

$$\leq \operatorname{Var} S \leq \sum_{k=1}^{2p-1} \frac{(-1)^{k+1}}{k!} \mathbb{E} J_k - \frac{1}{(2p)!} \mathbb{E} K_{2p},$$

Analogy: $X \sim N(0,1)$

$$\frac{1}{2}(\mathbb{E} f''(x))^2 \leq \mathbb{E} (f'(x))^2 - \operatorname{Var} f(x)$$

$$\leq \frac{1}{2} \mathbb{E} (f''(x))^2.$$

Why do I want variance bounds

(in addition to concentration considerations)

Ex: Longest Common Subsequences in Random Words

Let $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ be two $\perp\!\!\!\perp$ sequences of iid r.v.s with values in a finite alphabet $\mathcal{A}_m = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$.

Let LC_n be the length of the longest common subsequences of $X_1 \cdots X_n$ and

$Y_1 \cdots Y_n$

, i.e., $LC_n = \max\{k \in \{1, \dots, n\}$ s.t. $\exists$
$1 \le i_1 < i_2 < \cdots < i_k \le n$
$1 \le j_1 < j_2 < \cdots < j_k \le n$

s.t. $X_{i_s} = Y_{j_s}, \ \forall s \}$.

Binary case: $m = 2$, $\alpha_1 = 0$ and $\alpha_2 = 1$

X-word 0 0 1 1 0 0 1 1 1 1 0 0

Y-word 0 1 0 0 0 1 0 0 1 1 1 1

Superadditivity

$$\frac{\mathbb{E} LC_n}{n} \longrightarrow \gamma_m^* .$$

γ_m^* unknown, in the binary case, uniform

$\gamma_2^* \gtrsim 0.8$

So $\mathbb{E}\, LC_n \approx n\gamma_m^*$

What about

$\mathrm{Var}\, LC_n$?

$LC_n = S(X_1, \dots, X_n; Y_1, \dots, Y_n)$

Efron-Stein:

$$\mathrm{Var}\, LC_n \leq \frac{1}{2}\mathbb{E}\sum_{k=1}^{2n}\left(LC_n - LC_n^{(k)}\right)^2$$

But in See we just changed one r.v. by an $\perp\!\!\!\perp$ copy, i.e., we changed the kth letter.

$$\left(LC_n - LC_n^{(k)}\right)^2 \leq 1.$$

$$\mathrm{Var}\, LC_n \leq n.$$

What about lower bound? Is it true that $\operatorname{Var} LC_n \geq C n$?

Ans: Yes known in many "biased cases" but unknown even in the binary uniform case. Known in the binary case $p = \mathbb{P}(X_1 = 1)$ minuscule.

Can the previous lower bounds be useful?

Another related representation of $\operatorname{Var} S$ (Bordenave, Lugosi, Shinotovskiy) (Chatterjee)

$X_1, X_2, \dots, X_n \perp\!\!\!\perp$ r.v.

$$S : \mathbb{R}^n \longrightarrow \mathbb{R}$$

and $S^{(k)} = S(X_1, \dots, X_{k-1}, \hat{X}_k, X_{k+1}, \dots, X_n)$

$\hat{X}_k$ an $\perp\!\!\!\perp$ copy of X_k

More generally, for $\alpha \subset \{1, \dots, n\}$

$S^{(\alpha)} = S(X_1, \dots, X_n)$ expect that X_k is replaced by $\hat{X}_k$ for all $k \in \alpha$.

$$B_{k} = \mathbb{E}\, \frac{1}{n!} \sum_{i \in \mathcal{S}_n} S\left(S^{i_1,\dots,i_{k-1}} - S^{i_1,\dots,i_k}\right)$$

For $k=1$

$S^{i_1,\dots,i_{k-1}} = S.$

$\mathcal{S}_n$: symetric group.

$$\sum_{k=1}^{n} B_{k} = \mathbb{E} \sum_{k=1}^{n} \frac{1}{n!} \sum_{i \in \mathcal{S}_n} S\left(S^{i_1,\dots,i_{k-1}} - S^{i_1,\dots,i_k}\right)$$

telescopic

$$= \mathbb{E} S^2 - (\mathbb{E} S)^2$$

$$= \operatorname{Var} S$$

Another representation of the B_k's

$$\mathbb{E}\,(S - S^{i_k})\left(S^{i_1,\dots,i_{k-1}} - S^{i_1,\dots,i_k}\right)$$

$$= 2\,\mathbb{E}\, S\left(S^{i_1,\dots,i_{n-1}} - S^{i_1,\dots,i_n}\right)$$

Because $\mathbb{E}\, S^{\alpha} S^{\beta} = \mathbb{E}\, S\, S^{\alpha\Delta\beta}$.

Using this fact and introducing the notation

$$\Delta_{\ell_2} S = S - S^{\ell_2} \text{ and for } \ell_2 \neq \ell$$

$$\Delta_{\ell_2,\ell} S = \Delta_{\ell_2}(\Delta_{\ell} S)$$

$$B_{\ell_2} = \mathbb{E}\,\frac{1}{2n!}\sum_{i\in\mathfrak{S}_n} (\Delta_{i_{\ell_2}} S)(\Delta_{i_{\ell_2}} S)^{i_1,\dots,i_{\ell_2-1}}$$

$$\geq 0$$ (from properties of conditional expectation)

Link with the J_k and K_k

Let $DB_k = B_k - B_{k+1}$, for $k \in \{1, \dots, n-1\}$

and let D^ℓ, $\ell \geq 1$, be its iterates.

Then, for all $k \in \{1, \dots, n\}$

$$\frac{J_k}{k!} = \binom{n}{k} D^{k-1} B_1$$

$$\frac{K_k}{k!} = \binom{n}{k} D^{k-1} B_{n-k+1}$$

Using these representations we can invert

$$B_{\ell_2} = \sum_{j=0}^{\ell_2-1} (-1)^j \frac{\binom{\ell_2-1}{j}}{\binom{n}{j+1}} \frac{J_{j+1}}{(j+1)!}$$

$$B_{\ell_2} = \sum_{j=0}^{n-\ell_2} \frac{\binom{n-\ell_2}{j}}{\binom{n}{j+1}} \frac{K_{j+1}}{j+1}$$

It is known (Berdenave Lugosi, Shirotorskiy) that $(B_k)_{1 \le k \le n}$ is non-increasing.

More can be said:

Theorem For $k \in \{1, \dots, n-1\}$ let $DB_k = B_k - B_{k+1}$ and let D^l, $l \geq 1$ be the l-th iteration of D. Then,

$$D^l B_k \geq 0$$

Consequence:

$$\operatorname{Var} LC_n \geq 2n B_{2n}$$

Theorem For $X_1, \dots, X_n, Y_1, \dots, Y_n$ iid $\operatorname{Bu}(p_0)$, $p_0 \neq 1/2$. Then for some $p \in (p_0, 1/2)$;

$$\limsup_{n \to +\infty} \frac{\operatorname{Var} LC_n}{n} > 0$$

Moral:

For upper bounds use J's, for lower bounds use K's and B's and hope.

Thank You.