Clark formula and logarithmic Sobolev inequalities for Bernoulli measures

Fuqing Gao[1] and Nicolas Privault[2]

Abstract - Using a Clark formula for the predictable representation of random variables in discrete time and adapting the method presented in [3] in the Brownian case, we obtain a proof of modified and L^1 logarithmic Sobolev inequalities for Bernoulli measures. We also prove a bound that improves these inequalities as well as the optimal constant inequality of [2].

Formule de Clark et inégalités de Sobolev logarithmiques pour les mesures de Bernoulli

Résumé - A l'aide d'une formule de Clark pour la représentation prévisible de variables aléatoire en temps discret et en adaptant la preuve présentée dans [3] dans le cas brownien, nous obtenons une preuve des inégalités de Sobolev logarithmiques (inégalité modifiée et inégalité L^1) pour les mesures de Bernoulli. Nous prouvons aussi une borne qui améliore ces inégalités ainsi que l'inégalité de constante optimale de [2].

Version française abrégée

Le calcul stochastique et la formule de Clark ont été utilisés avec succès pour la preuve d'inégalités de Sobolev logarithmiques sur l'espace des chemins [3] et sur l'espace de Poisson [1], [10]. Dans cette note, des méthodes similaires sont utilisées en temps discret pour la preuve d'inégalités de Sobolev logarithmiques pour les mesures de Bernoulli. Par une autre méthode nous obtenons aussi une amélioration de ces inégalités. Soit $(X_k)_{k\in\mathbb{N}}$ une famille de variables aléatoires indépendantes de Bernoulli, à valeurs dans $\{-1,1\}$ sur $\Omega = \{-1,1\}^{\mathbb{N}}$, avec $p_k = P(X_k = 1) > 0$ et $q_k = P(X_k = -1) > 0$, $k \in \mathbb{N}$. Dans [2], l'inégalité modifiée (2.1) a été prouvée pour F de la forme $F = f(X_0, \ldots, X_n)$, avec le gradient modulo 2

$$\nabla_k F = X_k \left(f(X_0, \ldots, X_{k-1}, -1, X_{k+1}, \ldots, X_n) - f(X_0, \ldots, X_{k-1}, 1, X_{k+1}, \ldots, X_n) \right),$$

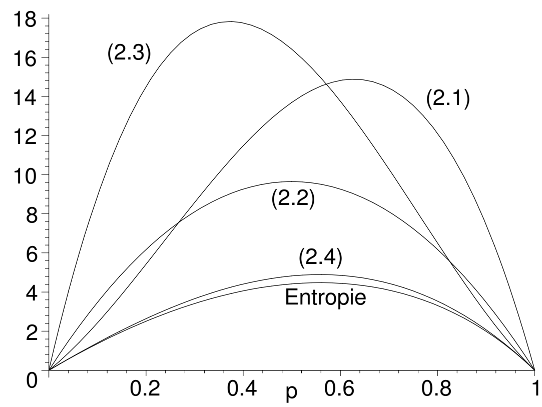
$k \in \mathbb{N}$. Dans cette Note nous utilisons l'opérateur $D_k = -X_k\sqrt{p_k q_k}\nabla_k$, qui satisfait une formule de Clark, pour prouver (2.1) ainsi que (2.2) qui est une inégalité L^1. Nous prouvons aussi l'inégalité

$$\text{Ent}\,[e^F] \le E\left[e^F \sum_{k=0}^{k=N} p_k q_k (\nabla_k F e^{\nabla_k F} - e^{\nabla_k F} + 1) \right] \tag{2.4}$$

[1]Department of Mathematics, Wuhan University, 430072 Wuhan, P.R. China. gaofq@public.wh.hb.cn This research was supported by the National Natural Science Foundation of China under grant No. 19971025.

[2]Département de Mathématiques, Université de La Rochelle, 17042 La Rochelle, France. nprivaul@univ-lr.fr

qui améliore (2.1), (2.2) et (2.3), comme illustré sur le graphe ci-dessous en dimension un, où l'entropie est représentée comme fonction de $p \in [0,1]$, avec $f(1) = 1$ et $f(-1) = 3.5$:



L'inégalité (2.4) est en fait une version discrète de l'inégalité optimale prouvée dans [10] sur l'espace de Poisson.

1 Clark formula for Bernoulli measures

In this section we describe a chaotic calculus and its associated predictable representation formula in discrete time, cf. [9]. We also refer to [6], [7], and [8] for other approaches to discrete chaotic calculus. Let $(X_k)_{k\in\mathbb{N}}$ be a family of i.i.d. Bernoulli random variables with $p_k = P(X_k = 1) > 0$, $q_k = P(X_k = -1) > 0$, $k \in \mathbb{N}$, and let

$$Y_k = \sqrt{\frac{q_k}{p_k}} 1_{\{X_k=1\}} - \sqrt{\frac{p_k}{q_k}} 1_{\{X_k=-1\}}, \quad k \in \mathbb{N},$$

so that $E[Y_k] = 0$ and $E[Y_k^2] = 1$, $k \in \mathbb{N}$. Let $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, $n \in \mathbb{N}$, and $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$. Let $(u_k)_{k\in\mathbb{N}}$ be a predictable process, i.e. u_k is $\mathcal{F}_{k-1}$-measurable, $k \in \mathbb{N}$, and assume that $(u_k)_{k\in\mathbb{N}}$ is square-integrable. Let $\ell^2(\mathbb{N})^{\circ n}$ denote the space of square-summable symmetric functions in n variables. Let

$$J_n(f_n) = \sum_{(k_1,\ldots,k_n)\in\Delta_n} f_n(k_1, \ldots, k_n) Y_{k_1} \cdots Y_{k_n},$$

$n \geq 1$, denote the discrete multiple stochastic integral of $f_n \in \ell^2(\mathbb{N})^{\circ n}$, where $\Delta_n = \{(k_1, \ldots, k_n) \in \mathbb{N}^n \ : \ k_i \neq k_j, \ 1 \leq i < j \leq n\}$. Every $F \in L^2(\Omega, P)$ has an orthogonal decomposition

$$F = E[F] + \sum_{n=1}^{\infty} J_n(f_n), \quad f_n \in l^2(\mathbb{N})^{\circ n}, \ n \geq 1.$$

We densely define the linear gradient operator $D : L^2(\Omega) \longrightarrow L^2(\Omega \times \mathbb{N})$ as

$$D_k J_n(f_n) = nJ_{n-1}(f_n(*, k)1_{\Delta_n}(*, k)), \quad k \in \mathbb{N}, \quad f_n \in \ell^2(\mathbb{N})^{\circ n}, \quad n \in \mathbb{N}. \tag{1.1}$$

From (1.1) and the expression of $J_n(f_n)$ we obtain the probabilistic interpretation of D_k as a finite difference operator:

$$D_k F = \sqrt{p_k q_k}(F_k^+ - F_k^-), \quad k \in \mathbb{N},$$

with $F = f(X_0, \ldots, X_n)$ and

$$F_k^+ = f(X_0, \ldots, X_{k-1}, +1, X_{k+1}, \ldots, X_n), \quad F_k^- = f(X_0, \ldots, X_{k-1}, -1, X_{k+1}, \ldots, X_n).$$

Proposition 1 *Let $F \in \mathrm{Dom}(D)$. We have the Clark formula*

$$F = E[F] + \sum_{k=0}^{\infty} E[D_k F \mid \mathcal{F}_{k-1}]Y_k. \tag{1.2}$$

Proof. The formula holds for $F = J_n(f_n)$:

$$\begin{aligned} J_n(f_n) &= n\sum_{k=0}^{\infty} J_{n-1}(f_n(*, k)1_{[0,k-1]^{n-1}}(*))Y_k = n\sum_{k=0}^{\infty} E[J_{n-1}(f_n(*, k)) \mid \mathcal{F}_{k-1}]Y_k \\ &= \sum_{k=0}^{\infty} E[D_k J_n(f_n) \mid \mathcal{F}_{k-1}]Y_k, \end{aligned}$$

and extends to $\mathrm{Dom}(D)$ by closability of D.

□

2 Logarithmic Sobolev inequalities

We start by recovering the modified logarithmic Sobolev inequality of [2] from the Clark representation formula in discrete time.

Theorem 1 *Let $F \in \mathrm{Dom}(D)$ with $F > \eta$ for some $\eta > 0$. We have*

$$\mathrm{Ent}\ [F] \leq E\left[\frac{1}{F}\|DF\|^2_{\ell^2(\mathbb{N})}\right]. \tag{2.1}$$

Proof. Let $S_n = E[F \mid \mathcal{F}_n]$, $0 \leq n \leq N$. The Clark formula (1.2) reads

$$S_n = S_{-1} + \sum_{k=0}^{k=n} u_k Y_k,$$

with $u_k = E[D_k F \mid \mathcal{F}_{k-1}]$, $0 \le k \le n \le N$, and $S_{-1} = E[F]$. Letting $f(x) = x \log x$ and using the bound

$$f(x+y) - f(x) = y \log x + (x+y) \log \left(1 + \frac{y}{x}\right) \le y(1 + \log x) + \frac{y^2}{x},$$

we have:

$$\begin{aligned}
\text{Ent } [F] &= E[f(S_N)] - E[f(S_{-1})] = E\left[\sum_{k=0}^{k=N} f(S_k) - f(S_{k-1})\right] \\
&= E\left[\sum_{k=0}^{k=N} f(S_{k-1} + Y_k u_k) - f(S_{k-1})\right] \\
&\le E\left[\sum_{k=0}^{k=N} Y_k u_k (1 + \log S_{k-1}) + \frac{Y_k^2 u_k^2}{S_{k-1}}\right] \\
&= E\left[\sum_{k=0}^{k=N} \frac{1}{E[F \mid \mathcal{F}_{k-1}]} (E[D_k F \mid \mathcal{F}_{k-1}])^2\right] \\
&\le E\left[\sum_{k=0}^{k=N} E\left[\frac{1}{F}(D_k F)^2 \mid \mathcal{F}_{k-1}\right]\right] = E\left[\frac{1}{F}\sum_{k=0}^{k=N} (D_k F)^2\right].
\end{aligned}$$

where we used the convexity of $(u, v) \mapsto v^2/u$ as in the Wiener and Poisson cases [3] and [1]. The extension of this inequality to $F \in \text{Dom}(D)$ relies on the closability of D.

□

The following L^1 inequality is present in [4] and [5], with applications to interacting random walks.

Theorem 2 *Let $F > 0$ be $\mathcal{F}_N$-measurable. We have*

$$\text{Ent } [F] \le E\left[\sum_{k=0}^{k=N} D_k F D_k \log F\right]. \tag{2.2}$$

Proof. Let $f(x) = x \log x$ and

$$\Psi(x, y) = (x+y) \log(x+y) - x \log x - (1 + \log x) y, \quad x,\ x+y > 0.$$

From the relation

$$\begin{aligned}
&Y_k u_k = Y_k E[D_k F \mid \mathcal{F}_{k-1}] \\
&\quad = q_k 1_{\{X_k=1\}} E[(F_k^+ - F_k^-) \mid \mathcal{F}_{k-1}] + p_k 1_{\{X_k=-1\}} E[(F_k^- - F_k^+) \mid \mathcal{F}_{k-1}] \\
&\quad = 1_{\{X_k=1\}} E[(F_k^+ - F_k^-) 1_{\{X_k=-1\}} \mid \mathcal{F}_{k-1}] + 1_{\{X_k=-1\}} E[(F_k^- - F_k^+) 1_{\{X_k=1\}} \mid \mathcal{F}_{k-1}],
\end{aligned}$$

we have, using the convexity of Ψ:

$$
\begin{aligned}
\text{Ent }[F] = E\left[\sum_{k=0}^{k=N} f\left(S_{k-1}+Y_k u_k\right) - f(S_{k-1})\right] \\
&= E\left[\sum_{k=0}^{k=N} \Psi(S_{k-1}, Y_k u_k) + Y_k u_k(1+\log S_{k-1})\right] \\
&= E\left[\sum_{k=0}^{k=N} \Psi(S_{k-1}, Y_k u_k)\right] \\
&= E\left[\sum_{k=0}^{k=N} p_k \Psi\left(E[F \mid \mathcal{F}_{k-1}], E[(F_k^+ - F_k^-)1_{\{X_k=-1\}} \mid \mathcal{F}_{k-1}]\right)\right. \\
&\quad \left. + q_k \Psi\left(E[F \mid \mathcal{F}_{k-1}], E[(F_k^- - F_k^+)1_{\{X_k=1\}} \mid \mathcal{F}_{k-1}]\right)\right] \\
&\leq E\left[\sum_{k=0}^{k=N} E\left[p_k \Psi\left(F, (F_k^+ - F_k^-)1_{\{X_k=-1\}}\right) + q_k \Psi\left(F, (F_k^- - F_k^+)1_{\{X_k=1\}}\right) \mid \mathcal{F}_{k-1}\right]\right] \\
&= E\left[\sum_{k=0}^{k=N} p_k 1_{\{X_k=-1\}} \Psi\left(F_k^-, F_k^+ - F_k^-\right) + q_k 1_{\{X_k=1\}} \Psi\left(F_k^+, F_k^- - F_k^+\right)\right] \\
&= E\left[\sum_{k=0}^{k=N} p_k q_k \Psi(F_k^-, F_k^+ - F_k^-) + p_k q_k \Psi(F_k^+, F_k^- - F_k^+)\right] \\
&= E\left[\sum_{k=0}^{k=N} D_k F D_k \log F\right].
\end{aligned}
$$

□

Th. 2 may also be proved by first using the bound

$$
f(x+y) - f(x) = y\log x + (x+y)\log\left(1+\frac{y}{x}\right) \leq y(1+\log x) + y\log(x+y),
$$

and then the convexity of $(u,v) \to v(\log(u+v) - \log u)$. The application of Th. 2 to e^F for $\mathcal{F}_N$-measurable F gives:

$$
\begin{aligned}
\text{Ent }[e^F] &\leq E\left[\sum_{k=0}^{k=N} p_k q_k \Psi(e^{F_k^-}, e^{F_k^+} - e^{F_k^-}) + p_k q_k \Psi(e^{F_k^+}, e^{F_k^-} - e^{F_k^+})\right] \\
&= E\left[\sum_{k=0}^{k=N} p_k q_k e^{F_k^-}(\nabla_K e^{\nabla_k F} - e^{\nabla_k F} + 1) + p_k q_k e^{F_k^+}(\nabla_k F e^{\nabla F} - e^{\nabla F} + 1)\right],
\end{aligned}
$$

which is not comparable to the optimal constant inequality of [2]:

$$
\text{Ent }[e^F] \leq E\left[e^F \sum_{k=0}^{k=N} p_k q_k (|\nabla_k F| e^{|\nabla_k F|} - e^{|\nabla_k F|} + 1)\right]. \tag{2.3}
$$

The inequality (2.4) below is better than (2.2) and (2.3). It also improves (2.1) from the bound

$$xe^x - e^x + 1 \le (e^x - 1)^2, \quad x \in \mathbb{R}.$$

Theorem 3 *Let F be $\mathcal{F}_N$-measurable. We have*

$$\text{Ent }[e^F] \le E\left[e^F \sum_{k=0}^{k=N} p_k q_k (\nabla_k F e^{\nabla_k F} - e^{\nabla_k F} + 1)\right]. \tag{2.4}$$

By tensorization the proof reduces to the following one-dimensional lemma.

Lemma 1 *For any $0 \le p \le 1$, $t \in \mathbb{R}$, $a \in \mathbb{R}$, $q = 1 - p$,*

$$\begin{aligned} & pte^t + qae^a - \left(pe^t + qe^a\right) \log \left(pe^t + qe^a\right) \\ & \le \; pq \left(qe^a \left((t-a)e^{t-a} - e^{t-a} + 1\right) + pe^t \left((a-t)e^{a-t} - e^{a-t} + 1\right)\right). \end{aligned}$$

Proof. Set

$$\begin{aligned} g(t) &= pq \left(qe^a \left((t-a)e^{t-a} - e^{t-a} + 1\right) + pe^t \left((a-t)e^{a-t} - e^{a-t} + 1\right)\right) \\ &\quad - pte^t - qae^a + \left(pe^t + qe^a\right) \log \left(pe^t + qe^a\right). \end{aligned}$$

Then

$$g'(t) = pq \left(qe^a (t-a)e^{t-a} + pe^t \left(-e^{a-t} + 1\right)\right) - pte^t + pe^t \log(pe^t + qe^a)$$

and $g''(t) = pe^t h(t)$, where

$$h(t) = -a - 2pt - p + 2pa + p^2 t - p^2 a + \log(pe^t + qe^a) + \frac{pe^t}{pe^t + qe^a}.$$

Now,

$$h'(t) = \frac{pq^2(e^t - e^a)(pe^t + (q+1)e^a)}{(pe^t + qe^a)^2},$$

which implies that $h'(a) = 0$, $h'(t) < 0$ for any $t < a$ and $h'(t) > 0$ for any $t > a$. Hence, for any $t \neq a$, $h(t) > h(a) = 0$, and so $g''(t) \ge 0$ for any $t \in \mathbb{R}$ and $g''(t) = 0$ if and only if $t = a$. Therefore, g' is strictly increasing. Finally, since $t = a$ is the unique root of $g' = 0$, we have that $g(t) \ge g(a) = 0$ for all $t \in \mathbb{R}$.

□

In the symmetric case $p_k = q_k = 1/2$, $k \in \mathbb{N}$, (2.4) reads

$$\text{Ent }[e^F] \le \frac{1}{2} E\left[\sum_{k=0}^{k=N} D_k F D \log F\right].$$

Let $U_n = (n + X_1 + \cdots + X_n)/2$, $F = \varphi(U_n)$, and $p_k = \lambda/n$, $k \in \mathbb{N}$, $\lambda > 0$. Then

$$E\left[e^F \sum_{k=0}^{k=n} p_k q_k (\nabla_k F e^{\nabla_k F} - e^{\nabla_k F} + 1)\right]$$
$$= \frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right) U_n e^{\varphi(U_n)}((\varphi(U_n) - \varphi(U_n - 1))e^{\varphi(U_n)-\varphi(U_n-1)} - e^{\varphi(U_n)-\varphi(U_n-1)} + 1)$$
$$+\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)(n - U_n)e^{\varphi(U_n)}((\varphi(U_n + 1) - \varphi(U_n))e^{\varphi(U_n+1)-\varphi(U_n)} - e^{\varphi(U_n+1)-\varphi(U_n)} + 1),$$

and as n goes to infinity we obtain from the Poisson limit theorem:

$$\text{Ent}\ [\varphi(U)] \leq \lambda E[e^{\varphi(U)}((\varphi(U + 1) - \varphi(U))e^{\varphi(U+1)-\varphi(U)} - e^{\varphi(U+1)-\varphi(U)} + 1)],$$

where U is a Poisson random variable with parameter λ. This corresponds to the sharp inequality of [10]. However, (2.4) does not improve the condition $\sup_{k\in\mathbb{N}} |\nabla_k F| \leq \beta$ for the deviation result of [2], since the conditions $\nabla_k F \leq \beta$ a.e. and $|\nabla_k F| \leq \beta$ a.e. are equivalent, $k \in \mathbb{N}$.

Acknowledgements. The second named author thanks L. Wu for his hospitality at Wuhan University, and for useful comments.

References

[1] C. Ané and M. Ledoux. On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probab. Theory Related Fields*, 116(4):573–602, 2000.

[2] S. G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[3] M. Capitaine, E.P. Hsu, and M. Ledoux. Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces. *Electron. Comm. Probab.*, 2:71–81 (electronic), 1997.

[4] P. Dai Pra, A.M. Paganoni, and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.*, 30(4):1959–1976, 2002.

[5] F. Gao and J. Quastel. Exponential decay of entropy in the random transposition and Bernoulli-Laplace models. Preprint, 2002.

[6] H. Holden, T. Lindstrøm, B. Øksendal, and J. Ubøe. Discrete Wick calculus and stochastic functional equations. *Potential Anal.*, 1(3):291–306, 1992.

[7] M. Leitz-Martini. A discrete Clark-Ocone formula. Maphysto Research Report No 29, 2000.

[8] P.A. Meyer. *Quantum Probability for Probabilists*, volume 1538 of *Lecture Notes in Mathematics*. Springer-Verlag, 1993.

[9] N. Privault and W. Schoutens. Discrete chaotic calculus and covariance identities. *Stochastics and Stochastics Reports*, 72:289–315, 2002. Eurandom Report 006, 2000.

[10] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.