Correction 1

1. Variance des statistiques d'ordre

(1) Pour $k \leq \lfloor \frac{n}{2} \rfloor$ et $1 \leq i \leq n$, on définit $X_{(k)}^{(i)}$ comme la $k^{\text{ième}}$ statistique d'ordre dans l'échantillon $X^{(i)} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. On remarque que lorsque $X_i = X_{(\ell)}$ pour $\ell \leq k$, alors $X_{(k)}^{(i)} = X_{(k+1)}$. Et quand $X_i = X_{(\ell)}$ pour $\ell \geq k+1$, alors $X_{(k)}^{(i)} = X_{(k)}$. L'inégalité d'Efron–Stein donne donc

$$\begin{aligned}\mathbf{Var}(X_{(k)}) &\leq \sum_{i=1}^{n} \mathbf{E}\left[\left(X_{(k)} - X_{(k)}^{(i)}\right)^2\right] \\ &\leq k\mathbf{E}\left[\left(X_{(k)} - X_{(k+1)}\right)^2\right] .\end{aligned}$$

Pour $k > \lfloor \frac{n}{2} \rfloor$, on définit $X_{(k)}^{(i)}$ comme la $(k-1)^{\text{ième}}$ statistique d'ordre dans l'échantillon $X^{(i)}$. Lorsque $X_i = X_{(\ell)}$ pour $\ell \leq k-1$, alors $X_{(k)}^{(i)} = X_{(k)}$. Et quand $X_i = X_{(\ell)}$ pour $\ell \geq k$, alors $X_{(k)}^{(i)} = X_{(k-1)}$. L'inégalité d'Efron–Stein donne alors

$$\begin{aligned}\mathbf{Var}(X_{(k)}) &\leq \sum_{i=1}^{n} \mathbf{E}\left[\left(X_{(k)} - X_{(k)}^{(i)}\right)^2\right] \\ &\leq (n-k+1)\mathbf{E}\left[\left(X_{(k)} - X_{(k-1)}\right)^2\right] .\end{aligned}$$

(2) Pour $X_1, \ldots, X_n$ i.i.d. de loi uniforme sur $[0,1]$, on a

$$(X_{(1)}, X_{(2)} - X_{(1)}, \ldots, X_{(n)} - X_{(n-1)}, 1 - X_{(n)}) \sim \mathrm{Dir}(1, \ldots, 1) .$$

En particulier, $X_{(k)} \sim \mathrm{Beta}(k, n+1-k)$ et $X_{(k)} - X_{(k-1)} \sim \mathrm{Beta}(1, n)$. Ainsi, en utilisant que

$$\mathbf{Var}\left(\mathrm{Beta}(a,b)\right) = \frac{ab}{(a+b)^2(a+b+1)} \qquad \text{et} \qquad \mathbf{E}\left[\mathrm{Beta}(a,b)^2\right] = \frac{a(a+1)}{(a+b)(a+b+1)} ,$$

on a

$$\mathbf{Var}(X_{(k)}) = \frac{k(n+1-k)}{(n+1)^2(n+2)} .$$

Pour $k \leq \lfloor \frac{n}{2} \rfloor$, $\mathbf{Var}(X_{(k)}) \asymp \frac{k}{n^2}$, et la borne d'Efron–Stein vérifie

$$k\mathbf{E}\left[\left(X_{(k)} - X_{(k+1)}\right)^2\right] = \frac{2k}{(n+1)(n+2)} \asymp \frac{k}{n^2} .$$

De même pour $k > \lfloor \frac{n}{2} \rfloor$, $\mathbf{Var}(X_{(k)}) \asymp \frac{n-k+1}{n^2}$, et la borne d'Efron–Stein vérifie

$$(n-k+1)\mathbf{E}\left[\left(X_{(k)} - X_{(k-1)}\right)^2\right] = \frac{2(n-k+1)}{(n+1)(n+2)} \asymp \frac{n-k+1}{n^2} .$$

Ainsi dans les deux cas, l'ordre de grandeur de la borne d'Efron–Stein est bien celui de la vraie variance.

2. Une inégalité isopérimétrique

Notons $f = \mathbb{1}_A$. Comme $1.\log 1 = 0.\log 0 = 0$, on a

$$\mathbf{Ent}(f) = \mathbf{E}[f \log f] - \mathbf{E}[f] \log \mathbf{E}[f] = \mu(A) \log \frac{1}{\mu(A)} .$$

D'autre part, par la sous-additivité de l'entropie,

$$\mathbf{Ent}(f) \leq \sum_{i=1}^{n} \mathbf{E}\left[\mathbf{Ent}^{(i)}(f)\right] ,$$

avec $\mathbf{Ent}^{(i)}(f) = -\mathbf{E}^{(i)}[f] \log \mathbf{E}^{(i)}[f]$. Remarquons que

$$\mathbf{E}^{(i)}[f] = \mathbf{P}\left(X \in A \mid X^{(i)}\right) = \begin{cases} 1 & \text{si } X \in A \text{ et } \bar{X}^{(i)} \in A, \\ 0 & \text{si } X \notin A \text{ et } \bar{X}^{(i)} \notin A, \\ \frac{1}{2} & \text{si } \mathbb{1}_{X \in A} \neq \mathbb{1}_{\bar{X}^{(i)} \in A}. \end{cases}$$

Ainsi

$$\sum_{i=1}^{n} \mathbf{E}\left[\mathbf{Ent}^{(i)}(f)\right] = \frac{1}{2}\log(2) \sum_{i=1}^{n} \mathbf{P}\left(\mathbb{1}_{X \in A} \neq \mathbb{1}_{\bar{X}^{(i)} \in A}\right) = \frac{1}{2}\log(2)\mathbf{I}(A)\,.$$

On obtient donc

$$\mathbf{I}(A) \geq 2\mu(A) \log_2 \frac{1}{\mu(A)}\,,$$

où $\log_2(x) = \frac{\log(x)}{\log(2)}$. Pour $A = \{x \in \{0,1\}^n,\ x_1 = \cdots = x_k = 1\}$ avec $k \in [\![1, n]\!]$, on a $\mu(A) = 2^{-k}$. De plus, pour $i > k$, $\mathbf{I}_i(A) = 0$, et pour $i \leq k$,

$$\mathbf{I}_i(A) = \mathbf{P}(X_1 = \cdots = X_{i-1} = X_{i+1} = \cdots = X_k = 1) = 2^{-k+1}\,.$$

Ainsi

$$\mathbf{I}(A) = k.2^{-k+1} = 2\mu(A) \log_2 \frac{1}{\mu(A)}\,.$$

3. Constante de Sobolev logarithmique pour la Bernoulli

Comme suggéré, prenons f donnée par $f(x) = z^{\sum x_i}$ avec $z = \frac{1-p}{p}$. En remarquant que $pz^2 + 1 - p = z$, on a

$$\mathbf{E}[f^2] = \mathbf{E}\left[z^{2\sum_{i=1}^n X_i}\right] = \mathbf{E}\left[z^{2X_1}\right]^n = \left(pz^2 + 1 - p\right)^n = z^n\,,$$

et

$$\begin{aligned} \mathbf{E}[f^2 \log(f^2)] &= \mathbf{E}\left[z^{2\sum_{i=1}^n X_i} \sum_{j=1}^{n} 2X_j \log(z)\right] \\ &= 2n \log(z) \mathbf{E}\left[X_1 z^{2\sum_{i=1}^n X_i}\right] \\ &= 2n \log(z) \mathbf{E}\left[X_1 z^{2X_1}\right] \mathbf{E}\left[z^{2\sum_{i=2}^n X_i}\right] \\ &= 2n \log(z) p z^2 z^{n-1} \\ &= 2n \log(z) p z^{n+1}\,. \end{aligned}$$

Ainsi

$$\mathbf{Ent}(f^2) = 2n\log(z)pz^{n+1} - nz^n \log(z) = nz^n \log(z)(2pz - 1) = nz^n \log(z)(1-2p)\,.$$

D'autre part,

$$\sum_{i=1}^{n} \mathbf{E}\left[\left(f(X) - f(\bar{X}^{(i)}\right)^2\right] = n\mathbf{E}\left[z^{2\sum_{i=2}^n X_i}(z-1)^2\right] = n(z-1)^2 z^{n-1}\,.$$

Il suffit donc de montrer que

$$z\log(z)(1-2p) = c(p)p(1-p)(z-1)^2\,,$$

ce qui se vérifie facilement.

4. Inégalité de Sobolev logarithmique et concentration de la mesure

(1) On peut supposer $x < y$. Par convexité de $t \mapsto e^t$, on a

$$e^{y/2} - e^{x/2} = \int_{x/2}^{y/2} e^t \mathrm{d}t = \frac{1}{2}(y-x)\frac{e^{x/2}+e^{y/2}}{2} \cdot$$

En utilisant l'inégalité $(a+b)^2 \leq 2(a^2+b^2)$, on a alors

$$\left(e^{y/2} - e^{x/2}\right)^2 \leq \frac{(y-x)^2}{16} \cdot 2\left(e^x + e^y\right) = \frac{(y-x)^2}{8}\left(e^x + e^y\right) .$$

(2) L'inégalité de Sobolev logarithmique appliquée à la fonction $x \mapsto e^{\frac{\lambda f(x)}{2}}$ donne

$$\mathbf{Ent}\left(e^{\lambda Z}\right) \leq \frac{1}{2}\sum_{i=1}^n \mathbf{E}\left[\left(e^{\frac{\lambda f(X)}{2}} - e^{\frac{\lambda f(\bar{X}^{(i)})}{2}}\right)^2\right] .$$

En utilisant l'inégalité de la question (1), on a

$$\begin{aligned}\mathbf{Ent}\left(e^{\lambda Z}\right) &\leq \frac{\lambda^2}{16}\sum_{i=1}^n \mathbf{E}\left[\left(f(X) - f(\bar{X}^{(i)})\right)^2\left(e^{\lambda f(X)} + e^{\lambda f(\bar{X}^{(i)})}\right)\right] \\ &= \frac{\lambda^2}{8}\sum_{i=1}^n \mathbf{E}\left[\left(f(X) - f(\bar{X}^{(i)})\right)^2 e^{\lambda f(X)}\right] ,\end{aligned}$$

où l'on a utilisé que $(X, \bar{X}^{(i)})$ et $(\bar{X}^{(i)}, X)$ ont la même loi. En utilisant l'hypothèse que pour tout $x \in \{0,1\}^n$,

$$\sum_{i=1}^n \left(f(x) - f(\bar{x}^{(i)})\right)^2 \leq 4v ,$$

on obtient

$$\mathbf{Ent}\left(e^{\lambda Z}\right) \leq \frac{\lambda^2 v}{2}\mathbf{E}\left[e^{\lambda f(X)}\right] .$$

(3) Pour $\psi(\lambda) = \log \mathbf{E}e^{\lambda(Z-\mathbf{E}Z)}$, on a

$$\psi'(\lambda) = \frac{\mathbf{E}\left[(Z-\mathbf{E}Z)e^{\lambda(Z-\mathbf{E}Z)}\right]}{\mathbf{E}\left[e^{\lambda(Z-\mathbf{E}Z)}\right]} = \frac{\mathbf{E}\left[Ze^{\lambda Z}\right] - \mathbf{E}[Z]\mathbf{E}\left[e^{\lambda Z}\right]}{\mathbf{E}\left[e^{\lambda Z}\right]} .$$

Ainsi

$$\lambda\psi'(\lambda) - \psi(\lambda) = \frac{\lambda\mathbf{E}\left[Ze^{\lambda Z}\right] - \mathbf{E}\left[e^{\lambda Z}\right]\log \mathbf{E}\left[e^{\lambda Z}\right]}{\mathbf{E}\left[e^{\lambda Z}\right]} = \frac{\mathbf{Ent}\left(e^{\lambda Z}\right)}{\mathbf{E}\left[e^{\lambda Z}\right]} .$$

On a donc

$$\frac{\lambda\psi'(\lambda) - \psi(\lambda)}{\lambda^2} \leq \frac{v}{2} ,$$

c'est-à-dire $\Psi'(\lambda) \leq \frac{v}{2}$ avec $\Psi(\lambda) = \frac{\psi(\lambda)}{\lambda}$. Par la règle de L'Hôpital, $\lim_{\lambda\to 0}\Psi(0) = 0$. Ainsi, par intégration de l'inégalité,

$$\Psi(\lambda) \leq \frac{v\lambda}{2} ,$$

c'est-à-dire $\psi(\lambda) \leq \frac{\lambda^2 v}{2}$.

(4) Pour tout $t \geq 0$ et $\lambda > 0$, on a par l'inégalité de Markov,

$$\mathbf{P}\left(Z - \mathbf{E}Z \geq t\right) = \mathbf{P}\left(e^{\lambda(Z-\mathbf{E}Z)} \geq e^{\lambda t}\right) \leq e^{-\{\lambda t - \psi(\lambda)\}} .$$

Or $\psi(\lambda) \leq \frac{\lambda^2 v}{2}$. Et comme l'inégalité est valable pour tout $\lambda > 0$, on a

$$\mathbf{P}\left(Z - \mathbf{E}Z \geq t\right) \leq e^{-\sup_{\lambda\geq 0}\left\{\lambda t - \frac{\lambda^2 v}{2}\right\}} .$$

La fonction $\lambda \mapsto \lambda t - \frac{\lambda^2 v}{2}$ est maximale en $\lambda_\star = \frac{t}{v}$, ce qui donne

$$\mathbf{P}\left(Z - \mathbf{E}Z \geq t\right) \leq \exp\left(-\frac{t^2}{2v}\right) .$$