Feuille 1 : correction

Exercice 1 (**Queue gaussienne**)
Soit $x > 0$. On a d'une part

$$\int_x^{+\infty} e^{-t^2/2}dt \leq \frac{1}{x}\int_x^{+\infty} te^{-t^2/2}dt = \frac{1}{x}e^{-x^2/2},$$

et d'autre part, par une intégration par parties,

$$\begin{aligned}\int_x^{+\infty} e^{-t^2/2}dt &= \left[-\frac{1}{t}e^{-t^2/2}\right]_x^{+\infty} - \int_x^{+\infty}\frac{1}{t^2}e^{-t^2/2}dt \\ &\geq \frac{1}{x}e^{-x^2/2} - \frac{1}{x^3}\int_x^{+\infty} te^{-t^2/2}dt \\ &= \left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-x^2/2}.\end{aligned}$$

■

Exercice 2 (**Inégalité de Hoeffding**)

1. Comme $Y \in [0,1]$, on a, par convexité, pour tout $\lambda \in \mathbb{R}$,

   $$e^{\lambda Y} \leq Ye^{\lambda} + 1 - Y.$$

   Ainsi, $\log \mathbb{E}e^{\lambda(Y-\mathbb{E}Y)} \leq \log\left(pe^{\lambda}+1-p\right) - \lambda p = \varphi(\lambda)$. Par le théorème de Taylor, il existe θ entre 0 et λ tel que

   $$\varphi(\lambda) = \varphi(0) + \lambda\varphi'(0) + \frac{\lambda^2}{2}\varphi''(\theta).$$

   Il suffit maintenant de remarquer que $\varphi(0) = \varphi'(0) = 0$ et que

   $$\varphi''(\theta) = \frac{p(1-p)e^{\theta}}{(pe^{\theta}+1-p)^2} \leq \frac{1}{4}.$$

2. Si X est dans $[a,b]$, la variable $Y = \frac{X-a}{b-a}$ est dans $[0,1]$. En appliquant le résultat de la question précédente, on a

   $$\log \mathbb{E}e^{\lambda(X-\mathbb{E}X)} = \log \mathbb{E}^{\lambda(b-a)(Y-\mathbb{E}Y)} \leq \frac{\lambda^2(b-a)^2}{8}.$$

   Par indépendance des variables (X_i), on a

   $$\log \mathbb{E}e^{\lambda(Z-\mathbb{E}Z)} = \sum_{i=1}^n \log \mathbb{E}e^{\lambda(X_i-\mathbb{E}X_i)} \leq \sum_{i=1}^n \frac{\lambda^2(b_i-a_i)^2}{8}$$

3. Notons $v = \sum_{i=1}^n (b_i-a_i)^2$. Pour tout $\lambda, t > 0$, on a, par l'inégalité de Markov,

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

   Comme c'est vrai pour tout $\lambda > 0$, on a

   $$\mathbb{P}\left(Z-\mathbb{E}Z \geq t\right) \leq e^{-\sup_{\lambda>0}\left\{\lambda t - \log \mathbb{E}\left[e^{\lambda(Z-\mathbb{E}Z)}\right]\right\}} \leq e^{-\sup_{\lambda>0}\left\{\lambda t - \frac{\lambda^2 v}{8}\right\}}.$$

   On vérifie que le supremum est atteint en $\lambda_t = \frac{t}{v}$, ce qui donne

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

■

Exercice 3 (**Tirages sans remise**)

■

Exercice 4 (**Lemme de Slutsky**) Soit $(s,t) \in \mathbb{R}^2$. On a par l'inégalité triangulaire,

$$\begin{aligned}\left|\mathbb{E}\left[e^{isX_n+itY_n}\right]-\mathbb{E}\left[e^{isX+ita}\right]\right| &= \left|\mathbb{E}\left[e^{isX_n}\left(e^{itY_n}-e^{ita}\right)\right]+e^{ita}\mathbb{E}\left[e^{isX_n}-e^{isX}\right]\right| \\ &\leq \mathbb{E}\left[\left|e^{itY_n}-e^{ita}\right|\right]+\left|\mathbb{E}\left[e^{isX_n}-e^{isX}\right]\right| .\end{aligned}$$

On a $\mathbb{E}\left[e^{isX_n}\right] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left[e^{isX}\right]$ et, comme la fonction $y \mapsto |e^{ity}-e^{ita}|$ est continue bornée, on a aussi $\mathbb{E}\left[|e^{itY_n}-e^{ita}|\right] \underset{n\to\infty}{\longrightarrow} 0$. ■

Exercice 5 (**Delta-méthode**)

1. Dire que g est dérivable en a signifie qu'il existe une fonction r qui tend vers 0 en a telle que pour tout $x \in \mathbb{R}$
$$g(x) = g(a) + (x-a)\left(g'(a) + r(x)\right) .$$
La fonction r est prolongeable par continuité en a, en posant $r(a) = 0$. On a alors $r(X_n) \xrightarrow{\mathbb{P}} 0$. En effet, $X_n \xrightarrow{\mathbb{P}} a$ (par le lemme de Slutsky, $X_n - a = \frac{1}{v_n}v_n(X_n-a)$ converge en loi vers $0.X = 0$, donc en probabilité), et r est continue en a, donc $r(X_n) \xrightarrow{\mathbb{P}} 0$. Ainsi, encore par le lemme de Slutsky,
$$v_n\left(g(X_n) - g(a)\right) = v_n(X_n-a)\left(g'(a) + r(X_n)\right) \rightsquigarrow g'(a)X .$$
2. Par le théorème central limite, on a
$$\sqrt{n}(\overline{X}_n - \lambda) \rightsquigarrow \mathcal{N}(0,\lambda) ,$$
et, par la question précédente, pour toute fonction g dérivable en λ, on a
$$\sqrt{n}\left(g(\overline{X}_n) - g(\lambda)\right) \rightsquigarrow \mathcal{N}\left(0, g'(\lambda)^2\lambda\right) .$$
Pour stabiliser la variance, il suffit donc de trouver une fonction g telle que $g'(\lambda)^2\lambda = 1$. La fonction $g : x \mapsto 2\sqrt{x}$ par exemple.

■

Exercice 6 (**Delta-méthode, ordre 2**)

1. Par la loi forte des grands nombres (X_i i.i.d et intégrables), $\overline{X}_n$ converge p.s. vers $\mathbb{E}(X_1) = 0$. De même, le TCL (X_i i.i.d et de carré intégrable) donne $\sqrt{n}\overline{X}_n \rightsquigarrow \mathcal{N}(0,1)$.
2. En se souvenant que $\cos' = -\sin$ et par la Delta méthode (cos dérivable en 0), on obtient $\sqrt{n}(\cos(\overline{X}_n)-1) \rightsquigarrow 0$. Ici la limite est dégénérée car p.s. constante. On en déduit uniquement que la vitesse de convergence de $\cos(\overline{X}_n)$ vers 1 est plus rapide que $1/\sqrt{n}$. Il faut donc pousser le développement à l'ordre suivant pour pouvoir être plus précis sur vitesse et limite.
3. Le développement limité de cos en 0 à l'ordre 2 donne $\cos(x) = 1 - x^2/2 + o(x^2)$. Autrement dit, il existe une fonction r tendant vers 0 en 0 telle que :
$$\cos(x) = 1 - \frac{x^2}{2}(1 + r(x)).$$
On a alors
$$\cos(\overline{X}_n) - 1 = -\frac{(\overline{X}_n)^2}{2}(1 + r(\overline{X}_n)),$$
soit
$$-2n(\cos(\overline{X}_n) - 1) = (\sqrt{n}\overline{X}_n)^2(1 + r(\overline{X}_n)).$$

Il reste maintenant à étudier la convergence du terme de droite avec comme outil le théorème de Slutsky. Tout d'abord, la fonction r est prolongeable par continuité en 0 de sorte que $r(0)=0$. Or $\overline{X}_n$ converge p.s. (et a fortiori en probabilité) vers 0, donc par le théorème de continuité, $r(\overline{X}_n) \xrightarrow{\mathbb{P}} 0$. De même, $\sqrt{n}\overline{X}_n \rightsquigarrow \mathcal{N}(0,1)$, donc par continuité $(\sqrt{n}\overline{X}_n)^2 \rightsquigarrow Y^2$ avec $Y \sim \mathcal{N}(0,1)$. Autrement dit, $(\sqrt{n}\overline{X}_n)^2 \rightsquigarrow \chi^2(1)$. Finalement, le théorème de Slutsky donne :

$$-2n(\cos(\overline{X}_n)-1) \rightsquigarrow \chi^2(1).$$

Bilan : la suite de v.a. $(\cos(\overline{X}_n))$ converge à vitesse $1/n$ vers sa limite, c'est-à-dire bien plus rapidement que pour le TCL classique.

■

Exercice 7 (**Propriétés des moyenne et médiane empiriques**)

1. (a) Soit f la fonction qui à tout réel t associe

$$f(t) = \mathbb{E}\left[(X_1-t)^2\right] = t^2 - 2\mathbb{E}(X_1)t + E(X_1^2).$$

Comme trinôme strictement convexe, la fonction f admet un unique minimum là où sa dérivée s'annule : $f'(t) = 2t - 2\mathbb{E}(X_1) = 0 \Leftrightarrow t = \mathbb{E}(X_1)$.

   (b) De la même manière que précédemment, $t \in \mathbb{R} \mapsto t^2 - 2\left(n^{-1}\sum_{i=1}^n X_i\right)t + n^{-1}\sum_{i=1}^n X_i^2$ admet un unique minimum en son point critique qui est $t = n^{-1}\sum_{i=1}^n X_i = \overline{X}_n$.

2. Rappelons que pour toute v.a. $Z \geq 0$, on a

$$\mathbb{E}[Z] = \int_0^\infty \mathbb{P}(Z>z)dz,$$

car l'égalité de Fubini donne

$$\int_0^\infty \mathbb{P}(Z>z)dz = \int_0^\infty \mathbb{E}\left[\mathbb{1}_{Z>z}\right]dz = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{Z>z}dz\right] = \mathbb{E}[Z].$$

   (a) En utilisant l'égalité ci-dessus pour $Z = |X_1 - t|$ et en remarquant que

$$\mathbb{P}(|X_1-t|>z) \quad = \quad \mathbb{P}\{X_1 - t > z\} + \mathbb{P}\{t - X_1 > z\},$$

on obtient pour tout $t \in \mathbb{R}$

$$\begin{aligned}
\mathbb{E}\left(|X_1-t|\right) &= \int_0^\infty \mathbb{P}(|X_1-t|>z)\,dz \\
&= \int_0^\infty \mathbb{P}(X_1-t>z)\,dz + \int_0^\infty \mathbb{P}(t-X_1>z)\,dz \\
&= \int_0^\infty (1-F(z+t))\,dz + \int_0^\infty F(t-z)\,dz \\
&= \int_t^\infty (1-F(u))\,du + \int_{-\infty}^t F(u)\,du.
\end{aligned}$$

Ainsi, en dérivant par rapport à t (ce qui est licite car F est continue), on obtient [1]

$$\frac{\partial}{\partial t}\mathbb{E}|X_1-t| = -(1-F(t)) + F(t) = 2F(t)-1.$$

Par suite, comme F est une bijection strictement croissante d'un voisinage de $x_{1/2}$ dans un voisinage de $1/2$, on a

$$2F(t)-1>0 \Leftrightarrow t > x_{1/2} \text{ et } 2F(t)-1<0 \Leftrightarrow t < x_{1/2}$$

donc la fonction $t \mapsto \mathbb{E}|X_1-t|$ atteint son unique minimum en $t = x_{1/2}$.

1. Étant données deux fonctions dérivables a et b, une fonction continue f et en appelant F une primitive de f, on a $\phi(t) = \int_{a(t)}^{b(t)} f(u)\,du = F(b(t)) - F(a(t))$. Ainsi, $\phi'(t) = b'(t)f(b(t)) - a'(t)f(a(t))$.

(b) La fonction $g : t \mapsto \frac{1}{n}\sum_{i=1}^n |X_i - t|$ est affine par morceaux et sa pente change à chaque X_i. En ordonnant les X_i de la façon suivante

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)},$$

et en notant $X_{(0)} = -\infty$ et $X_{(n+1)} = +\infty$, on remarque que la pente de la fonction g vaut $-n/n$ sur $]-\infty, X_{(1)}[$, $-(n-2)/n$ sur $[X_{(1)}, X_{(2)}[$, etc, puis $(n-2)/n$ sur $[X_{(n-1)}, X_{(n)}[$ et n/n sur $[X_{(n)}, \infty[$. Autrement dit, la pente vaut $(2(k-1)-n)/n$ sur l'intervalle $[X_{(k-1)}, X_{(k)}[$ (en excluant $X_{(0)}$ pour $k=1$). Ainsi, de deux choses l'une :

- si n est pair, la pente de g va s'annuler sur un intervalle, correspondant à $[X_{(n/2)}, X_{(n/2+1)}]$. Donc la fonction g est minimale sur l'intervalle $[X_{(n/2)}, X_{(n/2+1)}]$ (l'argmin est alors un intervalle).
- Si n est impair, la pente de g ne va pas s'annuler et cette fonction atteint son unique minimum en $X_{((n+1)/2)}$.

Cependant, puisque la médiane empirique est définie de façon unique par $X_{((n+1)/2)}$ si n est impair et $X_{(n/2)}$ si n est pair, elle minimise dans tous les cas la fonction g (c'est la plus petite valeur minimisant g).

■

Exercice 8 (**Lien médiane-espérance-écart-type**)

$$\begin{aligned} |m-\mu| &\leq \mathbb{E}|X-m| \quad \text{par Jensen} \\ &\leq \mathbb{E}|X-\mu| \quad \text{car la médiane est le minimiseur de } u \mapsto \mathbb{E}|X-u| \\ &\leq \left(\mathbb{E}|X-\mu|^2\right)^{1/2} = \sigma, \ \text{par Cauchy-Schwarz.} \end{aligned}$$

■

Exercice 9 (**Loi uniforme**)

1. (a) On peut commencer par remarquer que le vecteur $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ est uniformément distribué sur l'ensemble $\mathcal{E}_n = \{x \in [0,1]^n,\ x_1 \leq \cdots \leq x_n\}$. En effet, comme les variables $(X_1, \ldots, X_n)$ sont échangeables, en notant Σ la permutation aléatoire de $[\![1,n]\!]$ définie par $\Sigma(i) = j$ si et seulement si $X_{(i)} = X_j$, on a, pour toute fonction $\varphi : \mathbb{R}^n \to \mathbb{R}$ continue bornée,

$$\begin{aligned} \mathbb{E}\left[\varphi\left(X_{(1)}, \ldots, X_{(n)}\right)\right] &= \sum_{\sigma \in \mathcal{S}_n} \mathbb{E}\left[\mathbb{1}_{\Sigma=\sigma}\varphi\left(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\right)\right] \\ &= \sum_{\sigma \in \mathcal{S}_n} \mathbb{E}\left[\mathbb{1}_{X_{\sigma(1)} \leq \cdots \leq X_{\sigma(n)}}\varphi\left(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\right)\right] \\ &= n!\, \mathbb{E}\left[\mathbb{1}_{X_1 \leq \cdots \leq X_n}\varphi\left(X_1, \ldots, X_n\right)\right] \\ &= n! \int_{[0,1]^n} \mathbb{1}_{x \in \mathcal{E}_n}\varphi(x_1, \ldots, x_n) dx\,, \end{aligned}$$

ce qui montre que la loi est bien uniforme sur $\mathcal{E}_n$ (et au passage, que le volume de $\mathcal{E}_n$ vaut $\frac{1}{n!}$). Par un changement de variables immédiat, on obtient que le vecteur $(X_{(1)}, X_{(2)} - X_{(1)}, \ldots, X_{(n)} - X_{(n-1)})$ est uniformément distribué sur l'ensemble $\mathcal{E}'_n = \{x \in [0,1]^n,\ \sum_{i=1}^n x_i \leq 1\}$. En particulier ce vecteur (celui des écarts entre les statistiques d'ordre successives) est échangeable. Donc (en prenant le premier bout du segment $[0,1]$, celui de longueur $X_{(1)}$ et en le mettant à la fin), on obtient que $R_n = X_{(n)} - X_{(1)}$ a la même loi que $X_{(n-1)}$. Or, pour tout $k \in [\![1,n]\!]$ et $x \in [0,1]$, on a

$$\mathbb{P}(X_{(k)} \leq x) = \mathbb{P}\left(\mathrm{Bin}(n,x) \geq k\right)\,.$$

Ainsi

$$\mathbb{P}(X_{(n-1)} \leq x) = nx^{n-1}(1-x) + x^n\,.$$

En dérivant, on obtient que la densité est donnée par $x \mapsto n(n-1)n(n-1)x^{n-2}(1-x)\mathbb{1}_{[0,1]}(x)$. C'est la densité d'une loi Beta$(n-1,2)$. Pour le voir, on peut aussi remarquer que

$$\left(X_{(1)}, X_{(2)} - X_{(1)}, \ldots, X_{(n)} - X_{(n-1)}\right) \sim \left(\frac{E_1}{\sum_{i=1}^{n+1} E_i}, \frac{E_2}{\sum_{i=1}^{n+1} E_i}, \ldots, \frac{E_n}{\sum_{i=1}^{n+1} E_i}\right)\,,$$

où $E_1, \dots, E_{n+1}$ sont des variables aléatoires i.i.d. de loi exponentielle de paramètre 1. On obtient alors (par changement de variables),

$$R_n \sim X_{(n-1)} \sim \frac{\sum_{i=1}^{n-1} E_i}{\sum_{i=1}^{n+1} E_i} \,.$$

<u>Petit aparté sur les lois Gamma et Beta :</u>

Pour $p > 0$ et $\lambda > 0$, la loi Gamma $\Gamma(p, \lambda)$ est la loi dont la densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est donnée par

$$x \mapsto \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} \mathbb{1}_{[0,+\infty[}(x),$$

où

$$\Gamma(p) = \int_0^{+\infty} z^{p-1} e^{-z} \underset{\cdot}{z} \,.$$

Notons que la loi $\Gamma(1, \lambda)$ correspond à la loi exponentielle $\mathcal{E}(\lambda)$.

Pour $a > 0$ et $b > 0$, la loi Beta(a, b) est la loi dont la densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est donnée par

$$x \mapsto \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{[0,1]}(x),$$

où

$$B(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1} \underset{\cdot}{z} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \,.$$

Notons que la loi Beta$(1, 1)$ correspond à la loi uniforme sur $[0, 1]$.

— Si $X \sim \Gamma(p, \lambda)$ et $Y \sim \Gamma(q, \lambda)$ sont indépendantes, alors $Y + Z \sim \Gamma(p + q, \lambda)$.
  En particulier, si $E_1, \dots, E_n$ sont des variables i.i.d. de loi $\mathcal{E}(\lambda)$, alors $\sum_{i=1}^n E_i \sim \Gamma(n, \lambda)$.

— Si $X \sim \Gamma(p, \lambda)$ alors, pour $t > 0$, $tY \sim \Gamma\left(p, \frac{\lambda}{t}\right)$.

— Si $X \sim \Gamma(a, \lambda)$ et $Y \sim \Gamma(b, \lambda)$ sont indépendantes, alors $\frac{X}{X+Y} \sim$ Beta(a, b).

Ainsi, en utilisant les propriétés ci-dessus, on obtient bien $R_n \sim$ Beta$(n - 1, 2)$.

(b) Remarquons que $R_n \in [0, 1]$ p.s. Pour $\varepsilon \in]0, 1[$, on a

$$\mathbb{P}(1-R_n \geq \varepsilon) = \int_0^{1-\varepsilon} n(n-1)u^{n-2}(1-u)du = n\left[u^{n-1}\right]_0^{1-\varepsilon} - (n-1)\left[u^n\right]_0^{1-\varepsilon} = n\varepsilon(1-\varepsilon)^{n-1} + (1-\varepsilon)^n \,.$$

Ainsi $R_n \xrightarrow{\mathbb{P}} 1$, et comme $n\varepsilon(1 - \varepsilon)^{n-1} + (1 - \varepsilon)^n$ est le terme général d'une série convergente, on a même $R_n \xrightarrow{\text{p.s.}} 1$.

(c) Par ce qui précède, pour $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}(n(1-R_n) \leq x) &= \mathbb{1}_{]n,+\infty[}(x) + \mathbb{1}_{[0,n]}(x)\left(1 - x\left(1-\frac{x}{n}\right)^{n-1} - \left(1-\frac{x}{n}\right)^n\right) \\ &\underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\mathbb{R}_+}(x)\left(1 - xe^{-x} - e^{-x}\right) \,.\end{aligned}$$

La densité de la loi limite de $n(1 - R_n)$ est donc $x \mapsto xe^{-x}\mathbb{1}_{\mathbb{R}_+}(x)$ et l'on reconnaît la densité d'une loi Gamma$(2, 1)$. Cela aurait aussi pu se voir en utilisant que, par ce qui a été dit plus haut,

$$n(1-R_n) \sim \frac{E_1 + E_2}{\frac{1}{n}\sum_{i=1}^{n+1} E_i} \,.$$

Par la loi forte des grands nombres, $\frac{1}{n}\sum_{i=1}^{n+1} E_i \xrightarrow{\text{p.s.}} 1$. Donc, par le lemme de Slutsky, $n(1 - R_n) \rightsquigarrow E_1 + E_2 = \Gamma(2, 1)$.

2. (a) On a $\mathbb{E}[X_1] = \frac{\theta}{2}$, donc on peut proposer $\hat{\theta}_n = 2\overline{X}_n$.

(b) Par le TCL, on a

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightsquigarrow 2\mathcal{N}\left(0, \frac{\theta^2}{12}\right) = \mathcal{N}\left(0, \frac{\theta^2}{3}\right)$$

D'autre part, on peut remarquer que $(X_1, \ldots, X_n)$ a la même loi que $(\theta U_1, \ldots, \theta U_n)$, où $U_1, \ldots, U_n$ sont i.i.d. de loi uniforme sur $[0, 1]$. Ainsi $X_{(n)} \sim \theta U_{(n)}$. Par les mêmes arguments que ceux utilisés pour la question 1., on a

$$n(1 - U_{(n)}) \sim \frac{E_1}{\frac{1}{n}\sum_{i=1}^{n+1} E_i} \rightsquigarrow E_1 .$$

Donc $n\left(\theta - X_{(n)}\right) \rightsquigarrow \mathcal{E}(1/\theta)$. L'estimateur $X_{(n)}$ converge donc vers θ à vitesse $1/n$ alors que $\hat{\theta}_n$ converge bien plus lentement, à vitesse $1/\sqrt{n}$.

(c) Pour tout $t \in [0, 1]$,

$$\mathbb{P}_\theta\left(X_{(n)} \le t\theta\right) = t^n ,$$

Ainsi, en prenant $t = \alpha^{1/n}$, et comme $X_{(n)} \le \theta$ p.s., on obtient que l'intervalle

$$\left[X_{(n)} \, , \, \alpha^{-1/n} X_{(n)}\right]$$

est un intervalle de confiance de niveau $1 - \alpha$.

■

Exercice 10 1. Remarquons déjà que la variable $X_{(n)}$ prend ses valeurs dans $\{n, \ldots, N\}$. Pour tout $k \in \{n, \ldots, N\}$, on a

$$\mathbb{P}(X_{(n)} \le k) = \frac{\binom{k}{n}}{\binom{N}{n}} \cdot$$

Ainsi,

$$\mathbb{P}(X_n = k) = \frac{\binom{k}{n} - \binom{k-1}{n}}{\binom{N}{n}} = \frac{\binom{k-1}{n-1}}{\binom{N}{n}} \cdot$$

On a donc

$$\begin{aligned}
\mathbb{E}X_{(n)} &= \sum_{k=n}^{N} \frac{k\binom{k-1}{n-1}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=n}^{N} \frac{k!}{(n-1)!(k-n)!} \\
&= \frac{n}{\binom{N}{n}} \sum_{k=n}^{N} \binom{k}{n} = \frac{n}{\binom{N}{n}} \binom{N+1}{n+1} = \frac{n}{n+1}(N+1) \cdot
\end{aligned}$$

On en déduit l'estimateur sans biais $\frac{n+1}{n} X_{(n)} - 1$. D'autre part, on a, par échangeabilité des (X_i),

$$\mathbb{E}\overline{X}_n = \mathbb{E}X_1 = \frac{N+1}{2} ,$$

et l'on en déduit l'estimateur sans biais $2\overline{X}_n - 1$.

2. Pour tout $p \in]0, 1[$, on a

$$\mathbb{P}(X_{(n)} \le pN) = \frac{\binom{\lfloor pN \rfloor}{n}}{\binom{N}{n}} \le \left(\frac{\lfloor pN \rfloor}{N}\right)^n \le p^n .$$

Ainsi, en prenant $p = \alpha^{1/n}$, on a

$$\mathbb{P}\left(X_{(n)} \le N \le \alpha^{-1/n} X_{(n)}\right) \ge 1 - \alpha .$$

On peut aussi construire un intervalle de confiance à partir de $\overline{X}_n$. Pour cela, on peut utiliser le fait que l'inégalité de Hoeffding s'applique aussi aux sommes de variables issues de tirages sans remise. On a, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\overline{X}_n - \frac{N+1}{2}\right| > \varepsilon\right) \le 2\exp\left(-\frac{2(n\varepsilon)^2}{n(N-1)^2}\right) \le 2\exp\left(-\frac{2n\varepsilon^2}{(N+1)^2}\right) .$$

On cherche ε tel que le terme de droite soit égal à α :

$$2\exp\left(-\frac{2n\varepsilon^2}{(N+1)^2}\right) = \alpha \Leftrightarrow \varepsilon = \frac{(N+1)\sqrt{\ln\left(\frac{2}{\alpha}\right)}}{\sqrt{2n}}\,.$$

On obtient :

$$\mathbb{P}\left(\frac{N+1}{2} - \frac{(N+1)\sqrt{\ln\left(\frac{2}{\alpha}\right)}}{\sqrt{2n}} \leq \overline{X}_n \leq \frac{N+1}{2} + \frac{(N+1)\sqrt{\ln\left(\frac{2}{\alpha}\right)}}{\sqrt{2n}}\right) \geq 1-\alpha\,,$$

c'est-à-dire

$$\mathbb{P}\left(N \in \left[\frac{\overline{X}_n}{\frac{1}{2}+\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}}} - 1\,,\ \frac{\overline{X}_n}{\frac{1}{2}-\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}}} - 1\right]\right) \geq 1-\alpha\,.$$

La longueur du premier intervalle est

$$X_{(n)}\left(\alpha^{-1/n}-1\right) \leq N\left(\alpha^{-1/n}-1\right) = \frac{N\ln\left(\frac{1}{\alpha}\right)}{n} + o\left(\frac{1}{n}\right)\,,$$

tandis que celle du deuxième est

$$\overline{X}_n\left(\frac{1}{\frac{1}{2}-\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}}} - \frac{1}{\frac{1}{2}+\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}}}\right) \leq \frac{2N\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}}}{\frac{1}{4}-\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}} = 8N\sqrt{\frac{\ln\left(\frac{2}{\alpha}\right)}{2n}} + o\left(\frac{1}{\sqrt{n}}\right)\,.$$

Ainsi la longueur du premier décroît bien plus vite avec n, la taille de l'échantillon.

$\blacksquare$