Examen du cours de Probabilités

Mardi 25 Janvier 2022 (8h30-11h30)

Polycopié et notes de cours autorisés. Tout objet électronique est interdit.

Barême indicatif. L'ensemble de l'examen sera noté sur un barême entre 25 et 30 points, et les points seront approximativement répartis de la façon suivante : 1/4 pour l'exercice 1, 1/4 pour l'exercice 2 et 1/2 pour le problème. Il n'est donc pas nécessaire de tout faire pour avoir 20.

Exercice 1. Soit X une variable aléatoire à valeurs dans $[0,1]$ suivant la densité $-\ln(x)\mathbf{1}_{0<x<1}$.

1. Pour $u \in]0,1[$, calculer la densité de la loi de X^u.

Soit U une variable aléatoire indépendante de X suivant la loi uniforme sur $[0,1]$. On pose $Y = X^U$ et $Z = X^{1-U}$.

2. Calculer la loi de Y (on pourra s'aider du calcul fait à la question précédente), puis donner sans calcul celle de Z.
3. Calculer la loi de (Y,Z). Les variables aléatoires X et Y sont-elles indépendantes ?
4. On considère $E_1, E_2 \sim \mathcal{E}(1)$ indépendantes. Donner sans calcul la loi du couple $\left(E_1+E_2, \frac{E_1}{E_1+E_2}\right)$. Calculer la densité de la loi de e^{-E_1} et celle de $e^{-(E_1+E_2)}$. Puis retrouver sans calcul les résultats des questions 2 et 3.

⋆ ⋆ ⋆

Exercice 2. On considère une suite de variables aléatoires $(S_n)_{n\geq 1}$ à valeurs réelles qui converge en loi vers S. On considère une suite de variables aléatoires $(N_k)_{k\geq 1}$ à valeurs entières telles que $N_k \to +\infty$ en probabilité.

1. On suppose dans cette question uniquement que les suites S et N sont indépendantes. Montrer que pour une fonction $g : \mathbb{R} \to \mathbb{R}$ continue bornée, on a
$$\forall n \geq 1, \mathbb{E}[g(S_{N_k})] = \mathbb{E}[g(S_{N_k})\mathbf{1}_{N_k \leq n-1}] + \sum_{m=n}^{+\infty} \mathbb{E}[g(S_m)]\mathbb{P}(N_k = m).$$
Montrer alors que S_{N_k} converge en loi vers S lorsque $k \to \infty$. [Indication : on commencera par fixer $\varepsilon > 0$ et montrer $\exists n \in \mathbb{N}^*, \forall m \geq n, |\mathbb{E}[g(S_m)] - \mathbb{E}[g(S)]| \leq \varepsilon$.]

L'objectif de cet exercice est de montrer que cette propriété peut ne plus être vraie lorsque les suites S et N ne sont pas indépendantes. On considère une suite $(V_i)_{i\geq 1}$ i.i.d. telle que $\mathbb{P}(V_1 = 1) = \mathbb{P}(V_1 = -1) = 1/2$. Pour $n \geq 1$, on pose $S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n V_i$. On pose $N_0 = 0$ et puis, pour $k \geq 1$, $N_k = \inf\{n \geq N_{k-1}+1 : S_n \geq 0\}$ avec comme convention $\inf \emptyset = +\infty$.

2. Donner le comportement asymptotique de $S_n/\sqrt{n}$.
3. Montrer que S_n converge en loi vers S dont on précisera la loi. Quelle est la limite de $\mathbb{P}(S_n \geq 0)$? [Indication : On pourra utiliser que si X_n converge vers loi vers X et $\mathbb{P}(X = x) = 0$, alors, $\mathbb{P}(X_n \geq x) \to \mathbb{P}(X \geq x)$.]
4. On pose $S_0 = 0$ et on définit $\Sigma : \Omega \to \mathbb{N} \cup \{+\infty\}$ par $\Sigma = \sup_{n\geq 0}\sqrt{n}S_n$. Montrer que $\Sigma = \max(0, V_1 + \Sigma')$ où Σ' suit la même loi que Σ et est indépendante de V_1. En déduire que pour $n \in \mathbb{N}^*$, $\mathbb{P}(\Sigma = n) = \frac{1}{2}[\mathbb{P}(\Sigma = n-1) + \mathbb{P}(\Sigma = n+1)]$, puis que $\mathbb{P}(\Sigma = +\infty) = 1$.
5. En déduire avec la question précédente que $\mathbb{P}(\forall k \in \mathbb{N}, N_k < +\infty) = 1$ et $N_k \to +\infty$, p.s.
6. Que vaut $\mathbb{P}(S_{N_k} \geq 0)$? En conclure que S_{N_k} ne converge pas vers S lorsque $k \to \infty$.

Problème. (Sondages). On considère le cas d'un référendum dans un pays, c'est à dire un vote à deux issues : OUI ou NON. On suppose que la population est très grande et on souhaite faire un sondage parmi $n \in \mathbb{N}^*$ individus. On note X_i la réponse de la i-ème personne sondée : $X_i = 1$ si la réponse est OUI et $X_i = 0$ si c'est un NON. On suppose la suite $(X_i)_{i\geq 1}$ indépendante et identiquement distribuée de loi de Bernoulli de paramètre p. On note $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ la proportion de OUI.

1. Quel est le comportement asymptotique de $\bar{X}_n$ et de $\sqrt{n}(\bar{X}_n - p)$ lorsque $n \to \infty$?
2. Donner un intervalle de confiance à 95% sur la valeur de p construit uniquement à l'aide de n et $\bar{X}_n$.

On suppose que la population est divisée en deux groupes homogènes, A et B, qui représentent respectivement une proportion $\alpha \in]0,1[$ et $1-\alpha$ de la population. On suppose que le niveau d'abstention est le même dans chaque groupe, et on note $p^A \in]0,1[$ (respectivement $p^B \in]0,1[$) la probabilité qu'un individu du groupe A (resp. B) vote OUI. Ainsi, on souhaite estimer $p = \alpha p^A + (1-\alpha)p^B$, qui est la probabilité qu'un individu tiré au hasard vote OUI.

On effectue un sondage en interrogeant n^A individus du groupe A et n^B individus du groupe B. On note $(X_i^A)_{i\geq 1}$ (respectivement $(X_i^B)_{i\geq 1}$) une suite indépendante et identiquement distribuée de loi de Bernoulli de paramètre p^A (resp. p^B). Les suites $(X_i^A)_{i\geq 1}$ et $(X_i^B)_{i\geq 1}$ sont indépendantes entre elles. On note pour $n \geq 1$,

$$\bar{X}_n^A = \frac{1}{n}\sum_{i=1}^n X_i^A \text{ et } \bar{X}_n^B = \frac{1}{n}\sum_{i=1}^n X_i^B.$$

Enfin, on définit l'estimateur

$$M_n = \alpha \bar{X}^A_{\lfloor \beta n \rfloor} + (1-\alpha)\bar{X}^B_{n-\lfloor \beta n\rfloor},$$

avec $\beta \in]0,1[$, où $\lfloor x \rfloor$ désigne la partie entière de x. Ainsi, on interroge $n_A = \lfloor \beta n \rfloor$ personnes du groupe A, $n_B = n - \lfloor \beta n \rfloor$ personnes du groupe B (on suppose que n est assez grand pour que $n_A, n_B \geq 1$).

3. Donner le comportement asymptotique de M_n lorsque $n \to \infty$.
4. Montrer que $\sqrt{n}(M_n - p)$ converge en loi vers $\mathcal{N}(0, \sigma^2)$ lorsque $n \to \infty$, avec $\sigma^2 = \frac{\alpha^2}{\beta}p^A(1-p^A) + \frac{(1-\alpha)^2}{1-\beta}p^B(1-p^B)$.
5. Montrer qu'en choisissant $\beta = \alpha$, on a $\sigma^2 \leq p(1-p)$. Vaut-il mieux utiliser M_n ou $\bar{X}_n$ pour estimer p ?
6. Déterminer la valeur $\beta^* \in]0,1[$ qui minimise σ^2 pour p^A, p^B, α donnés, puis donner la valeur de $(\sigma^*)^2$ correspondante.
7. A l'aide de la formule obtenue pour β^*, construire une suite β_n^* s'exprimant en fonction de $\bar{X}_n^A$ et $\bar{X}_n^B$ telle que $\beta_n \to \beta^*$ p.s. (β_n est alors un estimateur de β^*).
8. Pour $a > 0$ et $n \in \mathbb{N}^*$ tel que $an \geq 1$, on définit

$$\hat{X}_n^A = \frac{1}{n}\sum_{i=\lfloor an\rfloor+1}^{\lfloor an\rfloor+n} X_i^A \text{ et } \hat{X}_n^B = \frac{1}{n}\sum_{i=\lfloor an\rfloor+1}^{\lfloor an\rfloor+n} X_i^B.$$

   Montrer que $\sqrt{n}(M_n^* - p)$ avec $M_n^* = \alpha \hat{X}^A_{\lfloor \beta^*_{\lfloor an\rfloor} n\rfloor} + (1-\alpha)\hat{X}^B_{n-\lfloor \beta^*_{\lfloor an\rfloor} n\rfloor}$ converge en loi vers $\mathcal{N}(0,(\sigma^*)^2)$ (on pourra utiliser le résultat de la question 1 de l'exercice 2). Cet estimateur consiste à utiliser an personnes pour estimer β^* puis a utiliser M_n. Il nécessite d'interroger $(1+a)n$ personnes et on prend en pratique a petit, par exemple $a \in]0,1/10]$.

Correction

Exercice 1.

1. Soit $f:\mathbb{R}\to\mathbb{R}$ bornée mesurable. On a
$$\mathbb{E}[f(X^u)] = \int_0^1 f(x^u)(-\ln(x))dx$$
On effectue le changement de variable $y = x^u$ qui est bijectif sur $[0,1]$. On obtient ($dy = ux^{u-1}dx = uy^{\frac{u-1}{u}}dx$)
$$\mathbb{E}[f(X^u)] = \int_0^1 f(y)(-\ln(y^{1/u}))\frac{dy}{uy^{\frac{u-1}{u}}} = \int_0^1 f(y)\frac{-y^{\frac{1}{u}-1}\ln(y)}{u^2}dy.$$
Ainsi X^u suit la loi de densité $\frac{-y^{\frac{1}{u}-1}\ln(y)}{u^2}\mathbf{1}_{]0,1[}(y)$ par le théorème de la fonction muette.
2. On a $\mathbb{E}[f(Y)] = \int_0^1\left(\int_0^1 f(x^u)(-\ln(x))dx\right)du = \int_0^1\left(\int_0^1 f(y)\frac{-y^{\frac{1}{u}-1}\ln(y)}{u^2}dy\right)du$. Par le théorème de Fubini,
$$\begin{aligned}\mathbb{E}[f(Y)] &= \int_0^1\left(\int_0^1\frac{-e^{(\frac{1}{u}-1)\ln(y)}\ln(y)}{u^2}du\right)f(y)dy\\ &= \int_0^1\left[e^{(\frac{1}{u}-1)\ln(y)}\right]_{u=0+}^{u=1}f(y)dy = \int_0^1 f(y)dy.\end{aligned}$$
Ainsi, X^U suit la loi uniforme sur $[0,1]$ par le théorème de la fontion muette. Par ailleurs, $1-U$ est suit la loi uniforme sur $[0,1]$ et est indépendante de X. Donc Z suit la même loi que X^U, c'est à dire la loi uniforme sur $[0,1]$.
3. Soit $f:\mathbb{R}^2\to\mathbb{R}$ bornée mesurable. On a
$$\mathbb{E}[f(Y,Z)] = \mathbb{E}[f(X^U, X^{1-U})] = \int_0^1\int_0^1 f(x^u, x^{1-u})(-\ln(x))dxdu.$$
On considère le changement de variable $\varphi :]0,1[^2 \to]0,1[^2$ défini par $\varphi(x,u) = (x^u, x^{1-u})$. Clairement, $(x^u, x^{1-u}) \in]0,1[^2$ si $(x,u)\in]0,1[^2$. Soient $(y,z)\in]0,1[^2$ tels que $y = x^u$ et $z = x^{1-u}$. Alors, $x = yz$ et $u = \frac{\ln(y)}{\ln(x)} = \frac{-\ln(y)}{-\ln(y)-\ln(z)}$. Ainsi $\varphi^{-1}(y,z) = \left(yz, \frac{-\ln(y)}{-\ln(y)-\ln(z)}\right)$, et on a clairement $\varphi^{-1}(]0,1[^2)\subset]0,1[^2$ ce qui prouve que φ est bijective. C'est un C^1-difféomorphisme, et
$$Jac(\varphi^{-1})(y,z) = \det\begin{pmatrix} z & y\\ \frac{\ln(z)}{y(\ln(y)+\ln(z))^2} & \frac{-\ln(y)}{z(\ln(y)+\ln(z))^2}\end{pmatrix} = \frac{1}{-\ln(y)-\ln(z)}.$$
Grâce au théorème de changement de variable, on obtient
$$\mathbb{E}[f(Y,Z)] = \int_0^1\int_0^1 f(y,z)(-\ln(yz))\frac{1}{-\ln(y)-\ln(z)}dydz = \int_0^1\int_0^1 f(y,z)dydz.$$
Cela prouve par le théorème de la fonction muette que Y et Z sont deux variables aléatoires uniformes sur $[0,1]$ indépendantes.

4. La proposition du cours sur les lois Gamma assure que $E_1+E_2 \sim \Gamma(2,1)$ est indépendante de $\frac{E_1}{E_1+E_2} \sim \mathcal{U}([0,1])$. On sait par le cours (chapitre Simulation) que $-\ln(U)$ a même loi que E_1 et donc e^{-E_1} a même loi que $U \sim \mathcal{U}([0,1])$. On calcule maintenant la loi de $e^{-(E_1+E_2)}$. Soit $f : \mathbb{R} \to \mathbb{R}$ bornée mesurable. On a, en utilisant la densité de la loi $\Gamma(2,1)$ et le changement de variable $y = e^{-x}$

$$\mathbb{E}[f(e^{-(E_1+E_2)})] = \int_0^\infty f(e^{-x}) x e^{-x} dx = \int_0^1 f(y)(-\ln(y)) dy.$$

Ainsi, $e^{-(E_1+E_2)}$ suit la même loi que X et est indépendante de $V = \frac{E_1}{E_1+E_2}$. Il vient $(e^{-(E_1+E_2)}, V) \overset{loi}{=} (X, U)$ puis

$$(X^U, X^{1-U}) \overset{loi}{=} (e^{-V(E_1+E_2)}, e^{-(1-V)(E_1+E_2)}) = (e^{-E_1}, e^{-E_2}),$$

ce qui prouve que (X^U, X^{1-U}) est un couple indépendant de lois uniformes sur $[0,1]$.

⋆ ⋆ ⋆

Exercice 2.

1. On a $g(S_{N_k}) = g(S_{N_k})\mathbf{1}_{N_k \le n-1} + \sum_{m=n}^\infty g(S_m)\mathbf{1}_{N_k=m}$. En utilisant le théorème de Fubini (g est bornée) et l'indépendance, on obtient l'égalité voulue.

   Considérons $\varepsilon > 0$. On a d'une part :

$$|\mathbb{E}[g(S_{N_k})\mathbf{1}_{N_k \le n-1}]| \le \mathbb{E}[|g(S_{N_k})|\mathbf{1}_{N_k \le n-1}] \le \|g\|_\infty \mathbb{P}(N_k \le n-1) \underset{k\to+\infty}{\to} 0,$$

   et donc $|\mathbb{E}[g(S_{N_k})\mathbf{1}_{N_k \le n-1}]| \le \varepsilon$ pour k suffisamment grand. D'autre part, $\mathbb{E}[g(S_m)] \to \mathbb{E}[g(S)]$ grâce à la convergence en loi. Ainsi, il existe n tel que pour $m \ge n$, $|\mathbb{E}[g(S_m)] - \mathbb{E}[g(S)]| \le \varepsilon$. Ainsi, on a par l'inégalité triangulaire

$$\left|\sum_{m=n}^\infty \mathbb{E}[g(S_m)]\mathbb{P}(N_k = m) - \mathbb{E}[g(S)]\mathbb{P}(N_k \ge n)\right| = \left|\sum_{m=n}^\infty (\mathbb{E}[g(S_m)] - \mathbb{E}[g(S)])\mathbb{P}(N_k = m)\right| \le \varepsilon.$$

   Par ailleurs, comme g est bornée et N_k converge en probabilité vers $+\infty$, $|\mathbb{E}[g(S)] - \mathbb{E}[g(S)]\mathbb{P}(N_k \ge n)| \le \varepsilon$ pour k suffisament grand. Cela montre que $|\mathbb{E}[g(S_{N_k})] - \mathbb{E}[g(S)]| \le 3\varepsilon$ pour k suffisamment grand, ce qui est le résultat voulu.

2. La variable V_1 est bornée donc intégrable. La loi forte des grands nombres assure que $S_n/\sqrt{n}$ converge p.s. vers $\mathbb{E}[V_1] = 0$.

3. On peut appliquer le TCL ($\mathbb{E}[V_1^2] = 1 < \infty$) et S_n converge en loi vers $S \sim \mathcal{N}(0,1)$. Comme $\mathbb{P}(S = 0) = 0$, on a $\mathbb{P}(S_n \ge 0) \to \mathbb{P}(S \ge 0) = 1/2$.

4. On a $\Sigma = \sup_{n\in\mathbb{N}} \sum_{i=1}^n V_i = \max(0, V_1 + \sup_{n\ge1} \sum_{i=2}^n V_i)$ avec pour convention $\sum_{i=n}^{n'}(\dots) = 0$ si $n' < n$. Comme $\Sigma' = \sup_{n\ge1} \sum_{i=2}^n V_i = \sup_{n\ge0} \sum_{i=2}^{n+1} V_i = \sup_{n\ge0} \sum_{i=1}^n V_{i+1}$ et que $(V_{i+1})_{i\ge1}$ a même loi que $(V_i)_{i\ge1}$ et est indépendante de V_1, on a l'égalité en loi demandée. Pour $n \in \mathbb{N}^*$, on a donc $\mathbb{P}(\Sigma = n) = \mathbb{P}(V_1 + \Sigma' = n) = \frac{1}{2}[\mathbb{P}(\Sigma = n-1) + \mathbb{P}(\Sigma = n+1)]$. Il vient $\mathbb{P}(\Sigma = n) - \mathbb{P}(\Sigma = n-1) = \mathbb{P}(\Sigma = n+1) - \mathbb{P}(\Sigma = n) = cte$ puis $\mathbb{P}(\Sigma = n) = \mathbb{P}(\Sigma = 0) + cte \times n$. Par positivité des probabilités, on a forcément $cte \ge 0$, et $\sum_{n\in\mathbb{N}} \mathbb{P}(\Sigma = n) \le 1$ implique que $\mathbb{P}(\Sigma = 0) = cte = 0$ et donc $\mathbb{P}(\Sigma = +\infty) = 1$.

5. Si $(u_n)_{n\in\mathbb{N}}$ est une suite telle que $u_n \in$ et $\sup_{n\in\mathbb{N}} u_n = +\infty$, alors la suite $n_0 = 0$, $n_{k+1} = \inf\{n \geq n_k + 1, u_n \geq 0\}$ est bien définie avec $n_k < \infty$ pour tout $k \geq 0$: sinon il existerait $K \in \mathbb{N}$ tel que $n_K < \infty$ et $n_{K+1} = +\infty$ ce qui implique que $\sup_{n\in\mathbb{N}} u_n = \sup_{n\leq N_k} u_n < \infty$. Comme $\mathbb{P}(\Sigma = +\infty) = 1$, on en déduit que, presque sûrement, la suite (N_k) est telle que $\forall k, N_k < \infty$. Comme $N_{k+1} \geq N_k + 1$, On a $N_k \to \infty$, p.s.
6. Par construction $S_{N_k} \geq 0$, donc $\mathbb{P}(S_{N_k} \geq 0) = 1$. Si S_{N_k} convergeait en loi vers S, on aurait comme précédemment $\mathbb{P}(S_{N_k} \geq 0) \to 1/2$, ce qui n'est pas le cas.

⋆⋆⋆

Problème.

1. On a $\bar{X}_n \to p$ presque sûrement par la loi forte des grands nombres et $\sqrt{n}(\bar{X}_n - p) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p(1-p))$ par le théorème de la limite centrale.
2. En utilisant le théorème de Slutsky, on a $\sqrt{\frac{n}{\bar{X}_n(1-\bar{X}_n)}}(\bar{X}_n - p) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ et donc
$$\left[\bar{X}_n - 1,96\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + 1,96\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\right]$$
est un intervalle de confiance asymptotique pour p à 95%.
3. En utilisant la loi forte des grand nombres, il vient que $\bar{X}_n^A \to p^A$ et $\bar{X}_n^B \to p^B$, presque sûrement. Comme $\beta \in]0,1[$, on a $\lfloor \beta n \rfloor \to \infty$ et $n - \lfloor \beta n \rfloor \to \infty$ lorsque $n \to \infty$. Ainsi, M_n converge presque sûrement vers $\alpha p^A + (1-\alpha)p^B = p$.
4. On a
$$\sqrt{n}(M_n - p) = \frac{\alpha}{\sqrt{\beta}}\sqrt{\beta n}(\bar{X}^A_{\lfloor \beta n \rfloor} - p^A) + \frac{1-\alpha}{\sqrt{1-\beta}}\sqrt{(1-\beta)n}(\bar{X}^B_{n-\lfloor \beta n \rfloor} - p^B).$$
Le théorème de la limite centrale assure que $\sqrt{\lfloor \beta n \rfloor}(\bar{X}^A_{\lfloor \beta n \rfloor} - p^A) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p^A(1-p^A))$ et $\sqrt{n - \lfloor \beta n \rfloor}(\bar{X}^B_{n-\lfloor \beta n \rfloor} - p^B) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p^B(1-p^B))$. Comme $\frac{\lfloor \beta n \rfloor}{\beta n} \to 1$ et $\frac{n-\lfloor \beta n \rfloor}{(1-\beta)n} \to 1$, on en déduit que $\sqrt{\beta n}(\bar{X}^A_{\lfloor \beta n \rfloor} - p^A) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p^A(1-p^A))$ et $\sqrt{(1-\beta)n}(\bar{X}^B_{n-\lfloor \beta n \rfloor} - p^B) \xrightarrow{\mathcal{L}} (0, p^B(1-p^B))$. Les suites étant indépendantes, il vient que le couple converge en loi vers un vecteur gaussien centré de matrice de covariance $\begin{bmatrix} p^A(1-p^A) & 0 \\ 0 & p^B(1-p^B) \end{bmatrix}$. Par conséquent, $\sqrt{n}(M_n - p)$ converge en loi vers $\mathcal{N}(0, \sigma^2)$ avec
$$\sigma^2 = \frac{\alpha^2}{\beta}p^A(1-p^A) + \frac{(1-\alpha)^2}{1-\beta}p^B(1-p^B).$$
5. En prenant $\beta = \alpha$, on a $\sigma^2 = \alpha p^A(1-p^A) + (1-\alpha)p^B(1-p^B)$. La fonction $x \mapsto x(1-x)$ étant concave, il vient que
$$p(1-p) \geq \alpha p^A(1-p^A) + (1-\alpha)p^B(1-p^B).$$
L'estimateur M_n a une variance asymptotique inférieure à celle de $\bar{X}_n$, il faut donc le préférer.

6. On a $\sigma^2 \to +\infty$ lorsque $\beta \to 0$ ou $\beta \to 1$, ce qui prouve l'existence d'un minimum β^*. Tout point critique satisfait

$$-\frac{\alpha^2}{\beta^2} p^A(1-p^A) + \frac{(1-\alpha)^2}{(1-\beta)^2} p^B(1-p^B) = 0.$$

Cela donne $\frac{\beta^*}{1-\beta^*} = \frac{\alpha}{1-\alpha}\sqrt{\frac{p^A(1-p^A)}{p^B(1-p^B)}}$ puis $\beta^* = \frac{\frac{\alpha}{1-\alpha}\sqrt{\frac{p^A(1-p^A)}{p^B(1-p^B)}}}{1+\frac{\alpha}{1-\alpha}\sqrt{\frac{p^A(1-p^A)}{p^B(1-p^B)}}} = \frac{\alpha\sqrt{p^A(1-p^A)}}{\alpha\sqrt{p^A(1-p^A)}+(1-\alpha)\sqrt{p^B(1-p^B)}}$ et on a

$$(\sigma^*)^2 = \left(\alpha\sqrt{p^A(1-p^A)} + (1-\alpha)\sqrt{p^B(1-p^B)}\right)^2.$$

7. La loi forte des grands nombre assure que $(\bar{X}_n^A, \bar{X}_n^B) \to (p^A, p^B)$, presque sûrement lorsque $n \to \infty$. La fonction $(p^A, p^B) \to \frac{\alpha\sqrt{p^A(1-p^A)}}{\alpha\sqrt{p^A(1-p^A)}+(1-\alpha)\sqrt{p^B(1-p^B)}}$ étant continue sur $]0,1[^2$, on en déduit que

$$\beta_n = \frac{\alpha\sqrt{\bar{X}_n^A(1-\bar{X}_n^A)}}{\alpha\sqrt{\bar{X}_n^A(1-\bar{X}_n^A)} + (1-\alpha)\sqrt{\bar{X}_n^B(1-\bar{X}_n^B)}}$$

converge presque sûrement vers β^*.

8. On a $M_n^* - p = \alpha(\hat{X}^A_{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^A) + (1-\alpha)(\hat{X}^B_{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^B)$ et donc

$$\sqrt{n}(M_n^* - p) = \frac{\alpha\sqrt{n}}{\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}}\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}(\hat{X}^A_{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^A) + \frac{(1-\alpha)\sqrt{n}}{\sqrt{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}}\sqrt{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}(\hat{X}^B_{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^B).$$

Par le théorème de la limite centrale, $\sqrt{n}(\bar{X}_n^A - p^A)$ converge en loi vers $\mathcal{N}(0, p^A(1-p^A))$ et comme $\sqrt{n}(\hat{X}_n^A - p^A)$ a même loi que $\sqrt{n}(\bar{X}_n^A - p^A)$, il vient que $\sqrt{n}(\hat{X}_n^A - p^A)$ converge en loi vers $\mathcal{N}(0, p^A(1-p^A))$. La suite $\beta_{\lfloor an \rfloor}$ étant indépendante de $\hat{X}_n^A$, il vient en utilisant la question 1 de l'exercice 2 que $\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}(\hat{X}^A_{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^A)$ converge en loi vers $\mathcal{N}(0, p^A(1-p^A))$. Comme $\frac{\alpha\sqrt{n}}{\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}} \to \frac{\alpha}{\sqrt{\beta^*}}$, le théorème de Slutsky assure que $\frac{\alpha\sqrt{n}}{\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}}\sqrt{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}(\hat{X}^A_{\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^A)$ converge en loi vers $\mathcal{N}(0, \frac{\alpha^2}{\beta^*}p^A(1-p^A))$. De la même façon, on montre que $\frac{(1-\alpha)\sqrt{n}}{\sqrt{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}}\sqrt{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor}(\hat{X}^B_{n-\lfloor \beta^*_{\lfloor an \rfloor} n\rfloor} - p^B)$ converge en loi vers $\mathcal{N}(0, \frac{(1-\alpha)^2}{1-\beta^*}p^B(1-p^B))$. Comme ces deux suites sont indépendantes, il vient que $\sqrt{n}(M_n^* - p)$ converge en loi vers $\mathcal{N}(0, \frac{\alpha^2}{\beta^*}p^A(1-p^A) + \frac{(1-\alpha)^2}{1-\beta^*}p^B(1-p^B)) = \mathcal{N}(0, (\sigma^*)^2)$.

⋆⋆⋆