IUFM de Midi-Pyrénées
Préparation au CAPES de Maths
2000-2001

Correction de l'épreuve blanche du 22 décembre 2000

I) Quelques propriétés des fonctions génératrices

a) On a $\sum_{i=0}^{+\infty} \mathbb{P}(Z=i)=1$ donc si $|s|\leq 1$ la série entière définissant G_Z est absolument convergente. Cette série est donc infiniment dérivable sur $]-1,1[$. De plus pour $s\in]-1,1[$ et $k\geq 0$

$$G_Z^{(k)}(s)=\sum_{i=k}^{+\infty}(i-k+1)\dots(i-1)i\mathbb{P}(Z=i)s^{i-k}. \tag{1}$$

En particulier $G_Z^{(k)}(0)=k!\mathbb{P}(Z=k)$ donc $\mathbb{P}(Z=k)=\frac{G_Z^{(k)}(0)}{k!}$.

b) Si $R>1$, l'équation (1) et vraie pour $s\in]-R,R[$ et $k\geq 0$. En particulier lorsque $s=1$ on obtient

$$G_Z^{(l)}(1)=\sum_{i=k}^{+\infty} i(i-1)\cdots(i-l+1)\mathbb{P}(Z=i)=\mathbb{E}[Z(Z-1)\cdots(Z-l+1)].$$

c) Puisque Z_1 et Z_2 sont des variables aléatoires indépendantes, on a pour $s\in[-1,1]$

$$G_{Z_1+Z_2}(s)=\mathbb{E}[s^{Z_1+Z_2}]=\mathbb{E}[s^{Z_1}s^{Z_2}]=\mathbb{E}[s^{Z_1}]\mathbb{E}[s^{Z_2}]=G_{Z_1}(s)G_{Z_2}(s).$$

d) On a pour $|s|<\theta^{-1}$

$$G_{Z_0}(s)=\sum_{k=0}^{+\infty}\theta^k(1-\theta)s^k=\frac{1-\theta}{1-\theta s}$$

et donc pour $l\in\mathbb{N}$

$$G_{Z_0}^{(l)}(s)=\frac{l!\theta^l(1-\theta)}{(1-\theta s)^{l+1}}.$$

Donc, puisque $R=\theta^{-1}>1$, en utilisant les résultats de la question b) on trouve

$$\mathbb{E}[Z_0(Z_0-1)\cdots(Z_0-l+1)]=G_{Z_0}^{(l)}(1)=\frac{l!\theta^l}{(1-\theta)^l}.$$

D'autre part d'après c) on a

$$G_{Z_0+\dots+Z_n}(s)=(G_{Z_0}(s))^{n+1}=\left(\frac{1-\theta}{1-\theta s}\right)^{n+1}. \tag{2}$$

Un calcul direct donne

$$G''_{Z_0+\dots+Z_n}(s)=\frac{(n+1)(n+2)\theta^2(1-\theta)^{n+1}}{(1-\theta s)^{n+3}}$$

et donc a) donne finalement

$$P(Z_0+\cdots+Z_n=2)=\frac{(n+1)(n+2)\theta^2(1-\theta)^{n+1}}{2}.$$

e) Convergence en loi et fonction génératrice.

i) On a vu que, si Z est une variable aléatoire à valeurs dans $\mathbb{N}$ alors $\mathbb{P}(Z=0)=G_Z(0)$. Donc

$$\lim_{n\to\infty} \mathbb{P}(U_n=0)=\lim_{n\to\infty} G_{U_n}(0)=G_U(0)=\mathbb{P}(U=0).$$

ii) Montrons par récurrence sur l que

$$H_{n,l}(s)=\sum_{j\geq l} \mathbb{P}(U_n=j)s^{j-l} \quad (s\in]-\varepsilon,\varepsilon[\setminus\{0\},\ n\in\mathbb{N}). \tag{3}$$

Remarquons tout d'abord que l'égalité (3) est vérifiée pour $l=0$. Supposons maintenant que cette égalité soit vérifiée pour $l=0,\ldots,L$ avec $L\geq 0$. On a alors, par construction pour $s\in]-\varepsilon,\varepsilon[\setminus\{0\}$,

$$H_{n,L+1}(s)=\frac{\sum_{j\geq L}\mathbb{P}(U_n=j)s^{j-L}-\mathbb{P}(U_n=L)}{s}=\sum_{j\geq L+1}\mathbb{P}(U_n=j)s^{j-L-1}.$$

Donc l'égalité (3) est vraie au rang $L+1$, elle est donc vraie sur $\mathbb{N}$. D'autre part on obtient facilement que

$$\lim_{s\to 0} H_{n,l}(s)=\mathbb{P}(U_n=l) \quad (l\geq 0,\ n\geq 0).$$

La série entière $H_{n,l}$ est donc bien définie en 0. En utilisant les mêmes arguments, on obtient pour $s\in]-\varepsilon,\varepsilon[\setminus\{0\}$

$$H_l(s)=\sum_{j\geq l}\mathbb{P}(U=j)s^{j-l} \text{ et } \lim_{s\to 0} H_l(s)=\mathbb{P}(U=l) \quad (l\geq 0).$$

Montrons maintenant par récurrence sur l'entier naturel l que pour $s\in]-\varepsilon,\varepsilon[$

$$\lim_{n\to\infty} H_{n,l}(s)=H_l(s). \tag{4}$$

Par hypothèse, l'égalité (4) est vraie pour $l=0$. Supposons qu'elle soit vraie pour $l=0,\ldots,L$ avec $L\geq 0$. On a alors au rang $L+1$ pour $s\in]-\varepsilon,\varepsilon[\setminus\{0\}$ et $n\geq 1$

$$H_{n,L+1}(s)=\frac{H_{n,L}(s)-\mathbb{P}(U_n=L)}{s}.$$

En utilisant l'hypothèse de récurrence et le fait que $H_{n,L}(0)=\mathbb{P}(U_n=L)$ on obtient

$$\lim_{n\to\infty} H_{n,L+1}(s)=\frac{H_L(s)-\mathbb{P}(U=L)}{s}=H_{L+1}(s).$$

Pour terminer la preuve par récurrence, il reste à montrer que

$$\lim_{n\to\infty} H_{n,L+1}(0)=H_{L+1}(0). \tag{5}$$

On a pour $|s| < \min(\varepsilon, 1/2)$

$$\begin{aligned} |H_{n,L+1}(0) - H_{L+1}(0)| &\leq |H_{n,L+1}(0) - H_{n,L+1}(s)| + |H_{n,L+1}(s) - H_{L+1}(s)| \\ &\quad + |H_{L+1}(s) - H_{L+1}(0)| \\ &\leq 4|s| + |H_{n,L+1}(s) - H_{L+1}(s)|. \end{aligned}$$

Soit $\eta > 0$, choisisons $s_0 < \min(\varepsilon, \eta/8, 1/2)$. Si n est assez grand, la quantité $|H_{n,L+1}(s_0) - H_{L+1}(s_0)|$ est majorée par $\eta/2$. Donc si n est assez grand $|H_{n,L+1}(0) - H_{L+1}(0)|$ est majoré par η. Puisque $\eta > 0$ est arbitraire on en déduit (5). La preuve précédente utilise principalement le fait que la suite de fonction $(H_{n,L+1})_{n\in\mathbb{N}}$ est équicontinue dans un voisinage de l'origine.

iii) Une conséquence directe de ii) est, pour $l \geq 0$,

$$\lim_{n\to\infty} H_{n,l}(0) = \lim_{n\to\infty} \mathbb{P}(U_n = l) = H_l(0) = \mathbb{P}(U = l).$$

Donc, la suite (U_n) converge en loi vers U.

f) En utilisant l'identité (2) avec $n-1$ et $\frac{\theta}{n}$ on obtient pour $s \in]-1, 1[$

$$G_{Z_1^{(n)}+\cdots+Z_n^{(n)}}(s) = \left(\frac{1-\frac{\theta}{n}}{1-\frac{\theta s}{n}}\right)^n.$$

Donc en utilisant la limite bien connue :

$$\lim_{n\to\infty}(1-\frac{x}{n})^n = e^{-x} \quad (x \in \mathbb{R}),$$

on obtient

$$\lim_{n\to\infty} G_{Z_1^{(n)}+\cdots+Z_n^{(n)}}(s) = e^{-\theta}e^{\theta s} = e^{\theta(s-1)}.$$

On reconnaît la fonction génératrice de la loi de Poisson de paramètre θ. Le résultat de la question f) nous permet de conclure que $Z_1^{(n)} + \cdots + Z_n^{(n)}$ converge en loi vers la loi de Poisson de paramètre θ.

II) Etude du régime stationnaire de la file

a) Quelques géneralités

i) Pour $n \geq 0$, il y a deux cas à considérer :

1) Sur $\{X_n \geq 1\}$ entre les instants n et $n+1$ la taille de la file augmente de un avec probabilité p ou diminue de un avec probabilité $1-p$. Donc sur cet événement on a

$$X_{n+1} = X_n + \varepsilon_n = (X_n + \varepsilon_n)^+, \quad \text{où} \quad \mathbb{P}(\varepsilon_n = 1) = 1 - \mathbb{P}(\varepsilon_n = -1) = p.$$

2) Sur $\{X_n = 0\}$ la file s'incrémente de un avec probabilité p ou reste vide avec probabilité $1-p$. Donc sur cet événement on peut aussi écrire $X_{n+1} = (X_n + \varepsilon_n)^+$.

Donc, dans le cas général on a bien

$$X_{n+1} = (X_n + \varepsilon_n)^+, \quad (n \in \mathbb{N}). \tag{6}$$

Où $(\varepsilon)_{n\in\mathbb{N}}$ est une suite de variables aléatoires identiquement distribuées à valeurs dans $\{-1, 1\}$ avec $\mathbb{P}(\varepsilon_1 = 1) = p$. De plus, puisque les arrivées et départs sont tous indépendants et sont aussi indépendants de X_0 la suite $(\varepsilon)_{n\in\mathbb{N}}$ est une suite i.i.d. indépendante de X_0.

ii) On a $X_1 = (X_0 + \varepsilon_0)^+ = (X_0 + \varepsilon_0)\mathbb{1}_{\{X_0+\varepsilon_0>0\}}$. On a aussi

$$S_0 - \inf_{-1\le j\le 0} S_j = X_0 + \varepsilon_0 - \inf(0, X_0 + \varepsilon_0) = (X_0 + \varepsilon_0)\mathbb{1}_{\{X_0+\varepsilon_0>0\}} = X_1.$$

L'égalité est donc établie pour $n = 0$. Soit $n \ge 0$, supposons maintenant que pour tout $k = 0, \ldots, n$ on ait

$$X_{k+1} = S_k - \inf_{-1\le j\le k} S_j. \tag{7}$$

Alors

$$\begin{aligned}
X_{n+2} = (X_{n+1} + \varepsilon_{n+1})^+ &= \max(X_{n+1} + \varepsilon_{n+1}, 0) \\
&= \max(S_{n+1} - \inf_{-1\le j\le n} S_j, S_{n+1} - S_{n+1}) \\
&= \max(\max_{-1\le j\le n}(S_{n+1} - S_j), S_{n+1} - S_{n+1}) \\
&= \max_{-1\le j\le n+1}(S_{n+1} - S_j) = S_{n+1} - \inf_{-1\le j\le n+1} S_j.
\end{aligned}$$

L'égalité (7) est vraie pour $k = n + 2$ et donc par récurrence elle est vraie pour tout $k \in \mathbb{N}$.

b) On suppose $p > \frac{1}{2}$.

i) On a par linéarité de l'espérance

$$\mathbb{E}(S_n) = \mathbb{E}(X_0) + \mathbb{E}\left(\sum_{j=0}^{n} \varepsilon_j\right) = m + \sum_{j=0}^{n} \mathbb{E}(\varepsilon_j) = m + (n+1)\mathbb{E}(\varepsilon_0).$$

D'autre part $\mathbb{E}(\varepsilon_0) = p-(1-p) = 2p-1$ donc $\mathbb{E}(S_n) = m+(n+1)(2p-1)$. Le calcul de la variance de S_n s'effectue en utilisant l'indépendance des variables et l'égalité var $\varepsilon_0 = 1 - (2p-1)^2 = 4p(1-p)$:

$$\text{var } S_n = \text{var } X_0 + \text{var } \left(\sum_{j=0}^{n} \varepsilon_j\right) = \sigma^2 + n\text{var } \varepsilon_0 = \sigma^2 + 4np(1-p).$$

D'après les calculs précédents, on a

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{S_n}{n}\right) = 2p - 1 \quad \text{et} \quad \lim_{n\to\infty} \text{var } \left(\frac{S_n}{n}\right) = 0.$$

Donc, puisque

$$\mathbb{E}\left[\left(\frac{S_n}{n}-[2p-1]\right)^2\right] = \text{var }\left(\frac{S_n}{n}\right)+\left[\mathbb{E}\left(\frac{S_n}{n}\right)-(2p-1)\right]^2,$$

en utilisant l'inégalité de Tchebycheff, on peut conclure que $\frac{S_n}{n}$ converge en probabilité vers $2p-1$. Soit $\epsilon>0$, avec $2p-1-\epsilon>0$ on a donc

$$\lim_{n\to\infty}\mathbb{P}\left[S_n<n(2p-1-\epsilon)\right] \leq \lim_{n\to\infty}\mathbb{P}\left[\left|\frac{S_n}{n}-(2p-1)\right|>\epsilon\right]=0.$$

Cela entraîne que S_n diverge en probabilité vers $+\infty$.

ii) Puisque $S_{-1}=0$ on a pour $n\geq 1$

$$X_{n+1}=S_n-\inf_{-1\leq j\leq n}S_j \geq S_n-S_{-1}=S_n. \tag{8}$$

En utilisant le résultat de la question précédente on en déduit que suite (X_n) diverge en probabilité vers $+\infty$. On peut interpréter ce dernier résultat de façon intuitive. Si le taux d'arrivée est plus important que le taux de départ la taille de la file croît indéfiniment.

c) On suppose que $p<\frac{1}{2}$.

i) Par construction on a, pour $n,x\in\mathbb{N}$,

$$\mathbb{P}(X_{n+1}=x+1|X_n=x)=p \ \text{ et } \ \mathbb{P}(X_{n+1}=(x-1)^+|X_n=x)=1-p.$$

La formule des probabilités totales donne alors, pour $n\in\mathbb{N}$,

$$\begin{aligned}\mathbb{P}(X_{n+1}=x) &= \mathbb{P}(X_{n+1}=x|X_n=x-1)\mathbb{P}(X_n=x-1)\\ &\quad +\mathbb{P}(X_{n+1}=x|X_n=x+1)\mathbb{P}(X_n=x+1)\\ &= p\mathbb{P}(X_n=x-1)+q\mathbb{P}(X_n=x+1) \quad (x\in\mathbb{N}^*), \qquad (9)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(X_{n+1}=0) &= \mathbb{P}(X_{n+1}=0|X_n=0)\mathbb{P}(X_n=0)\\ &\quad +\mathbb{P}(X_{n+1}=0|X_n=1)\mathbb{P}(X_n=1)\\ &= q\left[\mathbb{P}(X_n=0)+\mathbb{P}(X_n=1)\right]. \qquad (10)\end{aligned}$$

ii) Utilisant la question précédente, On a pour, $s\in]-1,1[\setminus\{0\}$ et $n\in\mathbb{N}$

$$\begin{aligned}G_{X_{n+1}}(s) &= \sum_{k=0}^{+\infty}s^k\mathbb{P}(X_{n+1}=k) \qquad (11)\\ &= q\left[\mathbb{P}(X_n=0)+\mathbb{P}(X_n=1)\right]\\ &\quad +\sum_{k=0}^{+\infty}s^k(p\mathbb{P}(X_n=x-1)+q\mathbb{P}(X_n=x+1))\\ &= q\left[\mathbb{P}(X_n=0)+\mathbb{P}(X_n=1)\right]\\ &\quad +psG_{X_n}(s)+\frac{q}{s}\left[G_{X_n}(s)-\mathbb{P}(X_n=0)-\mathbb{P}(X_n=1))\right]\\ &= qG_{X_n}(0)(1-\frac{1}{s})+(ps+\frac{q}{s})G_{X_n}(s). \qquad (12)\end{aligned}$$

En utilisant la question I.a) et l'égalité (10) on obtient

$$G_{X_{n+1}}(0) = \mathbb{P}(X_n = 0) = q\left[\mathbb{P}(X_n = 0) + \mathbb{P}(X_n = 1)\right] = q\left[G_{X_n}(0) + G'_{X_n}(0)\right].$$

D'autre part en utilisant (12) on a

$$\lim_{s\to 0} G_{X_{n+1}}(s) = \lim_{s\to 0}\left[qG_{X_n}(0)(1-\frac{1}{s}) + (ps+\frac{q}{s})G_{X_n}(s)\right] = q\left[G_{X_n}(0) + G'_{X_n}(0)\right].$$

iii) D'après la question I.d), on a pour $|s| < q/p$

$$G_{X_0}(s) = \frac{q-p}{q-ps} = \frac{1-2p}{q-ps}.$$

L'équation de récurrence (12) donne alors pour $s \neq 0$

$$\begin{aligned} G_{X_1}(s) &= qG_{X_0}(0)(1-\frac{1}{s}) + (ps+\frac{q}{s})G_{X_0}(s) \\ &= (q-p)(1-\frac{1}{s}) + (ps+\frac{q}{s})\frac{q-p}{q-ps} \\ &= \frac{1-2p}{s}\left[(s-1) + \frac{ps^2+q}{q-ps}\right] \\ &= \frac{1-2p}{q-ps} = G_{X_0}(s). \end{aligned}$$

La fonction génératrice caractérisant la loi, on en déduit que X_1 suit aussi la loi géométrique de paramètre $\frac{p}{q}$. Par récurrence, il en est de même pour X_n ($n \in \mathbb{N}$).

iv) Soit $s \in]-\varepsilon, \varepsilon[\setminus\{0\}$ fixé. Posons, $l(s) = \lim_{n\to\infty} G_{X_n}(s)$. L'équation de récurrence (12) entraîne que

$$l(s) = ql(0)(1-\frac{1}{s}) + (ps+\frac{q}{s})l(s).$$

Donc,

$$l(s) = \frac{ql(0)}{q-ps}.$$

D'autre part, puisque l'on suppose que $\varepsilon > 1$ on a $l(1) = \lim_{n\to\infty} G_{X_n}(1) = 1$. D'où $l(0) = (1-2p)/q$ et

$$l(s) = \frac{1-2p}{q-ps}.$$

On reconnaît la fonction génératrice de la loi géométrique de paramètre $\frac{p}{q}$. En utilisant le résultat de la question I.e) on peut conclure que la suite (X_n) converge en loi vers la loi géométrique de paramètre $\frac{p}{q}$. En fait, on peut montrer que cette convergence a lieu sans faire d'hypothèse sur la suite des fonctions génératrices. On dit que la loi géométrique est la loi stationnaire de la file. Quelle que soit la répartition de la file à l'instant initial, quand le temps grandit, la loi du nombre d'individus dans la queue tend vers la loi géométrique de paramètre $\frac{p}{q}$.

III) Estimation de p

a) On a, en utilisant la question I.d),

$$\mathbb{E}(\eta_N) = \frac{1}{N}\mathbb{E}\left(\sum_{k=0}^{N} X_j^k\right) = \mathbb{E}(X_j^1) = \frac{p}{q-p} = \frac{p}{1-2p}.$$

L'indépendance des variables aléatoires $(X_j^k)_{k=1,\dots N}$ donne

$$\text{var }(\eta_N) = \frac{1}{N^2}\text{var }\left(\sum_{k=0}^{N} X_j^k\right) = \frac{\text{var } X_j^1}{N}.$$

D'autre part, toujours en utilisant les résultats de I.d), on obtient

$$\text{var } X_j^1 = \frac{2!\left(\frac{p}{q}\right)^2}{\left(1-\frac{p}{q}\right)^2} + \frac{\frac{p}{q}}{1-\frac{p}{q}} - \left(\frac{\frac{p}{q}}{1-\frac{p}{q}}\right)^2 = \frac{p(1-p)}{(1-2p)^2}.$$

Donc

$$\text{var }(\eta_N) = \frac{p(1-p)}{N(1-2p)^2}.$$

Puisque les variables $(X_j^k)_{k=1,\dots N}$ admettent une espérance, la loi faible des grands nombre permet d'affirmer que η_N converge en probabilité vers $\mathbb{E}(X_j^1) = \frac{p}{1-2p}$. Donc, on a

$$f(p) = \frac{p}{1-2p}, \quad \left(p \in \left]0, \frac{1}{2}\right[\right).$$

Les variables $(X_j^k)_{k=1,\dots N}$ possèdent un moment d'ordre 2 fini. Le théorème de la limite central donne donc la convergence en loi de $\sqrt{N}(\eta_N - f(p))$ vers la loi gaussienne centrée de variance var X_j^1. Donc

$$\Delta^2(p) = \text{var } X_j^1 = \frac{p(1-p)}{(1-2p)^2}.$$

b) Pour $p \in]0, 1/2[$, on a

$$f'(p) = \frac{1}{(1-2p)^2} > 0, \quad \lim_{p\to 0^+} f(p) = 0 \quad \text{et} \quad \lim_{p\to \frac{1}{2}^-} f(p) = +\infty.$$

La fonction f est strictement croissante sur $]0, 1/2[$. C'est donc une bijection de $]0, 1/2[$ sur $]0, +\infty[$. Elle admet une fonction réciproque g qui est une bijection de $]0, +\infty[$ sur $]0, 1/2[$. De plus, g est dérivable sur $]0, +\infty[$ de dérivée strictement positive. g est donc continue et strictement croissante. Par définition on a pour tout $\tau > 0$ et $\theta \in]0, 1/2[$, $f(g(\tau)) = \tau$ et $g(f(\theta)) = \theta$. Un calcul direct donne par ailleurs

$$g(\tau) = \frac{\tau}{1+2\tau}, \quad (\tau \in]0, +\infty[).$$

c) Etude de $\widehat{p}_N = g(\eta_N)$.

i) On a vu que l'on a la relation $f(p) = \mathbb{E}(X_j^1)$ c'est-à-dire $p = g(\mathbb{E}(X_j^1))$. D'autre part, par la loi faible des grands nombres, η_N *estime* $\mathbb{E}(X_j^1)$. En effet, asymptotiquement cette variable aléatoire converge (en probabilité), vers $\mathbb{E}(X_j^1)$. L'idée est donc une utilisation empirique de l'égalité qui relie p et $\mathbb{E}(X_j^1)$ où l'on remplace $\mathbb{E}(X_j^1)$ par une variable aléatoire qui *l'estime*

ii) En invoquant la loi faible des grands nombres, on a vu que $\eta_N \geq 0$ converge, en probabilité, vers $f(p)$. La fonction g étant continue sur $\mathbb{R}^+$, la suite $\widehat{p}_N = g(\eta_N)$ converge en probabilité vers $g(f(p)) = p$. En effet, la continuité de f permet d'écrire :

$$\forall \delta > 0,\ \exists \epsilon_\delta > 0,\ \ |\eta_N - f(p)| \leq \epsilon_\delta \quad \Rightarrow \quad |g(\eta_N) - p| \leq \delta.$$

Soit $\delta > 0$, on a alors $\{|g(\eta_N) - p| > \delta\} \subset \{|\eta_N - f(p)| > \epsilon_\delta\}$. Donc

$$\mathbb{P}(|g(\eta_N) - p| > \delta) \leq \mathbb{P}(|\eta_N - f(p)| > \epsilon_\delta) \quad \text{et} \quad \lim_{N \to \infty} \mathbb{P}(|g(\eta_N) - p| > \delta) = 0.$$

iii) Ecrivons la formule de Taylor à l'ordre 2, pour $\tau > 0$ on a

$$g(\tau) - p = g(\tau) - g(f(p)) = g'\left[f(p)\right](\tau - f(p)) + \frac{g''(\tau^*)(\tau - f(p))^2}{2}$$

où τ^* appartient au segment non ordonné $)f(p), \tau($ donc en particulier $\tau^* > 0$. D'autre part, pour $\tau > 0$ on a

$$g'(\tau) = \frac{1}{(1+2\tau)^2}, \quad \text{et} \quad g''(\tau) = \frac{-4}{(1+2\tau)^3}.$$

De plus,

$$g'[f(p)] = \frac{1}{(1 + \frac{2p}{1-2p})^2} = (1-2p)^2.$$

En reportant ces expressions dans la formule de Taylor précédente, on obtient

$$g(\tau) - p = (1-2p)^2(\tau - f(p)) - \frac{2(\tau - f(p))^2}{(1+2\tau^*)^3}. \tag{13}$$

iv) D'après la question a), $\sqrt{N}(\eta_N - f(p))$ converge en loi vers une loi gaussienne centrée de variance $\Delta^2(p)$. Donc la suite $\sqrt{N}(1-2p)^2(\eta_N - f(p))$ converge en loi vers une loi gaussienne centrée de variance $(1-2p)^4\Delta^2(p)$.

d) L'inégalité de Tchebycheff appliquée à la variable positive $\sqrt{N}[\eta_N - f(p)]^2$ donne

$$\mathbb{P}\left(\sqrt{N}[\eta_N - f(p)]^2 \geq \epsilon\right) \leq \frac{\mathbb{E}\left[\sqrt{N}[\eta_N - f(p)]^2\right]}{\epsilon} = \frac{p(1-p)}{\epsilon\sqrt{N}(1-2p)^2}, \quad (\epsilon > 0).$$

Donc, pour $\epsilon > 0$, $\lim_{N\to\infty} \mathbb{P}\left(\sqrt{N}[\eta_N - f(p)]^2 \geq \epsilon\right) = 0$.
C'est-à-dire que $\sqrt{N}(\eta_N - f(p))^2$ converge en probabilité vers 0.

La formule (13) multipliée par $\sqrt{N}$ et appliquée à $\tau = \eta_N$ devient

$$\sqrt{N}(\widehat{p}_N - p) = \sqrt{N}(1-2p)^2(\eta_N - f(p)) - \frac{2\sqrt{N}(\eta_N - f(p))^2}{(1+2\tau_N^*)^3}. \tag{14}$$

Mais, pour $\tau > 0$, $2(1+2\tau)^{-3} \in]0,2[$ donc

$$-2\sqrt{N}(\eta_N - f(p))^2 \leq -\frac{2\sqrt{N}(\eta_N - f(p))^2}{(1+2\tau_N^*)^3} \leq 0.$$

Cette dernière inégalité entraîne que $-\frac{2\sqrt{N}(\eta_N - f(p))^2}{(1+2\tau_N^*)^3}$ converge en probabilité vers 0. Utilisant alors les points III.c.0) et III.c.iv) on conclut que $\sqrt{N}(\widehat{p}_N - p)$ converge en loi vers la loi normale centrée de variance

$$\Gamma^2(p) = (1-2p)^4\Delta^2(p) = p(1-2p)^2(1-p).$$

e) On a vu à la question précédente que $\sqrt{N}(\widehat{p}_N - p)$ converge en loi vers la loi normale centrée de variance $\Gamma^2(p)$. Donc,

$$\frac{\sqrt{N}(\widehat{p}_N - p)}{\Gamma(p)}$$

converge en loi vers une loi normale centrée réduite. De plus, puisque $\widehat{p}_N$ converge en probabilité vers p et que la fonction Γ est continue le point III.c.0) permet de conclure que

$$\frac{\sqrt{N}(\widehat{p}_N - p)}{\Gamma(\widehat{p}_N)} = \frac{\Gamma(\widehat{p}_N)}{\Gamma(p)} \frac{\sqrt{N}(\widehat{p}_N - p)}{\Gamma(p)},$$

converge également en loi vers une loi normale centrée réduite. Après lecture de la table de la loi normale standard, on obtient donc

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{\sqrt{N}(\widehat{p}_N - p)}{\Gamma(\widehat{p}_N)} \in [-1.96, 1.96]\right) = 0.95.$$

D'où

$$\lim_{N\to\infty} \mathbb{P}\left(p \in \left[\widehat{p}_N - 1.96\frac{\Gamma(\widehat{p}_N)}{\sqrt{N}}, \widehat{p}_N + 1.96\frac{\Gamma(\widehat{p}_N)}{\sqrt{N}}\right]\right) = 0,95.$$

En d'autres termes,

$$\left[\widehat{p}_N - 1.96\frac{\Gamma(\widehat{p}_N)}{\sqrt{N}}, \widehat{p}_N + 1.96\frac{\Gamma(\widehat{p}_N)}{\sqrt{N}}\right],$$

est un intervalle de confiance pour p de risque asymptotique 5%.