Exercice V

$(X_n)_{n\geq 1}$ sont indépendantes et suivent chacune une loi de de Bernoulli $\mathcal{B}(p)$, $p \in]0,1[$, donc S_n suit une loi binomiale $\mathcal{B}(n,p)$.

$$\mathbb{P}(S_n = k) = b_{n,k} = \begin{cases} \binom{n}{k} p^k q^{n-k} & q = 1-p,\ 0\leq k\leq n \\ 0 \quad \text{si} \quad k<0 \text{ ou } k>n. \end{cases}$$

$$Z_n = \frac{S_n - np}{\sqrt{npq}}$$

$\cancel{\mathbb{P}(Z_n}$ Soit $(a,b) \in \mathbb{R}^2$, $a \leq b$.

$$\mathbb{P}(a \leq Z_n \leq b) = \sum_{k=\lceil np + a\sqrt{npq}\rceil}^{\lfloor np + b\sqrt{npq}\rfloor} b_{n,k}$$

où $\lfloor x \rfloor = \max\{n \in \mathbb{Z} \quad n \leq x\}$ est la partie entière de x

et $\lceil x \rceil = \min\{n\in\mathbb{Z} \quad n \geq x\}$

On remarquera que puisque $p \in]0,1[$ il existe un n_0 tel que

$$\forall n \geq n_0 \quad \lfloor np + b\sqrt{npq}\rfloor \leq n$$
$$\lceil np + a\sqrt{npq}\rceil \geq 0.$$

Puisque seule nous intéresse $\lim\limits_{n\to+\infty} \mathbb{P}(a \leq Z_n \leq b)$ nous supposerons n choisi assez grand pour que ces deux conditions soient vérifiées.

On aura donc $\forall (n,k) \in$ Notre Domaine $\quad b_{n,k} = \binom{n}{k} p^k q^{n-k}$

Notons $\alpha_n = \lceil np + a\sqrt{npq}\rceil$ et $\beta_n = \lceil np + b\sqrt{npq}\rceil$

V.2

On a $\boxed{P(a \leq Z_n \leq b) = A_n = \sum_{k=\alpha_n}^{\beta_n} b_{n,k}}$

Remarques $\alpha_n \sim np$ en particulier $\lim_{n \to +\infty} \alpha_n = +\infty$

$\beta_n \sim nq$ $\quad n - \beta_n \sim np$ $\quad \lim_{n \to +\infty} n - \beta_n = +\infty$

Plus précisément :

$$\alpha_n = np + a\sqrt{npq} + \rho_n \quad \text{avec } \rho_n \in [0, 1[$$

Donc $\displaystyle \lim_{n \to +\infty} \frac{\alpha_n - np}{\sqrt{npq}} = a$

de même $\displaystyle \lim_{n \to +\infty} \frac{\beta_n - np}{\sqrt{npq}} = b$

On rappelle la formule de Stirling

$$\ln(n!) = n \ln n - n + \frac{1}{2}\ln n + \frac{1}{2}\ln 2\pi + r_n$$

avec $\displaystyle \lim_{n \to +\infty} r_n = 0$

Soit $\displaystyle R_n = \sup_{p \geq n} |r_p|$.

Or a $\displaystyle \lim_{n \to +\infty} R_n = 0$.

$\forall k \in [\alpha_n, \beta_n]$

$$\ln(k!) = k \ln k - k + \frac{1}{2}\ln k + \frac{1}{2}\ln(2\pi) + r_k$$

avec $|r_k| \leq R_k \leq R_{\alpha_n}$

et $\displaystyle \lim_{n \to +\infty} R_{\alpha_n} = 0$ car $\displaystyle \lim_{n \to +\infty} \alpha_n = +\infty$

On peut donc écrire

$$\ln(k!) = k\ln k - k + \frac{1}{2}\ln k + \frac{1}{2}\ln(2\pi) + o(1)$$

le $o(1)$ étant uniforme en k.

$(\forall \varepsilon\ \exists n_1\quad \forall n \geq n_1\ \forall k \in [\alpha_n, \beta_n] \qquad |o(1)| < \varepsilon)$

—

De même puisque $\lim\limits_{n\to+\infty} n - \beta_n = +\infty$ on peut écrire

$$\ln((n-k)!) = (n-k)\ln(n-k) - (n-k) + \frac{1}{2}\ln(n-k) + \frac{1}{2}\ln(2\pi) + o(1)$$

le $o(1)$ étant uniforme en k.

—

Plus précisément écrivons

$$k = np + u_k$$

alors

$$n-k = nq - u_k$$

et

$$u_k = O(\sqrt{n})$$

et reprenons les développement asymptotiques, avec $C = \sqrt{2\pi}$

$$\ln(k!) = (np+u_k)\ln(np+u_k) - np - u_k + \frac{1}{2}\ln(np+u_k) + \ln C + o(1)$$

$$\ln(k!) = (np+u_k)\left(\ln(np) + \ln\left(1+\frac{u_k}{np}\right)\right) - np - u_k + \frac{1}{2}\left(\ln(np) + \ln\left(1+\frac{u_k}{np}\right)\right) + \ln C + o(1) \quad (*)$$

On a $u_k = O(\sqrt{n})$ donc $\ln\left(1+\frac{u_k}{np}\right) \sim \frac{u_k}{np} = o(1)$.

$$\ln\left(1+\frac{u_k}{np}\right) = \frac{u_k}{np} - \frac{u_k^2}{2(np)^2} + O\left(\frac{1}{n^{3/2}}\right)$$

donc $np\ln\left(1+\frac{u_k}{np}\right) = u_k - \frac{u_k^2}{2np} + o(1)$ $\qquad u_k\ln\left(1+\frac{u_k}{np}\right) = \frac{u_k^2}{np} + o(1)$

Après simplification dans (*) à l'aide des deux développements ci-dessus on obtient ; <u>lorsque n tend vers $+\infty$, k restant dans $[\alpha_n, \beta_n]$</u>

$$\ln(k!) = np\ln n + np\ln p + \frac{1}{2}\frac{u_k^2}{np} + u_k\ln n + u_k\ln p - np + \frac{1}{2}\ln n + \frac{1}{2}\ln p + \ln C + o(1)$$

et de même

$$\ln((n-k)!) = nq\ln n + nq\ln q + \frac{1}{2}\frac{u_k^2}{nq} - u_k\ln n - u_k\ln q - nq + \frac{1}{2}\ln n + \frac{1}{2}\ln q + \ln C + o(1)$$

$$\cancel{\ln\left(q^{n-k}\right) = (np + u_k)}$$

$$\ln p^k = (np + u_k)\ln p.$$

$$\ln q^{n-k} = nq - u_k\ln q$$

<u>Finalement</u>

$$\ln(b_{n,k}) = \ln n! - \ln(k!) - \ln((n-k)!) + \ln(p^k) + \ln(q^{n-k})$$

Toutes simplifications faites

$$\ln(b_{n,k}) = -\frac{1}{2}\frac{u_k^2}{npq} - \frac{1}{2}\ln n - \ln C + o(1)$$

$$b_{n,k} = \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2n\pi pq}}(1 + o(1)) = \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2n\pi qp}} + \varepsilon_{n,k}$$

On rappelle que $o(1)$ tend vers 0 uniformément en k dans $[\alpha_n, \beta_n]$. $\left(\lim_{n\to+\infty} \sup_{k\in[\alpha_n,\beta_n]} |o(1)| = 0\right)$

Commençons par remarquer que

$$\forall k \in [\alpha_n, \beta_n] \quad \left| e^{-\frac{(k-np)^2}{2npq}} \right| \leq 1$$

donc $\quad c_{n,k} = o\left(\frac{1}{\sqrt{n}}\right)$

et $\quad \sum_{k=\alpha_n}^{\beta_n} c_{n,k} = o\left(\frac{\beta_n - \alpha_n}{\sqrt{n}}\right) = o(1)$

Donc, sous réserve qu'elle existe.

$$\boxed{\lim_{n\to+\infty} A_n = \lim_{n\to+\infty} \sum_{\alpha_n}^{\beta_n} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2n\pi qp}}} = \lim_{n\to+\infty} \Sigma_n$$

On compare maintenant avec une intégrale.

La fonction $x \mapsto e^{-\frac{(x-np)^2}{2npq}}$ est croissante sur $]0, np]$

décroissante sur $[np, +\infty[$.

Il y a plusieurs cas à étudier : $a < 0 < b$, $a = 0 < b$,

$a < b = 0$, $a < b < 0$, $0 < a < b$, $0 = a = b$.

Etudions le premier le cas $a < 0 < b$, cas qui est le plus complexe.

$$\Sigma_n = \sum_{k=\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2n\pi qp}} + \sum_{k=\lceil np \rceil}^{\beta_n} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2n\pi pq}} - \frac{\varepsilon}{\sqrt{2\pi npq}}$$

ou $\varepsilon = 0$ si np n'est pas entier $\quad \varepsilon = 1$ si $\lfloor np \rfloor = \lceil np \rceil$.

$\lim_{n\to\infty} \frac{\varepsilon}{\sqrt{2\pi npq}} = 0$. Interessons nous aux sommes.

Le traitement des deux est similaire. On se concentre sur la première.

Sur $[\alpha_n - 1, \lfloor np \rfloor]$ $x \mapsto e^{-\frac{(x-np)^2}{2npq}}$ est croissante donc

V.6

$$\int_{\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(x-np)^2}{2npq}}}{\sqrt{2\pi npq}} dx + \frac{e^{-\frac{(\alpha_n - np)^2}{2npq}}}{\sqrt{2\pi npq}} \leq \sum_{k=\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2\pi npq}} \leq \frac{1}{\sqrt{2\pi np}} \int_{\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(x-np)^2}{2npq}}}{\sqrt{2\pi npq}} dx + \frac{e^{-\frac{(\lfloor np \rfloor - np)^2}{2npq}}}{\sqrt{2\pi np}}$$

Clairement $\lim_{n \to +\infty} \frac{e^{-(\dots)^2}}{\sqrt{2npq}} = 0$

et $$\int_{\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(x-np)}{2npq}}}{\sqrt{2\pi npq}} dx = \int_{\frac{\alpha_n - np}{\sqrt{npq}}}^{\frac{\lfloor np \rfloor - np}{\sqrt{npq}}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

Or $\lim_{n \to +\infty} \frac{\alpha_n - np}{\sqrt{npq}} = a$ $\quad \lim_{n \to +\infty} \frac{\lfloor np \rfloor - np}{\sqrt{npq}} = 0$

Finalement
$$\lim_{n \to +\infty} \sum_{k=\alpha_n}^{\lfloor np \rfloor} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2\pi npq}} = \frac{1}{\sqrt{2\pi}} \int_a^0 e^{-\frac{y^2}{2}} dy$$

$$\lim_{n \to +\infty} \sum_{k=\alpha_n}^{\beta_n} \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2\pi npq}} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy$$

Soit $$\boxed{\lim_{n \to +\infty} \mathbb{P}(a \leq Z_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy.}$$

C'est le théorème "central limite", dans un cas particulier. Il s'étend à toute famille de variables aléatoires indépendantes de même lois, avec des hypothèses supplémentaires souvent vérifiées et que de nombreux mathématiciens travaillent à affaiblir.