

Appendice 2: Convergence des lois binomiales vers la loi de Poisson

Cet appendice montre une chose peu connue: c'est que la suite des lois binomiales de paramètres convenables converge vers une loi de Poisson, non seulement faiblement, mais aussi au sens de la convergence en norme de mesures. Cet appendice peut intéresser aussi les étudiants d'agrégation qui ont à traiter du sujet "lois binomiales, lois de Poisson".

Adoptons les notations suivantes: δ_a désigne la masse de Dirac en a ; si $m > 0$, on définit la loi de Poisson de moyenne m par

$$p_m(dx) = \sum_{n=0}^{\infty} e^{-m} \frac{m^n}{n!} \delta_n(dx)$$

et si $0 < p < 1$ on définit la loi de Bernoulli de moyenne p par

$$b_{1,p}(dx) = (1-p)\delta_0(dx) + p\delta_1(dx).$$

Si n est un entier > 0 , on définit la loi binomiale $b_{n,p}$ comme la n ème puissance de convolution de la loi de Bernoulli:

$$b_{n,p}(dx) = (b_{1,p})^{*n}(dx) = ((1-p)\delta_0 + p\delta_1)^{*n}(dx) = \sum_{k=0}^n C_n^k (1-p)^{n-k} p^k \delta_k(dx).$$

C'est un résultat simple et important que de constater que la suite de probabilités $(b_{n,m/n})_{n>m}$ converge faiblement vers p_m . En effet si $n \geq k$, alors

$$C_n^k \left(1 - \frac{m}{n}\right)^{n-k} \left(\frac{m}{n}\right)^k$$

est une fraction rationnelle en n et l'examen des termes de plus haut degré au numérateur et au dénominateur montre que

$$\lim_{n \rightarrow +\infty} C_n^k \left(1 - \frac{m}{n}\right)^{n-k} \left(\frac{m}{n}\right)^k = e^{-m} \frac{m^k}{k!}.$$

Toutefois, un résultat plus fort est vrai, puisque en fait $(b_{n,m/n})_{n>m}$ converge fortement vers p_m . S'agissant ici de probabilités concentrées sur l'ensemble \mathbf{N} des entiers, cette convergence forte est une convergence dans $l_1(\mathbf{N})$ et revient à affirmer que

$$\|b_{n,m/n} - p_m\| = \sum_{k=0}^n \left| C_n^k \left(1 - \frac{m}{n}\right)^{n-k} \left(\frac{m}{n}\right)^k - e^{-m} \frac{m^k}{k!} \right| + \sum_{k=n+1}^{\infty} e^{-m} \frac{m^k}{k!}$$

tend vers 0 si $n \rightarrow +\infty$. Nous allons montrer ce résultat de deux manières. Celle de Le Cam(1960) est courte et utilise une ingénieuse idée de couplage. Celle de Prohorov (1963) donne plus d'informations en montrant que $\|b_{n,m/n} - p_m\|$ est équivalente à un $\phi(m)/n$ et calcule explicitement $\phi(m)$.

Théorème 1: (Le Cam) Si $n > m > 0$ on a

$$\|b_{n,m/n} - p_m\| \leq \frac{4m^2}{n}.$$

Démonstration Posons pour simplifier $p = m/n \in]0,1[$ et considérons des variables aléatoires $(X_1, Y_1), \dots, (X_n, Y_n)$ de \mathbf{N}^2 indépendantes et de même loi m_p définie par

$$\begin{aligned} m_p(0,0) &= e^{-p} - p + pe^{-p} \\ m_p(0,1) &= p - pe^{-p} \\ m_p(1,1) &= pe^{-p} \\ m_p(n,0) &= \frac{p^n}{n!} e^{-p} \text{ si } n \geq 2, \end{aligned}$$

et $m_p(a,b) = 0$ ailleurs. Alors on constate facilement que X_i suit une loi de Poisson et que Y_i suit une loi de Bernoulli, toutes deux de moyenne p . Elles ne sont pas indépendantes, et satisfont à l'inégalité

$$\Pr(X_i = Y_i) = m_p(0,0) + m_p(1,1) = e^{-p} - p + 2pe^{-p} \geq 1 - 2p^2,$$

héritée du fait que $e^{-p} \geq 1 - p$ pour tout réel p . Notons pour simplifier $X = X_1 + \dots + X_n$, qui est donc de loi de Poisson p_m , et $Y = Y_1 + \dots + Y_n$, de loi binomiale $b_{n,p}$. Donc

$$\Pr(X \neq Y) \leq \Pr(\cup_{i=1}^n (X_i \neq Y_i)) \leq \sum_{i=1}^n \Pr(X_i \neq Y_i) \leq 2np^2.$$

Ensuite, si A est une partie de \mathbf{N} , on a l'inclusion d'évènements

$$(X \in A) = (X = Y \in A) \cup (Y \neq X \in A) \subset (Y \in A) \cup (X \neq Y)$$

qui entraîne $\Pr(X \in A) - \Pr(Y \in A) \leq \Pr(X \neq Y)$. Le raisonnement fait en échangeant les rôles de X et Y donne finalement

$$|\Pr(X \in A) - \Pr(Y \in A)| \leq 2np^2 = \frac{2m^2}{n}.$$

Pour terminer, on applique cette inégalité à l'ensemble $E = \{k \in \mathbf{N}; p_m(k) \geq b_{n,p}(k)\}$ puis à son complémentaire $E' = \mathbf{N} \setminus E$. Comme

$$\|b_{n,m/n} - p_m\| = (\Pr(X \in E') - \Pr(Y \in E')) + (\Pr(Y \in E) - \Pr(X \in E)) \leq \frac{4m^2}{n},$$

le résultat est montré.

Théorème 2: (Prohorov) Soit X une variable aléatoire de loi de Poisson p_m . Soit

$$\frac{\phi_n(m)}{n} = \|b_{n,m/n} - p_m\|.$$

Alors

$$\lim_{n \rightarrow +\infty} \phi_n(m) = \phi(m) = \frac{1}{2} \mathbf{E}(|X - (X - m)|).$$

Démonstration Notons $p_k = e^{-m} \frac{m^k}{k!}$ pour simplifier les notations. Observons d'abord que le résultat n'est pas si surprenant, car pour k fixé, en notant

$$a_k(n) = \frac{n}{p_k} \left(C_n^k \left(1 - \frac{m}{n}\right)^{n-k} \left(\frac{m}{n}\right)^k - p_k \right) \quad \text{si } 0 \leq k \leq n, \quad (1)$$

$$a_k(n) = -n \quad \text{si } n < k,$$

alors

$$\lim_{n \rightarrow +\infty} a_k(n) = \frac{1}{2}(k - (k - m)) \quad (2)$$

par un calcul standard et laborieux de développement limité (voir le détail de ce calcul dix lignes ci dessous). On est donc fondé de penser que

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{\infty} p_k |a_k(n)| = \frac{1}{2} \sum_{k=0}^{\infty} p_k |k - (k - m)|. \quad (3)$$

Le point délicat est alors de justifier cette interversion de limites. On va le faire par convergence dominée. L'idée pour cela est de considérer pour k fixé

$$\frac{1}{p_k} C_n^k \left(1 - \frac{m}{n}\right)^{n-k} \left(\frac{m}{n}\right)^k$$

comme une fonction f_k de $1/n$, en introduisant donc

$$f_k(h) = (1-h)(1-2h) \dots (1-(k-1)h)(1-mh)^{-k} \exp\left[m + \frac{1}{h} \log(1-mh)\right].$$

Cette fonction f_k est définie sur $h > -1/m$, et une reformulation de (2) est d'affirmer que $f'_k(0) = \frac{1}{2}(k - (k - m))$; on le voit ainsi:

$$\begin{aligned} f_k(h) &= [1 - (1+2+\dots+k-1)h + o(h)][1 + kmh + o(h)] \exp\left[m + \frac{1}{h}(-mh - mh/2 + o(h))\right] \\ &= \left[1 - \frac{k(k-1)}{2}h + o(h)\right][1 + kmh + o(h)] \left[1 - \frac{m}{2}h + o(h)\right] = 1 + \frac{1}{2}(k - (k - m))h + o(h). \end{aligned}$$

Fixons désormais $k_0 > m$. Pour $k > k_0$, soit

$$M_k = \max_{0 \leq h \leq \frac{1}{k}} f_k(h).$$

Montrons que $M = \sup_{k > k_0} M_k$ est fini. Pour cela, notons

$$K = \max_{0 \leq h \leq \frac{1}{k_0}} \exp\left[m + \frac{1}{h} \log(1 - mh)\right],$$

qui existe comme maximum d'une fonction continue sur un compact. Ensuite, si $0 \leq h \leq \frac{1}{k} \leq \frac{1}{k_0} < \frac{1}{m}$, alors $0 \leq 1 - jh \leq 1$ si $j = 1, \dots, k - 1$ et $(1 - mh)^{-k} \leq (1 - \frac{m}{k})^{-k}$. Donc

$$M_k \leq K \left(1 - \frac{m}{k}\right)^{-k}.$$

Or $\lim_{k \rightarrow +\infty} \left(1 - \frac{m}{k}\right)^{-k} = \exp(-m)$ est finie. Donc la suite $(M_k)_{k > k_0}$ est bornée et M est fini.

Soit maintenant

$$M'_k = \max_{0 \leq h \leq \frac{1}{k}} |f'_k(h)|.$$

Montrons que $M' = \sup_{k > k_0} k^{-3} M'_k$ est fini. Notons $G(h) = \frac{1}{h} \log(1 - mh)$ si $h \neq 0$ et $G(0) = -m$: G est donc continûment dérivable. Soit

$$K' = \max_{0 \leq h \leq \frac{1}{k_0}} |G'(h)|.$$

Alors

$$\frac{f'_k(h)}{f_k(h)} = \frac{km}{1 - mh} + G'(h) - \sum_{j=1}^{k-1} \frac{j}{1 - jh}.$$

Ensuite, si $0 \leq h \leq \frac{1}{k} \leq \frac{1}{k_0} < \frac{1}{m}$, alors

$$\left| - \sum_{j=1}^{k-1} \frac{j}{1 - jh} \right| \leq \sum_{j=1}^{k-1} \frac{j}{1 - \frac{j}{k}} \leq k \sum_{j=1}^{k-1} j = \frac{k(k-1)}{2},$$

et

$$\frac{km}{1 - mh} \leq \frac{km}{1 - \frac{m}{k}} \leq \frac{km}{k_0 - m}.$$

Donc

$$\left| \frac{f'_k(h)}{f_k(h)} \right| \leq \frac{km}{k_0 - m} + K' + \frac{k(k-1)}{2}.$$

Comme dans cet intervalle f_k est dans $]0, M]$, on en déduit

$$k^{-3} M'_k \leq \frac{M}{2} (1 - 1/k) + \frac{Mm}{k(k_0 - m)} + \frac{MK'}{k^3} \leq \frac{M}{2} + \frac{Mm}{k_0(k_0 - m)} + \frac{MK'}{k_0^3}.$$

M' est donc fini.

On peut alors terminer la démonstration du théorème: on a donc par la formule des accroissements finis pour $0 \leq k \leq n$:

$$|a_k(n)| = |n(f_k(1/n) - f_k(0))| = |f'_k(\theta/n)| \leq M' k^3$$

et pour $k > n$ $|a_k(n)| = n \leq k^3$. Soit M'' le maximum de 1 et M' . Observons que $\sum_{n=0}^{\infty} M'' k^3 p_k$ converge. On est donc dans les conditions d'application du théorème de la convergence dominée et donc (3) est démontré.

Voici quelques raffinements intéressants sur la fonction de Prohorov ϕ du théorème précédent:

Proposition 2: Pour $m > 0$ on note $p_k = e^{-m} \frac{m^k}{k!}$ et $p_{-1} = 0$. Soit r et R les racines du trinôme $P(x) = x - (x - m)$, avec $0 < r < R$, et soient a et A les parties entières de r et R . Alors

$$\phi(m) = m(p_a - p_{a-1} + p_{A-1} - p_A).$$

De plus, les deux fonctions de m définies par $q(m) = m(p_a - p_{a-1})$ et $Q(m) = m(p_{A-1} - p_A)$ sont continues, positives et tendent vers $\frac{1}{\sqrt{2\pi e}}$ à l'infini.

Démonstration Remarquons que les racines r et R existent, car le discriminant simplifié de P est $m + 1/4 > 0$, et qu'elles sont > 0 car de somme $2m + 1 > 0$ et de produit $m > 0$. Si X est une variable aléatoire de Poisson de moyenne m , alors $\mathbf{E}(X - (X - m)) = 0$. Donc, puisque $P(x)$ est positif si et seulement si $r < x < R$, on a

$$\begin{aligned} \phi(m) &= \frac{1}{2} \mathbf{E}(|P(X)|) = \frac{1}{2} \mathbf{E}(|P(X)| + P(X)) = \sum_{r < k < R} p_k P(k) = \sum_{k=a+1}^A p_k P(k) \\ &= m \sum_{k=a+1}^A (-p_{k-2} + 2p_{k-1} - p_k) = m \left(- \sum_{k=a-1}^{A-2} p_k + 2 \sum_{k=a}^{A-1} p_k - \sum_{k=a+1}^A p_k \right) \\ &= m(p_a - p_{a-1} + p_{A-1} - p_A). \end{aligned}$$

Définissons la fonction $m \mapsto r(m) = m + \frac{1}{2} - \sqrt{m + \frac{1}{4}}$ sur $[0, +\infty)$. Elle est continue, et sa dérivée $r'(m) > 0$. r est une bijection croissante de $[0, +\infty)$ sur lui-même de fonction réciproque $m = r^{-1}(x) = x + \sqrt{x}$. Donc $a \leq r < a + 1$ implique que

$$a + \sqrt{a} = r^{-1}(a) \leq m < r^{-1}(a + 1) = a + 1 + \sqrt{a + 1}.$$

Donc sur l'intervalle

$$I_n = \{m; a = n\} = [n + \sqrt{n}, n + 1 + \sqrt{n + 1})$$

la fonction q prend la valeur

$$q(m) = e^{-m} \frac{m^n (m - n)}{n!}.$$

Elle est donc bien positive. Sa limite à l'extrémité droite de I_n est bien $q(n + 1 + \sqrt{n + 1})$ ce qui montre sa continuité. Quant à sa limite à l'infini, c'est évident avec la formule de Stirling. La démonstration pour Q est entièrement analogue: sur l'intervalle

$$J_N = \{m; A = N\} = [N - \sqrt{N}, N + 1 - \sqrt{N + 1}),$$

la fonction Q prend la valeur

$$Q(m) = e^{-m} \frac{m^N (N - m)}{N!}.$$

Sur les intervalles I_n et J_N , les fonctions q et Q sont concaves.

Commentaires:

Il y a de nombreuses références sur cette question, (voir la bibliographie ci dessous, qui conduit à d'autres références) mais pas très accessibles un jour de concours d'agrégation. Dans Letac (1981), Problème IV 3 4ème et 5ème, on trouve la démonstration de Le Cam (1960).

Références:

LE CAM, L. "An approximation theorem for the Poisson binomial distribution". *Pacific J. Math.* 10, 1181-1197.

LETAC, G. (1982), *Intégration et Probabilités, Analyse de Fourier, Exercices corrigés*. Masson, Paris. (Seconde édition 1997).

PROHOROV, Ju. V. (1953), " Asymptotic behavior of the binomial distribution" (en russe). *Uspehi Matematicheskikh Nauk.* 8, 135-142.

VERVAAT, W. (1969), " Upper bounds for the distance in total variation between the binomial and the Poisson distribution". *Statistica Neerlandica*, 23, 79-86.

JOHNSON, N. J. and SIMMONS, G. (1971), "On the convergence of the binomial to Poisson distribution". *Annals of Math. Statist.* 49, 1735-1736.