264

Variables aléatoires discrètes. Exemples et applications.

I Exemples fondamentaux de lois discrètes [Ouvrard, chap I]

Déf 1 Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Une variable aléatoire $X : \Omega \to E$ est dite discrète si :
- $X(\Omega)$ est dénombrable.
- $X^{-1}(\{e\}) \in \mathcal{A}$ pour tout $e \in E$.

Déf 2 La loi P_X d'une variable aléatoire discrète $X : \Omega \to E$ (v.a.d) est l'unique probabilité sur E telle que
$P_X(\{k\}) = P(X^{-1}(\{e\}))$. (cette valeur est aussi notée $P(X=e)$).

① **Loi uniforme**

Si $X(\Omega) = [\![1,n]\!]$, X est de loi uniforme si pour tout $k \in [\![1,n]\!]$,
$$P(X=k) = \frac{1}{n}.$$

Exemple : La variable qui, à un lancer de dé équilibré renvoie le résultat obtenu est de loi uniforme.

② **Loi de Bernoulli**

Cas où X ne prend que les valeurs 0 et 1, et $p = P(X=1) \in]0,1[$.
p est appelé paramètre de la variable.

Exemple : On joue à pile ou face avec une pièce truquée, probabilité p d'obtenir face, la variable valant 1 si on obtient face et 0 si on obtient pile est de Bernoulli.

Note : si $p = \frac{1}{2}$, on a une loi uniforme.

③ **Loi binomiale**

X suit la loi binomiale $\mathcal{B}(n,p)$, avec $n \in \mathbb{N}$ et $p \in]0,1[$,
si pour tout $k \in [\![1,n]\!]$, $P(X=k) = \binom{n}{k} p^k q^{n-k}$, où $q = 1-p$.

Exemple : Le nombre de 6 obtenus lors de n lancers de dés suit la loi binomiale $\mathcal{B}(n, \frac{1}{6})$.

④ **Loi géométrique**

X suit la loi géométrique $\mathcal{G}_{\mathbb{N}^*}(p)$, avec $p \in]0,1[$,
si pour tout $k \in \mathbb{N}^*$, $P(X=k) = p\,q^{k-1}$, où $q = 1-p$.

Exemple : Le nombre de lancers de dés jusqu'au premier 6 suit la loi géométrique $\mathcal{G}_{\mathbb{N}^*}(\frac{1}{6})$.

⑤ **Loi hypergéométrique**

X suit la loi hypergéométrique $\mathcal{H}(n,p,A)$ où $p \in]0,1[$, $A \in \mathbb{N}$, $n \in [\![0,A]\!]$
si pour tout $\max(0, n-qA) \leq k \leq \min(pA, n)$, avec $q = 1-p$,
on a
$$P(X=k) = \frac{\binom{pA}{k}\binom{qA}{n-k}}{\binom{A}{n}}$$

Exemple : Dans un banc de A poissons, on en pêche n.
Si un poisson a une probabilité p d'être malade, le nombre de poissons pêchés malades suit la loi $\mathcal{H}(n,p,A)$.

Terminons avec un résultat utile sur les v.a.d.

Prop 3 Si X est une v.a.d, d'image E, et $f : E \to F$ une fonction quelconque,
Alors $f(X)$ est une v.a.d.

II Espérance d'une variable [Ouvrard, Chap V]

Déf 4 Soit $X : \Omega \to \mathbb{R}$ une v.a.d. telle que $\sum_{x \in X(\Omega)} |x| P(X=x) < \infty$,
on définit son espérance $E[X] = \sum_{x \in X(\Omega)} x\, P(X=x)$.

Quelques exemples

↝ Si X est de loi uniforme, d'image $\{x_1, \dots, x_n\}$, alors $\mathbb{E}[X] = \frac{1}{n} \sum_{i=1}^{n} x_i$.

↝ Si X est de Bernoulli, de paramètre p, alors $E[X] = p$.

⤳ Si X suit la loi $\mathcal{B}(n,p)$ (resp. $\mathcal{G}_{\mathbb{N}^*}(p)$)
alors $E[X] = np$ (resp. $E[X] = \frac{1}{p}$)

⤳ Si X suit la loi $\mathcal{H}(n,p,A)$, alors $E[X] = np$.

① Thm de transfert et propriétés de l'espérance

Thm 5 (transfert) Si $X : \Omega \to E$ v.a.d. et $f : E \to \mathbb{R}$,
alors $E[f(X)]$ est définie ssi $\sum_{x \in E} |f(x)|\, P(X=x)$,
et on a alors $E[f(X)] = \sum_{x \in E} f(x)\, P(X=x)$.

Applications • linéarité de l'espérance.
Si $a, b \in \mathbb{R}$, $E[aX+bY] = aE[X] + bE[Y]$.
• Calcul de la variance
Si on définit $var(X) = E[(X-E(X))^2]$, alors
$var(X) = E[X^2] - E[X]^2$.

exemples • X variable de Bernoulli de par. p : $var(X) = p(1-p)$.
• X de loi géométrique $\mathcal{G}_{\mathbb{N}^*}(p)$: $var(X) = \frac{1-p}{p^2}$.

Prop 6 (thm de Weierstrass) $\boxed{DEV}$
Soit $f : [0,1] \to \mathbb{C}$ une fonction continue, et ω son module de continuité uniforme, défini par $\omega : h \mapsto \sup\{|f(u)-f(v)|,\ |u-v| \leq h\}$.
Pour $n \in \mathbb{N}^*$, on considère le $n$ème polynôme de Bernstein de f :
$$B_n(f) = \sum_{k=0}^{n} \binom{n}{k} X^k (1-X)^{n-k} f\left(\frac{k}{n}\right)$$
Alors $(B_n(f))_n$ converge uniformément vers f sur $[0,1]$,
et $\|f - B_n(f)\|_\infty \leq C\, \omega\left(\frac{1}{\sqrt{n}}\right)$ pour une certaine constante C.

② Fonction génératrice

Déf 7 Soit $X : \Omega \to \mathbb{N}$ une v.a.d. d'espérance finie.
Pour tout $s \in [-1,1]$, la v.a.d. s^X est d'espérance finie,
et on note $G_X : s \mapsto E[s^X]$
la fonction génératrice de X.

Prop 8 • G_X est continue sur $[-1,1]$
et $\mathcal{C}^\infty$ sur $]-1,1[$.
• Pour tout $s \in [-1,1]$,
$$G_X(s) = \sum_{n \geq 0} s^n P(X=n).$$
G_X est donc la série génératrice de la suite $(P(X=n))_n$.
• G_X caractérise la ~~variable X~~ loi de la variable X : $G_X = G_Y$ ssi $P_X = P_Y$

Prop 9 X est d'espérance finie ssi G_X est dérivable à gauche en 1,
et alors $E[X] = G_X'(1^-)$.

exemples • Si X suit la loi de Bernoulli de paramètre p,
$G_X : s \mapsto 1 - p + ps$.
• Si X suit la loi géométrique $\mathcal{G}_{\mathbb{N}^*}(p)$,
$G_X : s \mapsto \frac{ps}{1-(1-p)s}$

III Variables indépendantes

① Conditionnement et indépendance

Déf 10 Soit $X : \Omega \to E$ une v.a.d., et $Y : \Omega \to F$ une v.a.d.
La loi de Y conditionnelle à l'événement $(X=x)$, où $P(X=x) > 0$,
est la probabilité $P_Y^{(X=x)}$ définie sur F par
$$P_Y^{(X=x)}(y) = \frac{P(Y=y)}{P(X=x)} \quad \text{pour tout } y \in F.$$

Note. La connaissance de P_X et des $P_Y^{(X=x)}$ pour $P(X=x) > 0$
détermine entièrement le couple (X,Y).

Déf 11 • X et Y sont dites indépendantes si pour tout $(x,y) \in E \times F$,
on a $P(X=x, Y=y) = P(X=x)\, P(Y=y)$.
• Une famille de variables $(X_i)_{i \in I}$ est dite mutuellement indépendante si pour toute famille $(X_j)_{j \in J}$ avec $J \subseteq I$ finie, et toute famille $(x_j)_{j \in J} \in \prod_{j \in J} X_j$,
$$P\left(\bigcap_{j \in J} X_j = x_j\right) = \prod_{j \in J} P(X_j = x_j)$$

~~Exemple Soit~~ **Prop 12** Si X et Y sont indépendantes, alors
$E[XY] = E[X] \cdot E[Y]$

2) Construction de modèle

En pratique, lors d'une expérience aléatoire, l'ensemble Ω est rarement donné, on connaît juste ~~[illegible]~~ certaines variables aléatoires définies sur Ω, avec quelques hypothèses d'indépendance et de conditionnement.
C'est au mathématicien de justifier l'existence de Ω.

Thm 13 Soient, pour tout $i \in \mathbb{N}$, (E_i, P_i) un espace probabilisé discret.
Il existe $(\Omega, \mathcal{A}, P)$ un espace muni de variables $X_i : \Omega \to E_i$ indépendantes de lois respectives P_i.

Exemple On considère un jeu où l'on tire à pile ou face jusqu'à obtenir face.
S'il a fallu k essais, on jette k dés, et on gagne si un seul 6 est sorti.
Si la pièce est truquée (probabilité p d'obtenir face), la probabilité de victoire est $\frac{6p}{(1+5p)^2}$.

3) Somme et variables i.i.d.

Déf 14 Des variables sont dites i.i.d si elles sont indépendantes de même loi.

Exemple Si $(X_i)_{i \in \mathbb{N}}$ sont des variables i.i.d de Bernoulli (paramètre p),
alors $\sum_{i=1}^{n} X_i$ est de loi binomiale $\mathcal{B}(n,p)$
et $S : \omega \longmapsto \inf\{n \in \mathbb{N}^* \mid X_n(\omega) = 1\}$ est de loi géométrique $\mathcal{G}_{\mathbb{N}^*}(p)$.

Thm 15 Soient $X_1, X_2 : \Omega \to \mathbb{R}$ v.a.d, alors $X_1 + X_2$ est discrète,
de loi $P_{X_1+X_2} = P_{X_1} * P_{X_2}$
où $P_{X_1} * P_{X_2}(k) = \sum_{i \in \mathbb{R}} P(X_1 = i)\, P(X_2 = k-i)$ pour $k \in \mathbb{R}$.

Exemples
- X_1 et X_2 des variables i.i.d uniformes sur $[\![0,n]\!]$.
La variable $Z = X_1 + X_2$ suit la loi triangulaire
$$P(Z=k) = \begin{cases} \frac{k+1}{(n+1)^2} & \text{si } 0 \leq k \leq n \\ \frac{2n-k+1}{(n+1)^2} & \text{si } n < k \leq 2n \end{cases} \quad \text{(voir annexe)}$$
- Si X_1 suit la loi $\mathcal{B}(n_1, p)$ et X_2 la loi $\mathcal{B}(n_2, p)$,
alors $X_1 + X_2$ suit la loi $\mathcal{B}(n_1 + n_2, p)$.

Corollaire Si X_1 et X_2 sont indépendantes, alors $G_{X_1+X_2} = G_{X_1} G_{X_2}$.

Exemple Si X est de loi binomiale $\mathcal{B}(n,p)$, alors $G_X(s) = (1-p+ps)^n$.

Le thm suivant est très utilisé dans l'étude des processus de Galton-Watson.

Thm 16 Soient (X_n) une suite de variables i.i.d à valeurs dans $\mathbb{N}$,
et T dans $\mathbb{N}$ aussi tq $(T, X_0, X_1, \dots)$ sont mutuellement indépendantes.
En posant $S_T = \sum_{i=1}^{T} X_i$, on a
$$G_{S_T} = G_T \circ G_{X_1}$$

IV Limites de variables

1) Loi hypergéométrique

Déf 17 Soit (X_n) une suite de v.a.d de même domaine et arrivée.
On dit que X_n converge en loi vers X (noté $X_n \xrightarrow{\mathcal{L}} X$)
si pour tout $x \in X(\Omega)$, $P(X_n = x) \xrightarrow[n \to \infty]{} P(X=x)$.

Prop 18 Si X_A est de loi hypergéométrique $\mathcal{H}(n,p,A)$,
alors X_A converge en loi vers une variable de loi binomiale $\mathcal{B}(n,p)$.

2) Loi de Poisson

Déf 19 Une v.a.d $X : \Omega \to \mathbb{N}$ suit la loi de Poisson $\mathcal{P}(\lambda)$, où $\lambda \in \mathbb{R}_+^*$ si
$$P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{pour tout } k \in \mathbb{N}.$$

Prop 20 On a alors
- $E[X] = \lambda = \mathrm{var}(X)$
- $G_X(s) = \exp(\lambda(s-1))$

Thm 21 (convergence de Lévy) Si (X_n) et X sont à valeurs dans $\mathbb{N}$,
X_n converge en loi vers X ssi G_{X_n} converge simplement vers G_X.

Prop 22 (évènements rares de Poisson)
Si X_n suit une loi binomiale $\mathcal{B}(n, p_n)$ où $p_n \sim \frac{\lambda}{n}$, avec $\lambda > 0$,
alors X_n converge en loi vers une variable de loi de Poisson $\mathcal{P}(\lambda)$.

Cela justifie que les variables suivantes sont souvent supposées suivre une loi de Poisson :
- Le nombre de clients d'un magasin pendant une journée.
- Le nombre de suicides d'enfants dans une année.
- etc.

Voici une version plus précise et générale de prop 22. DEV

Thm 23 (inégalité de Le Cam) Soit $(X_{1,n}, \dots, X_{n,n})_{n \geq 1}$ un tableau de variables de Bernoulli indépendantes, de paramètres respectifs $p_{k,n}$.
Si $S_n = \sum_{k=1}^{n} X_k$ et $\lambda_n = E[S_n]$, on a pour tout $A \subseteq \mathbb{N}$,
$$\left| P(S_n \in A) - \sum_{k \in A} \frac{\lambda_n^k e^{-\lambda_n}}{k!} \right| \leq \sum_{k=1}^{n} p_{k,n}^2$$

Ainsi, lorsque $\lambda_n \to \lambda$ et $\lim_n \sum_{k=1}^{n} p_{k,n}^2 = 0$,
S_n suit approximativement une loi de Poisson $\mathcal{P}(\lambda)$.

Inégalité de Le Cam

Mon développement

Soient $\lambda > 0$ et un entier $n > \lambda$. On note $B_{n,\lambda}$ la loi binomiale de paramètres n et $\frac{\lambda}{n}$ et P_λ la loi de Poisson de paramètre λ.

Commençons par quelques rappels. Les fonctions caractéristiques des lois $B_{n,\lambda}$ et P_λ sont données respectivement par : $\forall t \in \mathbb{R}$, $\widehat{B_{n,\lambda}}(t) = (1 + \frac{\lambda}{n}(e^{it} - 1))^n$ et $\widehat{P_\lambda}(t) = \exp(\lambda(e^{it} - 1))$. On en déduit que $(\widehat{B_{n,\lambda}})_{n\in\mathbb{N}}$ converge simplement sur $\mathbb{R}$ vers $\widehat{P_\lambda}$, donc que $(B_{n,\lambda})_{n\in\mathbb{N}}$ converge étroitement vers P_λ d'après le théorème de Lévy. En un certain sens, on va chercher à préciser la vitesse de convergence.
Remarquons qu'à l'aide des fonctions caractéristiques, on retrouve facilement la propriété $P_\lambda * P_\mu = P_{\lambda+\mu}$, qui nous sera utile plus tard.

Proposition. Soient $\lambda > 0$ et un entier $n > \lambda$. On a :

$$\sum_{k=0}^{+\infty} |B_{n,\lambda}(k) - P_\lambda(k)| \leq \frac{4\lambda^2}{n}$$

Pour tout $p \in [0, 1]$, on note μ_p le couplage défini par le tableau suivant :

$(0,0)$	$e^{-p} - p(1 - e^{-p})$
$(0,1)$	$p(1 - e^{-p})$
$(1,0)$	0
$(1,1)$	pe^{-p}
$(k,0)$	$e^{-p}p^k/k!$
$(k,1)$	0

pour tout $k \geq 2$. On vérifie facilement qu'on définit ainsi une loi de probabilité sur $\mathbb{N} \times \{0, 1\}$. De plus, si (X, Y) est un couple de variables aléatoires de loi μ_p, on vérifie facilement que X et Y suivent respectivement une loi de Poisson de paramètre p et une loi de Bernoulli de paramètre p.
On a de plus

$$\mathbb{P}(X = Y) = e^{-p} - p + 2pe^{-p} \geq -p + (1 + 2p)(1 - p) = 1 - 2p^2$$

en utilisant l'inégalité de convexité de l'exponentielle, donc

$$\mathbb{P}(X \neq Y) \leq 2p^2 . \tag{1}$$

Soient (X, Y) un couple de variables aléatoires dans $\mathbb{N}^2$ et $A \subset \mathbb{N}$.
On a

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in A,\ X = Y) + \mathbb{P}(X \in A,\ X \neq Y) \leq \mathbb{P}(Y \in A) + \mathbb{P}(X \neq Y)$$

et de même, $\mathbb{P}(Y \in A) \leq \mathbb{P}(X \in A) + \mathbb{P}(X \neq Y)$. Donc :

$$|\,\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)| \leq \mathbb{P}(X \neq Y) . \tag{2}$$

Soient $\lambda > 0$ et un entier $n > \lambda$. Soient $(X_1, Y_1), \dots, (X_n, Y_n)$ des variables aléatoires i.i.d. de loi $\mu_{\lambda/n}$ et $A \subset \mathbb{N}$. Puisque $X_1 + \dots + X_n$ suit la loi $B_{n,\lambda}$ et $Y_1 + \dots + Y_n$ suit la loi P_λ, on a :

$$\begin{aligned}
|B_{n,\lambda}(A) - P_\lambda(A)| &= |\mathbb{P}(Y_1 + \dots + Y_n \in A) - \mathbb{P}(X_1 + \dots + X_n \in A)| \\
&\leq \mathbb{P}(X_1 + \dots + X_n \neq Y_1 + \dots + Y_n) \quad \text{(d'après (2))} \\
&\leq \mathbb{P}(\exists i \in [\![1, n]\!],\ X_i \neq Y_i) \\
&\leq \sum_{i=1}^{n} \mathbb{P}(X_i \neq Y_i) \\
&\leq \sum_{i=1}^{n} 2\left(\frac{\lambda}{n}\right)^2 \quad \text{(d'après (1))} \\
&= \frac{2\lambda^2}{n}.
\end{aligned}$$

Pour conclure, il suffit d'appliquer cette inégalité aux ensembles $A_1 = \{k \in \mathbb{N} \mid B_{n,\lambda}(k) \geq P_\lambda(k)\}$ et $A_2 = \mathbb{N} \setminus A_1$. On obtient alors :

$$\begin{aligned}
\sum_{k=0}^{+\infty} |B_{n,\lambda}(k) - P_\lambda(k)| &= \sum_{k \in A_1} (B_{n,\lambda}(k) - P_\lambda(k)) + \sum_{k \in A_2} (P_\lambda(k) - B_{n,\lambda}(k)) \\
&= \sum_{k \in A_1} B_{n,\lambda}(k) - \sum_{k \in A_1} P_\lambda(k) + \sum_{k \in A_2} P_\lambda(k) - \sum_{k \in A_2} B_{n,\lambda}(k) \\
&= B_{n,\lambda}(A_1) - P_\lambda(A_1) + P_\lambda(A_2) - B_{n,\lambda}(A_2) \\
&\leq \frac{2\lambda^2}{n} + \frac{2\lambda^2}{n} = \frac{4\lambda^2}{n}
\end{aligned}$$

□

Références

J'ai utilisé [Let82, p. 95]. On peut aussi consulter [Ouv07, p. 220], [Dur10, p. 137], [Car07, p. 78], [Shi84, p. 345].

Leçons correspondantes

J'utilise ce développement pour la leçon 252. On peut également l'utiliser pour la leçon 249.

Remarques

- On peut détailler plus les calculs de fonctions caractéristiques et leurs conséquences dans l'introduction du développement.
- Pour un énoncé du théorème des évènements rares, voir [Ouv09, p. 310]. Plus généralement, la loi binomiale de paramètres (n, p_n) converge en loi vers la loi de Poisson de paramètre λ lorsque $\lim_{n \to +\infty} np_n = \lambda > 0$.
- Au lieu de dire « couplage », on peut simplement parler de loi d'un couple. On peut également construire le couplage « à la main » en revenant aux ω.

– Il est également possible d'utiliser le couplage μ_p défini par

$(0,0)$	$1-p$
$(0,1)$	$e^{-p}-(1-p)$
$(k,0)$	0
$(k,1)$	$e^{-p}p^k/k!$

pour tout $k \geq 1$. Dans ce cas, on obtient $\mathbb{P}(X \neq Y) \leq p^2$ et donc

$$\sum_{k=0}^{+\infty} |B_{n,\lambda}(k) - P_\lambda(k)| \leq \frac{2\lambda^2}{n} .$$

– La majoration optimale est en fait $\frac{2\lambda}{n}\min(2,\lambda)$.
– Il existe un résultat plus fort (mais difficile à trouver dans la littérature) : le théorème de Prohorov. Soit $\lambda > 0$, on a :

$$\lim_{n\to+\infty} n\sum_{k=0}^{+\infty} |B_{n,\lambda}(k) - P_\lambda(k)| = \frac{1}{2}\,\mathbb{E}(|X-(X-\lambda)^2|)$$

où X est une variable aléatoire de loi P_λ. Ce résultat se démontre à l'aide du théorème de convergence dominée pour les séries numériques.

Questions possibles

1. Montrer que $P_\lambda * P_\mu = P_{\lambda+\mu}$.
2. Donner la définition de la fonction caractéristique d'une variable aléatoire et donner ses propriétés.
3. Donner la définition d'un couplage.
4. Citer une autre approximation classique de loi.
5. En pratique, pour quelles valeurs des paramètres approche-t-on la loi binomiale par la loi de Poisson ?

Références

[Car07] Hervé CARRIEU : *Probabilité : exercices corrigés.* EDP Sciences, 2007.

[Dur10] Richard DURRETT : *Probability, theory and examples.* Cambridge University Press, 2010.

[Let82] Gérard LETAC : *Intégration et probabilités : analyse de Fourier et analyse spectrale. Exercices.* Masson, 1982.

[Ouv07] Jean-Yves OUVRARD : *Probabilités 1.* Cassini, 2007.

[Ouv09] Jean-Yves OUVRARD : *Probabilités 2.* Cassini, 2009.

[Shi84] Albert SHIRYAYEV : *Probability.* Springer, 1984.

Théorème de Bernstein.

2013 – 2014

Référence : Claude Zuily, Hervé Queffélec, *Analyse pour l'agrégation*, Dunod, 2006, p.518.

Théorème.
Soit $f : [0,1] \to \mathbb{C}$ une fonction continue. On pose

$$\omega(h) := \sup\{|f(u) - f(v)|, |u - v| \leq h\}$$

son module de continuité.
On considère

$$B_n(f,x) = B_n(x) := \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

le $n^{ième}$ polynôme de Bernstein de f.
Alors

(i) *(B_n) converge uniformément vers f sur $[0,1]$.*

(ii) $\|f - B_n\|_\infty \leq C\omega\left(\frac{1}{\sqrt{n}}\right)$ *où C est une constante.*

(iii) *L'estimation (ii) est optimale : il existe f lipschitzienne telle que $\|f - B_n\|_\infty \geq \frac{\delta}{\sqrt{n}}$ pour une constante $\delta > 0$.*

Démonstration. (i) Soit $x \in [0,1]$ et $(X_i)_{i\geq 1}$ une suite de variables de Bernoulli i.i.d de paramètre x. On note $S_n := X_1 + \cdots + X_n$.
On a alors

$$\mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right) = B_n(x) \text{ et } \mathbb{E}(f(x)) = f(x)$$

par théorème de transfert, d'où

$$\begin{aligned} |f(x) - B_n(x)| &= \left|\mathbb{E}\left(f(x) - f\left(\frac{S_n}{n}\right)\right)\right| \\ &\leq \mathbb{E}\left(\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right). \end{aligned}$$

f est continue sur $[0,1]$ compact donc uniformément continue par le théorème de Heine, donc $\omega(\delta)$ est défini pour tout $\delta > 0$ et $\lim_{\delta\to 0}\omega(\delta) = 0$. Par ailleurs, $\left|f(x) - f\left(\frac{S_n}{n}\right)\right| \leq 2\|f\|_\infty$ donc, pour $\delta > 0$,

$$\mathbb{E}\left|f(x) - f\left(\frac{S_n}{n}\right)\right| \leq \omega(\delta)\mathbb{P}\left(\left|x - \frac{S_n}{n}\right| \leq \delta\right) + 2\|f\|_\infty \mathbb{P}\left(\left|x - \frac{S_n}{n}\right| > \delta\right).$$

Par l'inégalité de Tchebychev, on a alors

$$\begin{aligned}\mathbb{P}\left(\left|x - \frac{S_n}{n}\right| > \delta\right) &\leq \frac{\operatorname{Var}\left(\frac{S_n}{n}\right)}{\delta^2} \\ &= \frac{\operatorname{Var}(S_n)}{n^2\delta^2} \\ &= \frac{x(1-x)}{n\delta^2} \\ &\leq \frac{1}{4n\delta^2}.\end{aligned}$$

On en déduit que, pour tout $x \in [0,1]$,

$$|f(x) - B_n(x)| \leq \omega(\delta) + \frac{\|f\|_\infty}{2n\delta^2},$$

d'où

$$\limsup_{n\to+\infty} \|f - B_n\|_\infty \leq \omega(\delta).$$

Le résultat vient de $\lim_{\delta\to 0}\omega(\delta) = 0$.

(ii) Affinons le résultat de convergence uniforme prouvé ci-dessus. On a d'abord

$$\mathbb{E}\left|f(x) - f\left(\frac{S_n}{n}\right)\right| \leq \mathbb{E}\omega\left(\left|x - \frac{S_n}{n}\right|\right).$$

Montrons que $\omega(\lambda h) \leq (\lambda+1)\omega(h)$:
ω est croissante et $\omega(h+k) \leq \omega(h)+\omega(k)$ donc, par récurrence, $\omega(nh) \leq n\omega(h)$ pour $n \in \mathbb{N}$. On a alors, pour $\lambda \in \mathbb{R}$, $\omega(\lambda h) \leq \omega(\lceil\lambda\rceil h) \leq \lceil\lambda\rceil\omega(h) \leq (\lambda+1)\omega(h)$.
On en déduit

$$\begin{aligned}|f(x) - B_n(x)| &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\mathbb{E}\left(\sqrt{n}\left|x - \frac{S_n}{n}\right| + 1\right) \\ &= \omega\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\left\|x - \frac{S_n}{n}\right\|_1\right) \\ &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\left\|x - \frac{S_n}{n}\right\|_2\right)\end{aligned}$$

par l'inégalité de Hölder. Or

$$\begin{aligned}\left\|x-\frac{S_n}{n}\right\|_2^2 &= \mathbb{E}\left(\left|x-\frac{S_n}{n}\right|^2\right)\\ &= \operatorname{Var}\left(x-\frac{S_n}{n}\right)+\left(\mathbb{E}\left(x-\frac{S_n}{n}\right)\right)^2\\ &= \frac{1}{n^2}nx(1-x)+\left(x-\frac{1}{n}nx\right)^2\\ &= \frac{x(1-x)}{n}.\end{aligned}$$

D'où

$$\begin{aligned}|f(x)-B_n(x)| &\le \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\sqrt{\frac{x(1-x)}{n}}\right)\\ &\le \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

D'où $\|f-B_n\|_\infty \le \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)$.

(iii) On pose $f : x \mapsto \left|x-\frac{1}{2}\right|$, on a $\omega(h) \le h$. Par ailleurs,

$$\begin{aligned}\|f-B_n\|_\infty &\ge \left|f\left(\frac{1}{2}\right)-B_n\left(\frac{1}{2}\right)\right|\\ &= \left|B_n\left(\frac{1}{2}\right)\right|\\ &= \mathbb{E}\left|\frac{S_n}{n}-\frac{1}{2}\right|\\ &= \frac{1}{2n}\mathbb{E}|2S_n-n|.\end{aligned}$$

D'où $\|f-B_n\|_\infty \ge \frac{1}{2n}\mathbb{E}|\varepsilon_1+\cdots+\varepsilon_n|$ avec $\varepsilon_j := 2X_j-1$ variables de Rademacher i.i.d. D'où

$$\begin{aligned}\|f-B_n\|_\infty &\ge \frac{1}{2n}\|\varepsilon_1+\cdots+\varepsilon_n\|_1\\ &\ge \frac{1}{2n\sqrt{e}}\|\varepsilon_1+\cdots+\varepsilon_n\|_2\end{aligned}$$

par inégalité de Khintchine.
Or $\|\varepsilon_1+\cdots+\varepsilon_n\|_2^2 = \operatorname{Var}(\varepsilon_1+\cdots+\varepsilon_n)+(\mathbb{E}(\varepsilon_1+\cdots+\varepsilon_n))^2 = n$. D'où

$$\|f-B_n\|_\infty \ge \frac{1}{2\sqrt{ne}} \ge \frac{1}{2\sqrt{e}}\omega\left(\frac{1}{\sqrt{n}}\right).$$

□

Détails supplémentaires

Proposition (Inégalité de Khintchine).
*Soit $\varepsilon_1, \ldots, \varepsilon_n$ des variables de Rademacher (*i.e.* valant ± 1 avec probabilité $\frac{1}{2}$) i.i.d. Soit $f \in vect_{\mathbb{R}}(\varepsilon_1, \ldots, \varepsilon_n)$.*
Alors $\|f\|_2 \leq \sqrt{e}\mathbb{E}(|f|)$.

Démonstration. On a $f = \sum_j a_j \varepsilon_j$ et on peut supposer $\|f\|_2^2 = 1 = \sum_j a_j^2$.
Posons $g := \prod_{j=1}^n (1 + i a_j \varepsilon_j)$. Alors pour presque tout x,

$$\begin{aligned}
|g(x)| &= \prod_{j=1}^n \sqrt{1 + a_j^2 \varepsilon_j^2(x)} \\
&= \prod_{j=1}^n \sqrt{1 + a_j^2} \\
&\leq \prod_{j=1}^n \sqrt{\exp(a_j^2)} \\
&= \sqrt{\exp\left(\sum a_j^2\right)} \\
&= \sqrt{e}.
\end{aligned}$$

D'où $\|g\|_\infty \leq \sqrt{e}$.
De plus, si $j \in \{1, \ldots, n\}$,

$$\begin{aligned}
\mathbb{E}(\varepsilon_j g) &= \mathbb{E}\left(\varepsilon_j (1 + i a_j \varepsilon_j) \prod_{k \neq j} (1 + i a_k \varepsilon_k)\right) \\
&= \mathbb{E}(\varepsilon_j (1 + i a_j \varepsilon_j)) \mathbb{E}\left(\prod_{k \neq j} (1 + i a_k \varepsilon_k)\right) \quad \text{par indépendance des } \varepsilon_j
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\varepsilon_j g) &= \mathbb{E}(\varepsilon_j (1 + i a_j \varepsilon_j)) \prod_{k \neq j} \mathbb{E}(1 + i a_k \varepsilon_k) \\
&= i a_j \text{ car } \mathbb{E}(\varepsilon_j) = 0.
\end{aligned}$$

Or $\mathbb{E}(fg) = \sum_{j=1}^n a_j \mathbb{E}(\varepsilon_j g)$, d'où $|\mathbb{E}(fg)| = \left| i \sum_{j=1}^n a_j^2 \right| = 1$.
Or

$$\|f\|_1 \geq \frac{|\mathbb{E}(fg)|}{\|g\|_\infty} \geq \frac{1}{\sqrt{e}}.$$

□