PROBABILITÉS ET SOMMES DE RIEMANN

www.daniel-saada.eu

Source : http://www.les-mathematiques.net/phorum/read.php?4,783579,783579#msg-783579

Soit $f(x) = \operatorname{int}\left(\frac{2}{x} - 2\operatorname{int}\left(\frac{1}{x}\right)\right)$ et $R_n(f) = \frac{1}{n}\sum_{k=1}^{n} f(k/n)$: on veut prouver la convergence des sommes de Riemann vers $I(f) = \int_0^1 f$. La fonction f est la fonction indicatrice de $A = \bigcup_1^{\infty} \left]\frac{1}{n+1}, \frac{2}{2n+1}\right]$; comme les $A_n = \left]\frac{1}{n+1}, \frac{2}{2n+1}\right]$ sont disjoints : $I(f)$ existe car $\sum_1^{\infty} m(A_n) < +\infty$, $f = \sum_1^{\infty} 1_{A_n}$.

On précise que int désigne la partie entière et $m(X)$ la mesure du borélien X de $[0,1]$.

Méthode 1

On utilise deux théorèmes :

1) Une fonction bornée est intégrable au sens de Riemann si et seulement si l'ensemble de ses discontinuités est de mesure nulle (caractérisation de Lebesgue des fonctions Riemann-intégrables[1]).

f est bornée et n'a qu'un nombre dénombrable de discontinuités : elle est donc intégrable au sens de Riemann.

2) Les sommes de Riemann d'une fonction intégrable au sens de Riemann convergent vers l'intégrale de la fonction[2].

La convergence de $R_n(f)$ vers $I(f)$ est donc assurée. Calculons maintenant $I(f) = \sum_1^{\infty}\left(\frac{2}{2n+1} - \frac{1}{n+1}\right)$.

D'abord $\sum_1^{N}\left(\frac{2}{2n+1} - \frac{1}{n+1}\right) = \sum_1^{N}\frac{2}{2n+1} - \sum_2^{N+1}\frac{1}{n} = \sum_3^{2N+1}\frac{2}{n} - \sum_2^{N}\frac{1}{n} - \sum_2^{N+1}\frac{1}{n} = 2\sum_{N+1}^{2N+1}\frac{1}{n} - 1 - \frac{1}{N+1}$.

Ensuite $\sum_{N+1}^{2N+1}\frac{1}{n} = \frac{1}{2N+1} + \frac{1}{N}\sum_1^{N}\frac{1}{1+n/N}$ a pour limite $\int_0^1 \frac{dx}{1+x} = \ln 2$ quand N tend vers l'infini (c'est une somme de Riemann classique !). On en déduit $I(f) = 2\ln 2 - 1$.

Méthode 2 (Daniel Saada)

1) Pour tout entier k , $R_n(1_{A_k}) \underset{n\to\infty}{\to} I(1_{A_k})$, car chaque A_k est un intervalle.

2) Par linéarité, $R_n\left(\sum_1^{N} 1_{A_k}\right) \underset{n\to\infty}{\to} I\left(\sum_1^{N} 1_{A_k}\right)$ pour tout entier N .

3) *Le passage à la limite*

On pose $f_N = \sum_1^{N} 1_{A_k}$ et on écrit $R_n(f) - I(f) = R_n(f_N) - I(f_N) + R_n(f - f_N) + I(f_N - f)$.

D'où $|R_n(f) - I(f)| \le |R_n(f_N) - I(f_N)| + R_n(f - f_N) + I(f - f_N)$.

[1] Jean-Marie Arnaudiès, *L'intégrale de Lebesgue*, Vuibert, 1997, § 2.16.

[2] http://perso.univ-rennes1.fr/jean-christophe.breton/Fichiers/Integration_Mesures.pdf

a) $I(f-f_N)=\sum_{N+1}^{\infty}m(A_n)\leq\dfrac{2}{2N+3}$ car $\sum_{N+1}^{\infty}A_n\subset\left]0,\dfrac{2}{2N+3}\right]$.

b) On se donne $\varepsilon>0$ et on fixe N tel que $2N+3\geq 2/\varepsilon$.

$R_n(f-f_N)=\dfrac{\theta_n}{n}$, où θ_n est un entier au plus égal à l'entier p_n défini par $\dfrac{p_n}{n}\leq\dfrac{2}{2N+3}<\dfrac{1+p_n}{n}$.

On a donc $R_n(f-f_N)\leq\dfrac{2}{2N+3}\leq\varepsilon$.

c) Pour $n\geq n_0$, $|R_n(f_N)-I(f_N)|\leq\varepsilon$ en vertu de **2)** car N est fixé.

Conclusion : $|R_n(f)-I(f)|\leq 3\varepsilon$ pour tout $n\geq n_0$, et donc $R_n(f)\to I(f)$.

Méthode 3 (Benoit Cloitre)

a) $nR_n(f)=\displaystyle\sum_{i=1}^{n}\left(\operatorname{int}\left(\frac{2n}{2i+1}\right)-\operatorname{int}\left(\frac{2n}{2i+2}\right)\right)$.

$R_n(f)=\dfrac{1}{n}\sum_{k=1}^{n}f(k/n)$, avec $f(k/n)=0$ si $\dfrac{2}{2i+1}<\dfrac{k}{n}\leq\dfrac{2}{2i}$ et $f(k/n)=1$ si $\dfrac{2}{2i+2}\leq\dfrac{k}{n}<\dfrac{2}{2i+1}$.

Le nombre d'entiers k vérifiant $\dfrac{2}{2i+2}\leq\dfrac{k}{n}<\dfrac{2}{2i+1}$ est exactement $\operatorname{int}\left(\dfrac{2n}{2i+1}\right)-\operatorname{int}\left(\dfrac{2n}{2i+2}\right)$, ce qui justifie la formule annoncée.

b) $nR_n(f)=-n+2n\displaystyle\sum_{1}^{\sqrt{2n}}\frac{(-1)^{j-1}}{j}+O(\sqrt{n})$.

On écrit $nR_n(f)$ sous la forme d'une somme alternée :

$$\sum_{i=1}^{n}\left(\operatorname{int}\left(\frac{2n}{2i+1}\right)-\operatorname{int}\left(\frac{2n}{2i+2}\right)\right)=\sum_{3}^{2n}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)=-n+\sum_{1}^{2n}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right).$$

On scinde $\displaystyle\sum_{1}^{2n}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)$ en $\displaystyle\sum_{1\leq j\leq\sqrt{n}}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)+\sum_{\sqrt{n}<j\leq n}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)$:

− comme $\operatorname{int}(x)=x+O(1)$, $\displaystyle\sum_{1\leq j\leq\sqrt{n}}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)=n\sum_{1\leq j\leq\sqrt{n}}\frac{(-1)^{j-1}}{j}+O(\sqrt{n})$;

− la deuxième somme est alternée avec des modules décroissants, aussi $\displaystyle\left|\sum_{\sqrt{n}<j\leq n}(-1)^{j-1}\operatorname{int}\left(\frac{2n}{j}\right)\right|=O(\sqrt{n})$ également.

L'égalité est donc prouvée.

c) $R_n(f)$ a pour limite $2\ln 2-1$.

Il suffit de se rappeler que $\displaystyle\sum_{1}^{\infty}\frac{(-1)^{j-1}}{j}=\ln 2$.

ORIGINE PROBABILISTE DE LA QUESTION

Soit $n \geq 2$. Pour $k \in \{1,2,...,n\}$, soit r_k le reste de la division de n par k. Le tirage suivant une loi uniforme, la probabilité p_n de l'événement $r_k \geq \frac{k}{2}$ est $R_n(f) = \frac{1}{n}\sum_{k=1}^{n} f(k/n)$. En effet,

$$r_k \geq \frac{k}{2} \text{ équivaut à } f(k/n) = 1 \text{ tandis que } r_k < \frac{k}{2} \text{ conduit à } f(k/n) = 0.$$

La limite de p_n existe et vaut $\ln(4) - 1 \approx 38,63\,\%$.

Équivalents de l'espérance et de la variance de la variable aléatoire $X_n : k \mapsto r_k$ (Claude Morin)

• Comme la distribution de X_n est difficile à atteindre, on partira de $E(X_n) = \frac{1}{n}\sum_{k=1}^{n} r_k$. D'où :

$E(X_n) = \frac{1}{n}\sum_{k=1}^{n}\left(n - k\,\text{int}\left(\frac{n}{k}\right)\right) = n - \sum_{1}^{n}\frac{k}{n}\text{int}\left(\frac{n}{k}\right)$; pour les raisons que j'ai dites, la somme de Riemann $\frac{1}{n}\sum_{1}^{n}\frac{k}{n}\text{int}\left(\frac{n}{k}\right)$ converge vers $\int_0^1 x\,\text{int}(1/x)dx = \frac{1}{2}\sum_{1}^{\infty}\left(\frac{1}{(p+1)^2} + \frac{1}{p(p+1)}\right) = \frac{\pi^2}{12}$.

Aussi, $\frac{1}{n}E(X_n) \to 1 - \frac{\pi^2}{12} \approx 0,18$.

Vérification. On calcule les premières valeurs de la somme des r_k : (0, 0, 1, 1, 4, 3, 8, 8, 12, 13) et on les soumet au moteur de recherche de l'OEIS qui renvoie la suite référencée **A004125**, dont l'URL est :

http://oeis.org/search?q=A4125&language=french&go=Chercher

Il y est bien affirmé que la moyenne des r_k est équivalente pour n infini à $n\left(1 - \frac{\pi^2}{12}\right)$.

• $E(X_n^2) = \frac{1}{n}\sum_{k=1}^{n}\left(n - k\,\text{int}\left(\frac{n}{k}\right)\right)^2 = n^2 - 2\sum_{1}^{n} k\,\text{int}\left(\frac{n}{k}\right) + \sum_{1}^{n}\frac{k^2}{n}\left(\text{int}\left(\frac{n}{k}\right)\right)^2$, d'où

$\frac{E(X_n^2)}{n^2} = 1 - \frac{2}{n}\sum_{1}^{n}\frac{k}{n}\text{int}\left(\frac{n}{k}\right) + \frac{1}{n}\sum_{1}^{n}\frac{k^2}{n^2}\left(\text{int}\left(\frac{n}{k}\right)\right)^2$ qui converge vers $1 - \frac{\pi^2}{6} + \int_0^1 x^2\left(\text{int}\left(\frac{1}{x}\right)\right)^2 dx$.

$$\int_0^1 x^2\left(\text{int}\left(\frac{1}{x}\right)\right)^2 dx = \sum_{1}^{\infty}\frac{p^2}{3}\left(\frac{1}{p^3} - \frac{1}{(p+1)^3}\right) = \frac{1}{3}\sum_{1}^{\infty}\left(\frac{2}{(p+1)^2} + \frac{1}{p(p+1)} - \frac{1}{(p+1)^3}\right)$$
$$= \frac{1}{3}\left(\frac{\pi^2}{3} - 2 + 1 - \sum_{1}^{\infty}\frac{1}{p^3} + 1\right) = \frac{\pi^2}{9} - \frac{1}{3}\sum_{1}^{\infty}\frac{1}{p^3} \approx 0,696.$$

Il est à noter que Maple, dans sa version 11 tout au moins, affiche $0,714$.

En conclusion : $\frac{E(X_n^2)}{n^2} \to 1 - \frac{\pi^2}{18} - \frac{1}{3}\sum_{1}^{\infty}\frac{1}{p^3} \approx 0,051$; comme $V(X) = E(X^2) - (E(X))^2$,

$$\frac{V(X_n)}{n^2} \to \frac{\pi^2}{9} - \frac{\pi^4}{144} - \frac{1}{3}\sum_{1}^{\infty}\frac{1}{p^3} \approx 0,0195.$$