Processus empirique basé sur des U-statistiques à deux échantillons

Davide Giraudo
Euler International Mathematical Institute (Saint-Pétersbourg)
en collaboration avec Herold Dehling et Olimjon Sharipov

Séminaire Probabilités, Statistique et Applications
Poitiers, 3 décembre 2020

U-statistique empirique à un échantillon

La convergence de la U-statistique empirique à un échantillon définie par

$$\frac{1}{n^{3/2}} \sum_{1 \leqslant i \neq j \leqslant n} \left(\mathbf{1}\left\{ g\left(X_i, X_j\right) \leqslant s \right\} - \mathbb{P}\left\{ g\left(X_i, X_j\right) \leqslant s \right\} \right)$$

a été traitée (Svetlana Borovkova, Robert Burton et Herold Dehling).

Si $(X_i)_{i \geqslant 1}$ est strictement stationnaires et vérifie certaines conditions de dépendance, alors la convergence a lieu dans $D[0,1]$ vers un processus gaussien dont la fonction de covariance est explicite.

U-statistique à deux échantillons

La convergence de la U-statistique à deux échantillons de noyau h défini par

$$T_n(h,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left(h(X_i, X_j) - \mathbb{E}\left[h(X_i, X_j)\right]\right)$$

a également été traitée (Herold Dehling, Roland Fried, Isabel Garcia et Martin Wendler) où $(X_i)_{i \geqslant 1}$ est strictement stationnaires et vérifie certaines conditions de dépendance.

La convergence a lieu dans $D[0,1]$ vers un processus gaussien dont la fonction de covariance est explicite.

Idée : on divise pour chaque $t \in [0,1]$ l'échantillon $X_1, \ldots, X_n$ en deux échantillons $X_1, \ldots, X_{[nt]}$ et $X_{[nt]+1}, \ldots, X_n$ et on considère des quantités du type $h(X, Y)$, où X est dans le premier échantillon et Y le second.

U-statistique empirique à deux échantillons

Pour $n \geqslant 1, 0 \leqslant t \leqslant 1, s \in \mathbb{R}$, on définit

$$e_n(s,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \mathbf{1}\{g(X_i, X_j) \leqslant s\},$$

où $g : \mathbb{R}^2 \to \mathbb{R}$ est une fonction mesurable et pour tout réel x, $[x]$ désigne l'unique entier k tel que $k \leqslant x < k+1$.

Lien avec les U-statistiques à deux échantillons:

$$T_n(h,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h(X_i, X_j) :$$

$e_n(s,t) = T_n(h_s, t)$ où $h_s(u,v) = \mathbf{1}\{g(X_i, X_j) \leqslant s\}$.

Objectif

Soit $(X_i)_{i\geqslant 1}$ une suite i.i.d. et

$$e_n(s,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \mathbf{1}\{g(X_i, X_j) \leqslant s\}.$$

Trouver une condition portant sur g et $(X_i)_{i\geqslant 1}$ telle que

$$e_n(s,t) - \mathbb{E}[e_n(s,t)] \to W(s,t)$$

en loi dans $D([-R,R]\times[0,1])$ vers un processus $W(\cdot,\cdot)$ et fournir une description de W.

Résultat

Theorem (Dehling, G., Sharipov (2020+))

Soit $(X_i)_{i\geqslant 0}$ *une suite i.i.d. et* e_n *défini par*

$$e_n(s,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \mathbf{1}\{g(X_i, X_j) \leqslant s\}$$

On suppose que g *est symétrique et que pour tout* $u \in \mathbb{R}$, *la variable aléatoire* $g(u, X_1)$ *a une densité* f_u *et que* $\sup_{u,x} f_u(x) < +\infty$. *Alors pour tout* $R > 0$,

$$e_n(s,t) - \mathbb{E}[e_n(s,t)] \to W(s,t) \text{ en loi dans } D([-R,R]\times[0,1]),$$

où $(W(s,t), s \in \mathbb{R}, t \in [0,1])$ *est un processus gaussien centré, dont la covariance est donnée pour* $0 \leqslant t \leqslant t' \leqslant 1$ *et* $s, s' \in \mathbb{R}$ *par :*

$$\mathrm{Cov}(W(s,t), W(s',t')) = t(1-t')(1-2t+2t')\, c_{s,s'},$$

où $c_{s,s'} = \mathbb{E}[h_{1,s}(X_1)\, h_{1,s'}(X_1)]$ *et*
$h_{1,s}(u) = \mathbb{P}\{g(u, X_1) \leqslant s\} - \mathbb{P}\{g(X_1, X_2) \leqslant s\}$.

Exemples

1. Soit $g(u, v) = u + v$ et supposons que X_1 aie une densité f bornée par M. La densité de $g(u, X_1) = u + X_1$ est $f_u(x) = f(x - u)$. De plus,
$$h_{1,s}(u) = \mathbb{P}\{X_1 \leqslant s - u\} - \mathbb{P}\{X_1 + X_2 \leqslant s\}.$$
2. Le choix $g(u, v) = u - v$ est possible (permet d'utiliser la symétrie de $X_i - X_j$).

Le cas non-symétrique

Les hypothèses sont les suivantes

1. la variable aléatoire $g(u, X_1)$ a une densité $f_{1,u}$ et $\sup_{u,x} f_{1,u}(x) < +\infty$.

2. La variable aléatoire $g(X_1, v)$ a une densité $f_{2,v}$ et $\sup_{u,x} f_{2,v}(x) < +\infty$.

La covariance du processus (gaussien centré) limite est donnée pour $t \leqslant t'$ par

$$\begin{aligned}\operatorname{Cov}(W(s,t), W(s',t')) = {} & t(1-t)(1-t')\,\mathbb{E}[h_{1,s}(X_0)\,h_{1,s'}(X_0)] \\ & + t(1-t')(t'-t)\,\mathbb{E}[h_{2,s}(X_0)\,h_{1,s'}(X_0)] \\ & + tt'(1-t')\,\mathbb{E}[h_{2,s}(X_0)\,h_{2,s'}(X_0)],\end{aligned}$$

où

$$h_{1,s}(u) = \mathbb{P}\{g(u, X_1) \leqslant s\} - \mathbb{P}\{g(X_1, X_2) \leqslant s\},$$

$$h_{2,s}(v) = \mathbb{P}\{g(X_1, v) \leqslant s\} - \mathbb{P}\{g(X_1, X_2) \leqslant s\}.$$

Exemples de fonctionnelles

1. Pour tout $R > 0$, la convergence en loi

$$\sup_{0 \leqslant t \leqslant 1} \sup_{-R \leqslant s \leqslant R} |e_n(s,t)| \to \sup_{0 \leqslant t \leqslant 1} \sup_{-R \leqslant s \leqslant R} |W(s,t)|$$

   a lieu.

2. Soit μ une mesure finie sur les boréliens de $\mathbb{R}$.

   Alors

$$\sup_{0 \leqslant t \leqslant 1} \int_{\mathbb{R}} \left(e_n(s,t) - \mathbb{E}\left[e_n(s,t)\right]\right)^2 d\mu(s) \to \sup_{0 \leqslant t \leqslant 1} \int_{\mathbb{R}} W(s,t)^2 \, d\mu(s).$$

Idée de démonstration : aperçu global

Les étapes principales sont les suivantes :

1. On considère le noyau $h_s : \mathbb{R}^2 \to \mathbb{R}$ défini par $h_s(u, v) = \mathbf{1}\{g(u, v) \leqslant s\}$. La décomposition d'Hoeffding's donne: une U-statistique contenant une partie linéaire + un terme dit dégénéré.

2. On établit la convergence des lois fini-dimensionnelles de la partie linéaire vers celles du processus W.

3. Puis on démontre que le processus associé à la partie linéaire converge vers W dans $D([-R, R] \times [0, 1])$ pour tout $R > 0$.

4. Enfin, on montre que la contribution de la partie dégénérée est négligeable.

Décomposition d'Hoeffding classique

Soit $h\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction mesurable et $(X_i)_{i\geqslant 1}$ une suite i.i.d. Soit

$$U_n = \sum_{1\leqslant i<j\leqslant n} h(X_i, X_j).$$

On définit $\theta := \mathbb{E}[h(X_1, X_2)]$,

$$h_1(x) = \mathbb{E}[h(x, X_1)] - \theta, \quad h_2(y) = \mathbb{E}[h(X_1, y)] - \theta,$$

$$h_3(x, y) = h(x, y) - h_1(x) - h_2(y) - \theta.$$

Alors

$$U_n = \binom{n}{2}\theta + \sum_{i=1}^{n-1}(n-i)\,h_1(X_i) + \sum_{j=2}^{n}(j-1)\,h_2(X_j) + \sum_{1\leqslant i<j\leqslant n} h_3(X_i, X_j)$$

et

$$\mathbb{E}[h_3(X_i, X_j) \mid X_1, \ldots, X_{j-1}] = \mathbb{E}[h_3(X_i, X_j) \mid X_{i+1}, \ldots, X_n] = 0.$$

Décomposition d'Hoeffding

Soit $\theta_s := \mathbb{P}\{g(X_1, X_2) \leqslant s\}$, $h_s(u, v) = \mathbf{1}\{g(u, v) \leqslant s\}$

$$h_{1,s}(u) = \mathbb{P}\{g(u, X_1) \leqslant s\} - \theta_s, \quad h_{2,s}(v) = \mathbb{P}\{g(X_1, v) \leqslant s\} - \theta_s,$$

et

$$h_{3,s}(u, v) = h_s(u, v) - h_{1,s}(u) - h_{2,s}(v) - \theta_s.$$

Alors

$$e_n(s, t) - \mathbb{E}[e_n(s, t)] = R_n(s, t) + W_n(s, t),$$

où

$$R_n(s, t) = \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h_{3,s}(X_i, X_j)$$

$$W_n(s, t) = \frac{n - [nt]}{n^{3/2}} \sum_{i=1}^{[nt]} h_{1,s}(X_i) + \frac{[nt]}{n^{3/2}} \sum_{j=[nt]+1}^{n} h_{2,s}(X_j).$$

Lois fini-dimensionnelles

Cramer-Wold : soient $0 = t_0 < t_1 < \dots t_d \leqslant t_{d+1} = 1$, $s_k \geqslant 1$, $1 \leqslant k \leqslant d$. On doit montrer que

$$\sum_{k,\ell=1}^{d} a_{k,\ell} W_n(s_\ell, t_k) \to \sum_{k,\ell=1}^{d} a_{k,\ell} W(s_\ell, t_k) \text{ en loi.}$$

On exprime $\sum_{k,\ell=1}^{d} a_{k,\ell} W_n(s_\ell, t_k)$ comme une combination linéaire de variables aléatoires indépendantes à laquelle on applique un théorème limite central pour les tableaux de variables aléatoires.
Soient

$$I_{n,u} = \{i \in \mathbb{N} \mid [nt_{u-1}] + 1 \leqslant i \leqslant [nt_u]\}, 2 \leqslant u \leqslant d,$$

$I_{n,1} = \{i \in \mathbb{N} \mid 1 \leqslant i \leqslant [nt_1]\}$ et $I_{n,d+1} = \{i \in \mathbb{N} \mid [nt_d] + 1 \leqslant i \leqslant n\}$.
Alors l'égalité suivante a lieu:

$$\sum_{k,\ell=1}^{d} a_{k,\ell} W_n(s_\ell, t_k) = \sum_{k,\ell=1}^{d} a_{k,\ell} \frac{n - [nt_k]}{n^{3/2}} \sum_{u=1}^{d+1} \mathbf{1}\{u \leqslant k\} \sum_{i \in I_{n,u}} h_{1,s_\ell}(X_i) +$$

$$\sum_{k,\ell=1}^{d} a_{k,\ell} \frac{[nt_k]}{n^{3/2}} \sum_{u=1}^{d+1} \mathbf{1}\{u \geqslant k+1\} \sum_{i \in I_{n,u}} h_{2,s_\ell}(X_i).$$

Convergence de la partie linéaire : outil

Theorem (Davydov, Zitikis)

Soit ξ_n, $n \geqslant 1$ des processus stochastiques définis sur $[0,1]^2$ et dont les trajectoires se trouvent dans $D\left([0,1]^2\right)$. Supposons que :

1. *$\xi_n \to \xi$ (fin. dim.) et ξ a des trajectoires continues ;*
2. *on peut exprimer $\xi_n = \xi_n^\circ - \xi_n^*$ où ξ_n° et ξ_n^* sont croissants, $\xi_n(0,0) = 0$;*
3. *il existe des constantes $\gamma \geqslant \beta > 2$, $c \in (0,\infty)$ telles que pour tout $n \geqslant 1$, si $\|(s,t) - (s',t')\|_\infty \geqslant 1/n$, alors*

$$\mathbb{E}\left[\left|\xi_n(s,t) - \xi_n(s',t')\right|^\gamma\right] \leqslant c\,\|(s,t) - (s',t')\|_\infty^\beta$$

4. *la convergence en probabilité suivante a lieu :*

$$\max_{1\leqslant i,j\leqslant n}\left|\xi_n^*\left(\frac{i}{n},\frac{j}{n}\right) - \xi_n^*\left(\frac{i-1}{n},\frac{j}{n}\right)\right| + \max_{1\leqslant i,j\leqslant n}\left|\xi_n^*\left(\frac{i}{n},\frac{j}{n}\right) - \xi_n^*\left(\frac{i}{n},\frac{j-1}{n}\right)\right| \to 0.$$

Alors $(\xi_n)_{n\geqslant 1}$ converge en loi vers ξ dans $D\left([0,1]^2\right)$.

Convergence de la partie linéaire

1. Condition sur $\mathbb{E}\left[\left|\xi_n(s,t)-\xi_n(s',t')\right|^{\gamma}\right]$: inégalité de Rosenthal et en définissant

$$a_s(u) := \mathbb{P}\left\{g(u,X_1)\leqslant s\right\}, \quad b_s(v) := \mathbb{P}\left\{g(X_1,v)\leqslant s\right\},$$

$$\|a_s - a_{s'}\|_\infty \leqslant C|s-s'|, \|b_s - b_{s'}\|_\infty \leqslant C|s-s'| \qquad (*)$$

2. Condition

$$\max_{1\leqslant i,j\leqslant n}\left|\xi_n^*\left(\frac{i}{n},\frac{j}{n}\right)-\xi_n^*\left(\frac{i-1}{n},\frac{j}{n}\right)\right|+\max_{1\leqslant i,j\leqslant n}\left|\xi_n^*\left(\frac{i}{n},\frac{j}{n}\right)-\xi_n^*\left(\frac{i}{n},\frac{j-1}{n}\right)\right|\to 0,$$

Le processus ξ_n^* est donné par

$$\xi_n^*(s,t) := \frac{[nt]}{n^{3/2}}\sum_{i=1}^{[nt]} a_s(X_i) + \frac{[nt]}{n^{3/2}}\sum_{j=1}^{[nt]} b_s(X_j) + 2\theta_{-R+2Rs}\frac{[nt]}{n^{1/2}},$$

où $\theta_s = \mathbb{P}\left\{g(X_1,X_2)\leqslant s\right\}$ et on utilise à nouveau (*).

Traitement de la partie dégénérée (1)

Rappelons que

$$R_n(s,t) = \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h_{3,s}(X_i, X_j),$$

où, par definition de $h_{3,s}$,

$$\|h_{3,s} - h_{3,s'}\|_\infty \leqslant C|s - s'|,$$

$$\mathbb{E}[h_{3,s}(X_i, X_j) \mid \sigma(X_1, \ldots, X_{j-1})] = 0$$
$$\mathbb{E}[h_{3,s}(X_i, X_j) \mid \sigma(X_k, k \geqslant i+1)] = 0.$$

On doit montrer que

$$\sup_{-R\leqslant s\leqslant R} \sup_{0\leqslant t\leqslant 1} |R_n(s,t)| \to 0 \text{ en probabilité.}$$

Traitement de la partie dégénérée (2)

Premier pas : supremum sur t pour un s fixé :

$$\sup_{0\leqslant t\leqslant 1} |R_n(s,t)| = \frac{1}{n^{3/2}} \max_{1\leqslant \ell\leqslant n-1} \left| \sum_{i=1}^{\ell} \sum_{j=\ell+1}^{n} h_s(X_i, X_j) \right|.$$

Soit $h_{i,j} := h_{3,s}(X_i, X_j)$ et pour n fixé, $S_\ell := \sum_{i=1}^{\ell} \sum_{j=\ell+1}^{n} h_{i,j}$, $1 \leqslant \ell \leqslant n-1$ et $S_0 = 0$. Alors pour $2 \leqslant \ell \leqslant n$,

$$\begin{aligned} S_\ell - S_{\ell-1} &= \sum_{i=1}^{\ell} \sum_{j=\ell+1}^{n} h_{i,j} - \sum_{i=1}^{\ell-1} \sum_{j=\ell}^{n} h_{i,j} \\ &= \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^{n} h_{i,j} + \sum_{j=\ell+1}^{n} h_{\ell,j} - \sum_{i=1}^{\ell-1} \sum_{j=\ell+1}^{n} h_{i,j} - \sum_{i=1}^{\ell-1} h_{i,\ell} \\ &= \sum_{j=\ell+1}^{n} h_{\ell,j} - \sum_{i=1}^{\ell-1} h_{i,\ell} \end{aligned}$$

donc

$$S_k = \sum_{\ell=1}^{k} (S_\ell - S_{\ell-1}) = - \sum_{1\leqslant i<\ell\leqslant k} h_{i,\ell} + \sum_{\ell=1}^{k} \sum_{j=\ell+1}^{n} h_{\ell,j}.$$

Traitement de la partie dégénérée (3)

Terme $\max_{1\leqslant k\leqslant n-1}\sum_{1\leqslant i<\ell\leqslant k} h_{i,\ell}$: on applique une inégalité exponentielle pour les U-statistiques dégénérées à noyau borné.

Terme $\max_{1\leqslant k\leqslant n-1}\sum_{\ell=1}^{k}\sum_{j=\ell+1}^{n} h_{\ell,j}$: il s'agit également d'un maximum d'une U-statistique, mais la suite considérée est $(X_n, X_{n-1}, \ldots, X_1)$ au lieu de $(X_1, \ldots, X_n)$.

On déduit que

$$\mathbb{P}\left\{\sup_{0\leqslant t\leqslant 1}|R_n(s,t)|>\varepsilon\right\}\leqslant C\exp\left(-c\sqrt{n}\varepsilon\right),$$

où les constantes c et C sont universelles (en particulier, indépendentes de s, n et ε).

Puis on découpe $[-R, R]$ en "petits"intervalles et on utilise l'inégalité

$$\|h_{3,s}-h_{3,s'}\|_\infty\leqslant K\,|s-s'|\,.$$

Extension du cas indépendant

On rappelle que

$$e_n(s,t) = \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left(\mathbf{1}\left\{g(X_i, X_j) \leqslant s\right\} - \mathbb{P}\left\{g(X_i, X_j) \leqslant s\right\}\right).$$

On se place dans le cadre suivant la suite $(X_i)_{i\in\mathbb{Z}}$ est

- strictement stationnaire ;
- β-mélangeante :

$$\beta(k) := \beta\left(\sigma\left(X_i, i \leqslant 0\right), \sigma\left(X_i, i \geqslant k\right)\right) \to 0,$$

où

$$\beta(\mathcal{A}, \mathcal{B}) := \frac{1}{2} \sup \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} \left|\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\,\mathbb{P}(B_j)\right| \right\},$$

et le supremum est pris sur les partitions $(A_i)_{i=1}^{I}$, $A_i \in \mathcal{A}$ et $(B_j)_{j=1}^{J}$, $B_j \in \mathcal{B}$.

Objectifs : donner une condition suffisante sur $(\beta(k))_{k\geqslant 1}$ et g garantissant la convergence de e_n dans $D([-R,R]\times[0,1])$ et une description du processus limite.

Résultat

Théorème (Dehling, G., Sharipov (2020+))

Soit $(X_i)_{i\in\mathbb{Z}}$ *une suite strictement stationnaire et* e_n *défini par*

$$e_n(s,t) := \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left(\mathbf{1}\left\{g(X_i, X_j) \leqslant s\right\} - \mathbb{P}\left\{g(X_i, X_j) \leqslant s\right\}\right).$$

On suppose que g *est symétrique et que pour tout* $u \in \mathbb{R}$, *la variable aléatoire* $g(u, X_1)$ *a une densité* f_u *et que* $\sup_{u,x} f_u(x) < +\infty$. *De plus, supposons que* $\beta(k) = O(k^{-p})$ *pour un* $p > 4$. *Alors pour tout* $R > 0$,

$$e_n(s,t) \to W(s,t) \text{ en loi dans } D([-R,R]\times[0,1]),$$

où $(W(s,t), s \in \mathbb{R}, t \in [0,1])$ *est un processus gaussien centré, dont la covariance est donnée pour* $s, s' \in \mathbb{R}$ *et* $0 \leqslant t \leqslant t' \leqslant 1$ *par :*

$$\mathrm{Cov}\left(W(s,t), W\left(s',t'\right)\right) = t\left(1-t'\right)\left(1-2t+2t'\right) \sum_{k\in\mathbb{Z}} \mathrm{Cov}\left(h_{1,s}(X_0), h_{1,s'}(X_k)\right),$$

où $h_{1,s}(u) = \mathbb{P}\left\{g(u, X_1) \leqslant s\right\}$.

Perspectives (1)

- Mettre au point un procédure statistique: trouver les quantiles des lois limites $\sup_{0\leqslant t\leqslant 1}\sup_{-R\leqslant s\leqslant R}|W(s,t)|$ et $\sup_{0\leqslant t\leqslant 1}\int_{\mathbb{R}}\left(e_n(s,t)-\mathbb{E}\left[e_n(s,t)\right]\right)^2 d\mu(s)\rightarrow$ $\sup_{0\leqslant t\leqslant 1}\int_{\mathbb{R}} W(s,t)^2\, d\mu(s)$, où W est le processus limite.

- Traiter d'autres exemples de suites dépendantes, comme $X_i = f\left(\left(\varepsilon_{i-j}\right)_{j\geqslant 0}\right)$, où $(\varepsilon_k)_{k\in\mathbb{Z}}$ est i.i.d., sous des conditions mettant en jeu la norme $\mathbb{L}^p$ de la densité conditionelle de $g(X_i, X_j)$ par rapport à la tribu engendrée par $\varepsilon_u, u\leqslant 0$.

Perspectives (2)

U-statistiques à k échantillons

- Pour $0 \leqslant t_1 < \cdots < t_k \leqslant 1$, on définit

$$T_{k,n,h}(t_1, \ldots, t_k) = \sum_{i_1=1}^{[nt_1]} \sum_{i_2=[nt_1]+1}^{[nt_2]} \cdots \sum_{i_k=[nt_k]+1}^{n} h(X_{i_1}, \ldots, X_{i_k}),$$

où $h\colon \mathbb{R}^k \to \mathbb{R}$ est une fonction mesurable et $(X_i)_{i \geqslant 1}$ vérifie certaines conditions de dépendance (mélange, fonction d'une suite i.i.d. par exemple).
Motivation : détecter un multiple changement de paramètre dans un échantillon.

- Versions empiriques : noyaux de la forme

$$h_s(x_1, \ldots, x_k) = \mathbf{1}\{g(x_1, \ldots, x_k) \leqslant s\}.$$

Perspectives (3)

Soit

$$e_n(g,s,t) = \frac{1}{n^{3/2}} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left(\mathbf{1}\left\{g\left(X_i, X_j\right) \leqslant s\right\} - \mathbb{P}\left\{g\left(X_i, X_j\right) \leqslant s\right\}\right).$$

On chercher à établir la convergence du processus $e_n(g,s,t)$, $g \in \mathcal{F}$, $s \in \mathbb{R}$, $t \in [0,1]$, où $\mathcal{F}$ est une classe de fonctions.