

Temps de mélange

Exercices 2

1 Distance en variation totale

On rappelle que si μ, ν sont deux mesures de probabilité sur Ω fini ou dénombrable, on définit

$$\|\mu - \nu\|_{TV} = \sup_{A \subset \Omega} \mu(A) - \nu(A).$$

On a vu en cours des caractérisations alternatives de la distance en variation totale.

Par ailleurs, si P désigne le noyau de transition d'une chaîne X irréductible sur Ω , et qui possède une unique mesure stationnaire π , on définit pour $t \geq 0$,

$$d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV},$$

et enfin pour $\varepsilon > 0$,

$$t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}, \quad t_{\text{mix}} = t_{\text{mix}}(1/4).$$

Exercice 1 Vérifier que $d_{TV}(\mu, \nu) = \|\mu - \nu\|_{TV}$ définit une distance sur l'espace $\mathbf{P}(\Omega)$ des mesures de probabilités sur Ω .

Exercice 2 Montrer que

$$\|\mu - \nu\|_{TV} = 1 - \sum_{x \in \Omega} \min\{\mu(x), \nu(x)\}.$$

Exercice 3 Soient $(X_n, n \geq 0)$ et X_∞ des variables aléatoires à valeurs dans \mathbb{Z} , $(\nu_n, n \geq 0)$ et ν_∞ leurs lois respectives. Montrer que

$$X_n \xrightarrow[n \rightarrow \infty]{\text{(loi)}} X_\infty \Leftrightarrow \|\nu_n - \nu_\infty\|_{TV} \xrightarrow[n \rightarrow \infty]{} 0.$$

Exercice 4 Soit $\Omega = \{1, 2\}$ et $P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$. Calculer $d(t)$ pour tout $t \geq 0$.

Exercice 5 Que vaut $\|\mu - \nu\|_{TV}$ lorsque μ et ν sont des lois de Bernoulli de paramètres respectifs p, q ?

Exercice 6 On fixe $p \in [0, 1]$. Soit μ la loi d'une variable Bernoulli de paramètre p et ν celle d'une variable de Poisson, également de paramètre p .

Montrer que

$$\|\mu - \nu\|_{\text{TV}} = p(1 - \exp(-p)).$$

Exercice 7 Si μ_1, μ_2 sont des mesures sur \mathbb{Z} on peut définir le produit de convolution

$$\mu_1 * \mu_2(x) = \sum_{y \in \mathbb{Z}} \mu_1(y) \mu_2(x - y), \quad x \in \mathbb{Z}$$

et la mesure produit (sur \mathbb{Z}^2)

$$\mu_1 \times \mu_2(x, y) = \mu_1(x) \mu_2(y), \quad (x, y) \in \mathbb{Z}^2.$$

Montrer alors que si $\mu_1, \mu_2, \nu_1, \nu_2$ sont des mesures sur \mathbb{Z} on a

$$\|\mu_1 * \mu_2 - \nu_1 * \nu_2\|_{\text{TV}} \leq \|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\|_{\text{TV}} \leq \|\mu_1 - \nu_1\|_{\text{TV}} + \|\mu_2 - \nu_2\|_{\text{TV}}.$$

Exercice 8 On considère des nombres $\{p_{n,m} \in [0, 1], n \in \mathbb{N}, 0 \leq m \leq n\}$ tels que

$$\lambda_n := \sum_{m=0}^n p_{n,m} \xrightarrow{n \rightarrow \infty} \lambda > 0, \quad \max_{0 \leq m \leq n} p_{n,m} \xrightarrow{n \rightarrow \infty} 0$$

On définit alors μ_n la loi de $\sum_{m=1}^n \xi_{n,m}$, où les variables $\xi_{n,m} \sim \text{Ber}(p_{n,m})$ sont indépendantes. On note par ailleurs ν_n la loi d'une variable de Poisson de paramètre λ_n , et ν celle d'une variable de Poisson de paramètre λ .

1. A l'aide des exercices précédents établir que

$$\|\mu_n - \nu_n\|_{\text{TV}} \leq 2 \sum_{m=0}^n p_{n,m} (1 - \exp(-p_{n,m})).$$

2. En déduire que lorsque $n \rightarrow \infty$, $\|\mu_n - \nu\|_{\text{TV}} \rightarrow 0$

3. Quel résultat classique cet exercice permet-il de généraliser ?

Exercice 9 Soit X, Y des variables binômiales de paramètres respectifs n, p et n, q avec $n \in \mathbb{N}^*$ et $0 \leq p \leq q \leq 1$. On note μ la loi de X , ν la loi de Y .

1. Montrer que si $k_0 = \left\lfloor n \frac{\log\left(\frac{1-p}{1-q}\right)}{\log\left(\frac{q(1-p)}{p(1-q)}\right)} \right\rfloor$, on a

$$\|\mu - \nu\|_{\text{TV}} = \mathbb{P}(X \leq k_0) - \mathbb{P}(Y \leq k_0).$$

2. Dans cette question on suppose que $p = 1/2$ et que pour un $c \in \mathbb{R}_+$, $q = q(n) = \frac{1}{2} + \frac{c}{2\sqrt{n}}$. On note μ et ν_n les lois respectives de X, Y_n . Montrer alors que

$$k_0 = \frac{n}{2} + \frac{c\sqrt{n}}{4} + O(1),$$

puis que

$$\lim_{n \rightarrow \infty} \|\mu - \nu_n\|_{\text{TV}} = \mathbb{P}(Z \in [-c/2, c/2]),$$

où Z est une variable dont on précisera la loi.

3. Sans faire les calculs, se convaincre qu'on pourrait faire un raisonnement similaire et obtenir des limites non dégénérées (ni 0 ni 1) lorsque $a \in (0, 1)$ est fixé, $p_n = a + \frac{c_1}{2\sqrt{n}}$ est fixé et $q_n = a + \frac{c_2}{2\sqrt{n}}$ (avec $c_1 < c_2$).
4. Qu'implique les questions précédentes lorsque p_n, q_n approchent la même valeur $a \in [0, 1]$ avec $|q_n - p_n| \ll \sqrt{n}$? Et lorsque p_n, q_n approchent la même valeur a avec $|q_n - p_n| \gg \sqrt{n}$?

Exercice 10 Soit X, Y des variables de Poisson de paramètres respectifs $\lambda, \lambda + \ell$ avec $\lambda > 0, \ell \geq 0$. On note μ la loi de X , ν la loi de Y .

1. Montrer que si $k_0 = \left\lfloor \frac{\ell}{\log(1 + \frac{\ell}{\lambda})} \right\rfloor$, on a

$$\|\mu - \nu\|_{TV} = \mathbb{P}(X \leq k_0) - \mathbb{P}(Y \leq k_0).$$

2. Dans cette question on va considérer l'asymptotique $\lambda \rightarrow \infty$. Montrer que si $\ell = c\sqrt{\lambda} + o(\lambda^{1/4})$, avec $c > 0$, on a

$$k_0 = \lambda + \left(\frac{c}{2}\right) \sqrt{\lambda} + o(\sqrt{\lambda}),$$

et en déduire que

$$\lim_{\lambda \rightarrow \infty} \|\mu - \nu\|_{TV} = \mathbb{P}(Z \in I(c)),$$

où on précisera la loi de Z , et l'intervalle $I(c)$.

Exercice 11

1. Montrer que pour tout $t \geq 0$

$$d(t) = \sup_{\mu \in \mathbf{P}(\Omega)} \|\mu P^t - \pi\|_{TV},$$

puis que

$$d(t) = \sup_{\mu \in \mathbf{P}(\Omega), A \subset \Omega} \sum_{x \in \Omega, y \in A} (\mu(x) - \pi(x)) P^t(x, y).$$

2. Montrer que $\bar{d}(t) := \max_{x, y \in \Omega} \|P^t(x, \cdot) - P^t(y, \cdot)\|_{TV}$ vérifie, pour tout $t \geq 0$,

$$\bar{d}(t) = \sup_{\mu \in \mathbf{P}(\Omega), \nu \in \mathbf{P}(\Omega)} \|\mu P^t - \nu P^t\|_{TV},$$

et en déduire une égalité similaire à celle obtenue pour $d(t)$ dans la question précédente.

2 Autres distances

Exercice 12 Soit, pour un $t \geq 1$, $\bar{d}_t^P(\mu, \nu) = \|\mu P^t - \nu P^t\|_{TV}$.

1. Montrer que \bar{d}_t^P est une pseudo-distance sur $\mathcal{P}(\Omega)$.

2. Donner un exemple de chaîne et d'un $t \geq 1$ pour laquelle \bar{d}_t^P n'est pas une distance.
3. Donner un exemple de chaîne pour laquelle \bar{d}_t est une distance pour tout $t \geq 1$.
4. Montrer que $t \rightarrow \bar{d}_t^P(\mu, \nu)$ est décroissante.
5. Montrer que $t \rightarrow d(t)$ est décroissante.

Exercice 13 Si π est la distribution stationnaire d'une chaîne X sur Ω , pour une fonction $f : \Omega \rightarrow \mathbb{R}$, on peut définir, pour $p \geq 1$, sa norme $\ell^p(\pi)$:

$$\|f\|_p = \left[\sum_{x \in \Omega} |f(x)|^p \pi(x) \right]^{1/p}.$$

sa norme $\ell^\infty(\pi)$ étant $\|f\|_\infty = \max_{x \in \Omega} |f(x)|$.

Le cas $p = 2$ est particulièrement intéressant puisque

$$\langle f, g \rangle_\pi := \sum_{x \in \Omega} f(x)g(x)\pi(x)$$

définit un produit scalaire sur l'espace $\{f : \Omega \rightarrow \mathbb{R}\}$.

1. Si ν et μ sont deux mesures de probabilité sur Ω que vaut

$$\left\| \frac{\nu}{\pi} - \frac{\mu}{\pi} \right\|_1 \quad ?$$

2. Soit $f : \Omega \rightarrow \mathbb{R}$. Etablir que $\|f\|_p$ croît avec $p \geq 1$.
3. Montrer que

$$4 \|P^t(x, \cdot) - \pi\|_{TV}^2 \leq \left\| \frac{P^t(x, \cdot)}{\pi} - 1 \right\|_2^2.$$

4. Montrer que

$$d(t) \leq \frac{1}{2} \max_{x, y \in \Omega} \left| \frac{P^t(x, y)}{\pi(y)} - 1 \right|.$$

Exercice 14 La distance de Hellinger entre deux probabilités sur Ω , μ et ν , est donnée par

$$d_H(\mu, \nu) = \sqrt{\sum_{x \in \Omega} (\sqrt{\mu(x)} - \sqrt{\nu(x)})^2}$$

1. Fixons $x_0 \in \Omega$ et introduisons $g_{x_0, t}(x) = \frac{P^t(x_0, x)}{\pi(x)}$, $x \in \Omega$. Etablir que

$$d_H(P^t(x_0, \cdot), \pi) = \|\sqrt{g_{x_0, t}} - 1\|_2.$$

2. Montrer que

$$\|\mu - \nu\|_{TV} \leq d_H(\mu, \nu).$$

3. En déduire que

$$t_{\text{mix}}(\varepsilon) \leq \inf\{t \geq 0 : \forall x_0 \in \Omega \quad d_H(P^t(x_0, \cdot), \pi) \leq \varepsilon\}.$$

3 Chaîne retournée dans le temps

Soit X est une chaîne de Markov irréductible, de distribution stationnaire π . On définit pour $x, y \in \Omega$,

$$\widehat{P}(x, y) := \frac{\pi(y)P(y, x)}{\pi(x)}.$$

Exercice 15

1. Montrer que \widehat{P} est bien définie, et que c'est le noyau de transition d'une chaîne \widehat{X} sur Ω , également irréductible. On note $\widehat{\mathbb{P}}_\mu$ la loi de cette chaîne issue de μ .
2. Montrer que π est également une distribution stationnaire de \widehat{X} .
3. Montrer que pour tous $t \geq 0, x_0, \dots, x_t \in \Omega$, on a

$$\mathbb{P}_\pi(X_0 = x_0, \dots, X_t = x_t) = \widehat{\mathbb{P}}_\pi(\widehat{X}_0 = x_t, \dots, \widehat{X}_t = x_0).$$

4. Que se passe-t-il lorsque X est réversible ?

Exercice 16 Dans cet exercice on suppose que X est une marche sur le groupe G dont la distribution d'un pas est dictée par la distribution μ . Précisément, le noyau P de X est tel que

$$P(g, hg) = \mu(h) \quad \forall g, h \in G.$$

On suppose que $\mathcal{H} = \{h : \mu(h) > 0\}$ engendre G . On note \widehat{P} le noyau de la chaîne retournée, $\widehat{d}(t) = \max_{x \in \Omega} \|\widehat{P}^t(x, \cdot) - \pi\|_{TV}$ et $\widehat{t}_{\text{mix}}(\varepsilon) := \inf\{t \geq 0 : \widehat{d}(t) \leq \varepsilon\}$.

1. Montrer que la chaîne retournée reste une marche sur le groupe G , et décrire la distribution $\widehat{\mu}$ correspondante.
2. Montrer que pour tout $g \in G, t \geq 0$ on a

$$P^t(\text{id}, g) = \widehat{P}^t(\text{id}, g^{-1}).$$

En déduire que pour tout $\varepsilon > 0, t_{\text{mix}}(\varepsilon) = \widehat{t}_{\text{mix}}(\varepsilon)$.

4 Couplages : deux exemples

Un couplage $\{(X_t, Y_t), t \geq 0\}$ avec noyau de transition P est tel que sous $\mathbb{P}_{x,y}$, pour tout $t \geq 0, X_t \sim P^t(x, \cdot), Y_t \sim P^t(y, \cdot)$, on note $\tau_{\text{couple}} := \inf\{t \geq 0 : X_t = Y_t\}$. On verra en cours que

$$d(t) \leq \overline{d}(t) \leq \max_{x,y \in \Omega} \mathbb{P}_{x,y}(\tau_{\text{couple}} > t).$$

Exercice 17 On considère le couplage suivant de deux marches paresseuses sur le tore discret $(\mathbb{Z}/n\mathbb{Z})^d$. Au temps $t + 1$:

– Indépendamment, et indépendamment des étapes précédentes, on tire

$$i_{t+1} \sim \text{Unif}\{1, \dots, d\}, \quad \xi_{t+1} = \begin{cases} 1 & \text{avec proba } 1/4 \\ -1 & \text{avec proba } 1/4, \quad B_{t+1} \sim \text{Ber}(1/2), \text{ et enfin} \\ 0 & \text{avec proba } 1/2 \end{cases}$$

$$\zeta_{t+1} = \begin{cases} 1 & \text{avec proba } 1/2 \\ -1 & \text{avec proba } 1/2 \end{cases}.$$

- Si $X_t^{i_{t+1}} = Y_t^{i_{t+1}}$, on pose $X_{t+1}^{i_{t+1}} = Y_{t+1}^{i_{t+1}} = X_t^{i_{t+1}} + \xi_{t+1}$.
- Si $X_t^{i_{t+1}} \neq Y_t^{i_{t+1}}$, on pose

$$X_{t+1}^{i_{t+1}} = X_t^{i_{t+1}} + B_{t+1}\zeta_{t+1}, \quad Y_{t+1}^{i_{t+1}} = Y_t^{i_{t+1}} + (1 - B_{t+1})\zeta_{t+1}.$$

- On laisse les autres coordonnées des deux chaînes inchangées.
 1. Montrer que $\{(X_t, Y_t), t \geq 0\}$ est un couplage de marches paresseuses sur le tore discret de dimension d .
 2. Soit $i \in \{1, \dots, d\}$ et $Z_t^i := X_t^i - Y_t^i, t \geq 0$. Montrer que $(Z_t^i, t \geq 0)$ est une chaîne de Markov sur $\mathbb{Z}/n\mathbb{Z}$ dont on précisera le noyau de transition.
 3. Estimer $\max_{x,y \in \Omega} \mathbb{E}_{x,y}[\tau_i]$ où $\tau_i = \inf\{t \geq 0 : Z_t^i = 0\}$.
 4. En déduire finalement que pour la marche paresseuse sur $(\mathbb{Z}/n\mathbb{Z})^d$ on a

$$t_{\text{mix}}(\varepsilon) \leq d^2 n^2 \left\lceil \log \left(\frac{1}{\varepsilon} \right) \right\rceil.$$

Exercice 18 On considère $(X_t, t \geq 0)$ la marche simple, paresseuse, sur l'arbre binaire enraciné de profondeur k . On note $|X_t|$ la profondeur de X_t .

1. Montrer que $(|X_t|, t \geq 0)$ reste une chaîne de Markov dont on déterminera le noyau de transition.
2. Déterminer les mesures stationnaires respectives de X , $|X|$.
3. Soit $\tau_j = \inf\{t \geq 0 : |X_t| = k\}$. Estimer $E = \mathbb{E}_0[\tau_k] + \mathbb{E}_k[\tau_0]$.
4. On définit $(X_t, Y_t)_{t \geq 0}$ itérativement, comme suit :
 - Si $|X_t| = |Y_t|$, alors X et Y effectuent au temps $t + 1$ un mouvement similaire dans l'arbre.
Par exemple si elles se trouvent à une profondeur $i \in \{1, \dots, k - 1\}$, alors avec proba $1/2$, les deux chaînes restent sur place, avec proba $1/6$, X_{t+1} est le parent de X_t et Y_{t+1} celui de Y_t , avec proba $1/6$, X_{t+1} est le premier descendant de X_t , Y_{t+1} celui de Y_t , enfin avec proba $1/6$, X_{t+1} est le deuxième descendant de X_t , Y_{t+1} celui de Y_t .
Si elles se trouvent toutes deux à profondeur k , avec proba $1/2$ elles restent toutes deux sur place, et avec proba $1/2$ elles remontent au parent correspondant.
Enfin si elles se trouvent toutes deux à profondeur 0 , on peut les faire évoluer ensemble.
 - Si $|X_t| \neq |Y_t|$, on tire $B_{t+1} \sim \text{Ber}(1/2)$. Si $B_{t+1} = 0$, X_{t+1} est l'un des voisins de X_t choisi uniformément, tandis que $Y_{t+1} = Y_t$, et si $B_{t+1} = 1$, $X_{t+1} = X_t$ tandis que Y_{t+1} est choisi uniformément parmi les voisins de Y_t .
 Montrer que $\{(X_t, Y_t), t \geq 0\}$ est un couplage de deux marches simples paresseuses sur l'arbre binaire.
5. Expliquer pourquoi $\mathbb{E}[\tau_{\text{couple}}] \leq E$. En déduire une borne supérieure sur $t_{\text{mix}}(\varepsilon)$ pour la marche simple paresseuse sur l'arbre binaire de profondeur k .