Théorie de l'information. Trois théorèmes de Claude Shannon.
Calvage & Mounet, 2022

Erratum (23 septembre 2022)

Je remercie Pascal Molin et des rapporteurs anonymes de la traduction anglaise pour certaines des corrections ci-dessous.

Page 9, preuve de la proposition 1.8. Les assertions *a*) et *b*) sont bien sûr les assertions 1. et 2. de la proposition.
Page 12, énoncé (P_3), lire : « Le complémentaire $\Omega \setminus A$ d'un événement A est un événement. »
Page 15, « l'ensemble des valeurs *possibles* de la variable aléatoire discrète X », « lorsque l'ensemble des valeurs *possibles* de X... »
Page 16, dernière ligne : « Cela prouve que $f(X)$ est une variable aléatoire discrète. »
Page 17, (3.4), lire : « Si X est certaine, de valeur a, alors X possède une espérance, égale à a. »
Page 23, ligne 4, effacer le bout de phrase « qui découle de ce que pour tous nombres réels x et y, ».
Page 26, item 2 de (4.6), lire « une espérance *relativement à* B ».
Page 29, question *d*) de l'exercice 5.6. Il faut ajouter l'hypothèse que X et Y sont indépendantes, et ajuster la solution page 187. [Remarque due à Pascal Molin]
Page 36, note de bas de page. Lire « Un troisième membre de la famille, Daniel... »
Page 63 La discussion des chaînes de Markov générales est incorrecte. À partir de « Lorsque ces conditions ne sont pas satisfaites... », lire :

Dans ce livre, nous ne considérerons pas d'exemples de chaînes de Markov pour lesquelles ces conditions ne sont pas satisfaites ; il peut cependant être utile d'indiquer en quelques mots comment ramener le cas général à cette situation.

On dit que deux états $a, b \in A$ « communiquent » s'il existe un chemin de a à b, ainsi qu'un chemin de b à a ; c'est une relation d'équivalence dans A. Fixons une « classe de communication » $C \subseteq A$ et considérons seulement les flèches du carquois initial qui relient deux états de cette classe ; par construction, on obtient un carquois connexe et la chaîne de Markov d'ensemble d'états C que l'on obtient a pour matrice de probabilités de transition une sous-matrice P_C et est irréductible.

Tous les états de C ont la même période, disons d, et la classe C se scinde naturellement en d classes $C_1, \dots, C_d$ telles que les flèches vont de C_1 à C_2, de C_2 à C_3, etc., et de C_d à C_1. La matrice stochastique P_C^d est la matrice de transition d'une chaîne de Markov d'états C et dont les classes de communication sont $C_1, \dots, C_d$. Chacune d'entre elles donne lieu à une chaîne de

Markov irréductible et apériodique par lesquelles on peut étudier la chaîne de Markov initiale.

Page 78, reprendre le dernier paragraphe de la preuve de la proposition 2.1 :

Ainsi, c_m est au plus égal au nombre de mots de longueur $\leqslant m$ dans l'alphabet B, si bien que

$$c_m \leqslant 1 + D + \cdots + D^m = \frac{D^{m+1} - 1}{D - 1} \leqslant D^{m+1}.$$

On a aussi $c_m = 0$ si $m > kN$. Alors,

$$\sum_{a \in A^k} D^{-\ell(C(a))} = \sum_m c_m D^{-m} \leqslant \sum_{m=1}^{kN} D^{m+1} D^{-m} = kND.$$

Par suite,

$$\sum_{a \in A} D^{-\ell(C(a))} \leqslant (kND)^{1/k}.$$

En faisant tendre k vers $+\infty$, on obtient l'inégalité voulue.

Page 80, énoncé de la proposition 2.4, lire « soit D un entier $\geqslant 2$. »

Page 89, ligne -6, phrase pas terminée : « puisque C *est un code préfixe.* » Ligne suivante : « puisque $C'(ab)$ est… »

Page 92, ligne -9 et suivantes, jusqu'à la fin de la démonstration : ce texte aurait dû être en petits caractères.

Pages 92–93, je préférerais aujourd'hui écrire la preuve autrement.

Fixons un nombre réel $\varepsilon > 0$ et choisissons $\delta > 0$ de sorte que

$$16\delta \mathbf{E}(|X_1|) t^{-2} < \varepsilon.$$

(...) En particulier, pour n assez grand, $E(X'_1 + \cdots + X'_n)^2 \leqslant \delta \mathbf{E}(|X_1|)^2 n^2$, et

$$\mathbf{E}((X'_1 + \cdots + X'_n)^2) = \mathbf{V}(X'_1 + \cdots + X'_n) + \mathbf{E}(X'_1 + \cdots + X'_n)^2 \leqslant 2\delta \mathbf{E}(|X_1|) n^2.$$

Appliquée à la variable aléatoire $(X'_1 + \cdots + X'_n)^2$, l'inégalité de Markov entraîne alors que

$$\mathbf{P}(|X'_1 + \cdots + X'_n| > nt/2) \leqslant 8\delta \mathbf{E}(|X_1|)/t^2 \leqslant \varepsilon/2. \tag{(4.2.1)}$$

pour tout entier n assez grand. (...)

Page 105, ligne 4 : « Puisque $t = \ldots$, on constate… »

Page 111, (6.6), ligne 3, lire : « sur un ensemble *fini* M ».

Page 141, corollaire (3.6). L'énoncé est un peu ambigu, il faut comprendre que l'ensemble des $t \in [0; T]$ tels que $f(t) \neq 0$ est fini ; par périodicité, on peut remplacer $[0; T]$ par n'importe quel intervalle borné.

Page 141, preuve du corollaire (3.6). Pour justifier la formule centrée, lire d'abord le dernier paragraphe de la preuve :

Ainsi, g est périodique, de période T ; elle est également continue. Puisque g est l'intégrale de f, elle est dérivable en tout point t où f est continue, de dérivée $g'(t) = f(t)$. Par suite, g est de classe $\mathscr{C}^1$ par morceaux. D'après l'exemple 2.7, les coefficients de Fourier de g sont donnés par

$$c_n(g) = \frac{T}{2\pi i n} c_n(f) = 0$$

pour $n \in \mathbf{Z} \setminus \{0\}$. Par conséquent, la fonction $g^* \colon t \mapsto g(t) - c_0(g)$ est continue et tous ses coefficients de Fourier sont nuls. D'après le théorème précédent, on a $g^* \equiv 0$, donc g est constante, donc nulle puisque $g(0) = 0$.

Comme $f(t) = g'(t)$ en tout point t où f est continue et que f est de classe $\mathscr{C}^1$ par morceaux, il n'y a qu'un nombre fini de $t \in [0; T]$ tels que $f(t) \neq 0$.

Page 142, ligne 3 de la démonstration du corollaire (3.7), lire : « pour tout entier $m \geqslant |n|$ »

Page 145, milieu de page, lire : « on a seulement $|f_k(t) - f(t)| \leqslant 2M$ » et remplacer $2hM^2$ par $8hM^2$ et $T/\varepsilon 4M^2$ par $T/\varepsilon 16M^2$ dans les formules suivantes.

Lire : « Soit g l'unique fonction périodique de période T telle que. . . ».

Page 159, ligne -12, lire : « de $\mathbf{R}$ dans $\mathbf{C}$ ».

Page 162, ligne -2, lire : « de période $2W$ ».

Page 164, line 5, lire : « de période $2W$ ».

Page 169, ligne -4, au lieu de « formule de Parseval », lire « formule de Plancherel ».

Page 176, exercice 8.7, au lieu de « localement intégrable », lire « continue ».

Page 182, ligne 4, au lieu de « condamné », lire « grâcié » (!).

Page 187, question *d*) de l'exercice 0-5-6. Changer la seconde phrase en :

Comme X et Y sont indépendantes, on a

$$\mathbf{P}(X = b \mid Z = a) = \frac{\mathbf{P}(X = b \text{ et } Y = a - b)}{\mathbf{P}(Z = a)} = \frac{\mathbf{P}(X = b)\mathbf{P}(Y = a - b)}{\mathbf{P}(Z = a)},$$

pour tous a, b tels que $\mathbf{P}(Z = a) > 0$. Comme X et Y suivent la même loi, il s'ensuit par symmétrie que

$$\mathbf{P}(X = b \mid Z = a) = \mathbf{P}(Y = b)\mathbf{P}(X = a - b)/\mathbf{P}(Z = a) = \mathbf{P}(Y = b \mid Z = a).$$

En particulier, on a $\mathbf{E}(X \mid Z = a) = \mathbf{E}(Y \mid Z = a)$.

Page 199, exercice 6.9, ligne 1 : lire « $H(X) = H(Y)$ ».

Page 212, ligne -7. Lire « Il y a trois cas pour les cases intérieures ».

Page 216, ligne 10, lire : « les fractions $2^{-\ell}$ ».

Page 222, dernière ligne, lire : « la loi de X ».

Page 223, ligne -4. Ajouter : « Notons $F(u, p)$ cette expression et étudions son comportement lorsque u varie, en supprimant p de la notation. »

Page 225, c), après la formule pour la matrice A, lire « du canal C_n ». Puis, ligne -5, avant « Pour calculer. . . », insérer : « Ainsi, A^n est la matrice de probabilités de transmission d'un canal symétrique binaire d'un certain paramètre p_n. »

Page 235, exercice 8.2, lignes 5–6, lire « de façon indépendante ».

Page 245, exercice 8.8, remplacer la dernière phrase (incomplète) du a) par « La fonction f est de classe $\mathscr{C}^1$ par morceaux, elle est continue en x. Le théorème de Dirichlet entraîne donc que $S_n(x)$ converge vers $\frac{4}{\pi}f(x) = \frac{4}{\pi}$. »
Page 245, ligne -3, remplacer « d'où » par « c'est-à-dire si ».
Page 247, ligne -6, au lieu de « $\pi/22$ », lire « $\pi/2$ ».
Page 250, dernière formule de l'exercice 8.11, supprimer le « t » qui suit « dy, ».