ERRATA ET COMPLÉMENTS

Ce document est clos : il sera remplacé par un ERRATA ET COMPLÉMENTS $2^{ème}$ édition

Page 13

L'opérateur Δ a l'étonnante propriété : $C = A\Delta B \Leftrightarrow A = C\Delta B \Leftrightarrow B = A\Delta C$.

Page 16, 4)

J'ai affirmé que dans un univers Ω infini non dénombrable existait un événement X indénombrable ainsi que son contraire (ou complémentaire). Voici comment on peut justifier cette assertion. D'après l'Hypothèse du continu, puisque $card\,\Omega > \aleph_0$, alors $card\,\Omega \geq \aleph_1 = card\,\mathbb{R}$. Il existe alors une injection f de $\mathbb{R}$ dans Ω : $X = f(\mathbb{R}^+)$ convient, car $\Omega - X$ contient déjà l'indénombrable $f\left(]-\infty, 0[\right)$. (Merci à Xavier Oudot.)

Comme exemple de tribus, on peut citer aussi l'ensemble des parties de $\mathbb{R}$ qui sont symétriques par rapport à O. Cette tribu est de plus stable par réunion et intersection quelconques. Une telle tribu est isomorphe à l'ensemble des parties d'un ensemble : considérer ses atomes définis par $A(\omega) = \left\{\bigcap T : \omega \in T \text{ et } T \text{ dans la tribu}\right\}$ puis montrer que les $A(\omega)$ *distincts* forment une partition de Ω.

Page 17 : tribus engendrées

Soit $\mathcal{T}$ la tribu engendrée par une famille $\mathcal{F}$, $\mathcal{T} = \mathcal{T}(\mathcal{F})$, et A un événement de Ω : la trace de $\mathcal{T}$ sur A est la tribu engendrée par $\mathcal{F} \cap A$ (ce sont deux sous-tribus de $\mathcal{P}(A)$). On a donc $\mathcal{T}(\mathcal{F}) \cap A = \mathcal{T}(\mathcal{F} \cap A)$: la trace d'une tribu engendrée par $\mathcal{F}$ est la tribu engendrée par la trace de $\mathcal{F}$.

Page 18, 3)

La tribu engendrée par les singletons d'un ensemble E a la propriété suivante : si un événement dénombrable A de cette tribu est l'intersection d'une suite A_n de la tribu, alors il existe n pour lequel A_n est dénombrable (passer aux complémentaires).

Page 22, l'inégalité d'Édith Kosmanek

L'inégalité universelle $|\,p(A \cap B) - p(A)p(B)\,| \leq 1/4$ est une conséquence d'une l'inégalité plus fine :

$$|\,p(A \cap B) - p(A)p(B)\,| \leq \sqrt{p(A)p(\overline{A})p(B)p(\overline{B})}\,.$$

Démonstration par l'intégrale (dans l'esprit des pages **182 – 183)**. Posons $f = 1_A$ et $g = 1_B$, abrégeons $\int_\Omega f\,dp$ en $\int f$ et $\int_\Omega g\,dp$ en $\int g$: il s'agit de prouver que $\left(\int fg - \int f \int g\right)^2 \leq \left(\int f - (\int f)^2\right)\left(\int g - (\int g)^2\right)$.

Ce sera une conséquence de $\left(\int fg - \int f \int g\right)^2 \leq \left(\int f^2 - (\int f)^2\right)\left(\int g^2 - (\int g)^2\right)$ car $f = f^2$ et $g = g^2$.

L'inégalité de Cauchy-Schwarz donnant $(\int_\Omega h\,dp)^2 \leq \int h^2$ pour toute h intégrable pour p (car p est de masse 1), on peut donc écrire $(\int_\Omega (f + tg)\,dp)^2 \leq \int (f + tg)^2$ pour tout réel t. On obtient alors un trinôme en t de signe constant sur R dont le discriminant réduit donne l'inégalité cherchée.

Démonstration ensembliste (**Frédéric PAUL**).

Posons $s(A,B)=p(A\cap B)-p(A)p(B)$ et $u(A)=p(A)p(\overline{A})$. Il est facile de vérifier que $s(\overline{A},B)=-s(A,B)$, d'où $s(A,B)=-s(\overline{A},B)=-s(B,\overline{A})=s(\overline{B},\overline{A})=s(\overline{A},\overline{B})$. De plus, s additive : $A\cap C=\varnothing$ implique $s(A+C,B)=s(A,B)+s(C,B)$, ce qui permet de retrouver $s(A,B)=s(\overline{A},\overline{B})$ car $s(\Omega,B)=0$.

Si $C=B-A$, alors $A\cap C=\varnothing$ et $A\cup B=A+C$; comme $s(C,B)\geq 0$, $s(A,B)\leq s(A+C,B)=s(A\cup B,B)\leq u(B)$. Ensuite, $-s(A,B)=s(\overline{A},B)\leq u(B)$, et donc $|s(A,B)|\leq u(B)$. Par symétrie, il vient $|s(A,B)|\leq u(A)$, d'où $|s(A,B)|\leq \min(u(A),u(B))$ et donc $\left(s(A,B)\right)^2\leq u(A)u(B)$.

Page 24, ligne 3

La justification de la σ-additivité de p se trouve en **0.7 4)** et non en **0.6**.

Page 24,d)

Comme il existe toujours une probabilité sur le couple $\left(\Omega,\mathcal{P}(\Omega)\right)$, il existera toujours au moins une probabilité sur $(\Omega,\mathcal{T})$ pour toute sous-tribu $\mathcal{T}$ de $\mathcal{P}(\Omega)$, la $\sigma-$additivité se conservant sur les sous-ensembles.

Page 24 e) : images d'une probabilité

On peut en dire un peu plus sur les mesures images. Soit m une mesure de probabilité sur $(\Omega_1,\mathcal{T}_1)$ et f une application *mesurable* de $(\Omega_1,\mathcal{T}_1)$ dans $(\Omega_2,\mathcal{T}_2)$. En posant $\mu(T_2)=m\left(f^{-1}(T_2)\right)$, on définit une mesure de probabilité sur $(\Omega_2,\mathcal{T}_2)$, de même masse. La loi p_X d'une variable aléatoire rentre dans ce cadre : $p_X(B)=p\left(X^{-1}(B)\right)$ pour tout borélien B de R.

Une fonction g sur Ω_2 est $\mu-$intégrable si et seulement si $g\circ f$ est $m-$intégrable et alors $\int_{\Omega_2} g\,d\mu=\int_{\Omega_1}(g\circ f)\,dm$; on observera que si g est une fonction indicatrice, l'égalité se réduit à la définition de μ.

Exemple. Soit m une mesure de probabilité sur $[0,2\pi]$ et f l'application continue $t\mapsto e^{it}$: μ est alors une mesure positive sur le cercle S ; $\int_S g\,d\mu=\int_{[0,2\pi]} g(e^{it})\,dm(t)$. Si m est la mesure de Borel sur $[0,2\pi]$, $dm(t)=dt$ et $\int_S g\,d\mu=\int_0^{2\pi} g(e^{it})\,d(t)$, μ étant la mesure longueur d'arc sur le cercle.

Page 25 : variables aléatoires

a) Pour souligner l'importance de la tribu $\mathcal{T}$ dans la définition d'une variable aléatoire X, on dit que X est $\mathcal{T}-$mesurable. Lorsque $\mathcal{T}$ est la tribu minimale $\{\Omega,\varnothing\}$, les seules variables aléatoires sont les fonctions constantes. Pour $\mathcal{T}=\mathcal{P}(\Omega)$, toute application X de Ω dans $\mathbb{R}$ est une variable aléatoire. X étant donnée, il existe toujours une tribu qui la rende aléatoire. La plus petite tribu qui rende X aléatoire est la tribu constituée par les $X^{-1}(B)$, on la note $\sigma(X)$ et on dit que c'est la tribu engendrée par X.

Lemme de Doob. Si Y est $\sigma(X)-$mesurable, on démontre que Y s'écrit $f\circ X$, avec f borélienne de R dans R : http://www.math.u-psud.fr/~jflegall/IPPA2.pdf, page 100 ; voir aussi **Page 297** de ce document.

Soit $\mathcal{T}$ une sous-tribu stricte de $\sigma(X)$, X n'est alors pas $\mathcal{T}-$mesurable; il est intéressant de se demander, parmi les Y qui sont $\mathcal{T}-$mesurables, laquelle est « la plus proche » de X : c'est le problème dit de *l'espérance conditionnelle*.

b) loi de probabilité

La loi d'une variable aléatoire X sur $(\Omega, \mathcal{T}, p)$ est une probabilité q sur $(\mathbb{R}, \mathcal{B}^1)$: réciproquement, q étant donnée sur $(\mathbb{R}, \mathcal{B}^1)$, existe-t-il X de loi q ? Réponse ici :

http://www.daniel-saada.eu/fichiers/22-existence_de_lois_uniformes.pdf

Page 26 : loi de probabilité et fonction de répartition d'une variable aléatoire réelle

D'abord, $p\left(]-\infty, x[\right) = F(x-)$ car $]-\infty, x[$ est la réunion croissante des $]-\infty, x-1/n]$.

Ensuite, je n'ai pas assez souligné que deux variables aléatoires X et Y qui ont la même fonction de répartition ont la même loi de probabilité : $F_X = F_Y$ implique $p_X = p_Y$. Par hypothèse, les probabilités p_X et p_Y, définies sur la tribu des boréliens de $\mathbb{R}$, coïncident sur tous les intervalles fermés $]-\infty, t]$ et donc sur les intervalles ouverts $]t, +\infty[$. Comme $]-\infty, t[$ est la réunion dénombrable des $]-\infty, t-1/n]$, p_X et p_Y sont égales sur tous les intervalles $]-\infty, t[$, et donc sur les $[t, +\infty[$.

L'intervalle $]a,b[$ étant la différence entre $]-\infty, a[$ et $]-\infty, b[$, $p_X\left(]a,b[\right) = p_Y\left(]a,b[\right)$: p_X et p_Y coïncident sur tous les intervalles ouverts de $\mathbb{R}$, donc sur tous les ouverts car tout ouvert de $\mathbb{R}$ est une réunion dénombrable disjointe d'intervalles ouverts. On en conclut (paragraphe **8.6**, page 132) que $p_X = p_Y$.

Page 41, 5)

Voici deux exemples remarquables de G_δ, empruntés à l'Analyse :

a) Si f est une fonction dérivable de $\mathbb{R}$ dans $\mathbb{R}$, l'ensemble des zéros de f' est un G_δ en vertu de **11.6.3.1** (page **178**) et parce que f' est la limite simple de la suite des fonctions continues $x \to \dfrac{f(x+1/n)-f(x)}{1/n}$.

Ce G_δ est un fermé quand f' est continue ; dans ce cas, la convergence de $\dfrac{f(x+1/n)-f(x)}{1/n}$ vers f' est uniforme sur tout compact (utiliser l'égalité des accroissements finis et la continuité uniforme de f' sur le compact).

b) L'ensemble des points de continuité d'une fonction définie sur un espace topologique E et à valeurs dans $\mathbb{R}$ est un G_δ de E (exercice 14, chapitre 6, *Topologie*, G. Choquet, *Masson*). Bien sûr, ce G_δ peut être vide, ou réduit à un point.

Je détaille maintenant la dernière ligne de la page.
L'intersection d'un ouvert U et d'un fermé F est à la fois un G_δ et un F_σ ; en effet, comme $U = \bigcup F_n$, $U \cap F = \bigcup (F_n \cap F)$, et comme $F = \bigcap U_n$, $U \cap F = \bigcap (U_n \cap U)$. Or [0,1[est l'intersection du fermé [0,1] et de l'ouvert]0,2[.

Page 43 : $(x \in E : f(x) + g(x) < a) = \bigcup_{r \in \mathbb{Q}} (x : f(x) < r) \cap (x : g(x) < a - r)$

Cette formule, citée également pages 217 et 224, n'a pas été démontrée. Voici une approche très simple. Commençons par décomposer $(x \in E : f(x) - g(x) < 0)$: entre deux réels figure un rationnel, donc

$$(x \in E : f(x) - g(x) < 0) = \bigcup_{r \in \mathbb{Q}} (x : f(x) < r) \cap (x : g(x) > r).$$

Changeons g en $-g$: $(x \in E : f(x) + g(x) < 0) = \bigcup_{r \in \mathbb{Q}} (x : f(x) < r) \cap (x : g(x) < -r)$.

Enfin, remplaçons g par $g - a$: $(x \in E : f(x) + g(x) < a) = \bigcup_{r \in \mathbb{Q}} (x : f(x) < r) \cap (x : g(x) < a - r)$.

Page 46, l'Axiome du choix

L'exemple le plus simple d'application est sans doute le suivant.

Soit f une surjection d'un ensemble E sur un ensemble de F. Puisque $f^{-1}(y)$ n'est vide pour aucun y de F, *choisissons* dedans un élément x et posons $g(y) = x$: on a ainsi construit une application g de F dans E telle que $f \circ g = id_F$. L'existence de g est très précieuse, voir **Page 272** du présent document.

Page 53, Chapitre 1

a) une propriété mérité d'être mentionnée :
si la réunion de deux algèbres $\mathcal{A}_1$ et $\mathcal{A}_2$ est une algèbre, alors soit $\mathcal{A}_1 \subset \mathcal{A}_2$, soit $\mathcal{A}_2 \subset \mathcal{A}_1$ (raisonner par l'absurde en supposant qu'existent A_1 dans $\mathcal{A}_1$ et A_2 dans $\mathcal{A}_2$ tels que A_1 ne soit pas dans $\mathcal{A}_2$ et A_2 ne soit pas dans $\mathcal{A}_1$).

Il en résulte que toute algèbre contenant deux algèbres $\mathcal{A}$ et $\mathcal{B}$ non emboîtées contient un événement qui n'est ni dans $\mathcal{A}$ ni dans $\mathcal{B}$.

b) atomes
Si A est un atome d'une tribu $\mathcal{T}$, toute variables aléatoire sur $(\Omega, \mathcal{T})$ est constante sur A. Si $\mathcal{T}$ est la tribu engendrée par une partition (A_n) de l'univers, toute variable aléatoire est constante sur chaque A_n.

Page 80, 1) : la proposition est énoncée de façon confuse. J'aurais du écrire

F est continue en x si et seulement si l'événement $X^{-1}(x)$ *est p-négligeable ;*
F sera donc continue sur $\mathbb{R}$ *si et seulement si* p_X *est diffuse.*

De façon générale, $p(X = x) = F(x) - F(x-)$.

Page 83, Complétion d'une tribu

Il y a équivalence entre : $X \in \widehat{\mathcal{T}}$, il existe $T \in \mathcal{T}$ tel que $X \Delta T$ soit négligeable, il existe N négligeable tel que $X \Delta N \in \mathcal{T}$. Cette assertion résulte des deux équivalences suivantes :

1) $X \in \widehat{\mathcal{T}} \Leftrightarrow$ il existe $T \in \mathcal{T}$ tel que $X \Delta T$ soit négligeable ;
2) $X \Delta T = N \Leftrightarrow X \Delta N = T \Leftrightarrow N \Delta T = X$.

Page 94, § 1, formule c) : il faut évidemment lire $p(A \Delta C) \le p(A \Delta B) + p(B \Delta C)$.

Page 103 : Tribus et familles monotones

Le concept de classe monotone serait du au mathématicien polonais Waclaw Sierpinski **(**1882 – 1969).

Page 104 : le théorème de la classe monotone

Voici un autre exemple d'emploi : si f intégrable sur R vérifie $\int_a^b f = 0$ pour tous réels a et b, alors f est nulle presque partout. Le plan de la preuve est le suivant :

a) on démontre d'abord que $\int_I f = 0$ pour tout intervalle I, soit avec le théorème de convergence monotone de Beppo-Levi soit avec le théorème de convergence dominée de Lebesgue. En particulier, $\int_R f = 0$.

b) on introduit la famille $\mathcal{M}$ des boréliens B de R tels que $\int_B f = 0$ et on prouve que $\mathcal{M}$ est une famille *monotone*.

c) puisque $\mathcal{M}$ contient les intervalles, $\mathcal{M}$ contient la famille monotone $\mathcal{M}'$ *engendrée* par les intervalles : comme les intervalles sont stables par intersection, $\mathcal{M}'$ est la tribu engendrée par les intervalles qui est la tribu des boréliens de R. Il en résulte que $\int_B f = 0$ pour tout borélien B de R.

d) soit $B_n = \{x \in R : f(x) > 1/n\}$: $0 = \int_{B_n} f \geq \frac{1}{n} m(B_n)$ et donc $m(B_n) = 0$. On en déduit que $\{x : f(x) > 0\}$ est de mesure nulle ; idem pour $\{x : f(x) < 0\}$, d'où f est nulle presque partout.

Page 105, 6.3

Soient X et X' deux variables aléatoires réelles définies sur $(\Omega, \mathcal{T}, p)$: pour établir que X et X' sont indépendantes il suffit d'établir que $p(X \in J \text{ et } X' \in J') = p(X \in J).p(X' \in J')$ pour tous les intervalles J et J'.

La démonstration se fait en deux étapes :

• soit J' un intervalle *fixé* et considérons l'ensemble des boréliens B de R vérifiant $p(X \in B \text{ et } X' \in J') = p(X \in B).p(X' \in J')$: cet ensemble contient tous les intervalles J de R, famille stable par intersection et qui engendre les boréliens ; il en résulte que $p(X \in B \text{ et } X' \in J') = p(X \in B).p(X' \in J')$ est vraie pour tout borélien B et tout intervalle J' car $B \mapsto p(X \in B \text{ et } X' \in J')$ et $B \mapsto p(X \in B).p(X' \in J')$ sont deux mesures de même masse.

• soit B un borélien *fixé* et envisageons l'ensemble des boréliens B' tels que $p(X \in B \text{ et } X' \in B') = p(X \in B).p(X' \in B')$: cet ensemble contient tous les intervalles J' et par le raisonnement précédent tous les boréliens B', aussi a-t-on $p(X \in B \text{ et } X' \in B') = p(X \in B).p(X' \in B')$ pour tous boréliens B et B', c'est la définition de l'indépendance.

Page 134, 8.6.4 : une erreur ?

A la deuxième ligne, il est affirmé que deux probabilités égales sur les boules de $\mathbb{R}^n$ seront égales sur les boréliens parce que « *tout ouvert est une réunion dénombrable de boules* » : l'argument donné est bancal car la réunion n'est pas disjointe. Il n'en reste pas moins que l'affirmation est vraie, et ceci pour *toute* norme sur $\mathbb{R}^n$. Consulter mon article : http://www.daniel-saada.eu/fichiers/26-Probabilites_sur_un_norme_de_dimension_finie.pdf

Page 134, 8.7

La démonstration faite prouve que l'approximation en mesure d'un borélien par un compact et un ouvert s'étend aux espaces métriques réunion dénombrable d'une suite de compacts.

Il en est de même pour les métriques qui sont à la fois *complets et séparables* : ce résultat est du à Ulam et date environ de 1939. En voici la preuve :

Comme E est à base dénombrable d'ouverts, E est la réunion d'une famille au plus dénombrable de boules ouvertes de rayon 1. Pour tout n, E est réunion d'une suite $(A_{n,i})_{i\in\mathbb{N}}$ de boules ouvertes de rayon 1/n.

On se donne $\varepsilon > 0$. Il existe un entier i_n tel que $m(\bigcup_{i=0}^{i_n} A_{n,i}) \geq m(E) - \varepsilon / 2^n$.

On introduit $B = \bigcap_{n\geq 1} \bigcup_{i=0}^{i_n} A_{n,i}$: B est borélien et $m(B) \geq m(E) - \varepsilon$ (passer au complémentaire).

B est précompact car pour tout n, B est contenu dans une famille *finie* de boules de rayon $1/n$.
L'adhérence $\overline{B}$ de B est compacte car précompacte et complète, et $m(\overline{B}) \geq m(B) \geq m(E) - \varepsilon$.

Ce qui est vrai pour E est vrai pour tout fermé F qui est complet et séparable : il existe K compact inclus dans F tel que $m(K) \geq m(F) - \varepsilon$. (Si besoin est, cliquez sur

http://www.daniel-saada.eu/fichiers/14-variables_aleatoires_a_valeurs_dans_un_metrique.pdf).

Page 148, simulation d'une loi de Poisson

L'algorithme figurant en haut de la page **149** souffre d'une coquille dans sa **ligne 3**. Il faut lire :

3. Si $P > e^{-\lambda}$ alors $X \leftarrow X + 1$ et aller en 2 (*et non en 3*)

J'en profite pour justifier l'algorithme. Quand $X = 0$, c'est que $P \leq e^{-\lambda}$ et comme P est uniforme sur [0,1], $p(X=0) = e^{-\lambda}$.
Appelons $(U_n)_{n\geq 1}$ la suite de variables aléatoires uniformes et indépendantes créée par le programme en ligne 2.
Ainsi, $X = 0$ signifie que $U_1 \leq e^{-\lambda}$, $X = 1$ veut dire que à $U_1 > e^{-\lambda}$ et $U_1.U_2 \leq e^{-\lambda}$, $X = n$
équivaut à

$$U_1.U_2...U_n > e^{-\lambda} \text{ et } U_1.U_2...U_n.U_{n+1} \leq e^{-\lambda}$$

ce qui prouve que X est une variable aléatoire sur la tribu commune où sont définies les U_i.

On démontre par récurrence sur n (http://www.daniel-saada.eu/fichiers/17-Densite_d_un_produit.pdf) que

$$p(U_1.U_2...U_n \leq a) = \frac{1}{(n-1)!}\int_0^a (-\ln x)^{n-1} dx \text{ quand } a \in [0,1].$$

En posant $t = -\ln x$, on déduit

$$p(U_1.U_2...U_n \leq e^{-\lambda}) = \frac{1}{(n-1)!}\int_\lambda^{+\infty} t^{n-1} e^{-t} dt = \sum_{k=0}^{n-1} e^{-\lambda}\frac{\lambda^k}{k!}$$ (voir **fin de la page 26 du livre**).

On remarque enfin que

$$p(U_1.U_2...U_n > e^{-\lambda} \text{ et } U_1.U_2...U_n.U_{n+1} \leq e^{-\lambda}) = p(U_1.U_2...U_n.U_{n+1} \leq e^{-\lambda}) - p(U_1.U_2...U_n \leq e^{-\lambda})$$

d'où $p(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}$.

Page 156, ligne 4 : il faut lire « dénombrable $\sum_n K_n$ **est** borélienne »

Page 159, 10.6.1

Une fonction numérique définie sur un segment est intégrable au sens de Riemann sur ce segment si et seulement si elle est bornée et continue presque partout : cette caractérisation est due à Lebesgue.
On pourra consulter l'ouvrage de Jean-Marie Arnaudiès, *L'intégrale de Lebesgue*, Vuibert, 1997, **§ 2.16**.

Exemple. La fonction indicatrice $1_{\mathbb{Q}}$ n'a aucun point de continuité, elle n'est donc intégrable au sens de Riemann sur aucun segment, tandis que $\int_a^b 1_{\mathbb{Q}}\, dm$ vaut évidemment 0.

Page 160, Théorème de Sard

a) Du fait qu'une fonction f à dérivée bornée transforme un négligeable en un négligeable (page **156**), le théorème de Sard est sans intérêt si l'ensemble des points à dérivée nulle est négligeable, par exemple dénombrable.
Il en va tout autrement quand les zéros de f' sont de mesure non nulle (voir par exemple le paragraphe **B** de http://www.daniel-saada.eu/fichiers/06-fonctions_mesurables_sur_la_tribu_des_singletons.pdf).

b) Évidemment, la mesure de $f\left((f')^{-1}(a)\right)$ est nulle pour tout réel a (raisonner sur $f(x)-ax$).

c) Le théorème de Sard est encore vrai si f est seulement supposée dérivable sur $[0,1]$.
(Rafik Imekraz et Dominique Malicet, **RMS 123-4**, 2012-2013, *Questions-Réponses*, **R742**).

Il est à noter que l'on perd en général le caractère borélien du négligeable $f(E)$. En effet, l'ensemble des zéros de f' est un G_δ en vertu de **11.6.3.1** (page **178**) et l'image d'un G_δ par une fonction continue n'est pas toujours borélienne.

Page 161 3) : Un théorème de Lebesgue sur les fonctions monotones

Le fait qu'une fonction monotone soit dérivable presque partout entraîne l'intéressante conséquence suivante : $g(x)=\int_a^x f(x)dx$ (f intégrable sur $\mathbb{R}$) est dérivable presque partout. En effet, la décomposition $f=f^+-f^-$ montre que g est la différence de deux fonctions croissantes, dérivables chacune en dehors d'un ensemble négligeable.

Page 176, 11.6

a) Il existe un critère très simple de mesurabilité quand $f(\Omega)$ est finie ou dénombrable : il faut et il suffit que pour tout réel x, $f^{-1}(x)$ soit dans la tribu. L'image réciproque par f de *toute* partie de $\mathbb{R}$ est alors dans la tribu.

b) J'ai glané sur le forum *Les-Mathématiques.net* l'intéressant exercice suivant : l'univers Ω étant muni de la tribu $\mathcal{T}$ engendrée par ses singletons, une fonction f de Ω dans $\mathbb{R}$ est $\mathcal{T}$ -mesurable si, et seulement si, il existe D dénombrable dans Ω tel que f soit constante sur $\overline{D}$; solution sur

http://www.daniel-saada.eu/fichiers/06-fonctions_mesurables_sur_la_tribu_des_singletons.pdf

Page 179 : toute f mesurable est limite d'une suite (h_n) de fonctions étagées

a) La fonction étagée h_n vaut $(k-1)M/n$ sur l'événement A_k, *et non* kM/n ; la formule donnant h_n est donc $h_n=\sum_{k=1}^{n}\frac{(k-1)M}{n}.1_{A_k}$; on a bien $0\le f(\omega)-h_n(\omega)\le M/n$.

b) Quand f n'est plus majorée, M n'existe plus, et il faut lire $h_n = \sum_{k=1}^{n.2^n} \frac{k-1}{2^n}.1_{\mathrm{A_k}} + n.1_{\mathrm{B}}$, avec les notations du texte ; à partir d'un certain rang, $f(\omega) \leq n$ et $0 \leq f(\omega) - h_n(\omega) \leq 1/2^n$.

Page 181, paragraphe 11.7

a) Soit m une mesure de probabilité discrète sur un univers Ω : il existe une partie dénombrable D de l'univers telle que $m(\Omega - D) = 0$. Posons $D = \{x_n\}_{n\in\mathbb{N}}$ et $p_n = m(x_n)$: une fonction $f : \Omega \to R$ est intégrable pour m si et seulement si $\sum_0^\infty p_n \, | f(x_n) | < +\infty$ et alors $\int_\Omega f \, dm = \sum_0^\infty p_n f(x_n)$.

Démonstration sur http://perso.univ-rennes1.fr/lionel.truquet/polyM1proba.pdf , page 42.

Exemple. Soit m sur $\Omega = [0, 2\pi[$ telle que : $m(\Omega) = 8, m(0) = m(\pi/2) = m(\pi) = m(3\pi/2) = 2$; soit f la fonction $x \mapsto | a\cos x + b\sin x |$, où a et b sont deux paramètres réels. Alors $\frac{1}{4}\int_{[0,2\pi[} f \, dm = | a | + | b |$.

b) Si m_1 et m_2 sont deux mesures de probabilité sur $(\Omega, \mathcal{T})$, il est légitime de se demander quelles sont les fonctions $\mathcal{T}$ -mesurables qui sont à la fois intégrables pour m_1 et m_2. La réponse est simple et belle :

$$\mathcal{L}^1(m_1) \cap \mathcal{L}^1(m_2) = \mathcal{L}^1(m_1 + m_2).$$

De plus, quand f est $(m_1 + m_2)$ -intégrable, $\int_\Omega f \, d(m_1 + m_2) = \int_\Omega f \, dm_1 + \int_\Omega f \, dm_2$.

En d'autres termes, l'application $(m, f) \to \int_\Omega f \, dm$ est linéaire en m et en f.

Page 185, le théorème de convergence monotone de Beppo-Levi

Voici un exemple d'emploi : l'existence et le calcul de l'intégrale triple $\int_0^1\int_0^1\int_0^1 \frac{1}{1-xyz} dxdydz$.

La fonction $F(x, y, z) = 1/(1 - xyz)$ est définie et continue sur le cube $[0,1]^3$ privé du sommet $(1,1,1)$, ensemble que nous nommerons D. Sur D, F est la limite de la suite *croissante* des fonctions $F_n(x, y, z) = \sum_0^n (xyz)^k$, et $\int_D F_n = \sum_0^n 1/(n+1)^3$. Il en résulte que F est intégrable sur D et que $\int_D F = \sum_1^\infty 1/n^3$.

Page 186, le théorème de convergence dominée de Lebesgue

Voici un exemple d'emploi : la fonction $f(t) = \int_{-1}^1 \cos(tx) \frac{dx}{\sqrt{1-x^2}}$ est continue sur $\mathbb{R}$.

La fonction f est définie sur $\mathbb{R}$ car $\left|\cos(tx)/\sqrt{1-x^2}\right|$ est majoré par $1/\sqrt{1-x^2}$ qui, bien que non bornée, est intégrable sur $[-1,1]$, son intégrale valant π. Soit (t_n) une suite de réels de limite t :

- $\cos(t_n x)/\sqrt{1-x^2}$ tend vers $\cos(tx)/\sqrt{1-x^2}$ pour x fixé dans $]-1,1[$;
- $\left|\cos(t_n x)/\sqrt{1-x^2}\right|$ est dominée par la fonction intégrable $1/\sqrt{1-x^2}$.

Il en résulte que $f(t_n)$ a pour limite $f(t)$.

Voici maintenant un contre-exemple : $g(x) = \int_0^{+\infty} xe^{-xt} dt$ vaut 1 pour $x > 0$ tandis que $g(0) = 0$.

Page 185 à 187, les théorèmes de Beppo-Levi et de Lebesgue

Ces deux théorèmes donnent deux conditions suffisantes pour que la limite f d'une suite f_n de fonctions intégrables soit intégrable :

- il suffit que la suite f_n soit monotone et la suite des intégrales $\int f_n$ bornée ;
- il suffit qu'il existe une fonction intégrable g majorant $|f_n|$.

Il est utile de connaître une condition plus large assurant l'intégrabilité de f, en perdant au passage l'égalité $\int f = \lim \int f_n$: il suffit que les intégrales $\int |f_n|$ soient bornées.

La décomposition $f = f^+ - f^-$ et l'égalité $|f| = f^+ + f^-$ montrent qu'il suffit de l'établir pour f, et donc pour des f_n, positives. Posons $g_n = \inf\{f_k : k \geq n\}$: les g_n sont définies sur tout l'univers, mesurables, et forment une *suite croissante de limite* f. Comme $g_n \leq f_n$, les intégrales des g_n sont majorées et donc f est intégrable, avec $\int f = \lim_n \int g_n \leq \sup_n \int f_n$.

En particulier, toujours dans le cas où f et les f_n sont positives, lorsque la suite des intégrales $\int f_n$ est *convergente*, f est intégrable et $\int f \leq L = \lim \int f_n$. En effet, à partir d'un certain rang, $\int f_n \leq L + \varepsilon$, aussi f est intégrable et $\int f \leq L + \varepsilon$; comme ε est arbitrairement voisin de zéro, $\int f \leq L$.

Application : la « dérivée » d'une fonction croissante est intégrable. On sait qu'une fonction F croissante sur $\mathbb{R}$ est dérivable presque partout (p. **161**). Sur l'ensemble où elle existe, la fonction positive $f = F'$ est la limite de la suite $f_n : x \to \dfrac{F(x+1/n) - F(x)}{1/n}$. Sur chaque $[a,b]$, les intégrales des f_n forment une suite convergente. En effet, $\int_a^b f_n(x)dx = n\left(\int_b^{b+1/n} F(x)dx - \int_a^{a+1/n} F(x)dx\right)$ a pour limite $F(b_+) - F(a_+)$.

Aussi F' est intégrable sur $[a,b]$ et $\int_a^b F' \leq F(b_+) - F(a_+)$. Quand F est de plus continue, $\int_a^b F' \leq F(b) - F(a)$; évidemment, il est fait grand cas des F continues croissantes telles que $\int_a^b F' = F(b) - F(a)$.

Page 187, paragraphe 11.9.1

J'aurais pu ajouter la propriété suivante : si f p-intégrable est > 0 sur A de la tribu et si $p(A) > 0$, alors $\int_A f dp > 0$.

En effet, si $A_n = \{\omega \in A : f(\omega) > 1/n\}$, A est la réunion croissante des A_n et donc $p(A)$ est la limite des $p(A_n)$. Il existe alors n tel que $p(A_n) > \frac{1}{2} p(A)$, d'où $\int_A f dp \geq \int_{A_n} f dp > \dfrac{p(A)}{2n}$.

Page 189 c) : dérivée d'une intégrale

Il a été rappelé que si $g(x) = \int_a^x f(t)dt$, $g'(x) = f(x)$ en tout point de continuité x de f. Cette condition suffisante de dérivabilité de l'intégrale g n'est nullement nécessaire : si $f(x) = 2x$ pour x non nul et $f(0) = 1$, $g(x)$ vaut néanmoins x^2 et $g'(0)$ existe bien que f ne soit pas continue en 0. Un exemple plus élaboré est donné dans http://www.daniel-saada.eu/fichiers/10-primitives.pdf : $g(x) = \int_0^x \sin(1/t)dt$ est dérivable en $x = 0$ alors que

$x \to \sin(1/x)$ ne possède aucun prolongement continu en 0. Plus fort encore : g est dérivable en tout x, sauf peut-être sur un ensemble de mesure nulle !

Page 191 : espérance d'une variable aléatoire

Calculons, à titre d'application, l'espérance de X uniforme sur $(0,1)$.

On divise l'intervalle en n parties égales. X étant diffuse, posons $T_k = X^{-1}\left(\frac{k}{n},\frac{k+1}{n}\right)$ et soit h_n la fonction étagée valant k/n pour k allant de 0 à $n-1$.

L'intégrale de h_n est par définition $\sum_{k=0}^{n-1} \frac{k}{n} p(T_k) = \sum_{k=0}^{n-1} k/n^2 = \frac{n-1}{2n}$, d'où $E(X) = 1/2$ par passage à la limite.

Page 196, ligne 6 de 12.4, il faut lire :

Si m (positive) est diffuse, il en est de même de toute mesure positive m' telle que $m' << m$.

Lorsqu'on a la fois $m' << m$ et $m << m'$, on dit que m et m' sont *équivalentes*.

Page 200 : le théorème de Radon-Nikodym

Quand Ω est un espace $\mathbb{R}^n$, la mystérieuse fonction f peut être illustrée comme suit. Soit A_r la boule de centre x et de rayon r, x étant fixé et r variant ; de l'égalité $m'(A_r) = \int_{A_r} f\,dm$, on peut déduire que $f(x)$ est la limite de $\frac{m'(A_r)}{m(A_r)}$ quand $r \to 0$. C'est facile à prouver quand x est un point de continuité de f, difficile à établir dans le cas général d'autant plus que l'égalité a lieu seulement presque partout.

Page 210 : produit de deux tribus

Soient X_1 la projection de $\Omega_1 \times \Omega_2$ sur Ω_1 et X_2 la projection de $\Omega_1 \times \Omega_2$ sur Ω_2.

La tribu produit est la plus petite tribu $\mathcal{T}$ sur $\Omega_1 \times \Omega_2$ qui rende *mesurables* ces deux projections.

En effet, pour que $X_1 : (\Omega_1 \times \Omega_2, \mathcal{T}) \mapsto (\Omega_1, \mathcal{T}_1)$ soit mesurable, il faut et il suffit que pour tout T_1 de $\mathcal{T}_1$, $(X_1)^{-1}(T_1) \in \mathcal{T}$, or $(X_1)^{-1}(T_1) = T_1 \times \Omega_2$; ainsi, la tribu $\mathcal{T}$ doit contenir tous les produits cartésiens $T_1 \times \Omega_2$, et pour que X_2 soit mesurable, $\mathcal{T}$ doit contenir également tous les $\Omega_1 \times T_2$. Il en résulte que $\mathcal{T}$ contient les $T_1 \times T_2 = (T_1 \times \Omega_2) \cap (\Omega_1 \times T_2)$.

Page 211, nouvelle description de la tribu borélienne de $\mathbb{R}^n$

On a prouvé que $\mathcal{B}^2 = \mathcal{B}^1 \otimes \mathcal{B}^1$ et on a affirmé que $\mathcal{B}^{n-1} \otimes \mathcal{B}^1 = \mathcal{B}^n$. Plus généralement, $\mathcal{B}^n \otimes \mathcal{B}^m = \mathcal{B}^{n+m}$, parce que $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ et que chaque $\mathbb{R}^k$ est un espace métrique séparable, c'est-à-dire contenant une suite dense. Plus généralement encore, la tribu borélienne d'un espace topologique E, notée $\mathcal{B}(E)$, étant la tribu engendrée par les ouverts de E, on a toujours $\mathcal{B}(E) \otimes \mathcal{B}(F) \subset \mathcal{B}(E \times F)$; si de plus E et F sont métrisables séparables, ou à base dénombrable d'ouverts (http://fr.wikipedia.org/wiki/Espace_%C3%A0_base_d%C3%A9nombrable), alors $\mathcal{B}(E \times F) = \mathcal{B}(E) \otimes \mathcal{B}(F)$. Justifions :

a) Par définition de la topologie produit, $U \times V$ est un ouvert de $E \times F$ si U est un ouvert de E et V un ouvert de F. Comme $\mathcal{B}(E) \otimes \mathcal{B}(F)$ est engendrée par les $U \times V$ (page **210**), $\mathcal{B}(E) \otimes \mathcal{B}(F) \subset \mathcal{B}(E \times F)$.

b) Par définition encore de la topologie produit, un ouvert W de $E\times F$ est une réunion de produits d'ouverts de chaque facteur, $W=\bigcup_i U_i\times V_i$; quand E et F sont à bases dénombrables, il en de même de $E\times F$ et W est alors une réunion *dénombrable* de $U_i\times V_i$. W appartient donc à la tribu produit $\mathcal{B}(E)\otimes\mathcal{B}(F)$, d'où l'inclusion inverse $\mathcal{B}(E\times F)\subset\mathcal{B}(E)\otimes\mathcal{B}(F)$.

Page 211 : produit de deux mesures

Quand $Z=(X,Y)$ avec X et Y indépendantes, $p_Z=p_X\otimes p_Y$, ce qui renforce l'intérêt des mesures produits. La réciproque est évidente si on se souvient qu'une probabilité q coïncide avec $p_X\otimes p_Y$ dès que $q(A\times B)=p_X(A).p_Y(B)$ (http://www.daniel-saada.eu/fichiers/24-Densites_d_un_couple.pdf).

Page 212 : mesure produit $m=m_1\otimes m_2$ et sections

Si A est dans la tribu produit $\mathcal{T}=\mathcal{T}_1\otimes\mathcal{T}_2$, on a démontré page **210** que toutes les sections A_x de A sont dans la tribu $\mathcal{T}_2$. Il est extrêmement intéressant de savoir qu'alors $m(A)=\int_{\Omega_1} m_2(A_x)$. Cette formule présente un double intérêt : elle montre que le calcul intégral est indispensable, elle explique pourquoi l'aire sous la courbe d'équation $y=f(x)$ s'obtient en intégrant $f(x)$ qui n'est autre que la longueur de la section de la surface en x.

Pour établir la formule, on introduit la famille $\mathcal{F}=\left\{A\in\mathcal{T} : x\to m_2(A_x) \text{ est m}_2\text{-intégrable et } m(A)=\int_{\Omega_1} m_2(A_x)\right\}$: c'est une famille monotone (page **103**) qui contient tous les pavés de $\mathcal{T}$; on en déduit que $\mathcal{F}$ est $\mathcal{T}$ tout entier en vertu du Théorème de la classe monotone (p. **104**). Évidemment, on a aussi $m(A)=\int_{\Omega_2} m_1(A_y)$.

Page 214 : les théorèmes de Fubini (1910)

Il s'agit de ramener une intégrale sur un espace produit Ω^2 (pour simplifier) à deux intégrales sur Ω. La formule de base est la suivante : $\int_{\Omega^2} f(x,y)d(p\otimes q)=\int_\Omega\left(\int_\Omega f(x,y)dp(x)\right)dq(y)=\int_\Omega\left(\int_\Omega f(x,y)dq(y)\right)dp(x)$, f devant être au minimum mesurable pour $p\otimes q$.

On remarquera que quand f est la fonction indicatrice d'un pavé, les égalités de Fubini se réduisent à la définition de la mesure produit. On remarquera aussi que la deuxième égalité $\int_\Omega\left(\int_\Omega f(x,y)dp(x)\right)dq(y)=\int_\Omega\left(\int_\Omega f(x,y)dq(y)\right)dp(x)$ est vraie quand $f(x,y)$ est de la forme $g(x)h(y)$.

Application 1. Montrons moins laborieusement l'égalité de **14.3** :

si $X(\Omega)\subset[a,b]$, alors $E(X)=a+\int_a^b\left(1-F_X(t)\right)dt$.

Pour la clarté de la démonstration, nous supposerons $a=0$. Comme $X(\omega)=\int_0^{X(\omega)}1dt=\int_0^b 1_{(t\leq X(\omega))}dt$,

$$E(X)=\int_\Omega Xdp=\int_\Omega X(\omega)dp(\omega)=\int_\Omega\left(\int_0^b 1_{(t\leq X(\omega))}dt\right)dp(\omega)=\int_0^b\left(1_{(t\leq X(\omega))}dp(\omega)\right)dt=\int_0^b p(X\geq t)\,dt\ .$$

Rappelons enfin que $\int_0^b p(X\geq t)\,dt=\int_0^b p(X>t)\,dt$ car $p(X=t)$ est presque toujours nulle (**page 80 du livre**).

Application 2 : espérance d'une variable aléatoire définie par une intégrale.

Soit X une variable aléatoire sur $(\Omega,\mathcal{T},p)$ définie par $X(\omega)=\int_{\Omega'} f(\omega,\omega')dp'(\omega')$ où $(\Omega',\mathcal{T}',p')$ est donné.

Alors, l'espérance de X, si elle existe, vaut

$$E(X)=\int_{\Omega}\left(\int_{\Omega'} f(\omega,\omega')dp'(\omega')\right)dp(\omega)=\int_{\Omega'}\left(\int_{\Omega} f(\omega,\omega')dp(\omega)\right)dp'(\omega')$$

et donc $E(X)=\int_{\Omega'} E\left(\omega'\mapsto f(\omega,\omega')\right)dp'(\omega')$.

Exemple. Soit X et Y deux variables aléatoires réelles. Grâce à la formule $|a-b|=\int_R \left(1_{a<t<b}+1_{b<t<a}\right)dt$, facile à vérifier, on a donc $|X(\omega)-Y(\omega)|=\int_R \left(1_{X(\omega)<t<Y(\omega)}+1_{Y(\omega)<t<X(\omega)}\right)dt$, d'où

$$E|X-Y|=\int_R E\left(1_{X(\omega)<t<Y(\omega)}+1_{Y(\omega)<t<X(\omega)}\right)dt\ .$$

Or $E(1_A)=P(A)$, aussi $E\left(1_{X(\omega)<t<Y(\omega)}\right)=P\left((X<t)\cap(Y>t)\right)$ et $E\left(1_{Y(\omega)<t<X(\omega)}\right)=P\left((X>t)\cap(Y<t)\right)$.

Si X et Y sont de plus indépendantes : $E|X-Y|=\int_R \left(P\left(X<t\right)P\left(Y>t\right)+P\left(X>t\right)P\left(Y<t\right)\right)dt$.

Si X et Y ont en outre même loi : $E|X-Y|=2\int_R P\left(X<t\right)P\left(X>t\right)dt$.

Page 215, *Un résultat sur les surfaces*

Le résultat est encore vrai si l'une des deux fonctions f ou g est dérivable, l'autre étant seulement continue.

(Rafik Imekraz et Dominique Malicet, **RMS 123-4**, 2012-2013, *Questions-Réponses*, **R742**).

Page 217, linéarité de l'espérance mathématique

1) La variable aléatoire somme $Z(x,y)=X(x)+Y(y)$**.**

Voici plus simple pour établir l'égalité $E(Z)=E(X)+E(Y)$: $Z=U+V$, avec $U(x,y)=X(x)$ et $V(x,y)=Y(y)$
Les fonctions U et V sont des variables aléatoires sur Ω^2 : par exemple, $(U\le a)=(X\le a)\times\Omega$, qui est dans la tribu produit. La loi de U est celle de X, la loi de V est celle de Y : par exemple,

$$q(U\le a)=q\left((X\le a)\times\Omega\right)=p(X\le a)\ .$$

On en déduit $E(U)=E(X)$, $E(V)=E(Y)$, $E(Z)=E(U)+E(V)=E(X)+E(Y)$.

De plus, les variables U et V sont indépendantes pour la probabilité produit, par exemple

$$q(U\le a \text{ et } V\le b)=q\left((X\le a)\times(Y\le b)\right)=p(X\le a).p(Y\le b)=q(U\le a).q(V\le b)\ .$$

Aussi la loi de Z est celle de $X+Y$ quand on suppose X et Y *indépendantes* ; en outre, on aura

$$V(Z)=V(U)+V(V)=V(X)+V(Y)\ .$$

2) Un exemple d'application.

La linéarité est intéressante quand la loi de X est compliquée et que l'on peut décomposer X en somme finie.

L'index d'un livre recense les occurrences de certains mots d'un livre, en procédant "par blocs" : si le mot apparaît aux pages 16, 19 à 22 et 33 à 45, on note "16, 19-22, 33-45" dans l'index, ce qui compte pour 3 blocs.
On suppose qu'un mot apparaît avec une probabilité p sur chaque page d'un livre de n pages, indépendamment des autres pages. On note X_n la variable aléatoire comptant le nombre de blocs d'un mot dans l'index. Il s'agit de calculer l'espérance de X_n en fonction de n et p. Voici ma solution.

La loi de $X_{n+1}-X_n$ est simple : $X_{n+1}-X_n$ ne prend que les valeurs 0 ou 1 ; $X_{n+1}-X_n$ ne prend la valeur 1 que si le mot est apparu à la page n+1 sans être apparu à la page n, événement de probabilité $p(1-p)$.
Il en résulte que $E\left(X_{n+1}-X_n\right)=p(1-p)=E\left(X_{n+1}\right)-E\left(X_n\right)$.
Comme $E\left(X_1\right)=p$, $E\left(X_{n+1}\right)=p+np(1-p)$ et donc $E\left(X_n\right)=np-(n-1)p^2$.

Page 219 : une inégalité remarquable sur l'espérance mathématique

L'inégalité $E\,|X+Y|\geq E\,|X|$, valable quand X et Y sont indépendantes et de même loi, est la conséquence d'une inégalité plus générale : $E\,|X+Y|\geq E\,|X-Y|$, que l'on découvrira sur

http://www.daniel-saada.eu/Notes/22-Inegalites-euclidiennes.pdf

En effet, l'identité $X=\dfrac{X-Y}{2}+\dfrac{X+Y}{2}$ implique $|X|\leq\left|\dfrac{X-Y}{2}\right|+\left|\dfrac{X+Y}{2}\right|$ et donc

$$E\left(|X|\right)\leq\frac{E\left(|X-Y|\right)+E\left(|X+Y|\right)}{2}\leq E\left(|X+Y|\right).$$

Page 224, 14.1 : exemples de fonctions boréliennes

J'ai donné quatre exemples de fonctions boréliennes, en voici un cinquième :

les fonctions continues sauf sur un ensemble fini ou dénombrable.

Soit D l'ensemble dénombrable des discontinuités de f, U un ouvert de R, $V=f^{-1}(U)$.

Soit $x\in V-\overset{\circ}{V}$, où $\overset{\circ}{V}$ est l'intérieur de V : $u=f(x)\in U$; comme U est ouvert, il existe un ouvert W tel que $u\in W\subset U$. On a $x\in f^{-1}(W)\subset V$: il en résulte que $f^{-1}(W)$ n'est pas ouvert, sinon x serait dans $\overset{\circ}{V}$, et donc que f n'est pas continue en x

V est donc la réunion de son intérieur, qui est un ouvert, et de points de discontinuités : $V=\overset{\circ}{V}+V\cap D$ est donc un borélien et f est borélienne comme annoncé.

On remarquera qu'en définitive, $V=f^{-1}(U)$ est la réunion de son intérieur avec un dénombrable : $f^{-1}(U)$ est donc un F_σ et f est de première classe de Baire ; de ce fait, f est limite simple d'une suite de fonctions continues.

Dans le langage des variables aléatoires :

toute fonction X de R dans R continue sauf sur un dénombrable est une variable aléatoire ;
de plus, X est limite simple d'une suite de fonctions continues.

En revanche, si les discontinuités de f ont la puissance du continu, f peut n'être pas borélienne <u>même si l'ensemble de ses discontinuités est de mesure nulle</u> (on dit que f est continue presque partout) :

http://www.daniel-saada.eu/fichiers/33-Fonctions-continues-presque-partout.pdf

Page 224, 14.2 : le théorème du transfert

1) Interprétation

Par définition, $p(X^{-1}(B))=p_X(B)$; en langage intégral, cette égalité s'écrit $\int_\Omega 1_{X^{-1}(B)}dp=\int_{\mathbb{R}}1_B dp_X$, mais comme

$1_{X^{-1}(B)} = 1_{B\circ X}$, il vient $\int_\Omega 1_{B\circ X}\, dp = \int_{\mathbb{R}} 1_B\, dp_X$. Le théorème du transfert est donc *l'élargissement* de cette égalité à une classe plus vaste de fonctions que les fonctions indicatrices.

2) Nouvelle démonstration

a) On montre d'abord que si les variables aléatoires X, de $(\Omega, \mathcal{T}, p)$ dans $\mathbb{R}$, et Y de $(\Omega', \mathcal{T}', p')$ dans $\mathbb{R}$ ont même loi, alors $\int_\Omega X dp = \int_{\Omega'} X' dp'$.

Décomposons X et Y en $X = X^+ - X^-$ et $Y = Y^+ - Y^-$: si X et Y ont même loi, il en est de même de leurs parties positives et négatives, et on aura donc

$$\int_\Omega X^+ dp = \int_{\Omega'} Y^+ dp' \text{ et } \int_\Omega X^- dp = \int_{\Omega'} Y^- dp'.$$

On peut donc supposer X et Y positives : X est limite de la suite *croissante* des fonctions étagées

$$h_n = \sum_{i=1}^{n.2^n} \frac{i-1}{2^n}.1_{\mathrm{A}_i} + n.1_{\mathrm{B}_n}$$

où $A_i = \left\{\omega \in \Omega : \frac{i-1}{2^n} \le X(\omega) < \frac{i}{2^n}\right\}$ pour i allant de 1 à $n.2^n$ et $B_n = \{\omega \in \Omega : X(\omega) \ge n\}$.

Par définition de l'intégrale, $\int_\Omega h_n dp = \sum_1^{n.2^n} \frac{i-1}{2^n} p(A_i) + np(B_n)$ et $\int_\Omega X\, dp = \lim_n \int_\Omega h_n dp$.

De même, Y est limite de la suite *croissante* des fonctions étagées $k_n = \sum_{i=1}^{n.2^n} \frac{i-1}{2^n}.1_{C_i} + n.1_{D_n}$,

où $C_i = \left\{\omega \in \Omega : \frac{i-1}{2^n} \le Y(\omega) < \frac{i}{2^n}\right\}$ pour i allant de 1 à $n.2^n$ et $D_n = \{\omega \in \Omega : Y(\omega) \ge n\}$.

Par définition encore, $\int_\Omega k_n dp = \sum_1^{n.2^n} \frac{i-1}{2^n} p'(C_i) + np'(D_n)$ et $\int_\Omega Y\, dp = \lim_n \int_\Omega k_n dp$.

Comme X et Y ont même loi, $p(A_i) = p'(C_i)$ pour tout i et $p(B_n) = p'(D_n)$ et donc

$$\int_\Omega h_n dp = \int_{\Omega'} k_n dp', \text{ puis } \int_\Omega X dp = \int_{\Omega'} Y dp'.$$

Plus généralement, si φ est une fonction mesurable de $\mathbb{R}$ dans $\mathbb{R}$, $\varphi \circ X$ est une variable sur $(\Omega, \mathcal{T}, p)$ et $\varphi \circ Y$ est une variable sur $(\Omega', \mathcal{T}', p')$. Comme $p_X = p_Y$ implique $p_{\varphi\circ X} = p_{\varphi\circ Y}$:

$$\int_\Omega (\varphi \circ X) dp = \int_{\Omega'} (\varphi \circ Y) dp'.$$

b) Considérons la fonction mesurable φ comme une variable aléatoire réelle sur $(\mathbb{R}, \mathcal{B}^1, p_X)$: sa loi est donnée par $p_\varphi = p_{\varphi\circ X}$ et donc $\int_\Omega (\varphi \circ X) dp = \int_{\mathbb{R}} \varphi . dp_X$.

3) Généralisation

Le théorème du transfert s'énonce ainsi : $\int_\Omega (f \circ X)\,dp = \int_R f\,d(p \circ X^{-1})$, où X est une variable aléatoire réelle sur $(\Omega, \mathcal{T}, p)$ et f mesurable, ou borélienne, de R dans R. Plus généralement, si X est mesurable de $(\Omega, \mathcal{T})$ dans $(\Omega', \mathcal{T}')$ et si f est mesurable de Ω' dans R, alors, sous réserve de convergence :

$$\int_\Omega (f \circ X)\,dp = \int_{\Omega'} f\,d(p \circ X^{-1}).$$

L'intégrale sur Ω a donc été *transférée* sur Ω'.

4) Complément

Quand une fonction φ de $\mathbb{R}$ dans $\mathbb{R}$ est intégrable pour la probabilité p_X, alors certes $\int_\Omega (\varphi \circ X)dp = \int_R \varphi dp_X$ mais rien n'a été dit sur le calcul effectif de $\int_R \varphi dp_X$. Si X a une densité f, nous allons prouver que pour toute φ intégrable selon p_X : $\boxed{\varphi.f \text{ est intégrable sur } \mathbb{R} \text{ pour la mesure de Lebesgue et } \int_R \varphi dp_X = \int_R \varphi(x) f(x)\,dx}$.

La démonstration tient en quatre points :

a) Les probabilités p_X et $B \to \int_B f$ sont deux probabilités sur la tribu borélienne de $\mathbb{R}$. Comme elles sont diffuses, on a $p_X(I) = F(\max I) - F(\min I) = \int_I f$ pour tout intervalle I. Il en résulte (par exemple **6.3, p 105**) que

$$p_X(B) = \int_B f \text{ pour tout borélien } B.$$

b) Les deux assertions sont vraies quand φ est étagée sur la tribu des boréliens de $\mathbb{R}$.
Soit $\varphi = \sum_1^n a_i.1_{B_i}$ avec $\mathbb{R} = \sum_1^n B_i$, chaque B_i étant un borélien. La fonction $\varphi.f$ valant $a_i f$ sur B_i, il en résulte que $|\varphi.f|$ est majorée sur $\mathbb{R}$ par $\max\limits_{1\le i\le n} |a_i|.f$ et donc que $\varphi.f$ est intégrable sur $\mathbb{R}$ comme f.
Deuxièmement, comme $\int_R \varphi dp_X = \sum_1^n a_i.p_X(B_i)$ et $\int_R \varphi(x)f(x)\,dx = \sum_1^n \int_{B_i} \varphi(x) f(x)\,dx = \sum_1^n a_i \int_{B_i} f(x)\,dx$, l'égalité des intégrales provient de la formule vue en **3)** : $p_X(B) = \int_B f$ pour tout borélien B.

c) Les deux assertions sont vraies quand φ est positive.
Il existe une suite croissante de fonctions étagées positives h_n de limite φ et $\int_R \varphi dp_X = \lim\limits_n \int_R h_n dp_X$.

La suite des fonctions *intégrables* $h_n f$ tend en croissant vers φf parce que f est positive ; d'autre part, la suite des intégrales $\int_R h_n(x) f(x)\,dx$ est majorée par $\int_R \varphi dp_X$ car $\int_R h_n(x) f(x)\,dx = \int_R h_n dp_X$.
Le théorème de Beppo-Levi (**11.8**, p 185) assure alors que $\int_R \varphi(x) f(x)\,dx$ existe et que

$$\int_R \varphi(x) f(x)\,dx = \lim_n \int_R h_n(x) f(x)\,dx.$$

Comme $\int_R h_n(x) f(x)\,dx = \int_R h_n\,dp_X$ d'après **a)**, $\int_R \varphi(x) f(x)\,dx = \int_R \varphi dp_X$.

d) Enfin, si φ n'est pas de signe constant, la décomposition $\varphi = \varphi^+ - \varphi^-$ et la linéarité de l'intégrale suffisent pour conclure. En particulier :

– pour $\varphi = id_{\mathbb{R}}$, $E(X) = \int_{\mathbb{R}} x.f(x)dx$, à condition bien sûr que $\int_{\mathbb{R}} |x|.f(x)dx < +\infty$;

– pour la fonction complexe $\varphi(x) = e^{itx}$, $E(e^{itX}) = \int_{\mathbb{R}} e^{itx} f(x)dx$ existe pour tout réel t, et si $t \to \left|E(e^{itX})\right|$ est intégrable sur $\mathbb{R}$, $f(x)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}E(e^{itx})dt$, c'est la formule d'inversion de Fourier : http://fr.wikipedia.org/wiki/Transform%C3%A9e_de_Fourier.

• Une formule pour l'espérance de $1/X$, $X>0$: $E(1/X)=\int_0^{+\infty} E(e^{-tX})dt$.

En effet, $E(e^{-tX})=\int_0^{+\infty} e^{-tx}f(x)dx$ et donc

$$\int_0^{+\infty} E(e^{-tX})dt=\int_0^{+\infty} dt\left(\int_0^{+\infty} e^{-tx}f(x)dx\right)=\int_0^{+\infty} f(x)dx\left(\int_0^{+\infty} e^{-tx}dt\right)=\int_0^{+\infty}\frac{f(x)}{x}dx=E(1/X).$$

Page 228, le Théorème d'Egorov

Exemple : la suite des fonctions continues $x\mapsto x^n$ ne converge pas uniformément sur l'univers $[0,1]$, mais sur tout intervalle $[0,1-\varepsilon]$.

Page 230, 14.6 : le Théorème de Lusin

Exemple : http://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_de_Lusin

Page 232, 14.7 : 3 erreurs

1) a) : c'est le théorème de Lusin **(14.6)** qu'il faut utiliser et non celui d'Egorov **.**
1) c) : à la deuxième ligne, il faut lire : en effet, soit y un élément de $K\cap(K+x)$ et non $K\cap(K+t)$.
2) : la ligne 5 doit être corrigée ainsi : $f(\lambda x)=\lim_n f(r_n x)=\lim_n r_n f(x)=\lambda f(x)$.

Page 236, d)

L'égalité $1_{\mathbb{Q}}(x)=\lim_{n\to\infty}\lim_{m\to\infty}\cos^{2m}(n!\pi x)$ est due au mathématicien allemand Alfred Pringsheim (1850-1941).

Page 239, 14.10

a) Nous avons dit **en PAGE 41 b)** que l'ensemble des points de continuité d'une fonction numérique était toujours un G_δ : le point remarquable est donc que ce G_δ *est dense*, et en particulier non vide.

b) Si une fonction f de R^2 dans R est continue en chacune de ses deux variables, f est de première classe de Baire sur R^2, autrement dit f est la limite simple d'une suite (f_n) de fonctions *continues* sur R^2.

(**Clément de Seguins Pazzis**, **RMS 123-4**, 2012-2013, **exercice d'Oral 233**.)

Exemple. $f(x,y)=\frac{xy}{x^2+y^2}$ prolongée par $f(0,0)=0$, qui est discontinue en l'origine, est la limite des fonctions continues $f_n(x,y)=\frac{n}{2}\int_{x-1/n}^{x+1/n} f(u,y)du=\frac{ny}{4}\ln\left(\frac{y^2+(x+1/n)^2}{y^2+(x-1/n)^2}\right)$, en convenant que $f_n(x,0)=0$ pour tout x.

Page 241

Dans les deux dernières lignes, f a malencontreusement remplacé f'. Il faut donc lire :
Comme $x+g_n$ a pour limite x_0, $f'(x+g_n)\to f'(x_0)$; or, $f'(x+g_n)=f'(x)$ et donc $f'(x)=f'(x_0)$.

Page 244 : la tribu du jeu infini de pile ou face

Voici deux nouvelles interprétations de la tribu $\mathcal{T}$ construite :

1) C'est la plus petite tribu rendant mesurables les applications coordonnées $X_n : (\omega_n)_{n\in\mathbb{N}^*} \mapsto \omega_n$.

2) Soit π_n la projection de Ω sur $\Omega_n = \{0,1\}^n$ définie par $\pi_n : (\omega_n)_{n\in\mathbb{N}^*} \mapsto (\omega_1,...,\omega_n)$. L'image réciproque par π_n de $\mathcal{P}(\Omega_n)$ est une tribu $\mathcal{T}_n$ sur Ω : ces tribus croissent avec n car $\pi_n^{-1}(A) = \pi_{n+1}^{-1}(A\times\{0,1\})$. La réunion des $\mathcal{T}_n$ est une algèbre qui n'est pas une tribu et $\mathcal{T}$ est la tribu engendrée par cette algèbre.

Remarque. $\mathcal{T}$ est de cardinal au plus $\mathfrak{c}$, car engendrée par une famille dénombrable.

Page 259, ligne 1

L'équivalent de $\ln(1-1/p_n)$ est évidemment $-1/p_n$, ce qui n'entache en rien le raisonnement qui suit.

Page 265, 2), a)

De l'existence d'une base $(e_i)_{i\in I}$ pour tout espace vectoriel V résulte l'existence d'une norme sur V.

En effet, tout x de V s'écrivant $\sum_J x_j e_j$, où *J* est une partie finie de I, il suffit de poser $\| x \| = \sum_J | x_j |$. Ce résultat est d'autant plus intéressant qu'on est incapable d'expliciter une norme sur $\mathbb{R}^{\mathbb{N}}$.

Page 267, ligne 6 en partant de la fin

On parle de la famille des suites vi = (cosn)i, alors qu'il faut lire $v_i = (\cos n)^i$.

Page 270, le théorème de Liapounov

Nous avons prouvé que si une probabilité p est sans atome sur $(\Omega,\mathcal{T})$, il existait une chaîne croissante d'événements $\left(E(t)\right)_{t\in[0,1]}$ telle que $p\left(E(t)\right) = t$.

Voici un autre exemple d'une telle chaîne que celui donné dans le livre : *s'il existe une variable aléatoire uniforme* X de $(\Omega,\mathcal{T},p)$ dans $(0,1)$, $E(t) = p(X \le t)$ convient.
Nous prouvons que réciproquement, l'existence d'une telle chaîne implique que p est sans atome, puis qu'il existe sur $(\Omega,\mathcal{T},p)$ une variable aléatoire uniforme à valeurs dans $(0,1)$.

a) Réciproque (démonstration **d'Alain Rémondière**)
On raisonne par l'absurde en supposant que la tribu contient une chaîne croissante d'événements $\left(E(t)\right)_{t\in[0,1]}$ telle que $p\left(E(t)\right) = t$ et que p a un atome A dans la tribu : $p(A) > 0$, $B \subset A$ implique $p(B) = 0$ ou $p(A)$.
On pose $A(t) = A \cap E(t)$, donc $p\left(A(t)\right) = 0$ ou $p(A)$ pour tout t, et on introduit

$$I = \left\{t \in [0,1] : p\left(A(t)\right) = 0\right\} \ , \ J = \left\{t \in [0,1] : p\left(A(t)\right) = p(A)\right\}.$$

I et J sont des intervalles, non vides, disjoints, de somme $[0,1]$, aussi $\sup I = \inf J$. On pose $t_0 = \sup I = \inf J$, $t_0 \in]0,1]$ car $t \in J$ implique $t \ge p(A)$ en vertu de l'inégalité $p\left(A(t)\right) \le t$.
On suppose $t_0 < 1$: à partir d'un certain rang n_0, $t_0 - 1/n_0$ et $t_0 + 1/n_0$ sont entre 0 et 1 et il est donc permis

d'envisager les événements $F=\bigcup_{n_0}^{\infty} E(t_0-1/n)\subset E(t_0)$ et $G=\bigcap_{n_0}^{\infty} E(t_0+1/n)\supset E(t_0)$.

Comme

$$F\cap A=\bigcup_{n_0}^{\infty} A(t_0-1/n) \quad \text{et} \quad G\cap A=\bigcap_{n_0}^{\infty} A(t_0+1/n)$$

on a, par monotonie, $p(F\cap A)=\lim_n p\left(A(t_0-1/n)\right)=0$ et $p(G\cap A)=\lim_n p\left(A(t_0+1/n)\right)=p(A)$.

Mais $p(G-F)=0$, car pour $n\geq n_0$, $G-F\subset\left(E(t_0+1/n)-\left(E(t_0-1/n)\right)\right)$ et $p(G-F)\leq 2/n$; on en déduit sans mal que $p(G\cap A)=0$: c'est la contradiction cherchée.

Reste le cas $t_0=1$: $F=\bigcup_2^{\infty} E(1-1/n)$ est de probabilité 1 et la décomposition $A=F\cap A+(\Omega-F)\cap A$ donne $p(A)=0$, ce qui est impossible.

b) Condition nécessaire et suffisante d'existence d'une loi uniforme

Soit $\left(E(t)\right)_{t\in[0,1]}$ une chaîne croissante d'événements telle que $p\left(E(t)\right)=t$. On peut remplacer $E(1)$ par Ω : l'ensemble des u tels que $\omega\in E_u$ n'est vide pour aucun ω. On pose alors $X(\omega)=\inf\left\{u\in[0,1]:\omega\in E_u\right\}$.

X est une variable aléatoire sur $(\Omega,\mathcal{T})$. Par exemple, $(X<\lambda)=\left\{\bigcup E(r):r\in[0,\lambda[\cap\mathbb{Q}\right\}$ comme on le vérifie par double inclusion, donc $(X<\lambda)$ est dans la tribu ; et alors, $(X\leq\lambda)=\bigcap_n (X<\lambda+1/n)$ est aussi dans la tribu.

X est uniforme sur $(0,1)$: les inclusions $E_r\subset(X<\lambda)\subset E_\lambda$ prouve que $p(X<\lambda)=\lambda$, d'où $p(X\leq\lambda)=\lim_n p(X<\lambda+1/n)=\lambda$. En conclusion :

$\boxed{\text{Il existe une loi uniforme sur } (\Omega,\mathcal{T},p) \text{ si et seulement si la probabilité } p \text{ est sans atome}}$

Page 272 : un exemple de chaînes (A_r)

Ce paragraphe est confus et sans doute faux. Je le corrige :

Supposons qu'il existe une probabilité *p* sur $\mathcal{B}^n$ pour laquelle toutes les sphères de centre 0 sont négligeables, y compris $\{0\}$ (il en est ainsi quand $n=1$ et p est diffuse).

Posons $F(r)=p(B(0,r))=p(B'(0,r))$: F est croissante et continue sur $\mathbb{R}^+$, à valeurs dans $[0,1]$.

La limite de F en $+\infty$, qui existe, est $p(\mathbb{R}^n)=1$ et $F(0)=p\{0\}=0$. Il en résulte que F est une surjection de $]0,+\infty[$ sur $]0,1[$. Soit G une fonction de $]0,1[$ dans $]0,+\infty[$ telle que $FoG=id$ et posons $E_t=B\left(0,G(t)\right)$ pour $t\in]0,1[$: les E_t sont croissants, appartiennent à $\mathcal{B}^n$, et $p(E_t)=F\left(G(t)\right)=t$.

Page 284 : le théorème de Nikodym

Si les P_n sont des probabilités sur une algèbre qui n'est pas une tribu, leur limite P peut ne pas être $\sigma-$ additive.

Voici un contre-exemple :

Soit $\mathcal{A}$ l'algèbre des parties de l'univers $\Omega=\mathbb{N}$ qui sont soit finies soit de complémentaire fini ; définissons P_n sur $\mathcal{A}$ par $P_n(E)=1$ si $n\in E$ et $P_n(E)=0$ si $n\notin E$. La limite P des P_n est définie par

$$P(E)=0 \text{ si } E \text{ est finie et } P(E)=1 \text{ si } \mathbb{N}-E \text{ est finie.}$$

P est additive mais n'est pas $\sigma-$ additive.

Page 292, la convergence en probabilité

a) Comme une probabilité est comprise entre 0 et 1, X_n tend vers X pour une probabilité p si et seulement si pour tous α et ε positifs il existe un entier au delà duquel $p(|X_n - X| > \alpha) \leq \varepsilon$ ou $p(|X_n - X| \leq \alpha) \geq 1-\varepsilon$.

b) Le réel $\alpha > 0$ étant fixé, la convergence pour p de X_n vers X s'exprime ainsi :

si E_n est l'événement $\{\omega : |X_n(\omega) - X(\omega)| \leq \alpha\}$, $p(E_n) \to p(\Omega)$.

Plus généralement, pour *tout* T de la tribu,

$$p(\{\omega \in T : |X_n(\omega) - X(\omega)| \leq \alpha\}) \to p(T).$$

En posant $F_n = \{\omega \in T : |X_n(\omega) - X(\omega)| \leq \alpha\}$, $F_n = E_n \cap T$ et donc déjà $p(F_n) \leq p(T)$ puisque $F_n \subset T$.

Comme $E_n = F_n + (E_n - F_n) \subset F_n + \overline{T}$, $p(E_n) \leq p(F_n) + p(\overline{T})$, mais $p(E_n) \geq p(\Omega) - \varepsilon$ à partir d'un certain rang, d'où $p(F_n) \geq 1 - \varepsilon + p(\overline{T}) = p(T) - \varepsilon$, ce qui prouve que $\lim_n p(F_n) = 1$.

c) Il résulte de la définition que la convergence simple (ou sûre) entraîne la convergence en probabilité. On en déduit que si une suite h_n de fonctions boréliennes de $\mathbb{R}$ dans $\mathbb{R}$ converge simplement vers une fonction h et si X est une variable aléatoire, $h_n \circ X$ converge vers $h \circ X$ pour toute probabilité p puisque $h_n \circ X$ tend vers $h \circ X$ en tout ω de Ω. (Le caractère borélien des h_n et de h assure que $h_n \circ X$ et $h \circ X$ sont des variables aléatoires.)

d) Si X_n tend vers X pour une probabilité p sur $(\Omega, \mathcal{T})$, alors $h \circ X_n \xrightarrow{p} h \circ X$ pour toute fonction h *continue* de $\mathbb{R}$ dans $\mathbb{R}$; si de plus les X_n sont identiquement distribuées, $h \circ X_n \xrightarrow{p} h \circ X$ pour toute fonction h *borélienne* : démonstrations sur http://www.daniel-saada.eu/fichiers/13-deux_resultats_sur_la_convergence_en_probabilite.pdf

e) Si p et q sont deux probabilités *équivalentes*, la convergence de $X_n \to X$ pour p équivaut à $X_n \to X$ pour q. (Source : http://www.les-mathematiques.net/phorum/read.php?12,663190).

Page 292, la convergence en loi

a) C'est parce que deux variables aléatoires X et Y ayant la même fonction de répartition ont la même loi de probabilité (voir plus haut **Page 26**) que cette convergence porte son nom, autrement on aurait dit convergence en fonction de répartition. Ceci dit, il faut rayer la préposition « sauf » de l'avant-dernière ligne et lire :

la convergence de $F_n(x)$ *vers* $F(x)$ *a lieu sur une partie dense et de complémentaire dénombrable*.

b) De plus, quand la variable aléatoire limite X est diffuse, sa fonction de répartition F est continue et de ce fait la convergence de F_n vers F est alors *uniforme* sur $\mathbb{R}$ en vertu d'un théorème de Dini :

http://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8mes_de_Dini

c) On a dit que lorsque F n'était pas continue en x, il se pouvait que $F_n(x)$ ne tende pas vers $F(x)$.
Il se peut aussi que $F_n(x) \to F(x)$ en tout point de discontinuité de F, comme le montre l'exemple suivant.
Soit Z uniforme sur $[0,1]$ et indépendante de X : $X_n = X - Z/n$ converge simplement, où sûrement, vers X, et donc X_n converge en loi vers X. On montre alors que $F_n(x) \to F(x)$ pour tout x, en calculant $F_n(x)$ qui vaut $n\int_x^{x+1/n} F(t)dt$ et en passant à la limite.

d) Une condition suffisante pour que X_n à densité f_n converge en loi vers X à densité f est que $f_n \to f$.
La démonstration figure en **19.4.3** page 298 : remplacer B par $]-\infty, x]$.

Page 293, 2)

Une suite (X_n) de variables aléatoires *discrètes* peut converger en loi vers une variable aléatoire X *diffuse.*
Par exemple, si X_n est la loi uniforme sur $\{1/n, 2/n, ..., n/n\}$, elle converge en loi vers la loi uniforme sur [0,1].
Quand la loi-limite X est discrète et prend ses valeurs dans $\mathbb{Z}$, $X(\Omega) \subset \mathbb{Z}$, sa fonction de répartition F est continue sur $\mathbb{R} - \mathbb{Z}$; si de plus $X_n(\Omega) \subset \mathbb{Z}$ pour tout n, il en résulte que pour tout k dans $\mathbb{Z}$, $p(X_n = k)$ tend vers $p(X = k)$ quand n tend vers l'infini puisque $p(X_n = k) = F_n(k + 1/2) - F_n(k - 1/2)$.
La réciproque est vraie : si les X_n et X prennent leurs valeurs dans $\mathbb{Z}$ et si $p(X_n = k) \to p(X = k)$ pour tout k de $\mathbb{Z}$, alors $X_n \to X$ en loi, c'est une conséquence de **18.2**.

Page 294, 3)

Supposons que $X(\omega) = \sum_1^\infty X_n(\omega)$ existe pour tout ω de Ω, chaque X_n étant une variable aléatoire réelle sur $(\Omega, \mathcal{T}, p)$: X est alors une variable aléatoire et $\sum_1^n X_i \to X$ sûrement.
Il en résulte que $\sum_1^n X_i \to X$ en loi.
Exemple. Soit B_n une suite indépendante de lois de Bernoulli de paramètre $1/2$ et soit $U = \sum_1^\infty B_n / 2^n$.
$\sum_1^n B_i / 2^i$ est une variable discrète uniforme qui converge en loi vers une variable uniforme sur $[0,1]$.
On en déduit que U, qui est aussi limite en loi des $\sum_1^n B_i / 2^i$, est uniforme : on a ainsi un procédé de simulation de lois uniformes.

Page 294, 4), a)

J'ai donné dans le livre une application de **a)** : si (X_n) est une suite de variables aléatoires *positives* qui converge en loi vers X, alors $(1 + X_n / n)^n$ converge en loi vers e^X.
En voici une autre : si (X_n) est une suite de variables aléatoires qui converge en loi vers X, alors $a_n.X_n$ converge en loi vers $a.X$ dès que la suite des réels a_n converge vers un réel *a non nul*.
En effet, si $a > 0$ par exemple, on peut supposer les $a_n > 0$, et alors $p(a_n.X_n \le t) = p(X_n \le t / a_n)$, d'où

$$\lim_n p(a_n.X_n \le t) = p(X \le t/a) = p(a.X \le t)$$

car si t est un point de continuité de F_{aX}, F_X est continue en t/a puisque $F_{aX}(t) = F_X(t/a)$.

Page 295, b)

On en déduit que lorsque F est continue en x, $p(X_n < x) \to p(X < x)$.
Quand F n'est pas continue en x, on démontre que

$$p(X < x) \le \liminf p(X_n < x) \text{ et } p(X \le x) \ge \limsup p(X_n \le x).$$

Prouvons par exemple $p(X < x) \le \liminf p(X_n < x)$.
Comme $p(]-\infty, x[) = F(x-)$ car $]-\infty, x[$ est la réunion croissante des $]-\infty, x - 1/n]$, il faut montrer que $F(x-) \le \liminf F_n(x-)$. L'ensemble des points de discontinuité de F étant au plus dénombrable, il existe une suite (x_n) de points de continuité croissante et de limite x. Donnons-nous un réel $\varepsilon > 0$.
Pour k fixé, il existe un entier n_k au delà duquel $F(x_k) \le F_n(x_k) + \varepsilon$, d'où $F(x_k) \le F_n(x-) + \varepsilon$.

On en déduit $F(x_k) \leq \inf_{n \geq n_k} F_n(x-) + \varepsilon \leq \varepsilon + \liminf F_n(x-)$; comme k est arbitraire, $F(x-) \leq \varepsilon + \liminf F_n(x-)$ et en faisant tendre ε vers 0, $F(x-) \leq \liminf F_n(x-)$.

Page 297 : un résultat de Kolmogorov

La démonstration faite prouve aussi que si (X_n) est une suite de variables aléatoires sur $(\Omega, \mathcal{T})$, l'ensemble des $\omega \in \Omega$ tels que $X_n(\omega)$ converge est dans la tribu. En effet, $X_n(\omega)$ converge si et seulement si $\forall \varepsilon > 0, \exists n, \forall k \text{ et } l \geq n, \left|X_k(\omega) - X_l(\omega)\right| \leq \varepsilon$ (critère de Cauchy), ou encore $\forall \varepsilon > 0, \exists n, \forall k \geq n, \left|X_k(\omega) - X_n(\omega)\right| \leq \varepsilon$. L'ensemble des $\omega \in \Omega$ pour lesquels $X_n(\omega)$ converge est donc $\bigcap_{N \geq 1} \bigcup_n \bigcap_{k \geq n} \left(| X_k - X_n | \leq 1/N\right)$ et de ce fait appartient à $\mathcal{T}$ (bien sûr, il peut être vide).

Appelons T l'ensemble des ω pour lesquels $X_n(\omega)$ converge et supposons T non vide mais distinct de Ω. Posons $X(\omega) = \lim_n X_n(\omega)$ quand $\omega \in T$ et prolongeons X à Ω en posant $X(\omega) = 0$ sur $\overline{T}$: alors, X est une variable aléatoire. En effet, soit B un borélien de R : si B ne contient pas 0, $X^{-1}(B) = (\lim X_n)^{-1}(B)$ est un élément de la tribu ; si $0 \in B$, $X^{-1}(B) = \overline{T} + (\lim X_n)^{-1}(B - \{0\})$ est aussi dans $\mathcal{T}$.