ERRATA ET COMPLÉMENTS DE LA DEUXIÈME ÉDITION

(édition du 14 juin 2014, **ajouts Page 245 et Page 247**)

Page 78 : Les questions que pose le théorème d'Ulam

L'existence d'une probabilité diffuse a été posée dans un cadre trop général pour recevoir une réponse.

Comme souvent en théorie de la mesure, des réponses peuvent être apportées quand l'univers est un espace topologique : si Ω est une espace métrique compact K sans point isolé, il existe une probabilité diffuse sur les boréliens de K (l'exemple le plus simple est $[0,1]$ où la mesure de Borel est diffuse).

Démonstration sur http://www.daniel-saada.eu/Notes/note%2025-existence-de-mesures-diffuses.pdf

Il n'est pas nécessaire que K soit métrisable, il suffit même que K soit localement compact :

http://www.les-mathematiques.net/phorum/read.php?12,930071

Pages 132 à 134 : Approximations en mesure des boréliens par les ouverts et les compacts

Les formules $p(B)=\inf\{p(U): U \textit{ ouvert et contenant } B\}=\sup\{p(F): F \textit{ fermé et contenu dans } B\}$

sont vraies pour tout borélien B d'un espace métrique E, il n'y a rien à changer à la démonstration.

En **8.7**, nous avons pu remplacer le fermé F par un compact K car R^n est une réunion dénombrable de compacts, à savoir les boules fermées de centre O et de rayon n .

Une mesure de probabilité m définie sur la tribu des boréliens de E est dite régulière si

$$m(B)=\inf\{m(U): U \textit{ ouvert et contenant } B\}=\sup\{m(K): K \text{ compact } \textit{et contenu dans } B\}$$

pour tout borélien B de E.

Si le métrique E est réunion dénombrable de compacts, toute mesure de probabilité sur E sera régulière.

Voici un autre cas : <u>il suffit que</u> E <u>soit complet et séparable pour sa distance</u> (comme l'est un espace R^n).

La démonstration qui suit est due à Stanislaw Ulam et date environ de 1939 (source : P. BILLINGSLEY, *Convergence of probability measures*, théorème 1.4.)

Comme E est à base dénombrable d'ouverts [1], E est 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 $m\left(\bigcup_{i=0}^{i_n} A_{n,i}\right)\geq m(E)-\varepsilon/2^n$.

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 aux complémentaires).

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 $m(K)\geq m(F)-\varepsilon$.

[1] Si besoin, consulter les paragraphes 2 et 3 de l'article http://www.daniel-saada.eu/fichiers/14-variables_aleatoires_a_valeurs_dans_un_metrique.pdf

Page 147 : simulation de variables aléatoires

Voici un algorithme pour simuler une variable X de densité f sous les conditions :

(i) $X(\Omega) \subset [0,1]$,

(ii) f est bornée sur $[0,1]$.

Cet algorithme est du à J. Von NEUMANN et date de 1951.

1. se donner un majorant M de f

2. $U \leftarrow random$ et $V \leftarrow random$

3. si $M.U > f(V)$ aller en 2

4. $X \leftarrow V$

Pour une démonstration, voir http://www.daniel-saada.eu/Notes/24-La-methode-de-Von-Neumann.pdf

Page 165 : Intégrale et Théorème de représentation de Riesz

On sait que toute mesure de probabilité m sur $(\Omega, \mathcal{T})$ se prolonge en une forme linéaire positive L sur l'espace vectoriel des fonctions numériques intégrables : $m(T) = \int_\Omega 1_T \, dm = L(1_T)$ pour tout $T \in \mathcal{T}$.

Si Ω est un espace topologique X et $\mathcal{T}$ la tribu $\mathcal{B}(X)$ de ses boréliens, L est une forme linéaire sur l'espace vectoriel des fonctions continues bornées et donc sur le sous-espace $C_K(\Omega)$ des fonctions continues à support compact.

Réciproquement, toute forme linéaire positive L sur l'espace vectoriel $C_K(\Omega)$ provient d'une *unique* mesure positive m sur $(\Omega, \mathcal{T})$ pourvu que Ω soit localement compact : c'est le théorème de représentation de Riesz.

On a donc $L(f) = \int_X f \, dm$ pour toute $f \in C_K(\Omega)$.

Démonstration (remarquable) sur http://li.perso.math.cnrs.fr/textes/Integration/RIESZ.pdf

Page 186 : le théorème de convergence dominée et la convergence en loi

On peut avoir $\int_\Omega | f_n - f | \, dm \to 0$ sans que $f_n \to f$ presque partout pour m ; il se peut même que $f_n(\omega)$ ne converge pas vers $f(\omega)$ et ceci pour tout $\omega \in \Omega$ (un exemple sera donné plus bas).

En revanche, $\int_\Omega | f_n - f | \, dm \to 0$ implique $f_n \to f$ presque partout pour m si f_n et f sont des fonctions de répartition, que nous noterons donc F_n et F, m étant la mesure de Lebesgue sur $\Omega = R$.

Nous allons prouver que la condition $\int_R | F_n(x) - F(x) | \, dx \to 0$ entraîne que $F_n(y) \to F(y)$ pour tout point de continuité y de F ; comme l'ensemble des points de discontinuité de F est au plus dénombrable, $F_n \to F$ pp.

En d'autres termes, la condition $\int_R | F_n(x) - F(x) | \, dx \to 0$ implique la convergence en loi.

On raisonne par l'absurde en supposant que $F_n(y)$ ne tend pas vers $F(y)$ pour un point de continuité y de F.

Il existe alors $\varepsilon > 0$ et une sous-suite de F_n, qu'on appellera encore F_n, telle que $| F_n(y) - F(y) | > \varepsilon$.

Comme F est continue en y, il existe $h>0$ tel que $|F(z)-F(y)|<\varepsilon/2$ si $z\in]y-h,y+h[$.
On écrit alors $F_n(z)-F(z)=F_n(z)-F_n(y)+F_n(y)-F(y)+F(y)-F(z)$.

Cas 1 : $F_n(y)-F(y)>\varepsilon$.
Alors, si $y<z<y+h$, $F_n(z)-F(z)\geq F_n(y)-F(y)+F(y)-F(z)$ car F_n est croissante
et donc $F_n(z)-F(z)\geq\varepsilon-\varepsilon/2$, puis $\int_R|F_n(x)-F(x)|dx\geq\int_y^{y+h}\varepsilon/2$.

Cas 2 : $F_n(y)-F(y)<-\varepsilon$.
Alors, si $y-h<z<y$, $F_n(z)-F(z)\leq F_n(y)-F(y)+F(y)-F(z)$ car F_n est croissante
et donc $F_n(z)-F(z)\leq-\varepsilon+\varepsilon/2$, puis $\int_R|F_n(x)-F(x)|dx\geq\int_{y-h}^{y}\varepsilon/2$.

Il en résulte que $\int_R|F_n(x)-F(x)|dx$ ne tend pas vers 0 puisque minorée par $h\varepsilon/2$.

<u>Un exemple de fonctions positives</u> f_n telles que $\int_R f_n\to 0$ et $f_n(x)$ ne tende vers 0 nulle part (Jean-Luc <u>VERLEY</u>).
Soit f_n la fonction indicatrice du $n^{ème}$ segment de la liste dénombrable :

$$[0,1],[0,1/2],[1/2,1],[0,1/3],[1/3,2/3],[2/3,1],...$$

On a $f_n\geq 0$, $\int_0^1 f_n(x)dx\to 0$ et cependant $f_n(x)$ ne tend vers 0 en aucun point de $[0,1]$.

Page 186 : le théorème de convergence dominée et la convergence de certaines suites réelles

Grâce à ce théorème, nous allons prouver :

si les suites $n\mapsto e^{itu_n}$ convergent pour tout réel t, alors la suite réelle u_n converge.

Démonstration de Denis Trotabas

Posons $l(t)=\lim_n e^{iu_nt}$: l est une fonction complexe de module 1, elle est mesurable comme limite simple de fonctions continues, elle est donc intégrable au sens de Lebesgue sur tout segment de R. De plus, pour tous a et b,

$$\int_a^b l(t)dt=\lim\int_a^b e^{itu_n}dt.$$

Si u_n n'est pas de Cauchy, on peut trouver $\varepsilon>0$ et deux sous suites u_n' et u_n'' de u_n telles que $|u_n'-u_n''|>\varepsilon$. Posons $v_n=u_n'-u_n''$: on a $|v_n|>\varepsilon$ (1).

La suite de fonctions $f_n(x)=e^{iv_nx}$ tend vers 1 car $f_n(x)=\dfrac{e^{iu_n'x}}{e^{iu_n''x}}$ et est dominée par 1, donc par Lebesgue l'intégrale de f_n de 0 à 2π tend vers 2π.

Par ailleurs cette même intégrale vaut $\dfrac{e^{2i\pi v_n}-1}{iv_n}$ qui tend vers 0 à cause de (1) : **contradiction**.

Application : convergence en loi d'une suite de variables aléatoires normales.

On suppose que $X_n=N(m_n,\sigma_n)\to X$ en loi : on va montrer que les suites (m_n) et (σ_n) convergent.
Pour ce faire, on utilise le théorème de Paul Lévy : les fonction caractéristiques $\phi_n(t)=E(e^{itX_n})$ convergent

simplement vers $\phi(t)=E(e^{itX})$. Puisque X_n a pour densité $f_n(x)=\dfrac{1}{\sigma_n\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-m_n}{\sigma_n}\right)^2}$, alors

$\phi_n(t)=\dfrac{1}{\sigma_n\sqrt{2\pi}}\int_R e^{itx}e^{-\frac{1}{2}\left(\frac{x-m_n}{\sigma_n}\right)^2}dx$ et on démontre que $\phi_n(t)=e^{im_nt-\sigma_n{}^2t^2/2}$.

Donc , pour tout réel t, $\phi_n(t)=e^{im_nt-\sigma_n{}^2t^2/2}\to\phi(t)=E(e^{itX})$; en passant aux modules, $e^{-\sigma_n{}^2t^2/2}\to|\phi(t)|$.

Comme ϕ est continue et $|\phi(t)|\leq\phi(0)=1$, on peut trouver $t\neq0$ tel que $1/2<|\phi(t)|\leq1$ et alors $\sigma_n{}^2$ a une limite finie positive ou nulle et donc $\sigma_n\to\sigma\geq0$.

Il en résulte que $e^{im_nt}\to\phi(t)e^{\sigma^2t^2/2}$ et on a vu alors que m_n converge ; si m est la limite, $\phi(t)=e^{imt-\sigma^2t^2/2}$.

En conclusion :

si $\sigma\neq0$, X_n tend en loi vers une loi normale $N(m,\sigma)$, c'est la **réciproque** du théorème de Paul Lévy ;

si $\sigma=0$, X_n tend en loi vers $X=m$.

Page 210, 4)

Voici un exemples de tribu produit en dehors des espaces R^n : la tribu $\mathcal{B}(R_+)\otimes\mathcal{P}(\mathbb{N})$.

Nous montrons que $\mathcal{T}=\mathcal{B}(R_+)\otimes\mathcal{P}(\mathbb{N})$ est égale à la famille des ensembles de la forme $\bigcup_0^\infty(B_n\times\{n\})$ où chaque B_n est un borélien de R_+.

a) Par définition, la tribu $\mathcal{T}$ est engendrée par les produits $B\times J$ où $\mathcal{B}\in\mathcal{B}(R_+)$ et $J\in\mathcal{P}(\mathbb{N})$.

Comme J est une réunion au plus dénombrable d'entiers, $\mathcal{T}$ est engendrée par les produits $B\times\{n\}$ et contient donc toutes les réunions $\bigcup_0^\infty(B_n\times\{n\})$. Il suffit donc de vérifier que l'ensemble de ces réunions forment une tribu $\mathcal{T}'$ de $R_+\times\mathbb{N}$ car $\mathcal{T}'\subset\mathcal{T}$ implique $\mathcal{T}'=\mathcal{T}$ puisque $\mathcal{T}$ est minimale.

Comme les $B_n\times\{n\}$ sont disjoints, nous remplacerons $A=\bigcup_0^\infty(B_n\times\{n\})$ par $A=\sum_0^\infty(B_n\times\{n\})$ (cette décomposition est unique car $B_n=\{x\in R_+:(x,n)\in A\}$).

La stabilité par réunion dénombrable de $\mathcal{T}'$ est évidente, une réunion dénombrable de boréliens étant borélienne.

La stabilité par complémentaire exige deux étapes :

b) Le complémentaire $\overline{B\times\{n\}}$ de $B\times\{n\}$ est dans $\mathcal{T}'$. En effet, $\overline{B\times\{n\}}=\overline{B}\times n+\sum_{p\neq n}R_+\times p$.

c) Le complémentaire de $A=\sum_0^\infty(B_n\times\{n\})$ est $\overline{A}=\bigcap_0^\infty\overline{B_n\times\{n\}}=\bigcap_n\bigcup_p(A_{n,p}\times p)$ où chaque $A_{n,p}$ est un borélien. Cette intersection de réunions est une réunion d'intersections soit vides (en général) soit égales à $\left(\bigcap_k A_{k,n}\right)\times n$: $\overline{A}=\sum_{n=0}^\infty\left(\left(\bigcap_k A_{k,n}\right)\times n\right)$ est bien dans $\mathcal{T}'$ et $\mathcal{T}'$ coïncide donc avec $\mathcal{T}$.

Remarque. La démonstration produite montre que ce résultat s'étend à toute tribu produit dont l'un deux des termes est $\mathcal{P}(\mathbb{N})$: $\Gamma \otimes \mathcal{P}(\mathbb{N})$ est égale à la famille des ensembles de la forme $\bigcup_0^\infty (G_n \times \{n\})$, avec $G_n \in \Gamma$ pour tout n.

Page 230 : Le théorème de Lusin

a) Une autre lecture du théorème de Lusin

La restriction de f à K étant continue, il existe une fonction g continue sur R^n qui prolonge f / K. On peut donc énoncer : pour tout $\varepsilon > 0$, il existe une fonction continue f_ε telle que $m\{x : f(x) \neq f_\varepsilon(x)\} \leq \varepsilon$. En effet, $\{x : f(x) \neq f_\varepsilon(x)\} \subset R^n - K$.

b) f est limite m -presque partout d'une suite de fonctions continues

D'après **a)**, pour tout n, il existe une fonction continue f_n telle que $m\{x : f(x) \neq f_n(x)\} \leq 1/2^n$.
Appelons E_n le borélien $\{x : f(x) \neq f_n(x)\}$: comme $\sum m(E_n) < +\infty$, le lemme de Borel-Cantelli assure que l'événement $N = \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty E_i$ est de mesure nulle pour m. En d'autres termes, si $x \in R^n - N$, x appartient à un nombre fini de E_n : de ce fait, à partir d'un entier n_x qui dépend de x, $f_n(x) = f(x)$ et donc f est la limite simple des f_n sur $R^n - N$.

Page 244 : la tribu du jeu de pile ou face

La tribu a été construite de manière ensembliste. Après examen, il s'avère qu'elle coïncide avec les boréliens de l'espace topologique produit $\mathcal{C} = \{0,1\}^{\mathbb{N}^*}$, qu'on appelle espace de Cantor. On pourra consulter :

http://fr.wikipedia.org/wiki/Topologie_produit

La probabilité p du jeu est donc une mesure borélienne. Il en résulte que sur tout espace topologique (compact) K homéomorphe à $\mathcal{C}$, il existe une probabilité diffuse q. En effet, il suffit de poser $q(B) = p\left(f(B)\right)$ pour tout borélien B de E, f étant une homéomorphie de E sur $\mathcal{C}$.

Page 245 : la probabilité du jeu

a) La probabilité diffuse p que nous avons construite sur les boréliens de $\{0,1\}^{\mathbb{N}^*}$ a une autre propriété : pour tout ouvert non vide U de $\{0,1\}^{\mathbb{N}^*}$, $p(U) > 0$. En effet, U contient une boule ouverte $B(a,r)$ et donc une boule ouverte $B(a,1/2^n)$ pour n assez grand. Cette deuxième boule contient l'événement

$$E = \{\omega \in \mathcal{C} : \omega_i = a_i \ pour \ i = 1,2,...,n\}$$

dont la probabilité est $1/2^n$ par définition de p. Aussi a-t-on $p(U) \geq 1/2^n$.

b) Bien entendu, pour tout $\lambda \in]0,1[$, on montrerait l'existence d'une probabilité diffuse p_λ telle que $p_\lambda(P_n) = \lambda$ (avec les notations du texte) et $p_\lambda(F_n) = 1 - \lambda$.

Page 247 : la loi forte des grands nombres

1) Je me cite :

L'ensemble E des suites convergentes de $\Omega = \{0,1\}^{\mathbb{N}^*}$ est dénombrable puisqu'il s'agit des suites constantes à partir d'un certain rang. Pour la probabilité diffuse p qui a été construite, cet ensemble, qui appartient à la tribu, est négligeable.

Toutefois, si la caractère dénombrable de E n'est pas accessible à un auditoire, voici une démonstration directe.

Comme $E = \bigcup_{m=1}^{\infty} \bigcap_{j=m}^{\infty} \{\omega_j = \omega_m\}$, il suffit de prouver que pour tout m, $E_m = \bigcap_{j=m}^{\infty} \{\omega_j = \omega_m\}$ est négligeable.

E_m se décompose en la réunion $\left\{\bigcap_{j=m}^{\infty} \{\omega_j = 1\}\right\} \cup \left\{\bigcap_{j=m}^{\infty} \{\omega_j = 0\}\right\}$; or, pour tout entier $k \in \mathbb{N}$,

$P\left(\bigcap_{j=m}^{\infty} \{\omega_j = 1\}\right) \le P\left(\bigcap_{j=m}^{m+k} \{\omega_j = 1\}\right) \le (1/2)^{k+1}$, ce qui prouve que $P\left(\bigcap_{j=m}^{\infty} \{\omega_j = 1\}\right) = 0$.

On montrerait de même que $P\left(\bigcap_{j=m}^{\infty} \{\omega_j = 0\}\right) = 0$, ce qui donne bien $P(E) = 0$.

2) Puisque l'événement $E = \left\{\omega \in \Omega : \frac{1}{n}\sum_1^n \omega_i \to \frac{1}{2}\right\}$ est de probabilité 1, tout événement dans le complémentaire de E est négligeable. Il en est donc ainsi pour $A_t = \left\{\omega \in \Omega : \frac{1}{n}\sum_1^n \omega_i \to t\right\}$ où $t \in [0,1] - 1/2$. De plus :

a) A_0 a la puissance du continu

En effet, A_0 contient toutes les suites (ω_n) telles que $\omega_n = 0$ pour n non carré parfait.

Rappelons que A_0 est dense (dernière ligne de la **page 63** du livre).

b) A_t est en bijection avec A_0

D'abord, A_t est non vide : la suite ω définie par $\omega_n = [nt + t] - [nt]$ est dans A_t ($[x]$ est la partie entière de x).

Ensuite, pour toute $s \in A_0$, la suite $(s + \omega)$ modulo 2 est dans A_t car le nombre de 2 qui se transforment en 0 est « petit ». De façon précise, si $\rho = s + \omega$ modulo 2 :

$$(\omega_1 + \ldots + \omega_n) - (s_1 + \ldots + s_n) \le \rho_1 + \ldots + \rho_n \le (\omega_1 + \ldots + \omega_n) + (s_1 + \ldots + s_n)$$

Il en résulte que $s \mapsto s + \omega$ modulo 2 est une bijection de A_0 sur A_t et que A_t a la puissance du continu.

Autre démonstration. Si au lieu de jeter une pièce équilibrée, on obtient Pile avec la probabilité t, $0 < t < 1$, la loi forte des grands nombres montre que c'est A_t qui est maintenant de probabilité 1 : il en résulte que l'ensemble A_t est infini non dénombrable puisque la probabilité est diffuse.

Il existe donc dans Ω une infinité non dénombrable de négligeables disjoints non dénombrables.

(Merci à **Roberto Pinciroli** et à **Alain Rémondière**.)

Page 269 : une probabilité diffuse sur $(R^n, \mathcal{B}^n)$ n'a pas d'atome

D'abord la référence à **8.6** donnée en fin de ligne 4 est erronée : il faut lire **8.7**.

Voici maintenant une autre démonstration, plus courte, qui utilise la deuxième affirmation de **8.7** :

$$p(B) = \inf\{p(U) : U \text{ ouvert contenant } B\} \text{ pour tout borélien } B.$$

Partons du compact K de l'atome A tel que $p(K) = p(A)$: pour tout x de K, $p(x) = 0$, et donc il existe un ouvert U_x contenant x tel que $p(U_x) < p(A)$. Comme $U_x \cap K \subset A$ et $p(U_x \cap K) < p(A)$, $p(U_x \cap K) = 0$. K étant compact, il est recouvert par une famille *finie* de U_x et alors $p(K) = 0$, contradiction.

Page 284 : le théorème de Nikodym

Supposons que l'univers Ω soit un espace compact K et $\mathcal{T}$ la tribu de ses boréliens : alors, pour toute fonction f mesurable et bornée sur K, $\int_K f\,dp_n \to \int_K f\,dp$ (par hypothèse, on a seulement $\int_K 1_B\,dp_n \to \int_K 1_B\,dp$ pour tout borélien B). Il suffit de l'établir pour $f \geq 0$, car la décomposition $f = f^+ - f^-$ entraînera la convergence de $\int_K f\,dp_n$ vers $\int_K f\,dp$ si f est de signe quelconque.

Soit donc f mesurable positive et majorée : f est limite *uniforme* d'une suite *croissante* de fonctions étagées (page 179 du livre).

Donnons-nous $\varepsilon > 0$: il existe h étagée telle que $0 \leq f - h \leq \varepsilon$ sur K tout entier.

En écrivant $\int_K f\,dp - \int_K f\,dp_n = \int_K f\,dp - \int_K h dp + \int_K h dp - \int_K h dp_n + \int_K h dp_n - \int_K f\,dp_n$,

on a déjà $\left|\int_K f\,dp - \int_K f\,dp_n\right| \leq 2\varepsilon + \left|\int_K h dp - \int_K h dp_n\right|$.

h étant fixée et étant une combinaison linéaire *finie* de fonctions indicatrices, $\left|\int_K h dp - \int_K h dp_n\right| \leq \varepsilon$

à partir d'un certain rang, ce qui achève la démonstration.