249 - Suites de variables de Bernoulli indépendantes.

« Le jeu de pile ou face, dont le principe est si simple, possède un très grand caractère de généralité et conduit, lorsqu'on l'étudie en détail, aux mathématiques les plus élevées » (Émile Borel, 1924, cité dans l'introduction de [Les]).

I) <u>Suites de variables de Bernoulli indépendantes</u> [Les] + [Pag]

1) <u>Loi de Bernoulli et jeu du pile ou face</u>

Déf : Soit X une v.a. de Omega dans R, et p dans [0,1]. X suit une loi de Bernoulli de paramètre p si elle prend les valeurs 0 et 1 avec P(X=1)=p, P(X=-1)=1-p [Les 12]

Prop : si X suit une Ber(p) alors E(X)=p, V(X)=p(1-p) [Les 12]

Exemple fondamental : jeu du pile ou face. On lance une pièce telle que la probabilité de tomber sur face est p (si p=1/2, la pièce est dite équilibrée, biaisée sinon). Omega={pile,face}. X :Omega->R, X(F)=p, X(P)=1-p. X suit une lo ide Bernoulli de paramètre p.

Exemple : le loto, un pile ou face biaisé. La probabilité d'avoir les 6 bons numéros est égale à 1/13.983.416 [Pag 113]

2) <u>Généralités sur l'indépendance</u>

Déf : évènements indépendants dans leur ensemble [Les 8]

Exemple : les jets consécutifs d'une pièce de monnaie forment des événements indépendants.

Paradoxe de Galton : on lance trois pièces équilibrées en même temps. Parmi ces trois pièces, deux indiquent un résultat identique. Il y a alors une chance sur deux pour que la 3^e pièce indique le même résultat. Donc la proba que les 3 pièces indiquent la même face est ½. Pourtant, P(PPP)+P(FFF)=1/8+1/8=1/4 *(c'est parce que quand on a trouvé ½, on a fait P(les pièces indiquent la même face)=P(il existe deux pièces indiquant la même face)*P(la 3^e pièce indique la même face)=1*1/2, mais ces deux évènements ne sont pas indépendants)*

Déf : évènements indépendants deux à deux [Les 8]

Prop : indépendants dans leur ensemble => indépendants deux à deux [Les 8]

C-ex : (Bernstein) [Les 8]

Déf : v.a. indépendantes [Les 9]

Prop : X et Y indépendantes =W E(XY)=E(X)E(Y) et V(X+Y)=V(X)+V(Y) [Les 10]

3) <u>Jeu du pile ou face répété, loi binomiale</u>

Prop : soit (X_i) une suite de v.a. de Bernoulli de paramètre p, indépendantes. Posons S_n=X_1+…+X_n. Alors P(S_n=k)=…. On dit alors que S_n suit une loi binomiale de paramètres (n,p) [Les 12]

Prop : espérance et variance de loi binomiale [Les 12]

En pratique : on peut considérer cette suite de variables indépendantes comme les résultats de lancers successifs d'une pièce telle que la probabilité de tomber sur face est p, et S_n est alors le nombre de fois qu'on est tombé sur face après n lancers. Ainsi, la probabilité de tomber k fois sur face après n lancers est …

4) <u>Attente d'un événement, loi géométrique</u>

Problème : on décide de lancer une pièce plusieurs fois, jusqu'à ce qu'elle tombe du côté face. Quelle est la durée du jeu ?

Prop : soit T le nombre de parties jouées. Alors $P(T=k)=p(1-p)^{k-1}$. On dit que T suit une loi géométrique de paramètre p.

Prop : espérance et variance de la loi géom.

Appl : il faut jouer au loto en moyenne 336 siècles pour gagner le maximum [Pag 120]

Prop : si on lance une pièce de proba $p>0$ de manière répétée, toute séquence S de pile ou face de longueur r apparaîtra presque sûrement *(cas où $p=1/2$: la proba que S apparaisse dans le premier bloc de r lettres est de $1/(2^r)$. La proba qu'elle n'apparaisse pas dans le premier bloc est donc $1-1/(2^r)$. La proba qu'elle n'apparaisse pas dans les k premiers blocs est $(1-1/(2^r))^k$ qui tend vers 0).*

Appl : le singe savant.

II) <u>Comportement asymptotique de Sn</u> [Les]

- La loi faible des grands nombre va nous dire que la probabilité que S_n fluctue autour de sa moyenne np avec un ordre de grandeur de n est très faible. On va pouvoir estimer grossièrement cette proba. Autrement dit, S_n/n tend en proba vers un truc de dégénéré.
- Le théorème des grandes déviations permet d'estimer plus précisément cette proba.
- Le TCL nous dit que $S_n/\sqrt{n}$ tend vers un truc non dégénéré, c'est-à-dire que les fluctuations de Sn autour de np sont de l'ordre de $\sqrt{n}$. Le TCL va nous permettre d'étudier asymptotiquement l'ampleur de ces fluctuations, au lieu de les majorer comme précédemment [Les 21]

1) <u>Loi faible des grands nombres</u>

Prop : inégalités de Markov, de Tchebychev [Les 6-7]

Th : loi faible des grands nombres + majoration [Les 13] *(démo : inégalité de Tchebychev, qui est un corollaire de celle de Markov)*

Appl : th de Weierstrass [Les 14]

2) <u>Raffinement de la loi faible</u>

Amélioration : définition de h+, théorème [Les 16]

Ex : exemples numériques de la proba de s'éloigner de la moyenne théorique [Les 18]

Prop : DL de h+ pour estimer la vitesse de convergence [Les 18]

Prop : encadrement optimal [Les 18] *(formule de Stirling)*

3) <u>Théorème central limite</u>

Déf : loi normale, densité

Théorème : TCL [Les 21]

Appl : $P[\sqrt{p(1-p)n}\cdot a+np<S_n<\sqrt{p(1-p)n}\cdot b+np]\to\Phi(b)-\Phi(a)$ (si on veut trouver $P(A<S_n<B)$, on fixe a et b pour tomber sur A et B).

Ex : $P(S_{100}>60)$ vaut environ 0.02, et $P(S_{1000}>540)$ vaut environ 0.006 [Les 25]

Pb : quand est-ce que cette approximation est valable ?
C'est vrai pour n grand, et ça converge d'autant plus vite que p est proche de 1/2. Lesigne : quand $np(1-p)>18$ [Les 25].

Th : (Berry Esseen) [Feller – Introduction volume 2 – BU maths 519.2]
Inegalité de Berry Esseen pour les Bernoulli : l'écart entre les fonctions de répartition est majoré par C*(1-2pq)/sqrt(npq), C vaut environ 0,7.
Dans le cas du pile ou face avec pièce équilibré, l'écart est majoré par C/sqrt(n). n=100 : 0,07. N=1000 : 0,022

Application : intervalles de confiance. On veut donner une fourchette encadrant la proportion de « oui » telle que la proportion réelle soit dans la fourchette avec une proba de 0,95.

4) Loi forte des grands nombres

Th : loi forte des grands nombres

III) Marche aléatoire sur Z

1) Définitions [Les]

Déf : (Xn) une suite de v.a de Bernoulli centrée indépendantes de paramètre p (p différent de 0 ou 1) (P(Xn=1)=p ; P(Xn=-1)=1-p). On pose S_0=0 et S_n=sum(X_k,k=0,..,n). Alors on dit que S_n est une marche aléatoire aux plus proches voisins sur Z, dont le point de départ est 0.

Rq : une marche aléatoire modélise le trajet d'un promeneur sur un trajet fixé qui effectue à chaque instant entier un pas de longueur fixé vers l'avant ou vers l'arrière (avec une proba p ou 1-p).
Une marche aléatoire modélise aussi un jeu de pile ou face : on part d'une fortune X0, on gagne 1 euro à chaque face et on perd 1 euro à chaque pile [Les 85]

Problèmes : lors d'une marche aléatoire, est-on sûr de revenir à notre point de départ ? De visiter tous les points ? De revenir infiniment souvent à son point de départ ? [Les 85]

Déf : un point de Z est dit récurrent s'il est visité presque sûrement un nombre infini de fois (ie P({Zn=x} infiniment souvent)=1). Un point est dit transitoire s'il est visité presque sûrement un nombre fini de fois (ie P({Zn=x} infiniment souvent)=0) *(Attention ! A priori, ces cas ne sont pas complémentaires. Un point peut a priori être visité une infinité de fois avec une proba ½ par exemple)*

2) Propriétés [BL] + [Tei]

Prop : soit (Zn) une marche aléatoire. Alors tous les points sont récurrents ou tous les points sont transitoires [BL 213] + [BL 210] *(la récurrence est une ppté de classe, et une chaîne de Markov est irréductible car chaque élément communique avec ses deux voisins, donc par transitivité, tous les éléments communiquent entre eux)*

Rq : on dira alors qu'une chaîne est récurrente ou transitoire suivant que tous ses points soient récurrents ou transitoires.

Prop : si p est différent de ½, Z_n->+oo ou –oo ps [Tei 79] *(csq de la loi forte des grands nombres)*

Cor : si p différent de ½, la marche est transitoire [Tei 79]

Prop : si p=½, la marche est récurrente [Tei 79] *(on montre que la marche est transitoire dés que d est plus grand que 3, en calculant des intégrales qui vont converger seulement si d est plus grand que 3).*

Prop : proba de ruine d'un joueur [Tei 79]

IV) Existence d'une suite de v.a. indépendantes de lois données [Ouv]

Th : Mu une loi de probas donnée. F sa fonction de répartition et G sa fonction quantile. X une v.a. suivant une loi unif sur [0,1]. Alors la loi de G(X) est mu [Ouv 29] *(remarquer que F(x)>t ssi x > G(t), en déduire que F est la fonction de rep de G(Y))*

Prop : Soit x dans [0,1[. Posons $R_0(x)=x$, $D_n(x)=[2R_{n-1}(x)]$ et $R_n(x)=2R_{n-1}(x)-D_n(x)$. Alors pour tout x dans [0,1[, $x=\sum(D_j(x)/2^j)$. De plus, les D_j sont des variables iid de loi Ber(1/2) [Ouv 54]

Prop : on peut construire une suite de va iid de loi uniforme dur [0,1] [Ouv p.58]

Cor : on peut donc construire une suite (V_n) de va uniformes sur [0,1] indépendantes, et une suite $(G_n(V_n))$ de v.a. de loi μ_n indépendantes [Ouv p.58]

Développements :
1 - Construction d'une suite de v.a. i.i.d. de lois données [Ouv 58] (**)
2 - Marche aléatoire sur Z [Teicher – Independence, interchangeability, martingales 79] (**)

Bibliographie :
Barbe & Ledoux – Probabilité
Lesigne – Pile ou face : une introduction aux théorèmes limites du calcul des probabilités
Ouvrard – Probabilités 2
Pagès & Bouzitat – En passant par hasard
Chow & Teicher – Probability theory

Autres références pour la partie IV :
Kallenberg 56
Revuz p.46
Billingsley 74
http://labomathlens.free.fr/Liens/ProbaM1/PROBA03.pdf

Ce que je n'ai pas mis :
- Dans le jeu de pile ou face, si X s'attribue un point à chaque « face » et Y un point à chaque « pile », il est fortement probable que l'un des deux soit en tête la grande majorité de la partie. Le nombre de fois où le score sera identique n'est pas proportionnel au nombre de lancers n mais à sqrt(n) [Les 41]
- Temps d'attente de r résultats « pile » (loi binomiale négative ?).
- Estimation des moyennes déviations (Lesigne).
- Marches aléatoires sur Z^d, d>1.
- Convergence d'une loi binomiale vers une loi de Poisson
- Partie sur le quitte ou double :

Exemple : le quitte ou double

Jeu : on mise une somme, on lance une pièce, si c'est face on gagne la sommé misée, si c'est pile on la perd.
Si on pose $X_k=1$ si c'est face et -1 si c'est pile, et $S_k=X_0+\sum(X_k)$, (S_k) est une marche aléatoire. $E(S_k)=x_0+k(2p-1)$.

Comment miser pour gagner ?
On double la mise à chaque partie perdue, jusqu'à ce qu'on gagne. Le gain final est alors la mise initiale.

Prop : on pose $T=\inf(n \text{ tq } X_n)=1$. Alors $T<\infty$ ps (la proba que T soit égal à k est de $(1-p)^k$).
Soit S la somme misée au final. $S=1+2+2^2+\ldots+2^{(T-2)}$. Si $p\le$ ½ alors $E(S)=\infty$: il faut être prêt à miser beaucoup.

Rapport du jury : Inexistant.