249 Suites de variables de Bernoulli indépendantes

On se place dans $(\Omega, \mathcal{F}, P)$ un espace de probabilité.

I) Comment obtenir une suite de Bernoulli indépendante ?

1) Les variables aléatoires de Bernoulli

Définition 1: Soit $p \in [0,1]$. Une variable aléatoire X suit une loi de Bernoulli $b(p)$ si $P(X=1) = p$ et $P(X=0) = 1-p$.

Remarques 2 – toute expérience à deux issues peut être modélisée par une variable de Bernoulli.
– un lancer de pièce équilibrée peut être modélisé par une Bernoulli $b(\frac{1}{2})$.

Proposition 3: Soit $X \sim b(p)$
$E[X] = p$; $Var(X) = p(1-p)$; fonction caractéristique :
$\forall t \in \mathbb{R}, \ \varphi_X(t) = 1-p+pe^{it}$.

Remarque 4: informatiquement, on peut simuler des variables aléatoires uniformes par congruence. On construit alors des variables de Bernoulli.

Proposition 5: Soit U une variable uniforme sur $[0,1]$, et $p \in [0,1]$. Alors $X := \mathbb{1}_{\{U \leq p\}}$ suit une Bernoulli de paramètre p.

2) Rappels sur l'indépendance

[BL] p73 **Définition 6**: Deux évènements A et B sont dits indépendants si $P(A \cap B) = P(A)P(B)$

[BL] p74 **Définition 7**: (i) Une famille quelconque d'évènements $A_i \in \mathcal{F}$, $i \in I$, est mutuellement indépendante si, pour tout J fini $\subset I$,
$$P\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} P(A_j)$$
(ii) Une famille quelconque de sous-tribus $\mathcal{A}_i \subset \mathcal{F}$, $i \in I$, est mutuellement indépendante si toute famille d'évènements $A_i \in \mathcal{A}_i$, $i \in I$, est mutuellement indépendante.

Exemple 8: on jette deux dés, un rouge et un bleu. On pose [BL] p75
$A = \{$le résultat du dé rouge est impair$\}$, $P(A) = \frac{1}{2}$
$B = \{$le résultat du dé bleu est impair$\}$, $P(B) = \frac{1}{2}$
$C = \{$la somme des deux dés est impaire$\}$, $P(C) = \frac{1}{2}$
A, B et C sont indépendants deux à deux, mais ne sont pas mutuellements indépendants. On a $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$ mais $P(A \cap B \cap C) = 0$ car $A \cap B \cap C = \emptyset$.

3) Une construction d'une suite de Bernoulli indépendantes.

[Ou2] p54 **Proposition 9**: développement dyadique d'un réel $x \in [0,1[$.
On définit : $R_0(x) = x$ et $\forall n \in \mathbb{N}^*, D_n(x) = [2R_{n-1}(x)]$,
$R_n(x) = 2R_{n-1}(x) - D_n(x)$. On a alors :
– $D_n(x) \in \{0,1\}$, $R_n(x) \in [0,1[$
– $\forall n \in \mathbb{N}^*$, $x = \sum_{j=1}^{n} \frac{D_j(x)}{2^j} + \frac{1}{2^n} R_n(x) = \sum_{j=1}^{+\infty} \frac{D_j(x)}{2^j}$

[Ou2] p55 DEV 1 **Proposition 10**: Soit l'espace probabilisé $([0,1[, \mathcal{B}_{[0,1[}, P)$ où P est la restriction de la mesure de Lebesgue à $[0,1[$. Sur cet espace, la suite $(D_n)_{n \in \mathbb{N}^*}$ est une suite de va indépendantes de même loi $b(\frac{1}{2})$. De plus, $\forall n \in \mathbb{N}^*$, la variable aléatoire R_n est de loi uniforme sur $[0,1[$, et les variables aléatoires R_n et $(D_1, \dots D_n)$ sont indépendantes

[Ou2] p58 **Corollaire 11** : Il existe une suite de variables $(U_k)_{k\in\mathbb{N}}$ uniformes sur $[0,1]$ et indépendantes.

– $\forall p\in[0,1]$, il existe une suite de Bernoulli indépendantes de paramètre p.

II) Lien entre les variables de Bernoulli et quelques autres lois

1) Loi de Rademacher

Si Y suit une loi de Bernoulli de paramètre $\frac{1}{2}$, $X=2Y-1$ suit une loi de Rademacher. On a $E[X]=0$, $Var(X)=1$.

2) Loi binomiale

Définition 12 : une loi binomiale de paramètre $n\in\mathbb{N}^*$ et $p\in[0,1]$ représente la loi de probabilité d'une variable aléatoire X donnant le nombre de succès remportés au cours de la répétition de n épreuves de Bernoulli indépendantes de paramètre p.

Si $X_i \sim b(p)$, alors $X \sim \sum_{i=1}^{n} X_i$.

Corollaire 12' : $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$; $E[X]=np$, $Var(X)=np(1-p)$

3) Loi géométrique

Définition 13 : Soit $(X_i)_{i\in\mathbb{N}}$ une suite iid de loi $b(p)$.
La loi géométrique de paramètre p $\mathcal{G}(p)$ est la loi de $\min\{i\in\mathbb{N}^*, X_i\neq 0\}$, c'est-à-dire la loi du rang du premier succès.

Si $X\sim\mathcal{G}(p)$, $P(X=k)=p(1-p)^{k-1}$; $E[X]=\frac{1}{p}$, $Var(X)=\frac{1-p}{p^2}$

4) Loi binomiale négative

[Ou1] **Définition 14** : Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de lois géométriques de paramètre p, indépendantes. Une variable aléatoire X suit une loi binomiale négative de paramètres $n\in\mathbb{N}^*$, $p\in[0,1]$, si $X \sim \sum_{i=1}^{n} X_i$

Remarque 15 : Une binomiale négative de paramètres (n,p) modélise la loi du nombre d'échecs rencontrés avant d'obtenir n succès dans la répétition d'épreuves de Bernoulli indépendantes.

On a : $P(X=k)=\binom{n+k-1}{k}p^n(1-p)^k$

$E[X]=n\frac{1-p}{p}$, $Var(X)=n\frac{1-p}{p^2}$

5) Loi hypergéométrique

[Ou1] **Définition 16** : X suit une loi hypergéométrique de paramètres n, r, r_1 si $\forall k\in\mathbb{N}_{[0,r_1]}$, $P(X=k)=\dfrac{\binom{r_1}{k}\binom{r-r_1}{n-k}}{\binom{r}{n}}$

On a alors $E[X]=n\frac{r_1}{r}$, $Var(X)=n\frac{r-n}{r-1}\frac{r_1}{r}\left(1-\frac{r_1}{r}\right)$

Remarque 17 : si d'une urne contenant r boules, dont r_1 boules rouges et $r-r_1$ boules blanches, on tire simultanément n boules, la loi du nombre de boules rouges obtenues est la loi hypergéométrique de paramètres n, r, r_1.

6) Loi de Poisson

Définition 18 : Une v.a. X suit une loi de Poisson de paramètre $\lambda>0$ si $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$. On a alors $E[X]=Var(X)=\lambda$.

[Ou2] p 327 **Théorème 19** (théorème des événements rares).
Soit, pour tout $n\in\mathbb{N}^*$, une famille finie $\{A_{n,j} \mid 1\leq j\leq M_n\}$ d'événements indépendants définis sur un espace probabilisé $(\Omega,\mathcal{A},P)$. On pose $P(A_{n,j})=p_{n,j}$ et on note : $S_n=\sum_{j=1}^{M_n}\mathbb{1}_{A_{n,j}}$

On suppose que M_n tend en croissant vers $+\infty$, que $\max_{1\leq j\leq M_n} p_{n,j} \xrightarrow[n\to+\infty]{} 0$ et que $\sum_{j=1}^{M_n} p_{n,j} \xrightarrow[n\to+\infty]{} \lambda>0$.

Alors $(S_n)_n$ converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$.

[Oul] p321 **Corollaire 20** (Théorème de Poisson) : Soit, $\forall n \in \mathbb{N}^*$, X_n une variable aléatoire de loi binomiale $\mathcal{B}(n, p_n)$. On suppose que $\lim_{n \to +\infty} n p_n = \lambda > 0$. Alors $X_n \xrightarrow{\mathcal{L}} \mathcal{P}(\lambda)$.

Application pratique 21 : Si X suit une binomiale (n,p), avec n grand et p petit, sa loi est approximativement $\mathcal{P}(p.n)$.

III) Quelques applications

1) Éléments de statistique pour le jeu de pile ou face.

[Dupont] • Trouver un estimateur avec l'étude du maximum de vraisemblance. On étudie un jeu de pile ou face avec une pièce suivant une loi $b(\theta)$. On note $X_1, \dots, X_n$ n mesures et $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ la moyenne empirique.

La vraisemblance est $L_n(X_1, \dots, X_n, \theta) = \theta^{n\overline{X_n}}(1-\theta)^{n(1-\overline{X_n})}$.

On montre que $\overline{X_n}$ réalise le maximum de cette fonction (en θ). $\overline{X_n}$ est un bon candidat d'estimateur de θ.

• Loi des grands nombres :

[BL] **Théorème 22** (LGN forte pour les Bernoulli) : Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de va iid de loi $b(\theta)$, pour $\theta \in [0,1]$. On note $S_n = \sum_{i=1}^{n} X_i$. Alors $\frac{S_n}{n} \longrightarrow \theta$ presque sûrement.

[Cadre, Vial] **Application 23** : L'estimateur du maximum de vraisemblance $\overline{X_n}$ est un estimateur fortement consistant de θ.

• TCL et intervalle de confiance

[Cadre, Vial] **Théorème 24** (TCL pour les Bernoulli). Avec les notations de (22) :

$$\frac{S_n - n\theta}{\sqrt{n\,\theta(1-\theta)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Application 25 : • $\forall a < b$, $P\left(a \le \sqrt{n}\,\frac{\overline{X_n} - \theta}{\sqrt{\theta(1-\theta)}} \le b\right) \xrightarrow[n \to +\infty]{} \int_a^b \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt$

• En appliquant Slutsky, comme $\overline{X_n} \to \theta$, on a :
$I_\alpha^n = \left[\overline{X_n} - \frac{q_\alpha}{\sqrt{n}}\sqrt{\overline{X_n}(1-\overline{X_n})}, \overline{X_n} + \frac{q_\alpha}{\sqrt{n}}\sqrt{\overline{X_n}(1-\overline{X_n})}\right]$ est un intervalle de confiance asymptotique de θ de niveau α, si q_α est le quantile d'ordre $1-\frac{\alpha}{2}$ d'une $\mathcal{N}(0,1)$, ie $P(\theta \in I_\alpha^n) \xrightarrow[n \to +\infty]{} 1-\alpha$

2) Une application en analyse

[ZQ] **Théorème 26** (Théorème de Weierstrass) : Soit $f : [0,1] \to \mathbb{R}$ continue. On lui associe ses polynômes de Bernstein définis $\forall n \geq 1$ par $B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$.

Alors $\forall n \geq 1$, $\exists C > 0$ indépendante de f, de n, telle que :
$\|f - B_n\|_\infty \le C\,\omega\left(\frac{1}{\sqrt{n}}\right)$ où ω est le module de continuité uniforme de f ($\omega(h) = \sup_{|x-y| \le h} \{|f(x) - f(y)|\}$, $h \in [0,1]$).

3) Marche aléatoire sur $\mathbb{Z}$

[BL] On se donne $(\varepsilon_i)_i$ une suite de lois de Rademacher de paramètre $p \in [0,1]$ (ie $P(\varepsilon_i = 1) = p$, $P(\varepsilon_i = -1) = 1-p$). On pose la marche aléatoire $X_n = \sum_{i=1}^{n} \varepsilon_i$. C'est une chaîne de Markov.

Définition 27 : On note $\tau_i^n = \inf\{k > \tau_i^{n-1}, X_k = i\}$ avec $\tau_i^0 = 0$.
Un point i est dit récurrent si $P(\tau_i^1 < +\infty \mid X_1 = i) = 1$. Il est dit transcient sinon.

Proposition 28 : 0 est un état récurrent si $p = \frac{1}{2}$. Il est transcient sinon.