NOM : LEROUVILLOIS Prénom : Vincent Jury :

Algèbre ← Entourez l'épreuve → Analyse

Sujet choisi : Suite de variables aléatoires de Bernoulli indépendantes. 249

Autre sujet : REF: Barbe-Ledoux, Ouvrard 1 et 2, Foata Fuchs, Durett

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire réelle.

I - Loi de Bernoulli et lois dérivées

[B-L]

1) Définition et propriétés

Def 1: On dit est une variable de Bernoulli de paramètre $p \in [0,1]$ si $\mathbb{P}(X=1) = p$ et $\mathbb{P}(X=0) = 1-p$. On note $\boxed{X \sim \mathcal{B}(p)}$

Rq 2: La loi de X (loi de Bernoulli) est $p\delta_1 + (1-p)\delta_0$

Ex 3: On peut modéliser un lancer de pile ou face équilibré par la réalisation d'une variable de Bernoulli de paramètre $\frac{1}{2}$, en notant 0 pour pile et 1 pour face.

Prop 4: Si $X \sim \mathcal{B}(p)$, alors $\mathbb{E}[X] = p$ et $Var(X) = p(1-p)$

Ex 5: Si $A \in \mathcal{A}$ est un évènement, $\mathbb{1}_A$ suit une loi de Bernoulli de paramètre $\mathbb{P}(A)$. En particulier,
$\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$, ce qui est utile dans de nombreux cas.

Prop 6: Si $(\Omega, \mathcal{A}, \mathbb{P}) = ([0,1], \mathcal{B}([0,1]), Leb)$, on peut générer une v.a de Bernoulli de paramètre p en posant $X = \mathbb{1}_{[0,p]}$
On a alors $X \sim \mathcal{B}(p)$

2) Somme de variables de Bernoullis indépendantes

Rappel 7: Une famille de variables aléatoires $(X_i)_{i \in I}$ réelles est indépendante si $\forall J \subseteq I$ fini, $\forall (A_j)_{j\in J} \in \mathcal{B}(\mathbb{R})$

$$\boxed{\mathbb{P}\left(\bigcap_{j\in J} X_j \in A_j\right) = \prod_{j\in J} \mathbb{P}(X_j \in A_j)}$$

[Ouv]

Rq 8: Si $(X_n)_{n\in\mathbb{N}}$ est à valeurs discrètes, les X_n sont indépendants si et seulement si $\forall n \in \mathbb{N}^*, \forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n)$$

(?) indép vs indép par paire

Def 9: Soit $X_1, \dots, X_n$ des variables de Bernoulli de paramètre p indépendantes. On note $S_n = \sum_{i=1}^n X_i$. La loi de S_n est une loi binomiale de paramètres (n,p)

$$\boxed{S_n \sim \mathcal{B}(n,p)}$$

Prop 10: Si $S_n \sim \mathcal{B}(n,p)$, $\forall k \in [\![0,n]\!]$ $\boxed{\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}}$
$\mathbb{E}[S_n] = np$, $Var(S_n) = np(1-p)$.

Ex 11: (Suite de Ex3) Si on lance n fois la pièce équilibrée, S_n correspond au nombre de faces observées.

(?) Quelles hyp.? Lois L^2

Prop 12: Par l'inégalité de Bienaymé-Tchebitchev, $\forall \varepsilon > 0$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} Var\left(\frac{S_n}{n}\right) = \frac{p(1-p)}{\varepsilon^2 n}$$

Coro 13: Si $(X_i)_{i\in\mathbb{N}}$ est une suite iid $\sim \mathcal{B}(p)$, alors si $S_n = \sum_{i=1}^n X_i$

$$\boxed{\frac{S_n}{n} \xrightarrow[n\to+\infty]{\mathbb{P}} p}$$

Résultat étendu à ... le dé qui le ...

Dev 1: [Ouv] (+ Prop 10 + Prop 12)

Appli 14: (Thm d'approximation de Weierstrass)
Soit $f : [a,b] \to \mathbb{R}$, continue. Il existe une suite de polynômes qui converge uniformément vers f sur $[a,b]$.

Coro 15: (Lemme de Riemann-Lebesgue) Soit $f \in L^1(\mathbb{R}/2\pi\mathbb{Z})$ alors , $c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t) e^{int} dt \xrightarrow[|n|\to+\infty]{} 0$

3) Loi géométrique

Def 16: Soit $(X_n)_{n\in\mathbb{N}}$ une suite iid $\sim \mathcal{B}(p)$, $p \in]0,1[$. On définit $T_1 = \inf\{i \in \mathbb{N} / X_i = 1\}$ et $T_{n+1} = \inf\{i > T_n / X_i = 1\}$.
La loi de T_1 est appelée loi géométrique de paramètre p

$\boxed{T_1 \sim Géo(p)}$ T_1 est le temps d'apparition du premier succès (par exemple la première face)

[F-F]

La loi de T_r est une loi binomiale négative de paramètre (r,p). $\boxed{T_r \sim \mathcal{B}(r,p)}$

Prop 17 : $\forall k \in \mathbb{N}^*, \mathbb{P}(T_1 = k) = p(1-p)^{k-1}$, $\forall k \in \mathbb{N}^*, \mathbb{P}(T_r = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$

De plus, les $(T_{i+1} - T_i)_{i \in \mathbb{N}}$ sont iid $\sim$ Géo (p).

II - Suites de variables de Bernoulli indépendantes, propriétés

[Dev 2]

1) Existence et génération de variables aléatoires indépendantes

Pour construire un nombre fini de Bernoullis de paramètre $\frac{1}{2}$, on peut prendre $\Omega = \{0,1\}^n$ muni de la probabilité uniforme et $\forall i \in [\![1,n]\!]$: $X_i : (x_1, \dots, x_n) \longmapsto x_i$.

Mais construire une suite iid $\sim \mathcal{B}(\frac{1}{2})$ est plus difficile...

Prop 18 : Soit $x \in]0,1[$. Si x n'est pas un nombre dyadique, x possède un unique développement dyadique : $\boxed{x = \sum_{i \geq 1} \frac{D_i(x)}{2^i}}$ avec $D_i(x) \in \{0,1\}$

Si x est dyadique, il possède exactement deux développements dyadiques dont un est fini et l'autre pas. On choisit $(D_n(x))_{n \in \mathbb{N}}$, le développement dyadique fini de x.

Prop 19 : Soit $(\Omega, \mathcal{A}, \mathbb{P}) = (]0,1[, \mathcal{B}(]0,1[), \text{Leb})$. Alors $(D_n)_{n \in \mathbb{N}^*}$, définie à la prop 18, est une suite de variables aléatoires iid $\sim \mathcal{B}(\frac{1}{2})$

Coro 20 : On peut construire une suite iid de pile ou face à partir de la mesure de Lebesgue sur $]0,1[$.

Réciproque 21 : Si on a construit une loi produit de pile ou face $(\{0,1\}^{\mathbb{N}^*}, \mathcal{A}, \mathbb{P} = (\frac{\delta_0 + \delta_1}{2})^{\otimes \mathbb{N}})$ où $\mathcal{A} = \sigma(\{A_1 \times \dots \times A_n \times \{0,1\}^{\mathbb{N} \setminus \{1,\dots,n\}} / n \in \mathbb{N}, A_i \in \mathcal{P}(\{0,1\})\})$ est la tribu produit et $\forall n \in \mathbb{N}^*$, $\forall x_1, \dots, x_n \in \{0,1\}^n$ $\boxed{\mathbb{P}(\{x_1\} \times \dots \times \{x_n\} \times \{0,1\}^{\mathbb{N} \setminus \{1,\dots,n\}}) = \frac{1}{2^n}}$ alors on peut construire la mesure de Lebesgue sur $[0,1]$ par : $X : \{0,1\}^{\mathbb{N}^*} \longrightarrow [0,1]$, $(x_i)_{i \in \mathbb{N}^*} \longmapsto \sum_{i \geq 1} \frac{x_i}{2^i}$

Prop 22 : Soit X une v.a réelle de fonction de répartition F, et U une v.a uniforme sur $]0,1[$. Posons $F^{\leftarrow}(u) = \inf\{x / F(x) \geq u\}$ $\forall u \in]0,1[$. On a $\boxed{F^{\leftarrow}(U)}$ qui a la même loi que X.

Coro 23 : Soit $(\mu_j)_{j \in \mathbb{N}^*}$ une suite de mesure de probabilités sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Alors il existe une suite de v.a réelles $(X_j)_{j \in \mathbb{N}^*}$ définie sur $(]0,1[, \mathcal{B}(]0,1[), \text{Leb})$ indépendantes et telle que $\forall j \in \mathbb{N}^*$ X_j est de loi μ_j.

q Preuve ? Prop 22 bijection $\mathbb{N} \to \mathbb{N}^2$

q Information faible

2) Propriétés d'une suite de Bernoulli iid

[B-L] + Dev 2

Thm 24 (Borel-Cantelli) Soit $(A_n)_{n \in \mathbb{N}}$, une suite d'événements

i) Si $\sum \mathbb{P}(A_n) < +\infty$, alors $\mathbb{P}(\limsup A_n) = 0$

ii) Si $(A_n)_{n \in \mathbb{N}}$ est indépendante et si $\sum \mathbb{P}(A_n) = +\infty$, alors $\mathbb{P}(\limsup A_n) = 1$

Appli 25 : Soit $(X_n)_{n \in \mathbb{N}^*}$ iid $\sim \mathcal{B}(p)$, $p \in]0,1[$ et $(\varepsilon_1, \dots, \varepsilon_n) \in \{0,1\}^n$, un mot en base 2, alors p.s, $(\varepsilon_1, \dots, \varepsilon_n)$ apparaît une infinité de fois dans la suite de pile ou face

i.e si $A_j = (X_{j+1}, \dots, X_{j+n}) = (\varepsilon_1, \dots, \varepsilon_n)$, alors $\mathbb{P}(\limsup A_n) = 1$.

q Ça sert à quoi ? interprétation

Coro 26 : L'ensemble des mots finis en base 2 étant dénombrable, on a

$\boxed{\text{presque sûrement, tout mot fini apparaît une infinité de fois dans un pile ou face}}$

Appli 27 : Soit $(X_n)_{n \in \mathbb{N}^*}$ iid $\sim \mathcal{B}(\frac{1}{2})$ une suite de pile ou face.

[Dev]

Posons $\ell_k = \max\{m \in [\![0,k]\!] / X_k, \dots, X_{k+m-1} = (1, \dots, 1)\}$, $\forall k \in \mathbb{N}^*$

et $L_n = \max_{1 \leq k \leq n} \ell_k$, le nombre maximum de face consécutif jusqu'au $n^{\text{ème}}$ lancer

alors $\boxed{L_n / \log_2(n) \xrightarrow[n \to +\infty]{p.s} 1}$

q Idée preuve ? Borel-Cantelli mais hard

3) Convergence de suites de Bernoulli

Prop 28 : Soit $(X_n)_{n \in \mathbb{N}}$ une suite de Bernoulli indépendante telle que $\forall n$, $X_n \sim \mathcal{B}(p_n)$ alors

- $X_n \xrightarrow{\mathbb{P}} 0 \iff p_n \xrightarrow[n \to \infty]{} 0$
- $X_n \xrightarrow[n \to \infty]{p.s} 0 \iff \sum_{n \in \mathbb{N}} p_n < +\infty$

q Et pour la CV en loi ? On regarde les fct caractéristiques

[Ouv 1]

Prop 29 : Soit $(S_n)_{n \in \mathbb{N}^*}$ une suite de variable aléatoire / $S_n \sim \mathcal{B}(n, p_n)$

Si $n p_n \xrightarrow[n \to \infty]{} \lambda \in \mathbb{R}_+^*$, alors $\boxed{S_n \xrightarrow[n \to \infty]{loi} \mathcal{P}oi(\lambda)}$ ie $\mathbb{P}(S_n = k) \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}$

Rq 30 : Ce résultat est utilisé pour approximer la loi binomiale lorsque n est grand ($n \geq 30$) et p_n petit ($p_n \leq 0,1$). La $\boxed{\text{loi de Poisson}}$ modélise donc l'apparition d'événements rares

On peut insister en étant quantitatif

Comment p l'obtenir avec TCL (thm 35) ?
(on suppose Var(X_1) connue)

[Ouv1]

III - Estimations statistique du paramètre p

1) Loi des grands nombres

Thm 31: (Loi faible des grands nombres) Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a indépendantes, de même loi. Si $\mathbb{E}[|X_1|^2]<+\infty$, alors on a

$$\boxed{\sum_{i=1}^{n} X_i/n \xrightarrow[n\to+\infty]{L^2,\ \mathbb{P}} \mathbb{E}[X_1]}$$

Coro 32: Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements indépendants, de même probabilité, alors

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{A_i} \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{P}(A_1)$$

Cela appuie la cohérence du modèle probabiliste comme "fréquence d'apparition"

Illustrer Glivenko-Cantelli

Thm 33: (Loi forte des grands nombres) Soit $(X_n)_{n\in\mathbb{N}}$ une suite iid telle que $\mathbb{E}[|X_1|]<+\infty$, alors

$$\boxed{\sum_{i=1}^{n} X_i/n \xrightarrow{p.s} \mathbb{E}[X_1]}$$

Coro 34: Si $(X_1,\dots,X_n)$ est un n-échantillon $(X_1,\dots,X_n$ iid$)$ de loi $\mathcal{B}(p)$ avec p le paramètre inconnu à estimer, alors l'estimateur du maximum de vraisemblance est bien défini et vaut $\hat{p}_n = \sum_{i=1}^{n} X_i/n$

$\hat{p}_n$ est fortement consistant au sens où $\hat{p}_n \xrightarrow[n\to+\infty]{p.s} p$

2) Théorème de la limite centrale et intervalles de confiance

Thm 35: Soit $(X_n)_{n\in\mathbb{N}}$ une suite iid de v.a / $\mathbb{E}[|X_1|^2]<+\infty$, alors

$$\boxed{\frac{\sum_{i=1}^{n} X_i - n\mathbb{E}[X_1]}{\sqrt{n\,\mathrm{Var}(X_1)}} \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1)}$$

Coro 36: Si $(X_n)_n$ iid $\sim\mathcal{B}(p)$, alors $\sqrt{n}\,\frac{(\hat{p}_n-p)}{\sqrt{p(1-p)}} \xrightarrow{loi} \mathcal{N}(0,1)$

Avec le lemme de Slutsky, on a de plus :

$$\sqrt{n}\,\frac{(\hat{p}_n-p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \longrightarrow \mathcal{N}(0,1)$$

q Peut-on l'écrire dans le cas de Bernoulli ?

On utilise $\hat{p}_n(1-\hat{p}_n) \le 1/4$

Coro 37: Asymptotiquement, $I_n = \left[\hat{p}_n - 1{,}96\frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}},\ \hat{p}_n + 1{,}96\frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}\right]$ est un intervalle de confiance de niveau 95% ie : $\mathbb{P}(p\in I_n) = 0{,}95$ (*)

IV - Application en probabilité : Marche aléatoire et Ruine du joueur

Def 38: Soit $(X_n)_n$ une suite iid $\mathcal{B}(p)$ et $Y_n = 2X_n - 1$ $(\mathbb{P}(Y_n=1)=p, \mathbb{P}(Y_n=-1)=1-p)$

On définit la marche aléatoire sur $\mathbb{Z}$ partant de 0 par $S_n = \sum_{i=1}^{n} Y_i$

Prop 39: (Comportement asymptotique)

- Si $p<\frac{1}{2}$, $S_n \xrightarrow[n\to\infty]{p.s} -\infty$, si $p>\frac{1}{2}$, $S_n \xrightarrow[n\to\infty]{p.s} +\infty$ (Transient) ZDV
- Si $p=\frac{1}{2}$, on a $\forall n\in\mathbb{N}^*$, $\mathbb{P}(S_{2n}=0) = \frac{1}{2^{2n}}\binom{2n}{n} \underset{n\to\infty}{\sim} \frac{1}{\sqrt{n\pi}}$ q Preuve ?

et donc $\mathbb{P}(\limsup S_{2n}=0) = 1$, S_n repasse infiniment souvent par 0 (Récurrent)

q Et sur $\mathbb{Z}^2$, $\mathbb{Z}^3$? Preuve (en adaptant $\mathbb{Z}^3$) projeté

Problème de la ruine du joueur

Un joueur possède une somme d'argent $s_0\in\mathbb{N}^*$. A chaque étape, il parie 1 et gagne 1 ou perd 1 avec même probabilité $\frac{1}{2}$. Il désire obtenir la somme d'argent N avec $0<s_0<N$. Si à un moment, il n'a plus d'argent ou qu'il a atteint son objectif N, le jeu s'arrête.

Prop 40: Soit $(X_n)_{n\in\mathbb{N}}$ iid $\sim\mathcal{B}(1/2)$ et $Y_n = 2X_n - 1$. $Y_n \sim \mathcal{U}(\{-1;1\})$

Posons $S_n = s_0 + \sum_{i=1}^{n} Y_i$, la fortune du joueur à l'instant n.

On note $T = \inf\{n>0 / S_n = 0$ ou $S_n = N\}$ et $G = \{T<+\infty\}\cap\{S_T = N\}$, qui correspond à la probabilité de gagner en temps fini T.

i) $T<+\infty$ p.s

ii) $\mathbb{P}(G) = \frac{s_0}{N}$

DEV 2

(*). On peut ainsi tester si une pièce de monnaie est équilibrée ou non $(p=\frac{1}{2}?)$

- Cela s'applique également au sondage.

Intro : Simplicité des v.A Bern mais c'est riche — Colo 23 → v.a 99

Dirac → Bernoulli → Binomial, Poisson (grands nombres), Géom. (premier succès) → $\mathcal{U}([0,1])$

etc...

Appli : Stats (trouver le param., intervalle de confiance), Ruine du joueur

---

Dév 2 : $T = \min\{n \in \mathbb{N}, S_n \in \{0, N\}\}$. 1) $\mathbb{P}(T < +\infty) = 1$ → Borel Cantelli

avec $A_p = (X_{pN+1} = \dots = X_{pN+N} = 1)$. On a $\bigcup_{p \in \mathbb{N}} A_p \subset T$

$\mathbb{P}(A_p) = \frac{1}{2^N}$ TG ΣDV → $\mathbb{P}\left(\bigcup_{p\in\mathbb{N}} A_p\right) = 1$ → $\mathbb{P}(T < +\infty) = 1$.

$\leq \mathbb{P}(T) \leq 1$

indépendants

2) $\xi_n = S_{n \wedge T} \xrightarrow{p.s} \xi = \begin{cases} S_T & \text{si } T < +\infty \\ \dots & \text{sinon} \end{cases}$ $\mathbb{E}(\xi) = N\,\mathbb{P}(G)$

et $\mathbb{E}(\xi) = \lim_{n \to +\infty} \mathbb{E}(\xi_n)$ par CVD

$\mathbb{E}(\xi_n) = \dots = \mathbb{E}(\xi_0)$ par réc.

1) Quand on enlève N, que se passe-t-il ?
Propr des marches aléatoires → p.s, on va tt perdre

2) Que se passe-t-il pour 1 pièce biaisée ? → On a 1 autre martingale

3) pièce biaisée + enlève N. Que se passe-t-il ? → dépend de $\mathbb{P}(\text{gain}) \gtrless 1/2$

Exos | Loi forte des gds nombres pour v.a Bernoulli ? $\frac{S_n}{n} \xrightarrow{p.s} p$ (On suppose moment d'ordre 4) → on l'a vu en L3 proba

(Supposées centrées) $\mathbb{P}\left(\frac{S_n}{n} > \varepsilon\right) \leq \mathbb{P}\left(\frac{S_n^4}{n^4} > \varepsilon^4\right) \leq \frac{\mathbb{E}(S_n^4)}{n^4 \varepsilon^4}$

$=$