Coïncidence dans des séries d'évolution individuelles indépendantes

Errances d'un groupe sur une excursion

Gabriel Lang

23 mars 2024

Le sujet

- On observe n individus indépendants pendant un temps T.
- A un instant $t \in \mathbb{R}$, chaque individu a deux états possibles : inactif ($X_i(t) = 0$) ou actif ($X_i(t) = 1$).
- La suite d'états d'un individu est une chaîne de Markov homogène en temps continu.

Pour un individu :

- Partant de l'état 0, un seul saut est possible vers l'état 1.
- Et inversement.
- Temps exponentiel entre deux sauts. $\mathcal{E}(\lambda_{0\to1})$ et $\mathcal{E}(\mu_{1\to0})$.
- Cas symétrique : $\lambda_{0\to1} = \mu_{1\to0}$: Instants = Processus de Poisson.

La motivation

On s'intéresse aux durées continues pendant lesquelles les individus actifs représentent un nombre supérieur à a.

Applications :

- Modifications voisines d'un locus d'un gène, longues altérations continues communes à un nombre supérieur à a dans une cohorte (Modélisation de coïncidences dues au hasard).
- Saturation d'une antenne relais. On s'intéresse aux périodes où un nombre supérieur à a des usagers téléphone en même temps (Dimensionnement des antennes).

Objectif

On considère le cas où n est grand et on va établir deux résultats asymptotiques.

- Pour la somme des $X_i(t)$ (binomiale) renormalisée : $Z_n(t) = \frac{\sum(X_i(t)-1/2)}{\sqrt{n/4}}$. Z_n converge vers Z.
- Pour la distribution des durées d'excursion de Z_n au dessus d'un seuil a.

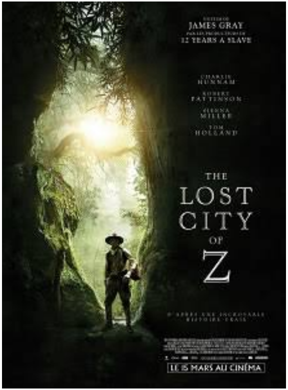
Le film

Le casting :

- S. Robin, initiateur de l'expédition, ayant mené des explorations préparatoires sur le sujet.
- G. Lang, amateur de vieux papiers.
- M-P. Etienne, cheffe d'expédition (évidemment!).
- L. Decreusefond, amateur de gros moteurs probabilistes.
- P. Valois, artisan-mineur probabiliste.

L'Aventure

Déterminer le processus Z



Après "Twelve years a slave"...

Trouver le cap

Sortie facile hé rondement menée

- $Z(t)$ est gaussien (TLC simple).
- Sa covariance est la même que celle d'un X_i.

Révélation : Z a été décrite et on doit sa diffusion à Ornstein et Uhlenbeck.

Trouver le cap

Sortie facile hé rondement menée

- $Z(t)$ est gaussien (TLC simple).
- Sa covariance est la même que celle d'un X_i.

Révélation : Z a été décrite et on doit sa diffusion à Ornstein et Uhlenbeck.
Ce qui achèvre du cap un franc choix.

Une première étape : le théorème de la limite centrale

Pour le processus Z_n

- La limite possible est identifiée. C'est une diffusion d'Ornstein-Uhlenbeck.
- Il faut montrer la convergence de Z_n par la tension.

S'approcher du but en traversant la jungle
On montre à la main que la suite Z_n est tendue en utilisant les vieux résultats de Billingsley.

Une rencontre
Un hélicoptère se pose ; L. Decreusefond rejoint le groupe à la recherche de Z mais par les airs. (Utilisation de la convergence des marches aléatoires vers le mouvement brownien, écriture sous forme de diffusion, théorème du point fixe).

Se rapprocher de la cité : Vitesse vers théorème central limite

On sait maintenant que Z_n tend vers Z. Le but est de connaître une borne de $d(Z_n, Z)$ en fonction de n.
On choisit une distance pratique (norme infinie sur $[0, 1]$ et la distance de Prokhorov sur Ω.)

Première expédition avec des vieilles cartes

On travaille sur les lois marginales finies en $(t_1, \ldots, t_k)$ et sur les oscillations des processus entre les temps des marginales.

- Borne sur la distance de Prokhorov entre marginales (Yurinskii 84).
- Borne sur les oscillations de Z_n (Strook et Varhadan).
- Borne sur les oscillations de Z (Garsia).

Très loin du but

La vitesse obtenue est $n^{-1/23}$.

Se rapprocher de la cité : Vitesse vers théorème central limite

On sait maintenant que Z_n tend vers Z. Le but est de connaître une borne de $d(Z_n, Z)$ en fonction de n.
On choisit une distance pratique (norme infinie sur $[0,1]$ et la distance de Prokhorov sur Ω.)

Première expédition avec des vieilles cartes

On travaille sur les lois marginales finies en $(t_1, \ldots, t_k)$ et sur les oscillations des processus entre les temps des marginales.

- Borne sur la distance de Prokhorov entre marginales (Yurinskii 84).
- Borne sur les oscillations de Z_n (Strook et Varhadan).
- Borne sur les oscillations de Z (Garsia).

Très loin du but

La vitesse obtenue est $n^{-1/23}$.**C'est nul**. (Pour les marches aléatoires, on obtient $n^{-1/4}$).

Passage vers les excursions

On veut un résultat sur les longueurs d'excursion

On passe d'une convergence en loi de Z_n vers Z à une convergence en loi des excursions Z_n vers les excursions de Z, si la fonction excursion est continue (delta-méthode).

La delta-méthode est le pont.

La rivière, première version

C'est comment qu'on traverse????

Notre pilote déclare "Vu de haut, y'a pas de pont" : Pour la distance choisie, la fonction excursion n'est pas clairement continue, la delta-méthode ne peut pas marcher :
Une excursion de longueur l au dessus du niveau x peut se couper en deux excursions de longueur $l/2$ pour le niveau $x + \varepsilon$ car la trajectoire de la diffusion est partout très irrégulière.

La rivière, deuxième version

Le pilote nous assure un passage avec des câbles (reformulation des excursions). Pas l'air très solide et la vitesse devient $n^{-1/46}$.

On se rapproche

En distance de Wasserstein, on a un résultat plus moderne pour la distance entre les lois marginales. Sans changer le reste, la vitesse obtenue est $n^{-1/10}$.

On se rapproche

En distance de Wasserstein, on a un résultat plus moderne pour la distance entre les lois marginales. Sans changer le reste, la vitesse obtenue est $n^{-1/10}$. **C'est moins nul**.

Un nouvel arrivant

Le visiteur inattendu

P. Valois entre en scène. Après avoir consulté les cartes, il propose : "Le mieux c'est de passer par en dessous, on utilise les cavités et on creuse". Il s'agit d'évaluer précisément et simplement tous les termes sans utiliser les bornes générales déjà connues. Et d'arriver à la bonne vitesse.

Un nouvel arrivant

Le visiteur inattendu

P. Valois entre en scène. Après avoir consulté les cartes, il propose : "Le mieux c'est de passer par en dessous, on utilise les cavités et on creuse". Il s'agit d'évaluer précisément et simplement tous les termes sans utiliser les bornes générales déjà connues. Et d'arriver à la bonne vitesse.
Cela fait nectre un certain découragement.

Mais Marie Pierre le suit dans la grotte avec confiance.

Le retour du pilote

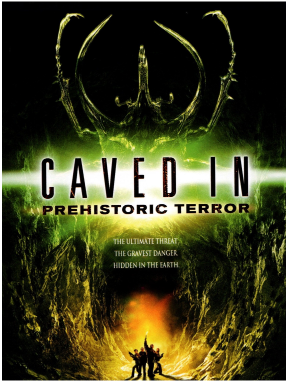
Laurent montre que, moyennant le pilotage d'un gros appareil (le rotor gyre à fond), on peut arriver directement au bon endroit. Il s'agit d'utiliser une représentation de l'espace de Cameron Martin, de maîtriser un peu de calcul de Malliavin. Et de redescendre sur notre résultat qui devient une simple application. Il n'y a pas l'ombe d'un doute.

Le retour du pilote

Laurent montre que, moyennant le pilotage d'un gros appareil (le rotor gyre à fond), on peut arriver directement au bon endroit. Il s'agit d'utiliser une représentation de l'espace de Cameron Martin, de maîtriser un peu de calcul de Malliavin. Et de redescendre sur notre résultat qui devient une simple application. Il n'y a pas l'ombe d'un doute.

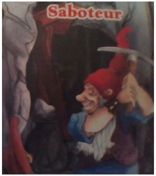
Nous ne sommes pas du tout capables de suivre l'argument..

Caved in


CAVED IN
PREHISTORIC TERROR
THE ULTIMATE THREAT.
THE GRAVEST DANGER.
HIDDEN IN THE EARTH.

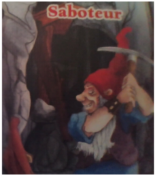
La crypte

Pierre continue à creuser ; après que Kubilius l'a détourné vers le mouvement brownien, il corrige vers le bon résultat (Ornstein Uhlenbeck).



La crypte

Pierre continue à creuser ; après que Kubilius l'a détourné vers le mouvement brownien, il corrige vers le bon résultat (Ornstein Uhlenbeck).



Nouvelle vitesse
Pour la distance de Prokhorov, la vitesse obtenue est $n^{-1/4} \ln(n)$.
C'est parfait.
Et la rivière ?

Passage de la rivière

Pierre pose son burin. Il sort une hachette et dit : "Bon, je vais construire un pont". Il attaque un arbre.

Passage de la rivière

Pierre pose son burin. Il sort une hachette et dit : "Bon, je vais construire un pont". Il attaque un arbre.

Définition d'une excursion longue au dessus du niveau a :
Partir d'un temps t de la diffusion au dessus de $Z(t) > x$ et de chercher le dernier passage de x avant t et la première après t.
Pour des niveaux proches, la longueur d'excursion se comporte de façon lipschitzienne. Ce qui la rend compatible avec l'usage de la distance de Prokhorov.
La vitesse de convergence des excursions sera parfaite.

Publier

Baliser le chemin
Marie-Pierre et Gabriel reprennent les notes et tentent d'en tirer un article raisonnable ($<$60 pages).
L'aspect pratique
Pour utiliser le résultat, il faut savoir simuler rapidement la distribution des excursions de la diffusion Z. Deux nouveaux explorateurs ont anticipé l'appel ; Pierre et le jeune Maxime propose des algorithmes efficaces pour simuler le premier temps d'atteinte de niveaux extrêmes pour la diffusion d'Ornstein-Uhlenbeck.

Du lourd

Importance d'un langage commun

Du lourd

Importance d'un langage commun

Sur un marché, un membre du groupe s'adresse à un paysan local en montrant un sac :

" Riz ?

- No sé. Arroz ? "

Du lourd

Importance d'un langage commun

Sur un marché, un membre du groupe s'adresse à un paysan local en montrant un sac :

" Riz ?

- No sé. Arroz ? "