UNIVERSITÉ DE NANTES

X22M010 : Probabilités appliquées et Statistiques

Modélisation par Chaîne de Markov

Exercice 1

On lance une pièce équilibrée : les résultats du lancer forme une suite IID $(Y_n)_{n\in\mathbb{N}}$ à valeurs dans $\{0,1\}$. Pour tout $n \geq 1$ on note $X_n = Y_n + Y_{n-1}$.

1. Calculer $\mathbb{P}(X_3 = 0 \mid X_2 = 1, X_1 = 0)$ et $\mathbb{P}(X_3 = 0 \mid X_2 = 1)$.
2. Est ce que $(X_n)_{n\geq 1}$ est une chaîne de Markov ?

Exercice 2

On joue à la roulette. Il y a 18 numéros rouges, 18 numéros noirs, et un numéro vert le numéro zéro. On joue rouge en misant 1€ à chaque fois. On commence avec 5€ et on s'arrête si on a 10€ ou si on est ruiné. X_n désigne la fortune après n coups. Modéliser $(X_n)_{n\in\mathbb{N}}$ comme une chaîne de Markov en précisant son graphe de transition, sa matrice de transition, et sa loi initiale.

Exercice 3

Un zoo a reçu six gorilles, trois mâles et trois femelles répartis au hasard en deux cages de trois singes. Le directeur presbyte, incapable de discerner les sexes, décide de favoriser leur reproduction en permutant chaque semaine deux pensionnaires pris au hasard, un dans chaque cage. X_n est le nombre de guenons présentes la semaine n dans la première cage. Modéliser $(X_n)_{n\in\mathbb{N}}$ comme une chaîne de Markov en précisant son graphe de transition et sa matrice de transition.

Exercice 4

Pour le climat à Nantes $P = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$. Montrer que si $X_0 = 1$ ps, ie on part d'un jour où il fait beau, alors

$$\mathbb{P}(X_n = 1) = \frac{1}{2}(1 + 2^{-n})$$

(ne pas oublier que $\mathbb{P}(X_n = 2) = 1 - \mathbb{P}(X_n = 1)$). Qu'advient il de la loi de X_n quand n est grand ?

Exercice 5

On considère une chaîne de Markov homogène de matrice de transition

$$P = \begin{pmatrix} 7/20 & 3/20 & 1/4 & 1/4 \\ 3/10 & 1/4 & 7/20 & 1/10 \\ 1/4 & 1/4 & 7/20 & 3/20 \\ 3/10 & 1/4 & 1/4 & 1/5 \end{pmatrix}$$

1. Quelle est la loi conditionelle de X_{62} sachant $X_{60} = 3$.

2. On suppose que la chaîne est issue de la loi uniforme $\lambda = (1/4, 1/4, 1/4, 1/4)$. Déterminer la loi à l'instant 2.

Exercice 6

Un employé se rend chaque matin à pied de son appartement à son bureau et fait le contraire le soir. Il dispose en tout de 3 parapluies, certains chez lui, les autres au bureau. A Rennes, ville peu ensoleillée, il pleut 2 fois sur 3 lorsqu'il fait le trajet, et ce indépendamment du passé. X_n est le nombre de parapluies à son domicile lorsqu'il le quitte le matin. Modéliser $(X_n)_{n\in\mathbb{N}}$ comme une chaîne de Markov en précisant son graphe de transition et sa matrice de transition.

Exercice 7

Un magasin suisse vend des horloges : pour des raisons de place, il ne peut pas en stocker plus de 3. Le gérant a constaté que, en une journée, les probabilités de demande de 0, 1, 2 ou au moins 3 horloges sont respectivement :

$$p = [p_0, p_1, p_2, p_{\geq 3}] = [0.3, 0.4, 0.2, 0.1].$$

Chaque soir, il peut en commander à nouveau, qui seront disponibles en magasin le lendemain matin. On dit que l'Helvète applique une méthode $(i, j) 0 \leq i < j \leq 3$ s'il passe commande lorsqu'à la fermeture il lui reste un nombre inférieur ou égal à i en stock afin d'en avoir j en magasin le matin suivant. X_n est le nombre d'horloges dans le magasin le soir à la fermeture.

1. Méthode (2, 3) : donner la matrice de transition.

2. Même questions avec les méthodes (1, 3) et (0, 3).

Exercice 8

Considérons un système qui peut se trouver dans 3 états différents notés 1, 2, 3 à chaque unité de temps. Supposons que l'on puisse décrire l'évolution de l'état du système par une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice de transition

$$Q = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.1 & 0.8 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}$$

On ne connait pas l'état du système à l'instant 0 mais on sait qu'il a une chance sur 2 de s'être trouvé dans l'état 1 à l'instant 0 et une chance sur 2 de s'être trouvé dans l'état 2.

1. Quelle est la loi de X_3 ?

2. A l'instant 1, on a observé que le système se trouvait dans l'état 2. Quelle est la probabilité qu'il se trouve dans l'état 3 à l'instant 4 ?

Espérances et classes de communications pour les chaînes de Markov

Loi et espérance

Exercice 1

On considère le modèle de jeu de roulette étudié dans l'exercice 2 de la feuille 1. Quelle est, en moyenne, la fortune du joueur après trois parties ?
Note pour le correcteur : on a déjà calculé P et P^3 se calcule facilement à la main

Exercice 2

En reprenant le modèle du climat rennais étudié en Cours Magistral et à l'exercice 4 de la feuille 1, calculer le nombre moyen de jour de beau temps dans le mois qui suit un jour de beau temps.

Exercice 3

Une unité de production comprend 2 machines automatiques, qui fonctionnent indépendamment l'une de l'autre. Chaque machine fonctionne toute la journée avec la probabilité p ou bien tombe en panne durant la journée avec probabilité $1-p$. L'unité de production possède 1 technicien travailleur de nuit qui peut réparer au plus une machine tombée en panne et la remettre en marche pour le lendemain. On considère la variable aléatoire X_n égale au nombre de machines en panne le matin du n-ième jour.

1. Ecrire le graphe de transition et la matrice de transition de ce modèle.
2. Pour $p = 1/3$, calculer P^n.
3. Sachant que les deux machines sont opérationnelles le matin du premier jour, qu'une machine opérationnelle produit 1000 boulons par jour et une machine en panne n'en produit aucun, calculer le nombre moyen de boulons produits au bout de 30 jours.

Classes de communication et irréductibilité

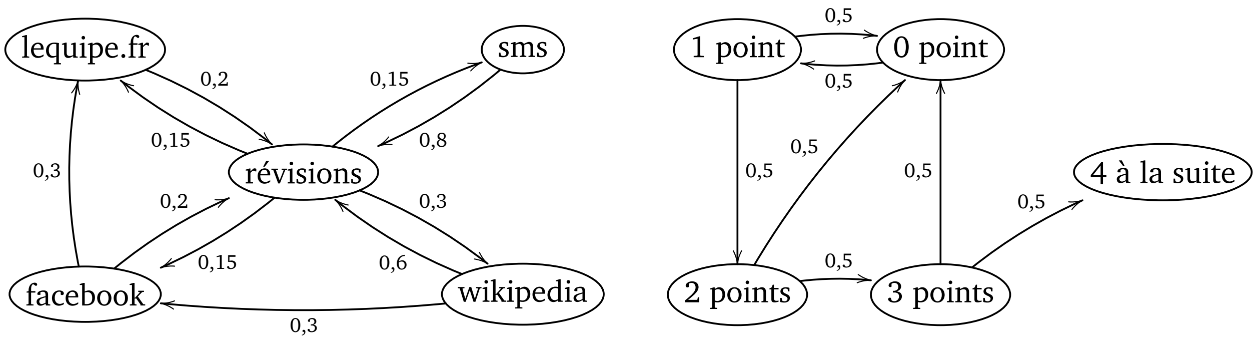
Exercice 4

On considère les chaînes de Markov dont les matrices de transition sont les suivantes :

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

Pour chacune de ces chaînes de Markov, écrire le graphe de transition et déterminer les classes de communication, en précisant celles sont ouvertes et celles qui sont fermées.

Exercice 5

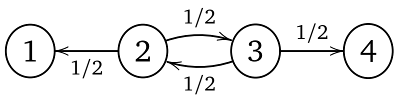


Pour chacun de ces diagrammes,

1. compléter les graphes avec les auto-arêtes manquantes et écrire les matrices de transitions correspondantes ;
2. déterminer les classes de communication, en précisant celles qui sont ouvertes et celles qui sont fermées.

Exercice 6

On considère la chaîne de Markov correspondant au diagramme suivant



1. Décomposer la chaîne en classes de communication. Que peut-on dire des états 1 et 4?
2. Ecrire la matrice de transition P de la chaîne.
3. On appelle mesure invariante toute mesure de probabilité λ telle que $\lambda P = \lambda$. Vérifier que λ est une mesure invariante si et seulement si λ s'écrit
$$\lambda = \begin{pmatrix} \alpha & 0 & 0 & 1-\alpha \end{pmatrix},$$
avec $\alpha \in [0,1]$.

Exercice 7

Soit $p \in]0,1[$. Montrer que la matrice

$$\begin{pmatrix} p & 1-p \\ 0 & 1 \end{pmatrix}$$

n'est pas irréductible mais possède une unique mesure invariante.

Exercice 8

Prendre la matrice de l'exercice XXIII p30 de [Falconnet] On considère la matrice de transition suivante :

$$P = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

1. Déterminer les classes de communication de ce processus en précisant lesquelles sont ouvertes ou fermées ;

2. Montrer qu'aucune probabilité invariante ne charge les états 2 et 4.

3. Déterminer les deux probabilités invariantes pour P.

X22M010 : Probabilités appliquées et Statistiques

Probabilités et temps d'absorption. Théorème ergodique

Exercice 1

JC est en prison et a 3 euros. Il peut sortir en versant une caution de 8 euros. Un garde accepte de faire une série de paris avec lui. Si JC parie A euros, il gagne A euros avec la probabilité $\alpha = 0.4$ et perd A euros avec la probabilité $1 - \alpha$.

1. Trouvez la probabilité qu'il gagne 8 dollars avent de perdre tout son argent si
   (a) Il mise 1 euro à chaque fois
   (b) Il mise à chaque fois le maximum, mais pas plus que le nécessaire pour obtenir 8 euros.
2. Quelle stratégie donne à JC la plus grande chance de sortir de prison ?

Exercice 2

On considère une chaîne de Markov homogène dont les états sont numérotés de 1 à 5. Sa matrice de transition est :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 3/20 & 1/10 & 3/10 & 1/5 & 1/4 \\ 1/5 & 3/20 & 3/20 & 7/20 & 3/20 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Déterminer les classes de communication de cette chaîne.
2. Calculer la probabilité que la chaîne de Markov, issue de l'état 4, réussisse à atteindre l'état 2.
3. Calculer cette même probabilité en supposant cette fois que la loi initiale de la chaîne est

$$\lambda = (0, 0, 1/3, 1/3, 1/3, 0).$$

Exercice 3

Un parieur possède 2 euros et désire obtenir 10 euros le plus vite possible. Il peut jouer au jeu suivant. Une pièce équilibrée est lancée. S'il parie sur le bon côté, il gagne une somme égale à sa mise, sinon il perd sa mise. Le parieur décide d'une statégie assez simpliste dans laquelle il mise tout son argent s'il a moins de 5 euros, et juste assez pour obtenir 10 euros s'il gagne sinon.
Soit $X_0 = 2$ et X_n son capital après n lancers. Montrez que le parieur obtiendra ses 10 euros avec une probabilité de $1/5$. Quel est le nombre moyen de lancers avant que le parieur ne perde tout ou parte avec son capital de 10 euros ? Pouvez vous trouver une meilleure stratégie ?

Exercice 4

On considère une chaîne de Markov sur l'espace d'états a, b, c, d de matrice de transition

$$\begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix}$$

1. Calculer la probabilité que la chaîne soit dans l'état b après 10 transitions, sachant qu'après 8 transitions elle est dans l'état a.
2. Montrer que la chaîne admet une unique probabilité invariante π que l'on déterminera.
3. Calculer pour tout état initial i, le temps moyen $k_i = \mathbb{E}_i[H_a]$ d'atteinte de l'état a : $H_a = \inf\{n \geq 0 : X_n = a\}$.
4. Calculer le temps moyen de retour en a, $\mathbb{E}_a[T_a]$ avec $T_a = \inf\{n \geq 1 : X_n = a\}$. Vérifier votre calcul en utilisant les résultats de la question 3.
5. On suppose que on vous propose un jeu pour lequel lorsque l'on se trouve dans l'état i, on se voit payer la somme $f(i)$ euros, avec $f(a) = 0, f(b) = f(c) = 1, f(d) = -2$. Devez vous accepter de jouer ?

Exercice 5

On considère une machine outil qui perce des trous dans une pièce de métal. Elle met 45 secondes pour percer un trou. Lorsqu'elle est bien réglée, elle perce toujours le trou parfaitement au centre de la pièce. Malheureusement, après chaque opération, elle a deux pour cent de chances de se dérégler. Lorsqu'elle est déréglée, elle perce un trou sur cinq hors du centre, et elle reste déréglée. Heureusement, on scanne toutes les pièces, et si on a un trou décentré, on rerègle la machine (on la réaligne), une opération qui prend 45 secondes : pendant cette période la machine ne perce aucun trou bien sûr.

1. Modéliser la vie de la machine outil par une chaîne de Markov sur un espace à 3 états à choisir parmi les quatre états suivants:
   - a: la machine est bien réglée et elle vient de percer un trou centré
   - b: la machine est bien réglée et elle vient de percer un trou décentré
   - c: la machine est déréglée et elle vient de percer un trou centré
   - d: la machine est déréglée et elle vient de percer un trou décentré
2. On suppose que l'on fait tourner cette machine sur un temps très long. Quelle est la proportion de pièces défectueuses produites ? Quel est le nombre moyen de pièces défectueuses produites sur une période d'une heure ?
3. Répondre aux questions précédentes, en supposant que le temps de réglage est maintenant de six minutes.

Exercice 6

Un robot collecteur de boîtes de soda peut se trouver dans 3 états suivant son niveau de batterie:

- 1 : batterie chargée

- 2 : batterie faible
- 3 : batterie à plat

Lorsque la batterie est chargée, elle se trouve faible la période suivante avec la probabilité α et chargée avec la probabilité $1-\alpha$. Lorsque la batterie est faible, elle se trouve à plat la période suivante avec la probabilité 4α et faible avec la probabilité $1-4\alpha$ (on suppose $0<\alpha<1/4$). Lorsque la batterie est à plat, on recharge le robot qui se retrouve ensuite avec sa batterie chargée.

1. Décrire l'évolution du robot par une chaîne de Markov dont on écrira explicitement la matrice de transition P. Quelle sont les classes de communication ? Montrer qu'il existe une unique probabilité invariante π que l'on déterminera en fonction de α.
2. On suppose que le coût de recharge de la batterie est $c=-3$ euros, et que par unité de temps, lorsqu'il n'est pas à plat, le robot collecte des boîtes pour une valeur de r euros. Montrer que $g_n=\frac{1}{n}S_n=\frac{1}{n}(R(X_0)+\cdots+R(X_n))$, gain temporel moyen sur n périodes, converge quel que soit l'état de départ vers une quantité $gm(r,\alpha)$.
3. Montrer qu'il existe une valeur $r_c(\alpha)$ telle que $gm(r,\alpha)>0$ si et seulement si $r>r_c(\alpha)$. C'est la valeur de collecte à partir de laquelle il est rentable d'installer le robot.

Exercice 7

On considère la chaîne de Markov $(X_n)_{n\geq 0}$ sur $I=\{1,2,3,4\}$ de matrice de transition

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

1. Montrer que la chaîne est irréductible, positive récurrente.
2. Trouver toutes les probabilités invariantes.
3. Montrer que $\frac{1}{n}(X_0+X_1+\ldots+X_{n-1})$ et $\frac{1}{n}(X_0^2+X_1^2+\ldots+X_{n-1}^2)$ convergent presque sûrement et déterminer leurs limites.

Exercice 8

Une chanteuse d'opéra doit donner une longue série de concerts. Son tempérament d'artiste la pousse à vouloir tous les soirs arrêter les concerts, et ce avec une probabilité de $\frac{1}{2}$. Une fois qu'elle a décidé d'arrêter, elle ne chantera pas de nouveau jusqu'à ce que l'organisateur la convainque de son admiration. Pour cela il lui envoie des fleurs chaque jour jusqu'à ce qu'elle revienne. Des fleurs coutant x milliers d'euros, $0\leq x\leq 1$, amènent une réconciliation avec la probabilité $\sqrt{x}$. L'organisateur fait 750 euros de bénéfice à chaque représentation donnée. Combien doit-il dépenser en fleurs ?

UNIVERSITÉ DE NANTES

X22M010 : Probabilités appliquées et Statistiques

Statistiques Descriptives

Exercice 1

Un atelier réalise le séchage de boues d'origine industrielle. Il obtient à la fin du processus des déchets mesurée en kg. On a observé les poids suivants après le traitement de plusieurs échantillons de 100 kg:

4.7	4.3	4.5	4.9	4.2	4.7
4.0	4.2	5.0	3.9	4.6	4.6
4.8	4.4	4.2	4.6	4.3	4.9
4.0	4.5	4.1	4.4	4.3	4.3

1. Constuire le tableau des effectifs et des fréquences
2. Définir la médiane, les quartiles et les calculer. Tracer le boxplot.
3. On établit 3 classes pour les modalités en prenant comme bornes 4.25 et 4.65. Construire le tableau des effectifs et en faire une représentation graphique.
4. Supposons que la 9ème valeur soit 50.0 et non 5.0. Que devient alors le boxplot ?

Exercice 2

On lance 100 fois un dé et on note les résultats

face	1	2	3	4	5	6
nombre d'occurences	13	16	18	16	13	24

1. Proposer deux manières de réprésenter graphiquement ces données. Faire une des deux représentations graphiques.
2. Tracer précisément le boxplot.

Exercice 3

Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5mn. Sur 100 observations de 5mn, on obtient les résultats suivants :

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1

1. Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures.
2. Calculer la moyenne et l'écart-type de cette série.

3. Déterminer la médiane, les quartiles et tracer le box-plot.

4. Etudier la symétrie de la série.

Exercice 4

La distribution des demandeurs d'emploi selon le sexe et la classe d'âge dans une localité est la suivante :

âge	Hommes	Femmes
[16,26[	280	160
[26,40[	310	360
[40,50[	240	120
[50,60[	420	530
[60,65[	70	50

1. Tracer sur une même figure les deux diagrammes en barre.

2. Déterminer les quartiles de la variable associant à chaque demandeur d'emploi masculin son âge. Même question pour les demandeurs d'emploi de sexe féminin.

3. Commentaires.

Estimation

Exercice 5

On dispose d'un n échantillon $(X_1,\dots,X_n)$ d'une loi de Poisson de paramètre $\lambda > 0$.

1. Montrer que la moyenne empirique $\bar{X}_n$est un estimateur sans biais de λ.

2. Montrer que $\bar{X}_n$ converge presque sûrement et dans L^2 vers λ lorsque $n \to +\infty$.

3. Calculer la variance de X_1 et déduisez en un autre estimateur de λ.

Exercice 6

1. **Préliminaire** Soit $g : E \to \mathbb{R}$ avec E une partie finie de $\mathbb{R}$. Montrer qu'il existe un polynôme $P[X]$ tel que pour tout x dans E, $g(x) = P(x)$.

2. Soit $E_1,\dots,E_n$ des parties finies de $\mathbb{R}$, $E = E_1 \times E_2 \times \dots \times E_n$, et $g : E \to \mathbb{R}$. Montrer qu'il existe un polynôme à n variables P tel que pour tout $(x_1,\dots,x_n)$ dans E: $g(x_1,\dots,x_n) = P(x_1,\dots,x_n)$. On dispose d'un n échantillon $(X_1,\dots,X_n)$ d'une loi de Bernoulli de paramètre $p \in]0,1[$ inconnu.

3. Calculer pour tous entiers positifs $m_1,\dots,m_n$
$$\mathbb{E}\left[X_1^{m_1}\cdots X_n^{m_n}\right].$$

4. Montrer qu'il n'existe pas d'estimateur sans biais de $\frac{1}{p}$.

X22M010 : Probabilités appliquées et Statistiques

Estimation et Tests

Exercice 1

On souhaite mesurer l'influence de l'alcool sur le temps de réaction au volant. Sur un échantillon aléatoire de 30 chauffeurs, le temps de réaction a été observé en laboratoire avec et sans consommation d'alcool (les 30 chauffeurs ont été réparti aléatoirement). Les temps de réactions en secondes ont été rapportés dans le tableau suivant :

Sans	0.68	0.64	0.68	0.82	0.58	0.80	0.72	0.65	0.84	0.73	0.65	0.59	0.78	0.67	0.65
Avec	0.73	0.62	0.66	0.92	0.68	0.87	0.77	0.70	0.88	0.79	0.72	0.60	0.78	0.66	0.68

1. Peut-on affirmer qu'il y a une influence de l'alcool sur le temps de réaction, avec un niveau de 5% ? Les valeurs de la fonction de répartition $\mathbb{P}(S \leq k)$ d'une binômiale $S \sim \mathscr{B}(14, 0.5)$ sont données ci dessous: 6.10351562e-05, 9.15527344e-04, 6.46972656e-03, 2.86865234e-02, 8.97827148e-02, 2.11975098e-01, 3.95263672e-01, 6.04736328e-01, 7.88024902e-01, 9.10217285e-01, 9.71313477e-01, 9.93530273e-01, 9.99084473e-01, 9.99938965e-01, 1.00000000e+00

2. Donner des intervalles de confiance au niveau 95 % pour les temps de réaction sans et avec alcool. On donne les moyennes et écarts types empiriques suivants:

$$\bar{x}_n = 0.7, \sigma_n = 0.08, \bar{x}'_n = 0.74, \sigma'_n = 0.09$$

   Commentez.

Solution de l'Exercice 1

1. On fait un test du signe, en supposant que l'alcool fait augmenter le temps de réaction comme H_1 :$N = 14$ et $D_+ = 11$ et $p = \mathbb{P}(S \geq D_+) = 0.03$ On rejette au niveau 5%

2. On obtient les intervalles de confiance suivants : sans alcool $[0.66, 0.74]$ avec alcool $[0.69, 0.78]$. Donc au niveau 90 % on ne rejette pas.

Exercice 2

On considère un très grand bassin avec M poissons. On ne peut pas tous les compter, cependant on aimerait estimer M. On prélève, une fois, K poissons, diférents, et on les marque.

1. Une première startégie consiste à prélever, avec remise, n poissons : le nombre de poissons marqués est S_n.

(a) Montrer que $\hat{M}_n = \frac{nK}{S_n}$ est un estimateur consistant de M.

(b) Montrer que $\hat{M}_n$ est biaisé.

2. Une seconde stratégie consiste à pêcher des poissons, avec remise, jusqu'à tomber sur un poisson marqué. On note T_1 le nombre de poissons qu'il a fallu pêcher. On recommence à pêcher des poissons et on note T_2 le nombre de poissons nécessaires. On continue jusqu'à disposer d'un échantillon $(T_1, \ldots, T_n)$.

   (a) Quelle est la loi de T_1 ?

   (b) Donner un estimateur consistant $\bar{M}_n$ de M. Quel est son biais ?

3. Donner un intervalle de confiance pour $\hat{M}_n$ au niveau 95%.

4. Comparer les deux stratégies.

Solution de l'Exercice 2

UNIVERSITÉ DE NANTES

X22M010 : Probabilités appliquées et Statistiques

On done les quantiles suivants de la loi normale :

$$q_{0.975} = 1.96, q_{0.95} = 1.64, q_{0.9} = 1.28, q_{0.995} = 2.57, q_{0.97} = 1.88$$

Exercice 1. Un échantillon de 10 000 personnes sur une population étant donné, on sait que le taux moyen de personnes à soigner pour un problème de cholestérol élevé est de 7,5%. Donner un intervalle dans lequel on soit «sûr» à 95%, de trouver le nombre exact de personnes à soigner sur les 10 000.

Exercice 2. Un vol Marseille - Paris est assuré par un Airbus de 150 places ; pour ce vol des estimations ont montré que la probabilité pour qu'une personne confirme son billet est $p = 0.75$. La compagnie vend n billets, $n > 150$. Soit X la variable aléatoire «nombre de personnes parmi les n possibles, ayant confirmé leur réservation pour ce vol».

1. Quelle est la loi exacte suivie par X ?
2. Déterminer un nombre maximum de places n_m que la compagnie peut vendre pour que, à au moins 95%, elle soit sûre que tout le monde puisse monter dans l'avion, c'est-à-dire n_m tel que : $P[X > 150] \leqslant 0.05$? Utiliser dans un premier temps l'inégalité de Markov puis l'inégalité de Chebyshev.
3. Reprendre le même exercice avec un avion de capacité de 300 places ; faites varier le paramètre $p = 0.5$; $p = 0.8$.

Exercice 3. Une compagnie aérienne a demandé des statistiques afin d'améliorer la sûreté au décollage et définir un poids limite de bagages. Pour l'estimation du poids des voyageurs et du poids des bagages, un échantillon est constitué de 300 passagers qui ont accepté d'être pesés : on a obtenu une moyenne m_e de 68kg, avec un écart-type σ_e de 7 kg. Définir un intervalle de confiance pour la moyenne des poids des passagers.

Exercice 4. Soit $(y_1, .., y_N)$ avec $N = 50$ l'observation d'un échantillon de taille N d'une loi de Poisson. On suppose que la moyenne empirique des y_i est 12 et que la variance empirique est 12. Construire un intervalle de confiance de niveau 95% pour le paramètre λ de la loi de Poisson.

Exercice 5. On s'intéresse au problème des algues toxiques qui atteignent certaines plages de France ; après étude on constate que 10% des plages sont atteintes par ce type d'algues et on veut tester l'influence de rejets chimiques nouveaux sur l'apparition de ces algues. Pour cela 50 plages proches de zones de rejet chimiques, sont observées ; on compte alors le nombre de plages atteintes par l'algue nocive : on constate que 10 plages sont atteintes par l'algue. Pouvez-vous répondre à la question «Les rejets chimiques ont-t-il modifié, de façon significative, avec le risque $\alpha = 0.05$, le nombre de plages atteintes ?»

Exercice 6. On interroge 1000 électeurs, 521 d'entre eux ont déclaré avoir l'intention de voter pour le candidat A. Indiquer avec une probabilité de 0,95 entre quelles limites se situe la proportion du corps électoral favorable à A au moment du sondage.

Exercice 7. La firme Comtec vient de développer un nouveau dispositif électronique. Avant de le mettre en production, on veut en estimer la fiabilité en termes de durée de vie. D'après le bureau de Recherche et Développement de l'entreprise, l'écart-type de la durée de vie de ce dispositif serait de l'ordre de 100 heures.
Déterminer le nombre d'essais requis pour estimer, avec un niveau de confiance de 95%, la durée de vie moyenne d'une grande production de sorte que la marge d'erreur dans l'estimation n'excède pas ± 50 heures. Même question pour une marge d'erreur n'excédant pas ± 20 heures.

Exercice 8. On effectue un sondage sur un échantillon de 10 000 personnes à la veille d'un référendum : 4903 d'entre elles s'apprêtent à voter oui, et 5097 à voter non. Quel risque d'erreur court-on en prédisant la victoire du non ?

Examen de Février 2018 : correction

Exercice 1

Soit une chaîne de Markov sur $E = \{1,2,3,4,5\}$ de matrice de transition :

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

1. Classifier les états.
2. Déterminer toutes les probabilités invariantes.
3. Quelle est la probabilité, partant de 1, d'atteindre $\{4,5\}$?

Correction

1. Les classes de communications sont les suivantes : $\{1\}$, qui est une classe ouverte, $\{2,3\}$ qui est une classe ouverte et $\{4,5\}$ qui est une classe fermée.

2. Une probabilité invariante est une mesure $\pi = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix}$ positive vérifiant $\pi P = \pi$ et $\sum_i \pi_i = 1$, avec P matrice de transition.

   Le système s'écrit
$$\left\{\begin{array}{cccccc} & & & & 0 & = \pi_1 \\ \frac{1}{2}\pi_1 & +\frac{1}{4}\pi_2 & +\frac{1}{2}\pi_3 & & & = \pi_2 \\ \frac{1}{2}\pi_1 & +\frac{1}{4}\pi_2 & +\frac{1}{2}\pi_3 & & & = \pi_3 \\ & \frac{1}{2}\pi_2 & & +\frac{1}{4}\pi_4 & +\frac{1}{2}\pi_5 & = \pi_4 \\ & & & \frac{3}{4}\pi_4 & +\frac{1}{2}\pi_5 & = \pi_5 \\ \pi_1 & +\pi_2 & +\pi_3 & +\pi_4 & +\pi_5 & = 1 \end{array}\right.$$

   Les trois premières équations forment un système en (π_1, π_2, π_3) qui donne facilement $\pi_1 = 0$, $\pi_2 = \pi_3 = 0$, ce qu'on aurait pu établir directement (les probabilités invariantes ne chargent pas les classes ouvertes).

   Les trois équations suivantes se résolvent alors facilement, avec deux équations équivalentes $\frac{3}{4}\pi_4 = \frac{1}{2}\pi_5$, on obtient $\pi_4 = \frac{2}{5}$, $\pi_5 = \frac{3}{5}$. Il y a donc une unique probabilité invariante $\begin{pmatrix} 0 & 0 & 0 & \frac{2}{5} & \frac{3}{5} \end{pmatrix}$.

3. La chaîne possède une unique classe fermée, $\{4,5\}$. La probabilité d'absorption dans cette unique classe fermée, partant de n'importe quel état, est donc 1.

Exercice 2

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov sur l'espace d'états $S=\{1,2\}$ de matrice de transition

$$P=\begin{pmatrix} 0,4 & 0,6 \\ 0,8 & 0,2 \end{pmatrix}.$$

1. Calculer $\mathbb{P}(X_7=1 \text{ et } X_5=2\,|\,X_4=1 \text{ et } X_3=2)$.
2. Calculer $\mathbb{E}[X_2|X_1=1]$.

Correction

1. D'après la définition des probabilités conditionnelles on a

   $$\mathbb{P}(X_7=1 \text{ et } X_5=2\,|\,X_4=1 \text{ et } X_3=2)=\mathbb{P}(X_7=1\,|\,X_5=2 \text{ et } X_4=1 \text{ et } X_3=2).\mathbb{P}(X_5=2\,|\,X_4=1 \text{ et } X_3=2).$$

   Comme (X_n) est une chaîne de Markov, on a d'après la définition et l'homogénéité

   $$\begin{cases} \mathbb{P}(X_7=1\,|\,X_5=2 \text{ et } X_4=1 \text{ et } X_3=2) & =\mathbb{P}(X_7=1\,|\,X_5=2)=\mathbb{P}(X_2=1\,|\,X_0=2) \\ \mathbb{P}(X_5=2\,|\,X_4=1 \text{ et } X_3=2) & =\mathbb{P}(X_5=2\,|\,X_4=1)=\mathbb{P}(X_1=2\,|\,X_0=1) \end{cases}$$

   Enfin $\mathbb{P}(X_2=1\,|\,X_0=2)=(P^2)_{21}=0,4.0,8+0,8.0,2=0,48$ et $\mathbb{P}(X_1=2\,|\,X_0=1)=P_{12}=0,6$. Donc $\mathbb{P}(X_7=1 \text{ et } X_5=2\,|\,X_4=1 \text{ et } X_3=2)=0,48.0,6=0,288$.

2. La loi de X_2 sachant $X_1=1$ est donnée par la première ligne de P : $\mathbb{P}(X_2=1|X_1=1)=P_{11}$ et $\mathbb{P}(X_2=2|X_1=1)=P_{12}$, d'où $\mathbb{E}(X_2|X_1=1)=1.P_{11}+2.P_{12}=0,4+2.0,6=1,6$.

UNIVERSITÉ DE NANTES

X22M010 : Probabilités appliquées et Statistiques

Examen de Mai 2018.

Documents non autorisés. Calculatrice, téléphone non autorisés. Toutes les réponses doivent être justifiées.
Barème indicatif : Exercice 1 (8 points), Exercice 2 (12 points).

Exercice 1

On modélise le climat de manière simpliste : soit il pleut soit il ne pleut pas.
On suppose que le fait qu'il pleuve ou non le lendemain ne dépend pas que du temps d'aujourdhui mais à la fois du temps d'aujourd'hui et de celui de la veille.
Plus précisément

- s'il a plu hier et aujourd'hui, alors il pleuvra demain avec la probabilité 0.8
- s'il a plu hier et pas aujourd'hui, alors il pleuvra demain avec la probabilité 0.3
- s'il n'a pas plu hier et qu'il pleut aujourd'hui, alors il pleuvra demain avec la probabilité 0.4
- S'il n'a plus ni hier ni aujourd'hui, alors il pleuvra demain avec la probabilité 0.2

1. Modéliser le climat comme une chaîne de Markov sur l' espace d'états

$$S = \{P, B\}^2 = \{(P, P), (P, B), (B, P), (B, B)\}$$

   L'état (P, B) signifie par exemple qu'il a plu hier et qu'il fait beau aujourd'hui. Vous devez préciser la matrice de transition P.

2. Y-a-t il une limite en temps long de la proportion du nombre de jours où il pleut ? Si oui calculer cette proportion limite en justifiant soigneusement votre réponse.

Solution de l'Exercice 1

1. On prend $S = \{(P, P), (P, B), (B, P), (B, B)\}$ avec P, B pour pluie et beau, $\omega = (i, j)$ signifie qu'il a fait i la veille et j aujourd'hui. Alors

$$P = \begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 \\ 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}$$

   (4 pts)

2. La chaîne est irréductible : Il faut le justifier. Par exemple:

$$PP \to PB \to BP \to PB \to BB \to BP \to PP\,.$$

donc la chaîne admet une unique probabilité invariante π(1 pts). Le théorème ergodique (1 pts) dit que la proportion limite du nombre de jours où il pleut est $p = \langle \pi, f\rangle = \pi((P,P)) + \pi((B,P))$ avec $f = \mathbf{1}_{(\{(P,P),(B,P)\})}$ car presque sûrement(2 pts)

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1} f(X_k) = \langle \pi, f\rangle\,.$$

On calcule (1 pts) $\pi = \frac{1}{15}(4,2,2,7)$ et donc $p = 0.4$.

Attention on ne peut pas appliquer de théorème de convergence vers l'équilibre qui ne donne pas le résultat voulu, car il dit que

$$\mathbb{P}(X_n \in \{(P,P),(B,P)\}) = \pi((P,P)) + \pi((B,P))\,.$$

Exercice 2

Un datacenter a N serveurs. Si l'un deux tombe en panne, alors il est remplacé par un neuf. Cela prend un jour pour retirer le serveur en panne, reconfigurer le neuf et le mettre en marche. Le temps est mesuré ici en jours.
Donc si au début du jour n il y a X_n serveurs en marche et que Y_n tombent en panne, alors au début du jour $n+1$ il y aura

$$X_{n+1} = N - Y_n$$

serveurs en fonctionnement.

1. On suppose qu'un serveur tombe en panne, indépendamment des autres serveurs, avec la probabilité $p \in]0,1[$. Montrer que

$$\mathbb{P}(Y_n = y \mid X_n = x) = \binom{x}{y} p^y (1-p)^{x-y}, \quad (y = 0, 1, \dots, x)\,.$$

2. Montrer que $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur l'espace d'états $S = \{0, 1, ..., N\}$ de probabilités de transition

$$p_{ij} = \binom{i}{N-j} p^{N-j}(1-p)^{i-N+j}\, \mathbf{1}_{(N-i\le j\le N)}\,.$$

3. Montrer que la chaîne admet une unique probabilité invariante π.

4. On pose $\mu = \sum_{i=0}^{N} i\pi(i)$. On suppose que X_0 suit la loi π. Quelle est la loi de X_n ?

   Montrer que pour tout $n \in \mathbb{N}$

   (a) $\mathbb{E}[X_n] = \mu$

   (b) $\mathbb{E}[Y_n] = p\mu$

   (c) $\mathbb{E}[X_{n+1}] = N - p\mu$.

   En déduire l'expression de μ en fonction de p et N.

5. Montrer que la chaîne admet une probabilité réversible λ et en déduire l'expression de π et un autre calcul de μ.

Solution de l'Exercice 2

1. Sachant $X_n = x$, Les x serveurs tombent en panne indépendamment avc la même probabilité : on est donc dans un schéma de succès échec et la loi conditionnelle de Y_n sachant $X_n = x$ est la loi binomiale $\mathscr{B}(x,p)$.(2 pts)

2. Y_n ne dépend que de X_n et pas de $X_{n-1}, X_{n-2}, \ldots$ (1 pts)et donc

$$\begin{aligned}\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)\\ = \mathbb{P}(Y_n = N-j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)\\ = \mathbb{P}(Y_n = N-j \mid X_n = i)\\ = \binom{i}{N-j} p^{N-j}(1-p)^{i-N+j}\,\mathbf{1}_{(N-i\le j\le N)}\end{aligned}$$

   (1 pts)

3. (1 pts) La chaîne est irréductible car pour tout i, $p_{iN} > 0$ et $p_{Ni} > 0$. Elle admet donc une unique probabilité invariante (1 pts)

4. (a) π est invariante donc $X_n \sim \pi P^n = \pi$ (1 pts) donc $\mathbb{E}[X_n] = \mathbb{E}[X_0] = \mu$.(1 pts)
   (b) On a $\mathbb{E}[Y_n \mid X_n = x] = xp$ car c'est l'espérance d'une binômiale. Ensuite on utilise $\mathbb{E}[Y_n] = \sum_x \mathbb{E}[Y_n \mid X_n = x]\mathbb{P}(X_n = x) = \sum xp\pi(x) = p\mu$. On peut aussi utiliser la formule des probabilités totales pour calculer $\mathbb{P}(Y_n = y)$ et puis $\mathbb{E}[Y_n] = \sum_y y\mathbb{P}(Y_n = y)$ pour retrouver bien sûr le résultat.(2 pts)
   (c) On a $X_{n+1} = N - Y_n$ et on utilise la linéarité de l'espérance(1 pts)

   On a $\mu = \mathbb{E}[X_{n+1}] = N - p\mu$ d'ou $\mu = \frac{N}{1+p}$. (1 pts)

5. On trouve en écrivant les équations de bilan détaillé $\lambda_i p_{ij} = \lambda_j p_{ji}$ que $\lambda(i) = Cp^{-i}\binom{N}{i}$ convient (1 pts). On en déduit la valeur de $C = (1+p^{-1})^N$ (1 pts), on a $\lambda = \mu$ par unicité de la proba invariante(1 pts) et on retrouve $\mu = \frac{N}{p+1}$(1 pts).

UNIVERSITÉ DE NANTES

X22M010 : Probabilités appliquées et Statistiques

Examen de Mai 2018.

Documents non autorisés. Calculatrice, téléphone non autorisés. Toutes les réponses doivent être justifiées.

Exercice 1

Etant donnés $p \in]0,1[$ et un entier $N \geq 2$, on considère une chaîne de Marov $(X_n)_{n\in\mathbb{N}}$ sur $S = \{1,\dots,n\}$ de matrice de transition P telle que

$$\begin{aligned} p_{i,i-1} &= p\,, \quad \text{pour } i = 2,3,\dots N \\ p_{i,i+1} &= \frac{1-p}{N-1}\,, \text{ pour } i = 1,2,3,\dots N-1 \\ p_{i,i} &= 1-p-\frac{1-p}{N-1}\,, \text{ pour } i = 2,3,\dots N-1 \\ p_{1,1} &= 1-\frac{1-p}{N-1}\,, \qquad p_{N,N} = 1-p\,. \end{aligned}$$

1. Montrer que la chaîne admet une unique probabilité invariante π.
2. Montrer que $\pi_2 = \frac{1-p}{p(N-1)}\pi_1$ et $\pi_3 = \frac{1-p}{p(N-1)}\pi_2$.
3. En déduire toutes les valeurs de π_i, pour $i = 1,\dots,N$.

Solution de l'Exercice 1

(8 pts)

1. D'après la première ligne $N \to N-1 \to \dots \to 1$ et d'après la seconde $1 \to 2 \to \dots \to N$. Donc la chaîne est irréductible et admet une unique probabilité invariante.
2. de l'équation$\pi P = \pi$ on tire $\pi_1 = \sum_j \pi_j p_{j,1} = p_{1,1}\pi_1 + p_{2,1}\pi_2$ donc $\pi_2 = \frac{1-p_{1,1}}{p_{2,1}}\pi_1$
3. Le pous simple est de chercher une proba réversible et on en troue une en posant $\pi_{i+1} = \rho\pi_i$ avec $\rho = \frac{1-p}{p(N-1)}$.

Exercice 2

Soit M un entier $M \geq 2$. Soit X une variable aléatoire discrète de loi uniforme qur $\{0,\dots,M\}$, c'est à dire $\mathbb{P}(X = k) = \frac{1}{M+1}$ pour $k \in \{0,\dots,M\}$.

1. Calculer $\mathbb{E}[X]$ et $var(X)$.
2. Soit $X_1,\dots,X_n$ IID de même loi que X. Proposer un estimateur sans biais $\hat{M}_n$ de M. DOnner son erreur quadratique moyenne.

---

3. Donner un intervalle de confiance I_n à 95 % pour l'estimation de M par $\hat{M}_n$ qui dépende de $\hat{M}_n$ et de la variance empirique σ_n^2 de l'échantillon $(X_1,\dots,X_n)$. On donne $q_{0.975} = 1.96$.

4. On pose $M_n = \max(X_1, X_2, \dots, X_n)$. Calculer pour tout $k \in \{0,\dots,M\}$ la probabilité $\mathbb{P}(M_n \leq k)$.

5. On considère l'intervalle de confiance $J_n = [M_n - 1, M_n]$ pour l'estimation de M par M_n. Calculer $\mathbb{P}(M \in J_n)$.

6. Comparer les deux estimateurs de M.

Solution de l'Exercice 2

1. $\mathbb{E}[X] = M/2$ et $var(X) = \frac{M(M+2)}{12}$