MAP 311 - Aléatoire - PC 1

Emmanuel Gobet

Feuille de PC disponible en avance sur le site `http://www.cmap.polytechnique.fr/~gobet/`.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : probabilité, événement, expérience, indépendance, conditionnement, formule de Bayes, lemme de Borel-Cantelli.

Probabilités

EXERCICE 1 - Les anniversaires communs. ($\star$)
Si n personnes sont réunies dans une pièce, quelle est la probabilité qu'aucune paire d'individus n'ait un anniversaire commun ? Déterminer $n_{\min}$ pour que cette probabilité soit inférieure à 0.5 (on pourra utiliser l'équivalence $\ln(1+x) \sim x$ pour $x \ll 1$).

EXERCICE 2 - Tirage de boules.
Une urne contient r boules rouges et b boules blanches.

1. Décrire l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ associé aux expériences suivantes :
   a) On tire k boules une à une, sans remise.
   b) On tire k boules une à une, avec remise.
2. Pour les deux expériences précédentes, calculer la probabilité que le nombre de boules blanches tirées soit égal à j, pour $j \in \mathbb{N}$.

EXERCICE 3 - Formule du crible. ($\star$)

1. Montrer la formule de Poincaré (formule du crible) :

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{1\le i\le n} \mathbb{P}(A_i) - \sum_{1\le i_1<i_2\le n} \mathbb{P}(A_{i_1}\cap A_{i_2}) \\ &\qquad + \cdots + (-1)^{n+1}\mathbb{P}(A_1\cap A_2\cap\cdots\cap A_n) \\ &= \sum_{p=1}^{n}(-1)^{p+1}\sum_{1\le i_1<\cdots<i_p\le n}\mathbb{P}(A_{i_1}\cap\cdots\cap A_{i_2}).\end{aligned}$$

   Indication : on pourra utiliser le fait que $1 - \mathbf{1}_{\bigcup_{i=1}^n A_i} = \prod_{i=1}^n (1-\mathbf{1}_{A_i})$, ou procéder par récurrence.
2. Soit S une permutation aléatoire sur n objets, de loi uniforme.
   a) Quelle est la probabilité que S ait au moins un point fixe ?
   b) Quelle est la probabilité que S ait exactement k points fixes ? Montrer que ceci est en général très proche de $\exp(-1)/k!$.
3. Une personne écrit n lettres et les ferme avant d'écrire les adresses sur les enveloppes. On suppose que tous les destinataires ont une adresse différente. Quelle est la probabilité pour qu'au moins une lettre arrive au bon destinataire ? Donner la limite de cette probabilité lorsque $n \to \infty$.

EXERCICE 4 - Indépendance deux à deux et indépendance mutuelle.

Soit $\Omega := \{\omega_1, \omega_2, \omega_3, \omega_4\}$ muni de la loi uniforme (i.e., chaque élément a probabilité 1/4). On définit les événements $A := \{\omega_1, \omega_2\}$, $B := \{\omega_1, \omega_3\}$ et $C := \{\omega_2, \omega_3\}$. Montrer que A, B et C sont indépendants deux à deux. Comparer $\mathbb{P}(A \cap B \cap C)$ et $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

EXERCICE 5 - Rôle de l'information. ($\star$)

L'exercice suivant est très classique et a de nombreuses variantes. Il illustre l'utilité d'une formulation mathématique rigoureuse pour éviter des pièges et des paradoxes dus à des raisonnements spécieux.

Une famille a deux enfants. Quelle est la probabilité pour que les deux enfants soient des filles sachant que l'enfant le plus jeune est une fille ? Quelle est la probabilité pour que les deux enfants soient des filles sachant que l'un des enfants est une fille ?

EXERCICE 6 - Prendre la bonne porte. ($\star$)

Lors d'un jeu télévisé, le candidat a le choix entre trois portes. Derrière l'une se trouve un cadeau, derrière les deux autres il n'y a rien. Le présentateur demande au candidat de choisir une porte. S'il choisit celle où se trouve le cadeau il le gagne, sinon il s'en retourne les mains vides.

Le candidat choisit sa porte. Pour faire monter le suspense, le présentateur décide d'ouvrir une autre porte, derrière laquelle il sait qu'il n'y a rien. Il redemande alors au candidat de choisir une porte. Quelle est la meilleure stratégie pour le candidat ? Garder la porte qu'il avait choisie, choisir au hasard une des deux portes restantes (à pile ou face), ou changer de porte ?

EXERCICE 7 - Test de qualité de production.

On considère une usine fabriquant des cartes électroniques. Lors de la production une carte sur 100 000 est défectueuse. En fin de production, on effectue un test pour savoir si la carte est défectueuse ou non. Pour les pièces défectueuses le test est positif dans 95 % des cas, pour les pièces correctes dans 1% des cas.

Le test indique qu'une pièce donnée est défectueuse (test positif). Quelle est la probabilité que la pièce soit effectivement défectueuse ?

EXERCICE 8 - Un exemple non trivial d'événements indépendants.

Dans l'expérience suivante, on dispose de trois dés à 6 faces, deux d'entre eux étant normaux et le 3ème pipé. Un dé est normal si chaque face a même probabilité d'apparaître, alors qu'un dé est pipé si la face 6 est trois fois plus probable que les autres.

1. On choisit au hasard un dé et on le lance. Quelle est la probabilité que la face i apparaisse ? Sachant que le dé a donné un 6, quelle est la probabilité qu'il soit pipé ?
2. On choisit maintenant deux dés parmi les trois disponibles et on les lance. On note $A = \{$le dé pipé fait partie des deux dés choisis$\}$.
   a) Calculer $\mathbb{P}(A)$. Quelle est la probabilité de A sachant que la somme des résultats des deux dés vaut 12 ?
   b) Même question avec une somme égale à 7.
   c) Soit $B = \{$la somme des résultats des deux dés lancés vaut 7$\}$. Montrer que A et B sont indépendants.

EXERCICE 9 - Mauvaise fouille.

On dispose de n sacs numérotés de 1 à n ; une balle a été mise au hasard dans l'un d'eux. Elle se trouve dans le sac i avec la probabilité p_i. Lorsqu'on fouille un sac, la probabilité de ne pas trouver la balle alors qu'elle s'y trouve vaut p (la fouille n'est pas assez minutieuse).

1. On fouille une première fois le sac 1 et on ne trouve pas la balle. Quelle est la probabilité qu'elle soit quand même dans le sac 1 ? dans le sac j $(j \neq 1)$?
2. On recommence la fouille du sac 1 et on ne trouve toujours pas la boule. Calculer de nouveau la probabilité qu'elle soit dans le sac 1, puis dans le sac j $(j \neq 1)$.

EXERCICE 10 - Un buveur impénitent décide de ne plus boire . . .

Un buveur impénitent décide d'essayer de ne plus boire. S'il ne boit pas un jour donné, alors la probabilité qu'il ne boive pas le lendemain est de 0,4. En revanche, s'il succombe un jour il ne boira pas le lendemain par remords avec une probabilité de seulement 0,2.

1. Quelle est la probabilité que ce buveur ne boive pas le n-ième jour ?
2. Que se passe-t-il quand n tend vers l'infini ? Interpréter.
3. On suppose à présent qu'à chaque Nouvel An, notre homme prend la résolution d'arrêter de boire. Il ne boit donc jamais le 1er janvier, indépendamment de son comportement de la veille, puis l'engrenage fatal le rattrape et il se comporte les jours suivants comme décrit précédemment. En utilisant le lemme de Borel-Cantelli, montrer qu'avec probabilité 1 il existe une année pendant laquelle il ne boira pas du tout, tandis que la probabilité qu'il existe une année à partir de laquelle il ne boira plus jamais est nulle. *(On supposera que notre buveur est immortel.)*

Evénements, tribus

EXERCICE 11 - Tribu engendrée par une famille finie.

1. Montrez qu'une tribu engendrée par une famille finie de parties est engendrée par une partition finie.
2. Décrire tous les éléments de cette tribu.
3. Montrez que ce résultat est faux si on remplace fini par dénombrable. (Utiliser la tribu borélienne de $\mathbb{R}$.)

EXERCICE 12 - Un exemple d'ensemble non mesurable.

Sur $\mathbb{R}$ on définit la relation d'équivalence $x \sim y$ si $x - y \in \mathbb{Q}$. En utilisant l'axiome du choix (si A est une fonction sur un ensemble I telle que $A(x) \neq \emptyset$ pour tout x de I, il existe une fonction f telle que $f(x) \in A(x)$ pour tout $x \in I$), construire un ensemble $A \in [0, 1[$ qui contient exactement un point de chaque classe d'équivalence. Supposons A mesurable, et soit $\alpha = \lambda(A)$ sa mesure de Lebesgue. Montrer que si $r, s \in \mathbb{Q}$ et $r \neq s$, alors $(A+s) \cap (A+r) = \emptyset$, où $A+x = \{y+x \,:\, y \in A\}$, et que $\lambda(A + s) = \lambda(A)$. Remarquer que

$$1 = \lambda\big([0,1]\big) \leq \lambda\Bigg(\bigcup_{r \in \mathbb{Q} \cap]-1,1[} (A + r)\Bigg) \leq \lambda\big([-1,2]\big) = 3.$$

En utilisant la σ-additivité de λ, montrer que cette inégalité conduit d'une part à $\alpha = 0$, d'autre part à $\alpha > 0$. Conclure.

MAP 311 - Aléatoire - PC 2

Emmanuel Gobet

Feuille de PC disponible en avance sur le site http://www.cmap.polytechnique.fr/~gobet/.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : variable aléatoire, loi, espérance, variance, fonction génératrice, indépendance de v.a., loi conditionnelle.

Variables aléatoires discrètes

EXERCICE 1 - Loi-produit uniforme.
Soient E et F des ensembles finis. Montrer qu'il est équivalent de supposer que X et Y sont deux v.a. indépendantes de lois uniformes sur E et F, et de supposer que le couple (X, Y) est de loi uniforme sur $E \times F$.

EXERCICE 2 - Caractérisation de la loi géométrique.
Trouver toutes les lois sur $\mathbb{N}$ satisfaisant la propriété suivante d'*absence de mémoire* :

$$\forall (i,j) \in \mathbb{N}^2 : \mathbb{P}(X > i + j | X > i) = \mathbb{P}(X > j).$$

EXERCICE 3 - Somme de Poisson.
Soit X_1, X_2 des variables aléatoires indépendantes de loi de Poisson de paramètres respectifs $\theta_1 > 0$ et $\theta_2 > 0$.

1. Calculer la loi de $X_1 + X_2$.
2. Calculer la loi de X_1 sachant $X_1 + X_2$. Reconnaître cette loi.
3. Calculer $\mathbb{E}(X_1 \,|\, X_1 + X_2)$.

EXERCICE 4 - Le vendeur de glaces.
Un vendeur de glaces installé dans la rue voit passer un nombre N de passants dans la journée. On suppose que la variable aléatoire N suit une loi de Poisson de paramètre $\lambda > 0$. On appelle G_i le nombre de glaces achetées par le i-ème passant. On fait l'hypothèse que les $N, G_1, G_2, \ldots$ sont indépendantes et que chaque passant achète une glace avec probabilité p et aucune glace avec probabilité $1 - p$. Quelle est la loi du nombre X de glaces vendues pendant une journée ?

EXERCICE 5 - Pari sportif OL-OM.
Au cours d'un match de foot, le nombre de buts B marqué par une équipe suit une loi géométrique de paramètre $p \in]0, 1[$, indépendamment de l'équipe adverse : $\mathbb{P}(B = i) = p(1 - p)^i$ pour $i \geq 0$.

1. En moyenne, l'équipe de l'Olympique de Massy marque 1 but par match, alors que pour l'Olympique de Longjumeau, la moyenne est de 2. Quels paramètres respectifs p_M et p_L choisir pour modéliser leur nombre de but marqué par match ?
2. Le prochain match oppose l'OL à l'OM. La cote pour l'*OL gagnant* est de $\frac{3}{2}$ (signifiant que pour 1 misé, on reçoit $\frac{3}{2}$, soit un gain total de $\frac{1}{2}$). Quelle est la cote pour *OM gagne ou fait match nul* ?

3. Est-il préférable de miser pour *OL gagnant* ?
   On cherchera à justifier son choix par des arguments probabilistes !

EXERCICE 6 - Quelques formules bonnes à retenir.

1. **($\star$)** Soit X une v.a. discrète à valeurs dans $\mathbb{N}$. Montrer que
$$\mathbb{E}(X) = \sum_{n \geq 0} \mathbb{P}(X > n).$$
2. Soient $X_1, \ldots, X_p$ des v.a. indépendantes de même loi, à valeurs dans $\mathbb{N}$. On pose $p_k = \mathbb{P}(X = k), k \in \mathbb{N}$ et $r_n = \sum_{k \geq n} p_k$. Montrer que
$$\mathbb{E}(\min(X_1, \ldots, X_p)) = \sum_{n \geq 1} r_n^p.$$

EXERCICE 7 - Minimum et maximum de v.a. géométriques.

Soient $X_1, \ldots, X_n$ des variables aléatoires géométriques de paramètre p :
$$\mathbb{P}(X_j = n) = p(1-p)^{n-1},\ n \in \mathbb{N}^*.$$

1. Déterminer la loi de $X = \min(X_1, \ldots, X_n)$ et $Y = \max(X_1, \ldots, X_n)$.
2. Déterminer la loi jointe de (X, Y).

EXERCICE 8 - Encore des boules.

On répartit au hasard n boules dans r cases. On définit les v.a. $(Y_i)_{1 \leq i \leq r}$ et $(X_i)_{1 \leq i \leq r}$ par
$$\begin{aligned} Y_i &= \text{nombre de boules dans la case } i \\ X_i &= 1 \text{ si la case } i \text{ est occupée, 0 sinon.} \end{aligned}$$

1. Quelle est la loi de Y_i ? de X_i ? Calculer $\mathbb{E}[X_i]$ et $\mathrm{Var}(X_i)$.
2. Pour $i \neq j$, calculer $\mathrm{Cov}(X_i, X_j)$.
3. Soit S_r le nombre de cases vides. Calculer $\mathbb{E}[S_r]$ et $\mathrm{Var}(S_r)$.

EXERCICE 9 - Nombre de succès.

Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d. (indépendantes et identiquement distribuées) suivant la loi de Bernoulli de paramètre $p \in]0, 1[$ et $S = X_1 + \ldots + X_n$ leur somme. Pour $s \in \{0, \ldots, n\}$, donner la loi conditionnelle de X_1 sachant $S = s$ et calculer $\mathbb{E}(X_1 \mid S)$.

EXERCICE 10 - Instant de succès.

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes, de même loi de Bernoulli de paramètre p. Pour tout $n \geq 1$, on définit par récurrence, $T_n = \inf\{k > T_{n-1} \,;\, X_k = 1\}$ si cet infimum est fini, $T_n = \infty$ sinon, et $T_0 = 0$.

1. Démontrer que les variables aléatoires T_1, $T_2 - T_1, \ldots, T_n - T_{n-1}, \ldots$ sont indépendantes et de même loi.
2. Calculer la loi de T_1 et sa fonction génératrice. En déduire la loi de T_n.

3. Calculer l'espérance et la variance de T_n de manière directe (formule du cours), puis en utilisant sa fonction génératrice.

EXERCICE 11 - Le jeu de la poule.

Trois joueurs A, B, C jouent une suite de parties suivant la règle suivante : chaque partie oppose deux joueurs et le perdant de la partie s'efface à la partie suivante au profit des deux autres joueurs. Les parties sont indépendantes entre elles et chaque participant y a la même probabilité $1/2$ de gagner ; les joueurs A et B commencent. Le jeu s'arrête la première fois qu'un joueur a gagné deux fois de suite ; ce joueur est le gagnant du jeu.

1. Quelle est la probabilité de chacun de gagner ?
2. Soit T la durée du jeu. Calculer la loi de T, sa fonction génératrice, son espérance et sa variance. Montrer que $\mathbb{P}(T < \infty) = 1$.

EXERCICE 12 - Le cambrioleur amnésique.

Un trousseau de a clés contient une seule clé pouvant ouvrir une porte donnée. Une personne tente d'ouvrir cette porte en essayant tour à tour les diverses clés du trousseau jusqu'à tomber sur la bonne. On note X le nombre d'essais nécessaires pour ouvrir la porte.
Déterminer la loi de X dans les cas suivants :

1. la personne n'essaie jamais d'ouvrir avec une clé déjà essayée en vain.
2. la personne n'a pas de mémoire et ne tient jamais compte des essais précédents.
3. la personne n'a qu'une mémoire limitée et fait chaque essai avec n'importe quelle clé autre que celle qu'elle vient d'essayer immédiatement en vain.

EXERCICE 13 - De longues études.

Pour avoir son diplôme, un étudiant doit réussir les examens des M cours qui sont proposés, et obtenir ainsi les M unités de valeur correspondantes. On suppose que l'étudiant a une probabilité $p \in]0, 1[$ de réussir chacun des examens. D'année en année, l'étudiant peut reporter les unités de valeur obtenues, et il ne passe alors que les examens des cours qu'il n'a pas réussis.

1. Soit X_n le nombre d'unités obtenues pendant les n premières années (si $X_n = M$, alors $X_m = M$ pour $m \geq n$). Calculer la loi de X_n.
2. Soit N le nombre total d'examens passés. Calculer $\mathbb{E}(N)$.
   Indication : comment s'exprime N en fonction des X_n ?
3. Soit T le nombre d'années nécessaires pour obtenir le diplôme. Calculer $\mathbb{E}(T)$.
   Indication : comment s'exprime $\{T > j\}$ en fonction des X_n ?

EXERCICE 14 - Cartes mémoires défectueuses.

La compagnie MicroHard produit des cartes mémoires pour micro-ordinateurs. Ces cartes comportent en particulier un composant C. On admet que chaque carte est défectueuse avec probabilité p, indépendamment des autres. D'autre part, si une carte est défectueuse, la probabilité que le composant C soit défectueux vaut q. On examine un lot de n cartes produites et on note N le nombre de celles qui sont défectueuses et K le nombre de celles dont le composant C est défectueux ($0 \leq K \leq N \leq n$).

1. Quelle est la loi de N ?
2. Quelle est la loi conditionnelle de K sachant $\{N = m\}$?
3. En déduire que la loi de K est une binomiale avec des paramètres à préciser.

MAP 311 - Aléatoire - PC 3

Emmanuel Gobet

Feuille de PC disponible en avance sur le site http://www.cmap.polytechnique.fr/~gobet/.
Les exercices marqués ($\star$) *sont corrigés dans le livre **Aléatoire** de S. Méléard.*

Mots-clés de la semaine : variable aléatoire réelle, fonction de répartition, espérance, mesure de Lebesgue, densité, simulation par inversion de la f.d.r., inégalités classiques (Markov, Bienaymé-Chebyshev, Jensen).

Lois et densités

EXERCICE 1 - Loi de Weibull.

Soit Z une variable aléatoire de loi exponentielle de paramètre 1. Pour tout $\alpha, \beta > 0$ on pose $X = \beta Z^{1/\alpha}$.

1. Quelle est la fonction de répartition de X ?
2. La loi de X admet-elle une densité ?

La loi de X est appelée *loi de Weibull de paramètre* (α, β). Elle est beaucoup utilisée (notamment pour modéliser des temps de panne) car en ajustant ses deux paramètres il est possible d'approcher correctement de nombreuses lois de v.a.r. positives.

EXERCICE 2 - Loi de Cauchy.

Une particule est émise dans le plan à partir du point origine O et dans la direction du vecteur aléatoire $\vec{u}$. L'extrémité de $\vec{u}$ est un point de loi uniforme sur le demi-cercle $\{x^2 + y^2 = 1,\, x > 0\}$. Après une trajectoire rectiligne, la particule heurte la droite $D = \{x = 1\}$ en un point M d'abscisse X.

1. ($\star$) Montrer X est égale à $\tan(\theta)$ avec θ de loi uniforme sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. En déduire la loi de X (fonction de répartition, densité, ...).
2. Montrer que $1/X$ a la même loi que X. (*Indication :* on pourra tirer parti des relations trigonométriques.)
3. Quelle est la probabilité que la particule heurte D dans le segment $I = \{|y| \leq 1\}$?
4. Le point O émet maintenant un nombre aléatoire poissonien de paramètre λ de particules. Chacune d'entre elles se comporte comme ci-dessus, indépendamment des autres. Quelle est la loi du nombre Z de particules heurtant I ?

EXERCICE 3 - Variables aléatoires normales et log-normales.

Posons $X = e^Y$, où Y est de loi $\mathcal{N}(\mu, \sigma^2)$. On dit que X est de loi Log-Normale.

1. Trouver
   a) la fonction de densité f de la v.a. X ;
   b) $\mathbb{E}(X)$ et $\mathrm{Var}(X)$.
2. *Un contre-exemple célèbre, d'après W. Feller.* On prend $\mu = 0$ et $\sigma = 1$. Fixons $|a| \leq 1$ et définissons
$$f_a(x) = (1 + a \sin(2\pi \log(x)))\, f(x).$$
   a) Montrer que f_a est une densité de probabilité. (*Indication :* On pourra écrire l'intégrale de f_a comme l'espérance d'une fonction de X.)

b) Montrer que f_a admet des moments de tout ordre qui ne dépendent pas de a.

c) En déduire qu'une variable aléatoire n'est pas caractérisée de manière unique par ses moments polynômiaux. *On peut montrer toutefois que c'est vrai dès que la variable est bornée.*

d) Trouver des fonctions $\varphi : \mathbb{R} \mapsto \mathbb{R}$ non nulles dont la projection orthogonale sur l'espace vectoriel des polynômes de degré n est nulle (projection au sens du produit scalaire induit par la loi log-normale de X), c'est-à-dire

$$\mathbb{E}\big[\varphi(X)X^n\big] = 0, \qquad \forall n \in \mathbb{N}.$$

Moments d'une v.a.r.

EXERCICE 4 - Moments et médiane d'une variable aléatoire réelle.

Soit X une variable aléatoire réelle à densité, de fonction de répartition F. Montrer que :

1. ($\star$) Si X est à valeurs positives et $k \geq 0$, alors

$$\mathbb{E}(X^{k+1}) = (k+1)\int_0^\infty x^k(1-F(x))\,dx.$$

On retiendra en particulier le cas $k = 0$ (comparer avec la formule donnée pour les v.a. discrètes).

2. Pour tout réel a, on a

$$\mathbb{E}(|X-a|) = \int_{-\infty}^a F(x)\,dx + \int_a^\infty (1-F(x))\,dx.$$

3. Pour quelle(s) valeur(s) de a la quantité $\mathbb{E}(|X-a|)$ est-elle minimale ?

EXERCICE 5 - Des inégalités importantes.

1. *Inégalité de Hölder.* Soit $p, q > 0$. Montrer que si $\frac{1}{p} + \frac{1}{q} = 1$, alors

$$\mathbb{E}\big[|XY|\big] \leq \mathbb{E}\big[|X|^p\big]^{1/p}\mathbb{E}\big[|Y|^q\big]^{1/q}.$$

Le cas $p = q = 2$ est aussi appelé *inégalité de Cauchy-Schwarz* et est à retenir absolument.

2. En déduire que $\mu(\lambda) = \ln \mathbb{E}[e^{\lambda X}]$ est convexe en λ. Pour éviter les problèmes d'intégrabilité, on supposera que X est bornée.

3. *Inégalité de Minkowski.* Montrer que si $p \geq 1$, $\mathbb{E}\big[|X+Y|^p\big]^{1/p} \leq \mathbb{E}\big[|X|^p\big]^{1/p} + \mathbb{E}\big[|Y|^p\big]^{1/p}$.
Inégalité de Chernov. Montrer que pour toute v.a. réelle X bornée et $\lambda \geq 0$,

$$\mathbb{P}(X \geq a) \leq e^{-(a\lambda - \mu(\lambda))},$$

où l'on a posé à nouveau $\mu(\lambda) = \ln \mathbb{E}[e^{\lambda X}]$.

EXERCICE 6 - L'espérance minimise (presque) les écarts d'ordre p.

On considère une v.a.r. X ayant des moments d'ordre $p \geq 1$ finis, et on note

$$m_p := \inf_{x\in\mathbb{R}} \Big[\mathbb{E}\big[|X-x|^p\big]\Big]^{1/p}.$$

1. Montrer que $x = \mathbb{E}[X]$ minimise m_p pour $p = 2$.
2. Plus généralement, montrer que $\mathbb{E}[X]$ minimise la norme L^p pour tout $p \geq 1$, à une constante 1/2 près :

$$\frac{1}{2}\left[\mathbb{E}\big[|X - \mathbb{E}[X]|^p\big]\right]^{1/p} \leq m_p \leq \left[\mathbb{E}\big[|X - \mathbb{E}[X]|^p\big]\right]^{1/p}.$$

Indication : On pourra utiliser l'inégalité de convexité $|x+y|^p \leq 2^{p-1}(|x|^p + |y|^p)$.

EXERCICE 7 - Une minoration utile.

Soit X une v.a. réelle de carré intégrable et $0 < a < \mathbb{E}[|X|]$. Etablir la minoration

$$\mathbb{P}(|X| \geq a) \geq \frac{\big(\mathbb{E}[|X|] - a\big)^2}{\mathbb{E}[X^2]}.$$

Simulation de v.a.r.

EXERCICE 8 - Générer une loi géométrique.

La fonction « rand » de Scilab vous permet de générer une variable aléatoire U de loi uniforme sur $[0, 1]$.

1. Comment générer à partir de U une variable X de loi exponentielle de paramètre $\lambda > 0$?
2. Soit $a > 0$. Quelle est la loi de $Y = \lfloor aX \rfloor + 1$?

EXERCICE 9 - Ecriture dyadique.

Tout réel $x \in [0, 1]$ peut s'écrire $x = \sum_{k \geq 1} y_k 2^{-k}$, où $y_k \in \{0, 1\}$: cette écriture est unique sauf sur l'ensemble dénombrable des nombres dyadiques (qui admettent deux développements, l'un avec $y_k \equiv 0$ pour k assez grand, l'autre avec $y_k \equiv 1$ pour k assez grand).

Soient Y_k une suite de v.a. indépendantes de loi de Bernoulli de paramètre 1/2 et définissons la v.a. X par

$$X := \sum_{k \geq 1} Y_k \, 2^{-k}.$$

L'objectif est de montrer de deux manières différentes que X suit la loi uniforme sur $[0, 1]$.

1. Montrer que pour tous $0 \leq a \leq b \leq 1$, $\mathbb{P}(a \leq X \leq b) = b - a$ (on pourra d'abord montrer que $\mathbb{P}(X \geq a) = 1 - a$ si a est dyadique). Conclure.
2. a) Calculer la *fonction caractéristique* de la loi uniforme sur $[0, 1]$, i.e. la fonction $t \mapsto \mathbb{E}[e^{itU}]$ où U est une v.a. de loi uniforme.

   b) Soit $X_n := \sum_{k=1}^n Y_k \, 2^{-k}$. Calculer la fonction caractéristique ϕ_n de X_n. En montrant par récurrence que

$$\sin\left(\frac{\lambda}{2^n}\right) \prod_{k=1}^n \cos\left(\frac{\lambda}{2^k}\right) = \frac{\sin \lambda}{2^n},$$

   déterminer la limite de ϕ_n lorsque $n \to \infty$. En admettant que la fonction caractéristique caractérise la loi d'une v.a.r., conclure.
3. En déduire un algorithme de simulation d'une v.a. de loi uniforme sur $[0, 1]$ avec une précision arbitrairement grande.

MAP 311 - Aléatoire - PC 4

Emmanuel Gobet

Feuille de PC disponible en avance sur le site http://www.cmap.polytechnique.fr/~gobet/.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : vecteur aléatoire, densité marginale, loi conditionnelle, loi gaussienne dans $\mathbb{R}^d$, espérance conditionnelle, covariance, indépendance, méthode de rejet, somme de variables indépendantes.

EXERCICE 1 - La cerise sur le gâteau.

On considère un gâteau circulaire avec une cerise sur le bord. On découpe le gâteau en deux parts en coupant suivant deux rayons choisis uniformément au hasard.

1. Avec quelle probabilité la part contenant la cerise est-elle plus petite que l'autre part ?
2. Quelle est la taille (longueur angulaire) moyenne de la part contenant la cerise ?

EXERCICE 2 - Principe de symétrie.

Posons $S_n = \sum_{i=1}^n X_i$, où les X_i sont des v.a. indépendantes de même loi, symétriques (X_1 et $-X_1$ ont même loi). Montrer que pour tout $y \geq 0$ et $N \in \mathbb{N}^*$,

$$\mathbb{P}\left[\sup_{n\leq N} S_n \geq y\right] \leq 2\,\mathbb{P}[S_N \geq y] = \mathbb{P}[|S_N| \geq y].$$

EXERCICE 3 - Matrice de covariance.

On considère un vecteur aléatoire $(X_1, \dots, X_n)$, de matrice de covariance K. Montrer que la matrice K n'est pas inversible si et seulement si les v.a. X_i sont linéairement dépendantes :

$$a_1X_1 + a_2X_2 + \cdots + a_nX_n = c \qquad \text{avec probabilité } 1$$

(avec $a_1, \dots, a_n$ non tous nuls).

EXERCICE 4 - Minimum de n exponentielles.

Soit $(T_k,\, 1 \leq k \leq n)$ une famille de v.a. réelles indépendantes telles que $P(T_k = x) = 0$ pour tout $1 \leq k \leq n$ et $x \in \mathbb{R}$. (Vérifier que c'est le cas si les variables sont à densité.)

1. Montrer que $P(T_i = T_j) = 0$ pour tout $1 \leq i < j \leq n$, puis que $P(\exists i \neq j : T_i = T_j) = 0$.

   Nous pouvons donc définir une v.a. M comme étant l'indice de la plus petite des variables T_k, $1 \leq k \leq n$, ainsi que

   $$S := \min_{1\leq k\leq n} T_k = T_M\,, \qquad \text{et} \qquad R_i = T_i - S\,, \quad 1 \leq i \leq n\,.$$

   Nous supposerons désormais que chaque T_k est une v.a. exponentielle de paramètre $\alpha_k > 0$.
2. Montrer que pour tout indice j et toutes fonctions positives f et g_i ($i \neq j$),

   $$\mathbb{E}\left(\mathbf{1}_{\{M=j\}} f(S) \prod_{i\neq j} g_i(R_i)\right) = \alpha_j \int f(t) e^{-(\alpha_1+\cdots+\alpha_n)t}\, dt \times \prod_{i\neq j} \alpha_i \int g_i(u) e^{-\alpha_i u}\, du\,.$$

   Indication : On cherchera à écrire directement la quantité à l'intérieur de l'espérance comme une fonction de $(T_1, \dots, T_n)$.

3. En déduire que $\{M = j\}$ est un événement indépendant de la v.a. S, que
$$\mathbb{P}(M = j) = \frac{\alpha_j}{\alpha_1 + \cdots + \alpha_n},$$
et que S est de loi exponentielle de paramètre $\alpha_1 + \cdots + \alpha_n$.
4. En déduire aussi que, conditionnellement à $\{M = j\}$, pour $i \neq j$ les v.a. S et R_i sont indépendantes et que chaque R_i suit une loi exponentielle de paramètre α_i.
5. **Application :** Trois personnes A, B et C arrivent en même temps pour téléphoner près de deux cabines téléphoniques. A et B occupent les cabines en premier et C remplace le premier sorti. On désigne par X_A, X_B et X_C les v.a. correspondant à la durée de la communication pour chaque personne. On suppose qu'il s'agit de v.a. de loi exponentielle de paramètre λ.
   a) Calculer la probabilité que C sorte le dernier.
   b) Donner la loi du temps total que C passe à attendre et à téléphoner.
   c) Déterminer la loi de l'instant du dernier départ.

EXERCICE 5 - Somme d'exponentielles.

Soit $X_1, \ldots, X_n$ des v.a. indépendantes de loi exponentielle de paramètre $\lambda > 0$.

1. ($\star$) Trouver par récurrence la densité de $S_n = \sum_{i=1}^{n} X_i$. On appelle la loi de S_n la loi Gamma de paramètres n et λ. On la note $\Gamma(n, \lambda)$.
2. Calculer la densité conditionnelle de X_i sachant S_n.
3. En déduire la valeur de $\mathbb{E}[X_i \,|\, S_n = t]$.
4. Retrouver ce résultat par un argument de symétrie.

EXERCICE 6 - Quelques identités remarquables sur les variables gaussiennes.

1. Pourquoi la méthode d'inversion de la fonction de répartition pose-t-elle problème lorsqu'on veut simuler une v.a. gaussienne ?
2. ($\star$) **Transformation de Box-Muller.** Soient ρ et θ deux variables aléatoires indépendantes, la première de loi exponentielle de paramètre $1/2$ et la seconde de loi uniforme sur $[0, 2\pi]$. Montrer que $\sqrt{\rho}\sin(\theta)$ et $\sqrt{\rho}\cos(\theta)$ sont deux v.a. gaussiennes $\mathcal{N}(0,1)$ indépendantes.

   En déduire une manière de simuler deux v.a. normales indépendantes à partir de deux v.a. de loi uniforme sur $[0, 1]$ indépendantes.
3. Soient X et Y deux v.a. normales centrées réduites indépendantes. Montrer que X/Y a la loi de Cauchy, c.à.d. qu'elle a pour densité
$$\frac{1}{\pi(1 + x^2)}.$$
4. **Seconde application :** Montrer que
$$Z = \frac{2XY}{\sqrt{X^2 + Y^2}} \qquad \text{et} \qquad W = \frac{X^2 - Y^2}{\sqrt{X^2 + Y^2}}$$
définissent à nouveau deux v.a. $\mathcal{N}(0,1)$ indépendantes.

EXERCICE 7 - Indépendance de v.a. réelles et racines d'équations.

Soient X et Y deux variables aléatoires réelles indépendantes dont les lois ont des densités données respectivement par

$$p_X(x) = \exp(-x)\mathbf{1}_{\{x\geq 0\}}, \qquad p_Y(y) = \mathbf{1}_{\{0<y<1\}}.$$

On considère l'équation en λ

$$\lambda^2 - 2\lambda X + Y = 0.$$

Calculer la probabilité que

1. l'équation ait une racine double ;
2. l'équation n'ait aucune racine réelle.

EXERCICE 8 - Roméo et Juliette.

Roméo et Juliette se sont donnés rendez-vous entre midi et 13h. Chacun des deux a promis de ne pas attendre l'autre plus de 15 minutes. On suppose qu'ils arrivent indépendamment à des instants uniformément distribués entre midi et 13h.

1. Quelle est la probabilité qu'ils se rencontrent ?
2. Roméo décide d'arriver à l'heure H. Quelle est la probabilité de la rencontre ?
3. Roméo arrive à l'heure H et ne trouve personne. Quelle est la probabilité de la rencontre ?

EXERCICE 9 - Espérance conditionnelle pour un vecteur gaussien.

Soit (X, Y) un vecteur gaussien non dégénéré. L'objectif est de montrer que

$$\mathbb{E}(Y|X) = a_0 + b_0 X, \tag{1}$$

pour des constantes a_0 et b_0 à évaluer.

1. Montrer que l'espérance conditionnelle $\mathbb{E}[Y|X]$ est la meilleure approximation de Y sachant X, au sens où il s'agit de la variable aléatoire Φ qui minimise $\mathbb{E}[(Y-\Phi)^2]$, dans l'ensemble de toutes les v.a. Φ de carré intégrable dépendant uniquement de X.
2. En utilisant le fait que le vecteur $(X, Y-(a+bX))$ est gaussien pour tout a et b, montrer (1).
3. En calculant la densité conditionnelle de Y sachant X, retrouver le résultat (1). Pour simplifier, on pourra supposer ici que (X, Y) est d'espérance nulle et de matrice variance-covariance $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, pour $|\rho| < 1$.

MAP 311 - Aléatoire - PC 5

Emmanuel Gobet

Feuille de PC disponible en avance sur le site http://www.cmap.polytechnique.fr/~gobet/.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : convergence en probabilité, en moyenne et presque sûre, lois faibles et fortes des grands nombres.

Convergence de variables aléatoires

EXERCICE 1 - Différents modes de convergence.

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.

1. Montrer que $(X_n)_{n\geq 1}$ converge en probabilité vers 0 si et seulement si $(\mathbb{P}(X_n > 0))_{n\geq 1}$ tend vers 0.
2. Montrer que $(X_n)_{n\geq 1}$ converge presque sûrement vers 0 ssi $\sum_{n\geq 1} \mathbb{P}(X_n > 0) < +\infty$.
3. On suppose que pour tout $n \geq 1$, X_n suit une loi binômiale $\mathcal{B}(n, p_n)$. Préciser si la suite $(X_n)_{n\geq 1}$ converge vers 0 en moyenne, en probabilité, presque sûrement, dans le cas où $p_n = \frac{1}{n^2}$ puis dans le cas où $p_n = \frac{1}{n^3}$.
4. On suppose que pour tout $n \geq 1$, X_n suit une loi de Poisson $\mathcal{P}(\alpha_n)$. Préciser si la suite $(X_n)_{n\geq 1}$ converge vers 0 en moyenne, en probabilité, presque sûrement, dans le cas où $\alpha_n = \frac{1}{n}$ puis dans le cas où $\alpha_n = \frac{1}{n^2}$.
5. On suppose que pour tout $n \geq 1$, la loi de X_n est donnée par
$$\mathbb{P}(X_n = n^2) = \beta_n \qquad \text{et} \qquad \mathbb{P}(X_n = 0) = 1 - \beta_n.$$
Préciser si la suite $(X_n)_{n\geq 1}$ converge vers 0 en moyenne, en probabilité, presque sûrement, dans le cas où $\beta_n = \frac{1}{n}$, $\beta_n = \frac{1}{n^2}$, puis dans le cas où $\beta_n = \frac{1}{n^3}$.

EXERCICE 2 - Majoration presque sûre.

Soit $(X_n)_{n\geq 0}$ une suite quelconque de variables aléatoires telles que $\mathbb{P}(|X_n| < \infty) = 1$ pour tout n. Montrer qu'il existe une suite $(c_n)_{n\geq 0}$ telle que $\lim_{n\to\infty} \frac{X_n}{c_n} = 0$ p.s.

EXERCICE 3 - Lois uniformes.

Soit $(U_n)_{n\geq 1}$ une suite de v.a. indépendantes, toutes de loi uniforme sur un intervalle $[a, b]$. Trouver les limites presque sûres de $\min_{i\leq n} U_i$ et $\max_{i\leq n} U_i$.

EXERCICE 4 - Lois de Cauchy.

Soit $(X_i)_{i\geq 1}$ une suite de v.a. i.i.d. de loi de Cauchy, i.e. de densité sur $\mathbb{R}$ égale à $p_X(x) = \frac{1}{\pi(1+x^2)}$. On admettra que pour tout $n \in \mathbb{N}$, $S_n := \frac{1}{n}\sum_{i=1}^n X_i$ suit également une loi de Cauchy (ceci se démontre facilement en utilisant les *fonctions caractéristiques*, introduites dans le prochain cours).

1. Que nous dit la loi des grands nombres sur la convergence de S_n lorsque $n \to \infty$?
2. Calculer la loi de $S_{2n} - S_n$. En déduire le comportement de $\mathbb{P}[|S_{2n} - S_n| > 1]$ lorsque $n \to \infty$.
Indication : on utilisera plusieurs fois la propriété énoncée ci-dessus, ainsi que la symétrie de X.

3. La suite $(S_n)_{n\geq 1}$ converge-t-elle en probabilité ?
4. Finalement, la suite $(S_n)_{n\geq 1}$ converge-t-elle presque sûrement ?

Loi des grands nombres

EXERCICE 5 - De l'art de gaspiller sa fortune.

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. strictement positives et d'espérance 1, telles que

$$\mathbb{P}(X_1 = 1) < 1 \qquad \text{et} \qquad \exists \alpha > 0 \text{ t.q. } \mathbb{P}(X_1 \geq \alpha) = 1.$$

La première condition garantit que X n'est pas la constante égale à 1 p.s., tandis que la seconde condition nous dit que X est p.s. bornée inférieurement par une constante $\alpha > 0$.

Posons $Y_n = \prod_{i=1}^n X_i$.

1. En remarquant que pour $x \in]0, \infty[\setminus\{1\}$, $\ln(x) < x - 1$, vérifier que $\mathbb{E}(\ln(X_1)) < 0$.
2. En étudiant le comportement de $\ln(Y_n)/n$ lorsque $n \to +\infty$, montrer que la suite Y_n converge presque sûrement vers une limite que l'on précisera.
3. Un joueur met tous les jours en jeu le dixième de sa fortune à pile ou face avec un ami. Quelle est l'évolution de cette fortune au bout d'un temps très long ?
4. Calculer $\mathbb{E}(Y_n)$. La suite Y_n converge-t-elle en moyenne ?

EXERCICE 6 - Loi des grands nombres déguisée.

Soit $(X_n)_{n\geq 1}$ une suite de v.a. i.i.d. de loi normale $\mathcal{N}(\mu, \sigma^2)$. Trouver la limite presque sûre de

$$\ln\left(\frac{1}{n}\sum_{i=1}^n e^{X_i}\right) \qquad \text{et} \qquad \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i^2}.$$

EXERCICE 7 - Calcul d'intégrales.

Pour $f : \mathbb{R} \to \mathbb{R}$ continue bornée et $\alpha > 0$, calculer les limites suivantes :

$$\lim_{n\to\infty} \int_{[0,1]^n} f\left(\frac{x_1 + \ldots + x_n}{n}\right) dx_1 \ldots dx_n \quad \text{et} \quad \lim_{n\to\infty} \sum_{k\geq 0} e^{-\alpha n}\frac{(\alpha n)^k}{k!} f\left(\frac{k}{n}\right).$$

EXERCICE 8 - Estimation de paramètres.

Un standard téléphonique reçoit chaque jour un nombre aléatoire d'appels, dont on suppose qu'il suit une loi de Poisson de paramètre $\theta > 0$ inconnu. On suppose en outre que d'un jour à l'autre, ces nombres sont indépendants. Afin de trouver le nombre optimal de standardistes pour subvenir aux besoins des clients, on cherche à estimer le paramètre θ grâce à la donnée d'une série $(J_n)_{1\leq n\leq N}$ de nombres d'appels journaliers.

1. Calculer l'espérance et la variance de $\frac{1}{N}\sum_{i=1}^N J_i$. Proposer un estimateur $\hat{\theta}_N$ de θ, i.e. une fonction des données $(J_i)_i$ qui donne une information sur la valeur du paramètre recherché.
2. On dit d'un estimateur qu'il est *sans biais* lorsque $\mathbb{E}[\hat{\theta}_N] = \theta$ (il est dit *biaisé* sinon). L'estimateur ci-dessus est-il biaisé ?
3. Calculer l'espérance de $\frac{1}{N}\sum_{i=1}^N J_i^2$. Proposer un nouvel estimateur $\tilde{\theta}_N$ de θ. Est-il biaisé ? Lequel de ces estimateurs vous semble le meilleur ?

MAP 311 - Aléatoire - PC 6

Emmanuel Gobet

Feuille de PC disponible en avance sur le site `http://www.cmap.polytechnique.fr/~gobet/`.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : fonctions caractéristiques, convergence en loi, théorème de la limite centrale, méthode de Monte Carlo.

Fonctions caractéristiques

EXERCICE 1 - Loi Gamma.

Pour tout $r > 0$, on pose $\Gamma(r) = \int_0^\infty e^{-t}t^{r-1}dt$ et on appelle loi $\Gamma_{r,\lambda}$ ($r > 0, \lambda > 0$) la loi de probabilité de densité

$$\gamma_{r,\lambda}(t) = \frac{\lambda^r}{\Gamma(r)} e^{-\lambda t} t^{r-1} \mathbf{1}_{\mathbb{R}_+}(t).$$

1. Calculer la fonction caractéristique de $\Gamma_{r,\lambda}$. En déduire l'espérance et la variance de cette loi.
2. Montrer que si X et Y sont des v.a indépendantes de loi respective $\Gamma_{r,\lambda}$ et $\Gamma_{s,\lambda}$, alors $X+Y$ a une loi $\Gamma_{r+s,\lambda}$. Pourrait-on aussi faire varier le paramètre λ ?
3. Soient $X_1, \ldots, X_n$ des v.a indépendantes de loi exponentielle de paramètre λ. Calculer la loi de $X_1 + \ldots + X_n$.
4. Soient $X_1, \ldots, X_n$ des v.a indépendantes de loi $\mathcal{N}(0,1)$. Montrer que X_1^2 a une loi Γ et en déduire la loi de $X_1^2 + \ldots + X_n^2$.
   On appelle cette loi la *loi du* χ^2 à n degrés de liberté ; elle a de nombreuses applications en statistiques (voir Proposition 6.2.8, *Aléatoire*).

EXERCICE 2 - Vecteurs gaussiens.

Soient $(X_1, \ldots, X_n)$ des variables aléatoires i.i.d gaussiennes centrées réduites. On pose $X = (X_1, \ldots, X_n)$.

1. Calculer la fonction caractéristique de X.
2. Soit $A \in M_n(\mathbb{R})$ une matrice inversible et $b \in \mathbb{R}^n$. Calculer la fonction caractéristique du vecteur aléatoire $AX + b$, ainsi que la densité de sa loi.
3. Soit $P \in M_n(\mathbb{R})$ une matrice orthogonale. Montrer que X et PX ont la même loi. En déduire que la loi de $\frac{X}{\|X\|}$ est une mesure de probabilité sur la sphère unité $S^{(n-1)}$ de $\mathbb{R}^n$ invariante par toute transformation orthogonale (cette propriété caractérise la mesure uniforme sur $S^{(n-1)}$).

EXERCICE 3 - Convergence en loi.

1. Etudier la convergence en loi de la suite $\left(\frac{X_n}{n}\right)_{n\geq 1}$, où X_n suit une loi géométrique de paramètre $p_n = \frac{\lambda}{n}$ et $\lambda > 0$ est fixé.
2. Soit X_n une v.a. de loi uniforme sur $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$.
   a) Trouver la limite en loi de la suite $(X_n)_{n\geq 1}$. On notera X une v.a. ayant cette loi.
   b) Montrer que $\mathbb{P}(X_n \in \mathbb{Q})$ ne converge pas vers $\mathbb{P}(X \in \mathbb{Q})$. Comparer avec la définition de la convergence en loi.

Théorème de la limite centrale

EXERCICE 4 - Stabilité des lois gaussiennes. ($\star$)

Considérons X_1 et X_2, deux v.a. indépendantes de même loi μ, de variance finie égale à σ^2 et qui satisfont la propriété suivante :

$$\textbf{(P)} \qquad \frac{X_1 + X_2}{\sqrt{2}} \text{ a encore la loi } \mu.$$

L'objectif est de montrer que nécessairement, μ est la loi normale $\mathcal{N}(0, \sigma^2)$.

1. Montrer que si μ est la loi normale $\mathcal{N}(0, \sigma^2)$, alors la propriété **(P)** est satisfaite.
2. On suppose maintenant que **(P)** est satisfaite.
   a) Montrer que $\mathbb{E}(X_1) = 0$.
   b) Prouver que si $X_1, X_2, \cdots, X_{2^n}$ sont des v.a. indépendantes de même loi μ, alors $\dfrac{X_1 + X_2 + \cdots + X_{2^n}}{\sqrt{2^n}}$ a la loi μ.
   c) Conclure que μ est la loi normale $\mathcal{N}(0, \sigma^2)$.

EXERCICE 5 - Dilemme du couturier.

Un couturier fabrique deux modèles de chemise, le modèle A à rayures et le modèle B à carreaux, pour 2000 clients. Chaque client choisit le modèle A avec probabilité $1/4$ et donc le modèle B avec probabilité $3/4$. Combien ce couturier doit-il fabriquer de chemises de type A pour qu'il y en ait suffisamment avec probabilité 90% ?

Estimation statistique

EXERCICE 6 - Estimation paramétrique de loi uniforme.

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. i.i.d. de loi uniforme sur $[0, a]$ avec $a > 0$. On suppose que la valeur de a n'est pas connue, et l'objectif est de l'estimer en utilisant les observations $(X_n)_{1 \le i \le n}$.

1. Que vaut $\mathbb{E}(X_1)$? En déduire un estimateur sans biais de a à partir de $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$.

   Quelle est sa vitesse d'estimation de a ?
2. On pose $M_n = \max\limits_{1 \le i \le n} X_i$. Calculer la loi de M_n et son espérance. Montrer que M_n converge $p.s$ vers a.
3. Quelle est limite en loi de $n(a - M_n)$? Des deux estimateurs précédents, lequel vous parait le plus intéressant ?

Méthode de Monte-Carlo

EXERCICE 7 - Réduction de variance. ($\star$)

Soit f une fonction monotone bornée sur $[0, 1]$. On souhaite calculer l'intégrale $I = \int_0^1 f(x)\, dx$ à l'aide d'un générateur de v.a. $(X_n)_{n \ge 1}$ i.i.d., de loi uniforme sur $[0, 1]$.

1. Comparer les performances des estimateurs

$$I_n = \frac{1}{2n} \sum_{i=1}^{2n} f(X_i) \qquad \text{et} \qquad I'_n = \frac{1}{2n} \sum_{i=1}^{n} \big\{ f(X_i) + f(1 - X_i) \big\},$$

qui requièrent tous deux $2n$ évaluations de la fonction f. *Indication :* on pourra commencer par montrer que si X et Y sont i.i.d. et uniformes sur $[0, 1]$, on a $\mathbb{E}[(f(X) - f(Y))(f(1 - X) - f(1 - Y))] \le 0$.

2. Donner pour le cas (d'école) $f(x) = x^2$ le nombre n minimal permettant d'obtenir avec 95% de chances une précision de calcul de 1%.

EXERCICE 8 - Echantillonnage d'importance.

On cherche à calculer par une méthode de Monte-Carlo l'intégrale $I = \int_{\mathbb{R}^d} f(x)p(x)\,dx$, où p est une densité de probabilité. Pour ce faire, on utilise une suite de tirages i.i.d. $(X_n)_{n\geq 1}$ de v.a. à valeurs dans $\mathbb{R}^d$ dont la loi a pour densité de probabilité q ne s'annulant pas. On note w le rapport (appelé *fonction d'importance*) $w(x) = \frac{p(x)}{q(x)}$ et on considère l'estimateur pondéré

$$I_n = \frac{1}{n}\sum_{i=1}^{n} w(X_i)f(X_i).$$

1. Vérifier la limite presque sûre de cet estimateur et donner une condition sous laquelle il vérifie le théorème de la limite centrale.
2. Sous l'hypothèse obtenue à la question précédente et en supposant que $f(x)p(x)$ est strictement positif pour tout x, trouver le choix optimal de la densité q (qui minimise la variance de I_n). Pourquoi ce choix est-il en général impossible en pratique ?

MAP 311 - Aléatoire - PC 7

Emmanuel Gobet

Feuille de PC disponible en avance sur le site http://www.cmap.polytechnique.fr/~gobet/.
Les exercices marqués ($\star$) *sont corrigés dans le livre* ***Aléatoire*** *de S. Méléard.*

Mots-clés de la semaine : intervalles de confiance, modèles dynamiques aléatoires, marches aléatoires, somme aléatoire de variables aléatoires indépendantes, processus de branchement

EXERCICE 1 - Intervalle de confiance.

On effectue une enquête durant une épidémie de grippe, dans le but de connaître la proportion p de personnes présentant ensuite des complications graves. On observe un échantillon représentatif de 400 personnes et pour un tel échantillon, 40 personnes ont présenté des complications.

1. Donner un intervalle de confiance pour p de niveau 95%.
2. On désire que la valeur estimée $\hat{p}$ diffère de la proportion inconnue exacte p de moins de 0.005 avec une probabilité égale à 95%. Combien de personnes notre échantillon devra-t-il contenir ?
3. Quel devrait être le risque pour obtenir le même intervalle qu'à la question précédente en conservant l'effectif $n = 400$? Quelle conclusion peut-on en tirer ?

EXERCICE 2 - Temps d'atteinte d'une marche aléatoire dans $\mathbb{Z}$.

Un joueur parie avec une pièce biaisée qui donne Pile avec probabilité $p \in]0,1[$ et Face avec probabilité $q = 1-p$: il la lance de manière répétée et indépendante et gagne 1€ si la pièce donne Pile et perd 1€ sinon.

Notons S_n ses gains cumulés à l'instant n, qu'on peut écrire à l'aide d'une marche aléatoire sur $\mathbb{Z}$ en posant $S_n = X_1 + \cdots + X_n$, où $(X_i)_{i\geq 1}$ est une suite de v.a. indépendantes de même loi donnée par $\mathbb{P}(X_1 = 1) = p$ et $\mathbb{P}(X_1 = -1) = q$.

Considérons maintenant le premier instant T_r où le joueur atteint un richesse $r \in \mathbb{N}^*$:

$$T_r = \inf\{n \geq 0 : S_n = r\}$$

avec la convention $T_r = +\infty$ si $S_n \neq r$ pour tout n.

1. Montrer que T_1 et $T_2 - T_1$ sont deux v.a. indépendantes de même loi.
2. Montrer que $\mathbb{P}(T_1 - 1 = k | S_1 = -1) = \mathbb{P}(T_2 = k)$ pour tout $k \in \mathbb{N}$.
3. En conditionnant par rapport aux valeurs de S_1, montrer que φ, la fonction génératrice de T_1, vérifie $\varphi(s) = ps + qs\varphi(s)^2$ pour $|s| \leq 1$. En déduire l'expression de φ.
4. Calculer $\mathbb{P}(T_1 < +\infty)$ et $\mathbb{E}(T_1)$ en distinguant les cas selon que $p <, =, > 1/2$.
5. Quelle est la loi de $M = \sup_{n\in\mathbb{N}} S_n$?

EXERCICE 3 - Marche aléatoire dans $\mathbb{Z}^3$.

Imaginons une particule de plancton se promenant de manière totalement aléatoire dans un océan immense. Pour simplifier, supposons qu'elle se déplace en sautant d'un point à l'autre de $\mathbb{Z}^3$ de la manière suivante : si $S_n \in \mathbb{Z}^3$ est sa position spatiale à l'instant n, alors $S_{n+1} = S_n + U_{n+1}$ où U_{n+1} est indépendante de $\{S_0, \ldots, S_n\}$ et suit une loi uniforme sur $\{(\pm1, \pm1, \pm1)\}$.

On suppose que $S_0 = (0,0,0)$.

1. ($\star$) Calculer $\mathbb{P}[S_n = (0,0,0)]$. Donner un équivalent de cette probabilité lorsque $n \to \infty$ (on se rappellera la formule de Stirling, $n! \sim e^{-n}n^n\sqrt{2\pi n}$).
2. ($\star$) Montrer qu'avec probabilité 1, la particule ne revient qu'un nombre fini de fois en $(0,0,0)$.
3. Pour $(x,y,z) \in \mathbb{Z}^3$, donner une majoration de $\mathbb{P}[S_n = (x,y,x)]$. Pour quel exposant α la suite $r_n = n^\alpha$ vérifie-t-elle $\sum_{n\geq 1} \mathbb{P}(|S_n| \leq r_n) < +\infty$.
4. En déduire que presque sûrement la particule « part à l'infini » et donner un minorant de sa vitesse de fuite.

EXERCICE 4 - Modèle AutoRégressif d'ordre 1 (AR(1)).

Soit $(Y_n)_{n\geq 1}$ une suite de v.a. i.i.d. de loi normale $\mathcal{N}(\mu, \sigma^2)$. Soient $a \in \mathbb{R}$ et X_0 une v.a. de loi normale $\mathcal{N}(\mu_0, \sigma_0^2)$, indépendante des $(Y_n)_n$. Pour tout $n \geq 1$, on définit la suite récurrente aléatoire $(X_n)_n$ par

$$X_n = aX_{n-1} + Y_n.$$

1. Montrer que $(X_0, \cdots, X_n)$ est un vecteur gaussien. Déterminer la loi de X_n et calculer $\mathrm{Cov}(X_k, X_n)$ pour $0 \leq k \leq n$.
2. A quelle condition sur a, la v.a. X_n converge-t-elle en loi ? Quelle est alors la loi limite ? Quelle est la loi de X_n si X_0 a la loi limite ?
3. Montrer que si $a \in]-1,1[$, le vecteur (X_n, X_{n+1}) converge en loi vers un vecteur gaussien dont on déterminera les paramètres.
4. Pour $a \in]-1,1[$, étudier la convergence en loi de la moyenne empirique $\frac{1}{n}\sum_{i=1}^n X_i$.

EXERCICE 5 - Suite récurrente aléatoire à deux états.

Soient $\alpha, \beta \in [0,1]$ et soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de loi uniforme sur $[0,1]$. Pour $X_0 \in \{0,1\}$ donnée, on définit $(X_n)_{n\geq 0}$ par

$$X_{n+1} = \mathbf{1}_{X_n=0}\mathbf{1}_{U_{n+1}\leq\alpha} + \mathbf{1}_{X_n=1}\mathbf{1}_{U_{n+1}\leq 1-\beta}, \qquad n \geq 0.$$

C'est une suite récurrente aléatoire au sens de la définition 7.4.1, à valeurs dans l'espace $E = \{0,1\}$. En particulier, elle vérifie la propriété de Markov (proposition 7.4.2) :

$$\mathbb{P}[X_{n+1} = x_{n+1}|X_n = x_n, \ldots, X_1 = x_1, X_0 = x_0] = \mathbb{P}[X_{n+1} = x_{n+1}|X_n = x_n]$$

pour tout $x_0, \ldots, x_{n+1}$.

1. Identifier les quantités $\mathbb{P}[X_{n+1} = x_{n+1}|X_n = x_n]$ à l'aide la matrice $Q := \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$.
2. Montrer par récurrence sur k que pour tout $n \in \mathbb{N}$ et $x_0, \ldots, x_k \in E$, on a
$$\mathbb{P}[X_n = x_0,\, X_{n+1} = x_1, \ldots,\, X_{n+k} = x_k] = \mathbb{P}[X_n = x_0]\, Q(x_0,x_1)Q(x_1,x_2)\cdots Q(x_{k-1},x_k).$$
En déduire que $\mathbb{P}[X_{n+k} = x_k|X_n = x_0] = [Q^k](x_0,x_k)$, où Q^k est la puissance k-ième de la matrice Q.
3. Que se passe-t-il dans le cas $\alpha = \beta = 0$? dans le cas $\alpha = \beta = 1$? On exclut dans la suite ces deux cas.
4. Montrer que les puissances de Q s'écrivent
$$Q^n = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} + \rho^n \begin{pmatrix} p & -p \\ p-1 & 1-p \end{pmatrix},$$
où $p = \alpha/(\alpha+\beta)$ et $\rho = 1-\alpha-\beta$.

5. En déduire que $\lim_{n\to\infty} Q^n(x,.) = \mu(.)$ pour tout x, où $\mu(\{0\}) = 1-p$ et $\mu(\{1\}) = p$. Quelle est la limite en loi de X_n, lorsque $n \to \infty$?
6. Vérifier que si X_0 est de loi μ, alors X_n est de loi μ pour tout n. On dira que μ est la *probabilité invariante* de la suite récurrente aléatoire.

EXERCICE 6 - Processus de Galton-Watson.

On considère une population évoluant de la façon suivante : à chaque génération, chaque individu donne naissance à un nombre aléatoire X de descendants suivant une loi géométrique sur $\mathbb{N}$ de paramètre $a = \frac{1}{1+m}$ avec $m > 0$ (i.e., $\mathbb{P}(X = k) = a(1-a)^k$). Les nombres de descendants de différents individus sont supposés indépendants. La génération suivante est alors constituée de tous ces nouveaux individus, sans leurs parents.

On note Z_n la taille de la population à la génération n et on suppose que $Z_0 = 1$.

1. 1) Soit $\phi_n(s) = \mathbb{E}(s^{Z_n})$ la fonction génératrice de Z_n. Montrer qu'il existe $\alpha > 0$ tel que
$$\frac{1}{1-\phi_1(s)} = 1 + \frac{\alpha}{1-s}.$$
2. Montrer que $\phi_n = \phi_1 \circ \phi_{n-1}$. En déduire l'expression de ϕ_n et calculer $\mathbb{E}[Z_n]$.
3. Soit $T = \inf\{n \geq 0 : Z_n = 0\}$ le *temps d'extinction de la population.* Calculer explicitement $\mathbb{P}(T \leq n)$.
Indication : on pourra relier le calcul à $\Phi_n(0)$.
4. Pour quelles valeurs de m, T est-il fini *p.s.* ? Que vaut alors $\mathbb{E}(T)$?
5. Dans le cas où $m > 1$, trouver la limite en loi de Z_n/m^n (*Indication : On pourra utiliser les fonctions caractéristiques*). Cette limite est-elle également en probabilité ? presque sûre ?