Math ECG 2. 2022-2023

Mathématiques Appliquées - F. Gaunard
http://frederic.gaunard.com
ENC Bessières, Paris 17e.

Concours Blanc n°2 - Math 2

Samedi 11 Mars
Durée : 4 heures

Le but de ce problème est d'étudier certaines quantités qui évoluent de manière aléatoire au cours du temps, mais dont l'évolution future est influencée par l'évolution passée. Il s'agit de ce qu'on appelle des *processus aléatoires renforcés*, qui possèdent de nombreuses applications, notamment en apprentissage automatique où les informations acquises au cours du temps orientent l'évolution future du processus d'apprentissage.

La première partie introduit et étudie quelques outils analytiques, la fonction gamma et la fonction beta. La deuxième partie de ce problème traite d'un cas où l'évolution du processus ne dépend que de la valeur précédente du processus. La troisième partie étudie un processus renforcé classique, l'urne de Pòlya, à la base de nombreuses applications. Les trois parties sont en gros indépendantes : dans la troisième partie, on utilise quelques résultats obtenus dans la première partie.

Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Pour une variable aléatoire X, on notera $E(X)$ son espérance et $\mathrm{Var}(X)$ sa variance lorsqu'elles existent.

Dans toutes les questions Python, on suppose les bibliothèques habituelles importées avec les alias usuels (np, rd, plt). Les pointillés peuvent éventuellement correspondre à plusieurs lignes d'instructions.

Première partie : Fonctions et lois gamma et beta.

Pour tout réel $a \geqslant 1$, on pose
$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$
La fonction est appelée *fonction gamma*.

1. On va montrer dans cette question que la fonction gamma est bien définie pour tout $a \geqslant 1$.

a. Soit $a \geqslant 1$ un réel fixé.

(i) Montrer que la fonction $x \mapsto x^{a-1}e^{x}$ est bien définie et continue sur $\mathbb{R}_+$.

(ii) En déduire que, pour tout $A > 0$, $\displaystyle\int_0^A x^{a-1}e^{-x}$ est une intégrale convergente.

b. Soit $a \geqslant 1$ un réel fixé.

(i) On définit la fonction h en posant $h(x) = x^{a-1}e^{-x/2}$ pour tout réel $x > 0$. Montrer que $\lim\limits_{x \to +\infty} h(x) = 0$.

(ii) Montrer qu'il existe un réel $A > 0$ tel que, pour tout $x \geqslant A$, $x^{a-1}e^{-x} \leqslant e^{-x/2}$.

(iii) En déduire que $\displaystyle\int_A^{+\infty} x^{a-1}e^{-x}dx$ est une intégrale convergente.

c. Conclure que pour tout $a \geqslant 1$, $\Gamma(a) = \displaystyle\int_0^{+\infty} x^{a-1}e^{-x}dx$ est bien définie.

2. **a.** Soit $a \geqslant 1$.

(i) Montrer que $\displaystyle\int_0^t x^a e^{-x}dx = -t^a e^{-t} + a\int_0^t x^{a-1}e^{-x}dx$ pour tout réel $t > 0$.

(ii) En déduire que pour tout $a \geqslant 1$, $\Gamma(a+1) = a\Gamma(a)$.

b. (i) Calculer $\Gamma(1)$.

(ii) Montrer que pour tout entier $n \geqslant 1$, on a $\Gamma(n) = (n-1)!$
(On rappelle que par convention, $0! = 1$.)

3. Soit $a \geqslant 1$ et soit g_a la fonction définie par $g_a(x) = \frac{1}{\Gamma(a)} x^{a-1}e^{-x}$ pour tout $x \geqslant 0$ et $g_a(x) = 0$ pour tout $x < 0$.

a. Vérifier que g_a est une densité de probabilité sur $\mathbb{R}$. La loi correspondante est dite *loi gamma de paramètre a.*

b. Soit X une variable aléatoire réelle de densité g_a.

(i) Calculer $E(X)$ en fonction de a.

(ii) Calculer $\mathrm{Var}(X)$ en fonction de a.

4. Soit $a \geqslant 1$ et $b \geqslant 1$ deux réels. On pose

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx.$$

a. Montrer que pour tout $t > 0$, on a $\displaystyle\int_0^t x^{a-1}(t-x)^{b-1}dx = B(a,b)t^{a+b-1}$.

b. Montrer que $B(a,b) = B(b,a)$.

c. Pour tout $a \geqslant 1$, calculer $B(a,1)$.

d. (i) Montrer que pour tout $a \geqslant 1$ et tout $b \geqslant 1$, $B(a,b+1) = \frac{b}{a}B(a+1,b)$.

(ii) Montrer par récurrence sur b que pour tous entiers naturels $a, b \in \mathbb{N}^*$, on a

$$B(a,b) = \frac{(b-1)!(a-1)!}{(a+b-1)!}.$$

On admettra que si X est une variable aléatoire de densité f et Y est une variable aléatoire de densité g et que X et Y sont indépendantes, alors la variable aléatoire $X+Y$ admet pour densité la fonction définie, pour tout $t \in \mathbb{R}$, par

$$h(t) = \int_{-\infty}^{+\infty} f(x)g(t-x)dx.$$

5. Soit $a \geqslant 1$ et $b \geqslant 1$ deux réels donnés et soit X une variable aléatoire de loi gamma de paramètre a (c'est-à-dire de densité g_a) et Y une variable aléatoire de loi gamma de paramètre b (c'est-à-dire de densité g_b), avec X et Y indépendantes. On note $h_{a,b}$ la densité de la variable aléatoire $X+Y$.

a. Montrer sans calculs que $h_{a,b}(t) = 0$ pour tout $t \leqslant 0$.

b. Montrer que pour tout $t > 0$, on a $h_{a,b}(t) = \frac{B(a,b)}{\Gamma(a)\Gamma(b)}t^{a+b-1}e^{-t}$.

c. Montrer que $\frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = 1$, c'est-à-dire que $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

d. Conclure que $X+Y$ est de loi gamma de paramètre $a+b$.

6. Soit $a \geqslant 1$ et $b \geqslant 1$ deux réels donnés. On définit la fonction $f_{a,b}$ en posant

$$f_{a,b}(x) = \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}$$

pour $x \in [0,1]$ et $f_{a,b}(x) = 0$ pour $x \notin [0,1]$.

a. Montrer que la fonction $f_{a,b}$ est une densité de probabilité sur $\mathbb{R}$. La loi correspondante est dite *loi beta de paramètres a et b.*

Soit Z une variable aléatoire de densité $f_{a,b}$.

b. Calculer $E(Z)$ en fonction de a et b.

c. Calculer $\mathrm{Var}(Z)$ en fonction de a et b.

Deuxième partie : Évolution de l'opinion d'un individu.

Dans cette partie, on étudie l'évolution de l'opinion (pour ou contre, modélisée par la valeur 0 ou 1) d'un individu à propos d'une question donnée. On considère la situation où l'individu change d'opinion d'un jour sur l'autre en fonction de circonstances aléatoires, avec une probabilité qui varie au cours du temps et qui peut notamment tendre vers 0 (si l'opinion tend à se figer au cours du temps).

Soit $(Z_i)_{i\geqslant 0}$ une suite de variables aléatoires indépendantes qui représentent les circonstances au temps successifs. On suppose que pour tout entier $i \geqslant 0$, Z_i suit une loi de Bernoulli de paramètre $p_i \in]0,1[$.

On considère une suite de variables aléatoires $(X_i)_{i\geqslant 0}$ (la variable aléatoire X_i représente l'opinion de l'individu au jour i) définie de la manière suivante :

- $X_0 = Z_0$.
- Pour tout $i \geqslant 1$, on pose $X_i = (1-Z_i)X_{i-1} + Z_i(1-X_{i-1})$.

Autrement dit, X_i est une variable aléatoire à valeurs 0 ou 1, égale soit à X_{i-1} dans le cas où $Z_i = 0$, soit à $1 - X_{i-1}$ dans le cas où $Z_i = 1$. Pour tout i, X_i suit donc une loi de Bernoulli et on note $\alpha_i = P(X_i = 1)$ son paramètre. On va d'abord chercher l'expression de ce paramètre.

7. **a.** Exprimer $E(X_i)$ en fonction de α_i.

b. Exprimer $\mathrm{Var}(X_i)$ en fonction de α_i.

c. (i) Montrer par récurrence que pour tout $i \geqslant 0$, la variable aléatoire X_i est une fonction de $(Z_0, \cdots, Z_i)$.

(ii) En déduire que pour tout $i \geqslant 1$, les variables aléatoires X_{i-1} et Z_i sont indépendantes.

d. Montrer que, pour tout $i \geqslant 1$, on a la relation de récurrence

$$\alpha_i = (1 - p_i)\alpha_{i-1} + p_i(1 - \alpha_{i-1}).$$

e. (i) Soit k un entier naturel. Montrer que si $\alpha_k = \frac{1}{2}$, on a $\alpha_{k+1} = \frac{1}{2}$.

(ii) On suppose **dans cette question** qu'il existe un entier $k \geqslant 0$ tels que $p_k = \frac{1}{2}$. Montrer que pour tout entier $i \geqslant k$, on a $\alpha_i = \frac{1}{2}$.

f. On suppose **dans cette question** que pour tout $i \geqslant 0$, on a $p_i = p$ où $p \in]0, 1[$ est un paramètre fixé tel que $p \neq \frac{1}{2}$.

(i) Montrer que pour tout $i \geqslant 1$, on a $\alpha_i - \frac{1}{2} = r\left(\alpha_{i-1} - \frac{1}{2}\right)$ avec $r = 1 - 2p$.

(ii) On pose, pour tout entier i, $\beta_i = \alpha_i - \frac{1}{2}$. Calculer β_i pour tout $i \geqslant 0$.

(iii) En déduire la valeur de α_i pour tout $i \geqslant 0$.

(iv) Discuter la valeur de $\lim\limits_{i \to +\infty} \alpha_i$ suivant les valeurs du paramètre p.

8. **a.** Écrire une fonction Python d'en-tête def simul_Z(p) : qui prend en argument un réel p et renvoie une simulation de Z_i (de paramètre $p = p_i$).

b. On suppose donnée une fonction p qui prend en argument un entier $i \geqslant 0$ et telle que p(i) renvoie p_i.
Recopier et compléter la fonction suivante qui prend en argument un entier $n \geqslant 0$ et renvoie une liste $[X_0, X_1, \cdots, X_n]$ correspondant à la réalisation de $(X_0, X_1, \cdots, X_n)$.

```
def Suite_X(n) :
  X=[Z(p(0))]
  for k in range(n) :
    ..........
  return X
```

9. Pour tout $i \geqslant 1$ on considère l'événement $A_i = [X_i = X_{i-1}]$.

a. Montrer que $(A_i)_{i \geqslant 1}$ forme une famille d'événements indépendants.

b. Déterminer $P(A_i)$.

c. Pour tous entiers $n \geqslant 0$ et $k \geqslant 1$, on pose l'événement $B_{n,k} = [X_i = X_n \text{ pour tout } n \leqslant i \leqslant n + k]$.

(i) Montrer que pour tout $n \geqslant 0$ et tout $k \geqslant 1$, on a $B_{n,k+1} \subset B_{n.k}$.

(ii) Montrer que $B_{n,k} = \bigcap\limits_{i=n+1}^{n+k} A_i$.

(iii) En déduire que $P(B_{n,k}) = \prod\limits_{i=n+1}^{n+k} (1 - p_i)$.

d. Pour tout entier $n \geqslant 0$, on considère l'événement $B_n = [X_i = X_n \text{ pour tout } i \geqslant n]$, qui peut être décrit comme "la suite $(X_i)_{i \geqslant 0}$ est constante à partir du rang n".

(i) Montrer que pour tout $n \geqslant 0$, on a $P(B_n) = \lim\limits_{k \to +\infty} P(B_{n,k})$.

(ii) Montrer que pour tout $n \geqslant 0$, on a $B_n \subset B_{n+1}$.

e. On considère l'événement $B = [\text{il existe un entier } n \text{ tel que pour tout } i \geqslant n, X_i = X_n]$. B est donc l'événement "la suite $(X_i)_{i\geqslant 0}$ est stationnaire" et on a $B = \bigcup_{n=1}^{+\infty} B_n$. Montrer que

$$P(B) = \lim_{n\to+\infty} P(B_n).$$

10. On suppose dans cette question que la série de terme général $(p_i)_{i\geqslant 0}$ est convergente.
On pose $x_i = -\ln(1-p_i)$.

a. Montrer que $\lim_{i\to+\infty} x_i = 0$.

b. Montrer que la série de terme général $(x_i)_{i\geqslant 0}$ est convergente.

c. En déduire que pour tout entier $n \geqslant 0$, on a $P(B_n) = \exp\left(-\sum_{i=n+1}^{+\infty} x_i\right)$.

d. Conclure que $P(B) = 1$.

Troisième partie : Urne de Pòlya.

On considère une urne, qui contient initialement r boules rouges et v boules vertes, où r et v sont deux entiers naturels non nuls donnés. On considère alors le processus suivant : *on tire une boule de l'urne, de manière aléatoire et uniforme, et on la remplace par deux boules de la même couleur.* On note R_n (respectivement V_n) le nombre de boules rouges (respectivement vertes) dans l'urne après n répétitions du processus.

Par exemple, si après n tirages, l'urne contient $R_n = 3$ boules rouges et $V_n = 5$ boules vertes, alors lors du $(n+1)$-ème tirage, on a une probabilité $\frac{3}{8}$ de tirer une boule rouge et $\frac{5}{8}$ de tirer une boule verte. Dans le cas où on tire une boule rouge, on la remplace par deux boules rouges de sorte que $R_{n+1} = R_n + 1$ (et $V_{n+1} = V_n$) ; dans le cas où on tire une boule verte, on la remplace par deux boules vertes, de sorte que $V_{n+1} = V_n + 1$ (et $R_{n+1} = R_n$).

11. **a.** Montrer que $R_0 + V_0 = r + v$.

b. Montrer que pour tout entier $n \geqslant 0$, $R_n + V_n = r + v + n$.

12. Pour tout $k \geqslant 1$, on note A_k l'événement "*la boule tirée lors de la k-ième répétition du processus est rouge*".

a. Montrer que pour tous entiers $n \geqslant 0$ et $i \in \{r, \cdots, r+n\}$, on a

$$P_{[R_n=i]}(A_{n+1}) = \frac{i}{r+v+n}.$$

b. Déduire que pour tout entier $n \geqslant 0$ et tout $i \in \{r, \cdots, r+n\}$, on a

$$P_{[R_n=i]}(R_{n+1} = i+1) = \frac{i}{r+v+n} \quad \text{et} \quad P_{[R_n=i]}(R_{n+1} = i) = \frac{r+v+n-i}{r+v+n}.$$

c. Recopier et compléter la fonction suivante qui prend en argument les deux entiers $r, v \geqslant 1$ et un entier $n \geqslant 0$ et renvoie une liste $[R_0, R_1, \cdots, R_n]$ correspondant à la réalisation de $(R_0, R_1, \cdots, R_n)$.

```
def Polya(r,v,n) :
  R=[r]
  for k in range(n) :
     ............
  return R
```

13. Soit $(s_n)_{n\geqslant 1}$ une suite d'entiers naturels croissante telle que $s_1 = 0$ ou $s_1 = 1$ et $s_{n+1} = s_n$ ou $s_{n+1} = s_n + 1$ [1].

1. Je trouve la définition de la suite $(s_n)_{n\geqslant 1}$ un peu ambiguë : j'ajouterais que la formule de récurrence peut changer avec n, il n'y a donc aucune raison que $s_{n+1} = s_n$ *pour tout* n ou que $s_{n+1} = s_n + 1$ *pour tout* n.

a. On veut montrer par récurrence

$$P((R_1 = r + s_1) \cap (R_2 = r + s_2) \cap \cdots \cap (R_n = r + s_n)) = \frac{(r + s_n - 1)!}{(r-1)!} \frac{(v + n - s_n - 1)!}{(v-1)!} \frac{(r + v - 1)!}{(r + v + n - 1)!}$$

(i) Montrer que la formule est vraie pour $n = 1$.

(ii) Montrer que

$$P_{[(R_1=r+s_1)\cap(R_2=r+s_2)\cap\cdots\cap(R_n=r+s_n)]}(R_{n+1} = r + s_{n+1}) = P_{[R_n=r+s_n]}(R_{n+1} = r + s_{n+1}).$$

(iii) Conclure

b. En déduire que $P((R_1 = r + s_1) \cap (R_2 = r + s_2) \cap \cdots \cap (R_n = r + s_n)) = \frac{B(r+s_n, v+n-s_n)}{B(r,v)}$, où $B(a, b)$ est défini dans la question **4.** On voit donc que cette quantité ne dépend pas des valeurs de $s_1, s_2, \cdots, s_{n-1}$.

c. (i) Montrer que pour tout $k \in \{0, \cdots, n\}$, il y a $\binom{n}{k}$ choix possibles de $s_1, s_2, \cdots, s_{n-1}$ tels que $s_n = k$.

(ii) Déduire que pour tout $k \in \{0, \cdots, n\}$ on a

$$P(R_n = r + k) = \binom{n}{k} \frac{B(r + k, v + n - k)}{B(r, v)}.$$

(iii) Conclure que

$$P(R_n = r + k) = \binom{n}{k} \int_0^1 x^k (1 - x)^{n-k} f_{r,v}(x) dx,$$

où $f_{r,v}(x) = \frac{1}{B(r,v)} x^{r-1}(1 - x)^{v-1}$ est la densité de la loi Beta de paramètres r et v définie dans la question **6.**

d. Montrer que pour tout entier $n \geqslant 1$ et tout $l \in \{r, \cdots, r + n\}$, on a

$$P(R_n \leqslant l) = \int_0^1 P(W_n^x \leqslant l - r) f_{r.v}(x) dx$$

où W_n^x est une variable aléatoire de loi binomiale de paramètres n et x.

Pour tout $x \in]0, 1[$ et tout $n \geqslant 1$, on considère dans le reste du problème W_n^x une variable aléatoire de loi binomiale de paramètres n et x.

14. **a.** Donner $E(W_n^x)$.

b. Montrer que $\text{Var}(W_n^x) \leqslant \frac{n}{4}$ pour tout $x \in]0, 1[$.

c. Montrer que pour tout $\varepsilon > 0$, on a $P(|W_n^x - nx| > \varepsilon) \leqslant \frac{n}{4\varepsilon^2}$.

d. Montrer que $P(W_n^x < nx - n^{2/3}) \leqslant \frac{1}{4} n^{-1/3}$ et que $P(W_n^x > nx + n^{2/3}) \leqslant \frac{1}{4} n^{-1/3}$.

e. En déduire que pour tout $m \in \{0, \cdots, n\}$, on a

$$P(W_n^x < m) \leqslant \frac{1}{4} n^{-1/3} \quad \text{pour tout } x \geqslant \frac{m}{n} + n^{-1/3},$$

$$P(W_n^x > m) \leqslant \frac{1}{4} n^{-1/3} \quad \text{pour tout } x \leqslant \frac{m}{n} - n^{-1/3}.$$

15. Soit $t \in]0, 1[$ et soit $(m_n)_{n \geqslant 1}$ une suite d'entiers naturels tels que $\lim_{n \to +\infty} \frac{m_n}{n} = t$.

a. Montrer que $\displaystyle\lim_{n \to +\infty} \int_{\frac{m_n}{n} + n^{-1/3}}^1 P(W_n^x \leqslant m_n) f_{r,v}(x) dx = 0.$

b. Montrer que $\displaystyle\lim_{n \to +\infty} \int_{\frac{m_n}{n} - n^{-1/3}}^{\frac{m_n}{n} + n^{-1/3}} P(W_n^x \leqslant m_n) f_{r,v}(x) dx = 0.$

c. Montrer que $\displaystyle\lim_{n \to +\infty} \int_0^{\frac{m_n}{n} - n^{-1/3}} P(W_n^x > m_n) f_{r,v}(x) dx = 0.$