



PRÉPASIMPLE

prepasimple.fr

HEC / ESSEC

Mathématiques — Option économique

2020

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Sujet

Partie I - Deux résultats généraux

On démontre dans cette partie deux résultats préliminaires, aux questions 5 et 6. Ces résultats seront utilisés dans la suite du sujet et pourront être admis.

Calcul d'une probabilité

Soient X et Y deux variables aléatoires sur un espace probabilisé, à densité et indépendantes. On note F_X et F_Y leurs fonctions de répartition. On suppose que Y est à valeurs positives et possède une densité f_Y dont la restriction à $[0, +\infty[$ est continue sur cet intervalle.

Pour tout $x \in \mathbb{R}_+$, on pose $H(x) = P([X \leq Y] \cap [Y \leq x])$.

1. a) Montrer que H est une fonction croissante sur $\mathbb{R}_+$ qui admet une limite finie en $+\infty$.

1. b) En utilisant la suite $(H(n))_{n \in \mathbb{N}}$, montrer que $\lim_{x \rightarrow +\infty} H(x) = P(X \leq Y)$. Que vaut $H(0)$?

Soit (u, v) un couple de réels positifs tels que $u < v$.

2. a) Montrer que $H(v) - H(u) = P([X \leq Y] \cap [u < Y \leq v])$, puis que :

$$F_X(u) \frac{F_Y(v) - F_Y(u)}{v - u} \leq \frac{H(v) - H(u)}{v - u} \leq F_X(v) \frac{F_Y(v) - F_Y(u)}{v - u}.$$

2. b) En déduire que, pour tout $x \in \mathbb{R}_+^*$, H est dérivable en x et $H'(x) = F_X(x)f_Y(x)$.

2. c) En conclure que, pour tout réel positif x , $H(x) = \int_0^x F_X(t)f_Y(t) dt$.

3. Montrer que $P(X \leq Y) = \int_0^{+\infty} F_X(t)f_Y(t) dt$.

4. En utilisant la fonction $K : x \mapsto P([X < Y] \cap [Y \leq x])$, on montrerait de même, et nous l'admettrons, que :

$$P(X < Y) = \int_0^{+\infty} F_X(t) f_Y(t) dt = P(X \leq Y).$$

Que peut-on en déduire pour $P(X = Y)$?

5. Application aux lois exponentielles

On suppose que U et V sont deux variables aléatoires indépendantes suivant des lois exponentielles de paramètres respectifs λ et μ , réels strictement positifs. Soit θ un réel positif ou nul.

5. a) Déterminer la fonction de répartition de la variable aléatoire $X = U - \theta$.

5. b) En déduire que, pour tout $\theta \geq 0$:

$$P(U - \theta \leq V) = 1 - \frac{\mu}{\lambda + \mu} e^{-\lambda\theta}.$$

Inégalité de Boole

6. On considère $(B_k)_{k \in \mathbb{N}^*}$ une famille d'événements d'un espace probabilisé.

6. a) Montrer par récurrence sur $n \in \mathbb{N}^*$ que

$$P\left(\bigcup_{k=1}^n B_k\right) \leq \sum_{k=1}^n P(B_k).$$

6. b) On suppose que la série

$$\sum_{k \geq 1} P(B_k)$$

converge. Montrer que

$$P\left(\bigcup_{k \geq 1} B_k\right) \leq \sum_{k=1}^{+\infty} P(B_k).$$

Partie II - Une compétition entre deux groupes

Dans toute la suite du sujet, $p \in]0, 1[$ et $q = 1 - p$.

On modélise une compétition entre deux groupes d'individus A et B avec les règles suivantes.

- Le groupe A doit résoudre une suite de problèmes $(P_k)_{k \geq 1}$ dans l'ordre des indices. Au temps $t = 0$, il commence P_1 , qui lui prend un temps représenté par X_1 . Une fois P_1 résolu, il aborde immédiatement P_2 , et X_2 est le temps consacré à sa résolution, et ainsi de suite. Pour tout $k \in \mathbb{N}^*$, X_k est le temps consacré à P_k par A .
- De même, B résout dans l'ordre une suite $(Q_k)_{k \geq 1}$; la résolution de Q_1 commence à $t = 0$, et Y_k est le temps consacré par B à Q_k .
- Sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ sont définies les suites $(X_k)_{k \geq 1}$ et $(Y_k)_{k \geq 1}$. Pour tout $k \in \mathbb{N}^*$, X_k suit la loi exponentielle $\mathcal{E}(p)$ et Y_k la loi exponentielle $\mathcal{E}(q)$. Pour tout $k \in \mathbb{N}^*$, les variables $X_1, \dots, X_k, Y_1, \dots, Y_k$ sont indépendantes.
- On établit la liste de tous les problèmes résolus dans l'ordre où ils le sont par les deux groupes. En cas de simultanéité temporelle de la résolution de deux problèmes, on place d'abord le problème résolu par A , puis celui résolu par B .

Pour tout $n \in \mathbb{N}^*$, U_n est la variable de Bernoulli associée à l'événement «le n -ième problème placé dans la liste est un problème résolu par A ». Par exemple, si le début de liste est $(P_1, P_2, Q_1, P_3, Q_2)$, alors $U_1 = 1, U_2 = 1, U_3 = 0, U_4 = 1$ et $U_5 = 0$.

Pour tout $n \geq 0$, S_n est le nombre de problèmes résolus par A présents dans la liste des n premiers problèmes résolus; en particulier $S_0 = 0$.

7. a) Que représente la variable aléatoire

$$\sum_{k=1}^n X_k$$

?

7. b) On suppose $X_1 = 5, X_2 = 2, X_3 = 3, X_4 = 2, Y_1 = 2, Y_2 = 2, Y_3 = 4, Y_4 = 2$.

Déterminer $U_1, \dots, U_7$. Peut-on aussi en déduire U_8 ?

7. c) Compléter le script Python suivant pour qu'il simule le jeu et, pour n et p donnés, affiche la liste des valeurs $U_1, \dots, U_n$.

La commande `rd.exponential(a)` simule une loi exponentielle d'espérance a , donc de paramètre $1/a$. `input` lit une chaîne saisie au clavier; `float` la convertit en réel, `int` en entier. `np.zeros(n, dtype=int)` crée un tableau de n zéros entiers. L'indexation Python commence à zéro; `print` affiche son argument.

Python

```
import numpy as np
import numpy.random as rd

p = float(input("p="))
n = int(input("n="))
q = 1 - p
U = np.zeros(n, dtype=int)
sommeX = rd.exponential(1 / p)
sommeY = rd.exponential(1 / q)
mini = min(sommeX, sommeY)
for k in range(n):
    if sommeX == ...:
        U[k] = ...
        sommeX = sommeX + rd.exponential(1 / p)
    else:
        sommeY = ...
```

```
mini = min(sommeX, sommeY)
```

...

7. d) Quelle(s) instruction(s) faut-il ajouter pour afficher la valeur de S_n ? La commande `np.sum(U)` renvoie la somme des éléments du tableau `U`.

8. Loi de U_n

Dans cette question, on démontre par récurrence sur $n \geq 1$ que $P(U_n = 1) = p$.

8. a) Montrer que $P(U_1 = 1) = P(X_1 \leq Y_1) = p$.

8. b) i. Montrer que, pour tout réel $x < 0$, $P_{[U_1=1]}(Y_1 - X_1 \leq x) = 0$.

8. b) ii. Soit $x \geq 0$. Établir que $P_{[U_1=1]}(Y_1 - X_1 \leq x) = \frac{1}{p}P(X_1 \leq Y_1 \leq X_1 + x)$, puis calculer cette probabilité conditionnelle.

8. c) On peut interpréter ce résultat en disant que la loi conditionnelle de $Y_1 - X_1$, sachant $[U_1 = 1]$, est une loi exponentielle. Quel est son paramètre ? Par analogie, quelle est la loi conditionnelle de $X_1 - Y_1$, sachant $[U_1 = 0]$? On n'attend pas une démonstration précise de cette dernière affirmation, mais un argument de bon sens pour justifier le résultat proposé.

8. d) On suppose $n \in \mathbb{N}^*$ et $P(U_n = 1) = p$. Dédurre de cette hypothèse et de la question précédente que $P_{[U_n=1]}(U_{n+1} = 1) = p$ et $P_{[U_n=0]}(U_{n+1} = 1) = p$.

On pourra préciser pourquoi, après chaque résolution, les délais jusqu'aux deux prochaines résolutions sont encore deux exponentielles indépendantes de paramètres p et q , conditionnellement aux groupes gagnants déjà observés.

8. e) Conclure.

9. On montrerait aussi par récurrence, et nous l'admettrons, que, pour tout $n \in \mathbb{N}^*$, les variables $U_1, \dots, U_n$ sont mutuellement indépendantes. En déduire la loi de S_n .

Soit $r \in \mathbb{N}$. On s'intéresse à la probabilité a_r de l'événement A_r : «il existe un entier $n \geq r$ tel que, lorsque n problèmes ont été résolus, le groupe A en a résolu r de plus que B ».

10. a) Justifier que $a_0 = 1$.

10. b) Montrer que, pour tout $r \geq 1$, $P_{[U_1=1]}(A_r) = P(A_{r-1})$ et $P_{[U_1=0]}(A_r) = P(A_{r+1})$.

10. c) En déduire que, pour tout $r \geq 1$, $a_{r+1} = \frac{1}{q}a_r - \frac{p}{q}a_{r-1}$.

10. d) En remarquant que $1 - 4pq = (1 - 2p)^2$, donner une expression de a_r en fonction de p, q et de deux constantes que l'on introduira.

11. Le cas $p \geq \frac{1}{2}$

Montrer que, dans les cas $p = \frac{1}{2}$ et $p > \frac{1}{2}$, la suite $(a_r)_{r \in \mathbb{N}}$ est constante égale à 1.

12. Le cas $p < \frac{1}{2}$

12. a) Soit $k \in \mathbb{N}$.

12. a) i. Établir que $A_{2k} = \bigcup_{i \geq k} [S_{2i} = i + k]$.

12. a) ii. Montrer que, pour tout $i \geq k$, $P(S_{2i} = i + k) = \binom{2i}{i+k} p^{i+k} q^{i-k}$.

12. a) iii. Après avoir donné la valeur de

$$\sum_{j=0}^{2i} \binom{2i}{j},$$

montrer que, pour tout entier $i \geq k$, $\binom{2i}{i+k} \leq 4^i$.

12. a) iv. En déduire que

$$\sum_{i=k}^{+\infty} P(S_{2i} = k + i) \leq \left(\frac{p}{q}\right)^k \frac{(4pq)^k}{1 - 4pq}.$$

12. b) Montrer, en utilisant l'inégalité de Boole de la question 6, que, si $p < \frac{1}{2}$,
 $\lim_{k \rightarrow +\infty} a_{2k} = 0$.

12. c) Conclure, en utilisant la question 10. d), que, si $p < \frac{1}{2}$, alors $a_r = (p/q)^r$ pour tout $r \in \mathbb{N}$.

On a ainsi établi dans les questions 11 et 12 que, pour tout $r \in \mathbb{N}$, $a_r = (p/q)^r$ si $p < 1/2$, et $a_r = 1$ si $p \geq 1/2$. Ce résultat pourra être admis et utilisé dans la suite du sujet.

Partie III - La blockchain et la stratégie de la double dépense

On utilise, dans cette partie, les notations et résultats de la partie II. Soit $n \geq 1$ un entier.

La blockchain est formée d'une suite de blocs, chacun associé à plusieurs transactions, qui contient l'historique de toutes les transactions effectuées depuis la création du bitcoin. Avant d'entrer dans la blockchain, un nouveau bloc doit être validé. Cette validation nécessite la mise en œuvre d'une très grande puissance de calcul pour résoudre un problème dépendant fortement du contenu du bloc et des blocs qui le précèdent. Les individus validant les blocs sont appelés mineurs.

Il est possible qu'à un instant donné coexistent sur le réseau deux blockchains valides et différentes. Dans ce cas, le réseau choisira celle qui comporte le plus de blocs et l'autre sera abandonnée.

Par prudence, lorsqu'un bloc est validé, il est recommandé d'attendre que $n - 1$ blocs le suivant soient aussi validés pour considérer que les transactions incluses dans le bloc soient honnêtes.

Un groupe de mineurs mal intentionnés, noté A , peut essayer de dépenser deux fois les mêmes bitcoins en procédant ainsi.

- Le groupe A demande la validation de l'achat d'un bien d'un montant de s bitcoins qu'il a en sa possession.
- Lorsque le bloc K incluant cette transaction est proposé à la validation sur le réseau, A modifie ce bloc en K' , qu'il ne diffuse pas, en remplaçant l'achat par une vente des s bitcoins en euros à son profit, par exemple. Il se met alors à la validation de ce nouveau bloc et crée ainsi une deuxième instance de la blockchain qu'il continue à développer sans la diffuser.
- Lorsque le groupe B , représentant l'ensemble des autres mineurs du réseau, a validé K ainsi que les $n - 1$ blocs suivants, le vendeur du bien considère que la transaction est valide et fournit le bien.
- Le groupe A attend alors d'avoir une blockchain plus longue que celle de B , qui est publique, pour la diffuser et donc invalider la blockchain publique et l'achat du bien. Le crédit en bitcoins du vendeur du bien est alors annulé.

On reprend et complète la modélisation précédente pour déterminer la probabilité que la stratégie de la double dépense réussisse et le choix de n pour que cette probabilité soit faible.

Une première phase du jeu de la partie II s'achève à l'instant aléatoire t où le problème Q_n est ajouté à la liste des problèmes résolus.

Le groupe A est déclaré vainqueur s'il se trouve un instant $t' \geq t$ où le nombre de problèmes résolus par A depuis le début du jeu est strictement supérieur au nombre de ceux résolus par B . On note G_n cet événement. On détermine la probabilité de G_n en fonction de n et p .

13. On s'intéresse d'abord à la loi de la variable aléatoire T_n , égale au nombre de problèmes résolus par A lorsque l'on place Q_n dans la liste.

13. a) Montrer que, pour tout $k \in \mathbb{N}$, $[T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0]$.

13. b) En déduire que $P(T_n = k) = \binom{n+k-1}{k} p^k q^n$.

14. a) En utilisant la formule des probabilités totales, établir que

$$P(G_n) = P(T_n \geq n+1) + \sum_{k=0}^n P(T_n = k) a_{n+1-k}.$$

14. b) Dans le cas $p \geq \frac{1}{2}$, en déduire que $P(G_n) = 1$.

14. c) Lorsque $p < \frac{1}{2}$, montrer de même que :

$$P(G_n) = 1 - \sum_{k=0}^n \binom{n+k-1}{k} (p^k q^n - p^{n+1} q^{k-1}).$$

15. Une meilleure expression de $P(G_n)$ lorsque $p < \frac{1}{2}$

Pour tout $x \in [0, 1]$ et $n \in \mathbb{N}^*$, on pose

$$u_n(x) = (1-x)^n \sum_{k=0}^n \binom{n+k-1}{k} x^k.$$

15. a) Vérifier que, pour tout $n \in \mathbb{N}^*$, $P(G_n) = 1 - u_n(p) + \frac{p}{q} u_n(q)$.

15. b) Pour tout $x \in [0, 1]$ et $n \in \mathbb{N}^*$, établir que :

$$u_{n+1}(x) = u_n(x) + (1-x)^n x^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} x \right).$$

15. c) En déduire que, pour tout $n \in \mathbb{N}^*$:

$$P(G_{n+1}) = P(G_n) - \left(1 - \frac{p}{q}\right) (pq)^{n+1} \binom{2n+1}{n+1}.$$

15. d) Montrer par récurrence que, pour tout $n \in \mathbb{N}^*$:

$$P(G_n) = \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^n \binom{2k-1}{k} (pq)^k.$$

16. Application à la sécurisation des transactions

Connaissant $p < \frac{1}{2}$, on cherche à limiter le risque de réussite de la stratégie mise en place par A.

16. a) Après avoir établi la formule $\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1}$ pour $k \in \llbracket 1, n \rrbracket$, écrire une fonction Python calculant les coefficients binomiaux.

L'opérateur Python `//` renvoie le quotient de la division entière ; on pourra l'utiliser lorsque le quotient est un entier.

16. b) Écrire un script Python qui détermine n_p , le plus petit entier $n \geq 1$ tel que $P(G_n) \leq \varepsilon$, pour $0 < p < \frac{1}{2}$ et $\varepsilon > 0$ saisis au clavier. On pourra utiliser la récurrence de 15. c).

Pour effectuer exactement les comparaisons au seuil avec des données décimales, on fournit `from fractions import Fraction: Fraction("0.1")` représente exactement $1/10$, et `Fraction(input("p="))` lit ainsi le nombre saisi. Ces objets s'utilisent avec les opérations usuelles d'addition, de soustraction, de multiplication, de division, de puissance entière et les comparaisons, en conservant des valeurs rationnelles exactes. La maîtrise de cette bibliothèque n'est pas exigible.

Pour $\varepsilon = 10^{-4} = 0,1$ pour mille et p variant de 10% à 32%, on obtient la représentation suivante de n_p en fonction de p :

