MATHÉMATIQUES I - ESSEC E 2018

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major-Prépa

Dans tout le sujet :

- On désigne par n un entier naturel, au moins égal à 2.
- X est une variable aléatoire à valeurs dans un intervalle $]0, \alpha[$, où α est un réel strictement positif. On suppose que X admet une densité f strictement positive et continue sur $]0, \alpha[$, et nulle en dehors de $]0, \alpha[$.
- On note F la fonction de répartition de X.
- $X_1, X_2, \ldots, X_n$ est une famille de variables aléatoires mutuellement indépendantes et de même loi que X.

On admet que toutes les variables aléatoires considérées sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

I - Lois des deux plus grands

On définit deux variables aléatoires Y_n et Z_n de la façon suivante.
Pour tout $\omega \in \Omega$:

- $Y_n(\omega) = \max\big(X_1(\omega), \ldots, X_n(\omega)\big)$ est le plus grand des réels $X_1(\omega), \ldots, X_n(\omega)$;
- $Z_n(\omega)$ est le "deuxième plus grand" des nombres $X_1(\omega), \ldots, X_n(\omega)$. Lorsque la plus grande valeur est présente plusieurs fois, $Z_n(\omega)$ et $Y_n(\omega)$ sont égaux.

1. Loi de Y_n.
   Soit G_n la fonction de répartition de Y_n.
   a) Pour tout réel x : l'événement $[Y_n \leqslant x]$ est réalisé si et seulement si les n variables aléatoires de l'échantillon dont on a pris le maximum, prennent elles-mêmes chacune une valeur inférieure ou égale à x, soit :
   $$\forall x \in \mathbb{R}, \quad [Y_n \leqslant x] = \bigcap_{i=1}^{n} [X_i \leqslant x]$$
   et par indépendance mutuelle des $(X_i)_{1 \leqslant i \leqslant n}$:
   $$\forall x \in \mathbb{R}, \quad G_n(x) = \mathbb{P}(Y_n \leqslant x) = \prod_{i=1}^{b} \mathbb{P}(X_i \leqslant x) = F(x)^n$$
   puisque les $(X_i)_{1 \leqslant i \leqslant n}$ suivent la même loi, celle de X.
   b) Comme X est une variable à densité : la fonction F est donc continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf peut-être en un nombre fini de points.
   Lorsqu'on multiplie n fois par elle-même cette fonction, ces caractéristiques sont conservées par produit : G_n est donc elle-même continue sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf peut-être en un nombre fini de points.
   On en déduit que Y_n est une variable à densité ; une densité g_n de Y_n est obtenue par dérivation de G_n sur $\mathbb{R}$ sauf peut-être en un nombre fini de points :
   $$G_n'(x) = n.F'(x).F(x)^{n-1} = n.f(x).F(x)^{n-1}$$
   est une formule valable partout où F est de classe $\mathcal{C}^1$, mais on peut en fait définir $g_n : \ x \mapsto n.f(x).F(x)^{n-1}$ sur tout $\mathbb{R}$ comme densité de Y_n.

c) D'après ce qui précède : la densité g_n de Y_n est, comme f, continue sur $]0;\alpha[$, nulle en dehors de $]0;\alpha[$, et :

$\forall x \in]0;\alpha[, \quad |x.g_n(x)| = n.x.f(x).F(x)^{n-1} \leqslant n.\alpha.f(x) \quad$ puisque $0 < x < \alpha$ et $0 \leqslant F(x) \leqslant 1$

(F est une fonction de répartition).

La convergence de l'intégrale $\displaystyle\int_0^\alpha f(x)\mathrm{d}x$ (qui vient du fait que f est une densité) assure alors, d'après le théorème de comparaison des fonctions positives et continues, que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x.g_n(x)\mathrm{d}x = \int_0^\alpha x.g_n(x)\mathrm{d}x$ est absolument convergente, donc que Y_n admet une espérance.

2. Loi de Z_n.

Soit H_n la fonction de répartition de Z_n.

a) Soit x un réel.

i. Soit $\omega \in \Omega$: $Z_n(\omega) \leqslant x$ si et seulement si le deuxième plus grand des n réels $X_1(\omega), X_2(\omega), \ldots, X_n(\omega)$ est inférieur ou égal à x : cela signifie bien que les $n-1$ valeurs les plus petites de l'échantillon sont inférieures ou égales à x. La plus grande peut aussi l'être, donc en effet : $Z_n(\omega) \leqslant x$ est vrai si et seulement si au moins $n-1$ des n éléments $X_1(\omega), \ldots, X_n(\omega)$, sont inférieurs ou égaux à x.

On en déduit que :

$$[Z_n \leqslant x] = \bigcup_{k=1}^{n} \Big(\bigcap_{i \neq k} [X_i \leqslant x] \cap [X_k > x] \Big) \cup \Big(\bigcap_{i=1}^{n} [X_i \leqslant x] \Big)$$

ii. L'union précédente est disjointe ; comme les $(X_i)_{1 \leqslant i \leqslant n}$ sont mutuellement indépendantes et de même loi :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}(Z_n \leqslant x) = \sum_{k=1}^{n} \prod_{i \neq k} \mathbb{P}(X_i \leqslant x) \times \mathbb{P}(X_k > x) + \prod_{i=1}^{n} \mathbb{P}(X_i \leqslant x)$$
$$= \sum_{k=1}^{n} F(x)^{n-1}.\big(1 - F(x)\big) + F(x)^n = n\big[1 - F(x)\big].\big[F(x)\big]^{n-1} + F(x)^n$$

b) Sous cette forme : la fonction H_n est continue sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points, puisque c'est le cas de F.

La variable aléatoire Z_n est donc une variable à densité, et une densité h_n de Z_n est définie sur $\mathbb{R}$, sauf peut-être en un nombre fini de points, par :

$$h_n(x) = H_n'(x) = n.\big[-F'(x)\big].\big[F(x)\big]^{n-1} + n\big[1 - F(x)\big].(n-1).F'(x).\big[F(x)\big]^{n-2} + n.F'(x).F(x)^{n-1}$$
$$= \big[-n.F(x) + n(n-1).\big(1 - F(x)\big) + n.F(x)\big].f(x).\big[F(x)\big]^{n-2}$$
$$h_n(x) = n(n-1).f(x).\big[1 - F(x)\big].\big[F(x)\big]^{n-2}$$

et cette formule peut être utilisée pour tout réel x pour définir une densité h_n de Y_n.

3. Simulation informatique.

On suppose que l'on a défini une fonction `Scilab` d'en-tête `function x = simulX(n)` qui retourne une simulation d'un échantillon de taille n de la loi de X sous la forme d'un vecteur de longueur n.

La fonction suivante généralise celle, plus classique de la recherche du plus grand élément d'un vecteur : les deux premiers éléments de la liste servent de valeur de référence, le premier test permet d'affecter le plus grand comme valeur initiale de `y`, et le second comme valeur initiale de `z`.

Puis chaque valeur suivante du vecteur `X` est examinée : si `X(k)` est supérieure à `y`, alors on a trouvé un nouveau maximum relatif : il devient la nouvelle valeur de `y`, et l'ancienne valeur de `y` devient le nouveau deuxième plus grand. Une autre possibilité est que `z` $<$ `X(k)` $\leqslant$ `y`, auquel cas `X(k)` devient le nouveau deuxième plus grand, le maximum restant inchangé. Sinon `X(k)` $\leqslant$ `z`, auquel cas rien ne change.

```
function [y,z] = DeuxPlusGrands(n)
    X = simulX(n)
    if X(1) > X(2) then
        y = X(1); z = X(2)
    else
        y = X(2); z = X(1)
    end
    for k = 3:n
        if X(k) > y
            z = y; y = X(k)
        else
            if X(k) > z then
                z = X(k)
            end
        end
    end
endfunction
```

4. Premier exemple : loi uniforme.

On suppose dans cette question que X suit la loi uniforme sur $]0;\alpha[$.

a) Une densité f de X est alors définie par : $\forall x \in \mathbb{R}, \begin{cases} \dfrac{1}{\alpha} & \text{si } 0 < x < \alpha \\ 0 & \text{sinon} \end{cases}$, et la fonction de répartition F est définie par : $\forall x \in \mathbb{R},\ F(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{x}{\alpha} & \text{si } 0 < x < \alpha \\ 1 & \text{si } x \geqslant \alpha \end{cases}$.

D'après les formules précédentes, une densité g_n de Y_n est alord définie par :

$$\forall x \in \mathbb{R}, \quad g_n(x) = n.f(x).\big[F(x)\big]^{n-1} = \begin{cases} \dfrac{n}{\alpha}.\Big(\dfrac{x}{\alpha}\Big)^{n-1} & \text{si } 0 < x < \alpha \\ 0 & \text{sinon} \end{cases}$$

Une densité h_n de Y_n est, elle, définie par :

$$\forall x \in \mathbb{R}, \quad h_n(x) = \begin{cases} \dfrac{n(n-1)}{\alpha}.(1 - \dfrac{x}{\alpha}).\Big(\dfrac{x}{\alpha}\Big)^{n-2} & \text{si } 0 < x < \alpha \\ 0 & \text{sinon} \end{cases}$$

b) La variable aléatoire Y_n admet, comme on l'a déjà prouvé, une espérance. Celle-ci est obtenue via la formule :

$$E(Y_n) = \int_{-\infty}^{+\infty} x.g_n(x)\mathrm{d}x = \frac{n}{\alpha^n}\int_0^\alpha x^n\mathrm{d}x = \frac{n}{\alpha^n}.\Big[\frac{x^{n+1}}{n+1}\Big]_0^\alpha = \frac{n.\alpha^{n+1}}{(n+1).\alpha^n}$$

soit : $$E(Y_n) = \frac{n\alpha}{n+1}$$

Il est clair que par une argumentation similaire à celle donnée pour Y_n, la variable aléatoire Z_n admet elle aussi une espérance définie par :

$$E(Z_n) = \int_{-\infty}^{+\infty} x.h_n(x)\mathrm{d}x = \frac{n(n-1)}{\alpha}\int_0^{\alpha} x\big(1-\frac{x}{\alpha}\big).\Big(\frac{x}{\alpha}\Big)^{n-2}\mathrm{d}x$$
$$= \frac{n(n-1)}{\alpha}.\Big(\frac{1}{\alpha^{n-2}}\int_0^{\alpha} x^{n-1}\mathrm{d}x - \frac{1}{\alpha^{n-1}}\int_0^{\alpha} x^{n}\mathrm{d}x\Big) = \frac{n(n-1)}{\alpha}.\Big(\frac{1}{\alpha^{n-2}}\Big[\frac{x^n}{n}\Big]_0^{\alpha} - \frac{1}{\alpha^{n-1}}\Big[\frac{x^{n+1}}{n+1}\Big]_0^{\alpha}\Big)$$
$$= \frac{n(n-1)}{\alpha}.\Big(\frac{\alpha^2}{n} - \frac{\alpha^2}{n+1}\Big) = \frac{n(n-1).\alpha^2}{\alpha n(n+1)} = \frac{(n-1).\alpha}{n+1}$$

5. Deuxième exemple : loi puissance.

On suppose dans cette question que la densité f est donnée par : $\quad f(x) = \begin{cases} \lambda\dfrac{x^{\lambda-1}}{\alpha^{\lambda}} & \text{si } x \in]0;\alpha[\\ 0 & \text{sinon} \end{cases}$

où λ est une constante strictement positive.
On dit que X suit la *loi puissance* de paramètres α et λ.

a) i. La fonction f est positive sur $\mathbb{R}$ tout entier ; elle est continue sur $]0,\alpha[$ puisqu'il s'agit, à un facteur constant près, d'une fonction de référence continue sur cet intervalle.
Elle est aussi continue et positive sur $]-\infty;0[$ et $]\alpha;+\infty[$ comme fonction constante nulle.
Enfin :
$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \frac{\lambda}{\alpha^{\lambda}}\int_0^{\alpha} x^{\lambda-1}\mathrm{d}x = \frac{\lambda}{\alpha^{\lambda}}.\Big[\frac{x^{\lambda}}{\lambda}\Big]_0^{\alpha} = \frac{\lambda}{\alpha^{\lambda}}.\frac{\alpha^{\lambda}}{\lambda} = 1$$
donc f est bien une densité de probabilité.

ii. La fonction de répartition F de X est définie par : $\quad \forall x \in \mathbb{R},\ F(x) = \displaystyle\int_{-\infty}^{x} f(t)\mathrm{d}t.$

On distingue alors plusieurs cas :

- Pour tout x de $]-\infty,0]$: $\quad F(x) = \displaystyle\int_{-\infty}^{x} 0\ \mathrm{d}t = 0$
- Pour tout x de $]0;\alpha[$: $\quad F(x) = \displaystyle\int_{-\infty}^{0} 0\ \mathrm{d}t + \int_0^{x} \lambda\frac{x^{\lambda-1}}{\alpha^{\lambda}}\mathrm{d}x = \frac{\lambda}{\alpha^{\lambda}}\Big[\frac{x^{\lambda}}{\lambda}\Big]_0^{x} = \Big(\frac{x}{\alpha}\Big)^{\lambda}$
- Pour tout x de $[\alpha;+\infty[$: $\quad F(x) = \displaystyle\int_{-\infty}^{0} 0\ \mathrm{d}t + \int_0^{\alpha} f(t)\mathrm{d}t + \int_0^{x} 0\ \mathrm{d}t = 1.$

iii. La densité f de X est nulle en-dehors de $]0;\alpha[$, et bornée sur $]0;\alpha[$; la variable aléatoire X admet donc une espérance donnée par :
$$E(X) = \int_{-\infty}^{+\infty} x.f(x)\mathrm{d}x = \frac{\lambda}{\alpha^{\lambda}}\int_0^{\alpha} x^{\lambda}\mathrm{d}x = \frac{\lambda}{\alpha^{\lambda}}.\Big[\frac{x^{\lambda+1}}{\lambda+1}\Big]_0^{\alpha} = \frac{\lambda}{\alpha^{\lambda}}.\frac{\alpha^{\lambda+1}}{\lambda+1} = \frac{\lambda\alpha}{\lambda+1}$$

b) i. La formule otenue en 1.b) permet de calculer l'expression de la densité g_n de Y_n dans cet exemple ; pour tout réel x :
$$g_n(x) = n.f(x).\big[F(x)\big]^{n-1} = \begin{cases} n.\lambda.\dfrac{x^{\lambda-1}}{\alpha^{\lambda}}.\Big(\dfrac{x}{\alpha}\Big)^{\lambda.(n-1)} & \text{si } x \in]0;\alpha[\\ 0 & \text{sinon} \end{cases} = \begin{cases} \lambda n.\dfrac{x^{\lambda n-1}}{\alpha^{\lambda n}} & \text{si } x \in]0;\alpha[\\ 0 & \text{sinon} \end{cases}$$
Au vu de la forme obtenue pour sa densité, on peut donc en conclure que Y_n suit une loi puissance de paramètres α et λn.

ii. La formule obtenue pour l'espérance de X peut alors être directement réutilisée pour obtenir celle de Y_n, en remplaçant λ par λn :
$$E(Y_n) = \frac{\lambda n\alpha}{\lambda n+1}$$

c) La variable Z_n admet de son côté une densité h_n définie cette fois, pour tout réel x, par :

$$\begin{aligned} h_n(x) &= n(n-1).f(x).\big[1-F(x)\big].\big[F(x)\big]^{n-2} \\ &= \begin{cases} n(n-1).\lambda.\dfrac{x^{\lambda-1}}{\alpha^{\lambda}}.\Big[1-\big(\dfrac{x}{\alpha}\big)^{\lambda}\Big].\Big(\dfrac{x}{\alpha}\Big)^{\lambda(n-2)} & \text{si } x\in]0;\alpha[\\ \quad 0 & \text{sinon} \end{cases} \\ &= \begin{cases} n(n-1).\lambda.\Big[\dfrac{x^{\lambda(n-1)-1}}{\alpha^{\lambda(n-1)}}-\dfrac{x^{\lambda n-1}}{\alpha^{\lambda n}}\Big] & \text{si } x\in]0;\alpha[\\ \quad 0 & \text{sinon} \end{cases} \end{aligned}$$

La fonction h_n est nulle en-dehors de $]0;\alpha[$, continue et bornée sur $]0;\alpha[$, donc Z_n admet une espérance donnée par :

$$\begin{aligned} E(Z_n) &= \int_{-\infty}^{+\infty} x.h_n(x)\mathrm{d}x = n(n-1)\lambda.\int_0^{\alpha}\Big(\frac{x^{\lambda(n-1)}}{\alpha^{\lambda(n-1)}}-\frac{x^{\lambda n}}{\alpha^{\lambda n}}\Big)\mathrm{d}x \\ &= n(n-1)\lambda.\Big(\frac{\alpha^{\lambda(n-1)+1}}{\alpha^{\lambda(n-1)}\big(\lambda(n-1)+1\big)}-\frac{\alpha^{\lambda n+1}}{\alpha^{\lambda n}(\lambda n+1)}\Big) \\ &= n(n-1)\lambda.\alpha.\Big(\frac{1}{\lambda(n-1)+1}-\frac{1}{\lambda n+1}\Big) = n(n-1)\lambda.\alpha.\frac{\lambda n+1-\lambda(n-1)-1}{\big(\lambda(n-1)+1\big)(\lambda n+1)} \end{aligned}$$

$$E(Z_n) = \frac{n(n-1)\lambda^2.\alpha}{\big(\lambda(n-1)+1\big)(\lambda n+1)}$$

Remarque : la loi uniforme sur $]0;\alpha[$ correspondant à la loi puissance de paramètres α et $\lambda=1$, on vérifie que cette formule est cohérente dans le cas particulier étudié à la fin de la question 4. (on retrouve bien $\dfrac{(n-1)\alpha}{n+1}$ pour l'espérance quand $\lambda=1$).

II - Un problème d'optimisation

On répond dans cette partie au problème d'optimisation suivant : trouver une fonction σ définie sur $]0;\alpha[$ et vérifiant les trois propriétés :

- σ est une bijection de $]0;\alpha[$ dans un intervalle $]0;\beta[$, avec β un réel strictement positif.
- σ est de classe $\mathcal{C}^1$ sur $]0;\alpha[$ et σ' est à valeurs strictement positives sur $]0;\alpha[$.
- On définit, pour tout $x\in]0;\alpha[$ et tout $y\in]0;\beta[$:

$$\gamma(x,y) = (x-y)G_{n-1}\big(\sigma^{-1}(y)\big)$$

Alors pour tout $x\in]0;\alpha[$, $\gamma(x,y)$ atteint son maximum lorsque $y=\sigma(x)$.

6. Analyse.

On suppose dans un premier temps qu'une telle fonction σ vérifiant ces trois propriétés existe.

a) C'est une question de cours : la fonction σ étant bijective, de classe $\mathcal{C}^1$ de $]0;\alpha[$ dans $]0;\beta[$, et de dérivée σ' strictement positive sur $]0;\alpha[$, la bijection réciproque σ^{-1} est de classe $\mathcal{C}^1$ sur $]0;\beta[$, et d'après la formule du cours pour la dérivée d'une bijection réciproque :

$$\forall x\in]0;\beta[,\quad \big(\sigma^{-1}\big)'(x) = \frac{1}{\sigma'\big(\sigma^{-1}(x)\big)}.$$

b) D'après ce qui précède, et au vu de son expression, la fonction de deux variables γ admet une dérivée partielle par rapport à la seconde variable définie par :

$$\forall x\in]0;\alpha[,\ \forall y\in]0;\beta[,\quad \partial_2(\gamma)(x,y) = -G_{n-1}\big(\sigma^{-1}(y)\big)+(x-y).\big(\sigma^{-1}\big)'(y).G'_{n-1}\big(\sigma^{-1}(y)\big)$$

$$= -G_{n-1}\big(\sigma^{-1}(y)\big) + \frac{x-y}{\sigma'\big(\sigma^{-1}(y)\big)}.g_{n-1}\big(\sigma^{-1}(y)\big)$$

c) La troisième hypothèse faite sur σ affirme que la fonction $k:\ y \mapsto (x-y).G_{n-1}\big(\sigma^{-1}(y)\big)$, atteint un maximum en $y=\sigma(x)$.

Comme σ^{-1} est, ainsi qu'on vient de le voir, de classe $\mathcal{C}^1$ sur $]0;\beta[$, à valeurs dans $]0;\alpha[$ sur lequel G_{n-1} est de classe $\mathcal{C}^1$: la fonction k est de classe $\mathcal{C}^1$ sur $]0;\beta[$, et comme elle atteint sur cet intervalle, un maximum en $y=\sigma(x)$, alors effectivement :

$$k'\big(\sigma(x)\big)=0 \iff \partial_2(\gamma)\big(x,\sigma(x)\big)=0$$

Or pour tout $x\in]0;\alpha[$, si on calcule explicitement la dérivée partielle d'ordre 2 au point $\big(x,\sigma(x)\big)$:

$$\partial_2(\gamma)\big(x,\sigma(x)\big) = -G_{n-1}\big(\sigma^{-1}(\sigma(x))\big) + \frac{x-\sigma(x)}{\sigma'\big(\sigma^{-1}(\sigma(x))\big)}.g_{n-1}\big(\sigma^{-1}(\sigma(x))\big)$$

$$= -G_{n-1}(x) + \frac{x-\sigma(x)}{\sigma'(x)}.g_{n-1}(x)$$

Et ainsi :

$$\partial_2(\gamma)\big(x,\sigma(x)\big)=0 \iff \frac{x-\sigma(x)}{\sigma'(x)}.g_{n-1}(x) = G_{n-1}(x)$$

$$\iff x.g_{n-1}(x) - \sigma(x).g_{n-1}(x) = \sigma'(x).G_{n-1}(x)$$

$$\iff \sigma'(x).G_{n-1}(x) + \sigma(x).g_{n-1}(x) = xg_{n-1}(x)$$

d) On reconnaît dans le membre de gauche, une formule de dérivation de la fonction produit $x\mapsto \sigma(x).G_{n-1}(x)$ qui est de classe $\mathcal{C}^1$ sur $]0;\alpha[$, de sorte que pour tout x de $]0;\alpha[$:

$$\int_0^x t.g_{n-1}(t)\mathrm{d}t = \int_0^x \big(\sigma'(t).G_{n-1}(t)+\sigma(t).g_{n-1}(t)\big)\mathrm{d}t = \int_0^x \big(\sigma.G_{n-1}\big)'(t)\mathrm{d}t$$

$$= \Big[\sigma(t).G_{n-1}(t)\Big]_0^x = \sigma(x).G_{n-1}(x)$$

puisque $G_{n-1}(0)$ est nulle. Par contre, pour tout $x\in]0;\alpha[,\ G_{n-1}(x)=F(x)^{n-1}>0$, ce qui permet effectivement de conclure :

$$\forall x\in]0;\alpha[,\quad \sigma(x) = \frac{1}{G_{n-1}(x)}\int_0^x tg_{n-1}(t)\mathrm{d}t \qquad (\star)$$

e) Dans l'intégrale précédente, on réalise une intégration par parties en posant :

$$u(t)=t \qquad\longrightarrow\quad u'(t)=1$$
$$v'(t)=g_{n-1}(t) \quad\longrightarrow\quad v(t)=G_{n-1}(t)$$

Les fonctions concernées sont continues et bornées sur $]0;\alpha[$, donc par intégration par parties (les intégrales sont faussement impropres en 0) : pour tout $x\in]0;\alpha[$,

$$\sigma(x) = \frac{1}{G_{n-1}(x)}.\left(\Big[t.G_{n-1}(t)\Big]_0^x - \int_0^x G_{n-1}(t)\mathrm{d}t\right) = \frac{x.G_{n-1}(x)}{G_{n-1}(x)} - \int_0^x \frac{G_{n-1}(t)}{G_{n-1}(x)}\mathrm{d}t$$

soit : $$\sigma(x) = x - \int_0^x \frac{G_{n-1}(t)}{G_{n-1}(x)}\mathrm{d}x \qquad (\star\star)$$

7. Synthèse.

On suppose à présent que σ est la fonction définie par l'une des deux égalités ($\star$) ou ($\star\star$).

a) Pour tout réel $x\in]0;\alpha[$ fixé : la fonction $t\mapsto\dfrac{G_{n-1}(t)}{G_{n-1}(x)}$ est bien définie et continue sur $[0;x]$; comme de plus la fonction de répartition G_{n-1} est croissante, et même strictement croissante sur $]0;\alpha[$ (puisque F l'est, voir le préambule du problème), on peut écrire :

$$\forall t\in]0;x[,\quad 0<G_{n-1}(t)<G_{n-1}(x)\iff 0<\frac{G_{n-1}(t)}{G_{n-1}(x)}<1.$$

Par continuité de la fonction évoquée sur $[0;x]$, et par strictes croissance et positivité de l'intégrale ($0<x$), on en déduit que :

$$\forall x\in]0;\alpha[,\quad 0<\int_0^x\frac{G_{n-1}(t)}{G_{n-1}(x)}\mathrm{d}t<\int_0^x 1\ \mathrm{d}t\iff 0<\int_0^x\frac{G_{n-1}(t)}{G_{n-1}(x)}\mathrm{d}t<x$$

ce qui donne bien :

$$\forall x\in]0;\alpha[,\quad 0<x-\int_0^x\frac{G_{n-1}(t)}{G_{n-1}(x)}\mathrm{d}t<x\iff 0<\sigma(x)<x$$

b) Sous sa forme ($\star$) : la fonction σ apparaît comme le produit de deux fonctions de classe $\mathcal{C}^1$ sur $]0;\alpha[$: la fonction $x\mapsto\dfrac{1}{G_{n-1}(x)}$ d'abord, bien de classe $\mathcal{C}^1$ sur $]0;\alpha[$ puisque G_{n-1} l'est, et la fonction $x\mapsto\displaystyle\int_0^x t.g_{n-1}(t)\mathrm{d}t$, bien de classe $\mathcal{C}^1$ sur $]0;\alpha[$ puisque la fonction $t\mapsto t.g_{n-1}(t)$ est continue sur cet intervalle.

D'après les règles de dérivation, on en déduit :

$$\forall x\in]0;\alpha[,\quad \sigma'(x)=-\frac{g_{n-1}(x)}{\big[G_{n-1}(x)\big]^2}.\int_0^x t.g_{n-1}(t)\mathrm{d}t+\frac{1}{G_{n-1}(x)}.xg_{n-1}(x)$$

$$=-\frac{g_{n-1}(x)}{G_{n-1}(x)}.\sigma(x)+\frac{xg_{n-1}(x)}{G_{n-1}(x)}$$

$$\forall x\in]0;\alpha[,\quad \sigma'(x)=\frac{g_{n-1}(x).\big(x-\sigma(x)\big)}{G_{n-1}(x)}$$

Comme les fonctions g_{n-1} et G_{n-1} sont strictement positives sur $]0;\alpha[$, on en déduit bien que le signe de $\sigma'(x)$ est celui de $x-\sigma(x)$.

Or comme on l'a démontré à la question précédente :

$$\forall x\in]0;\alpha[,\quad 0<\sigma(x)<x\Longrightarrow x-\sigma(x)>0\Longrightarrow\sigma'(x)>0.$$

La fonction σ' est donc strictement positive sur $]0;\alpha[$.

c) La fonction σ est ainsi de classe $\mathcal{C}^1$, donc continue sur $]0;\alpha[$, strictement croissante sur cet intervalle : d'après le théorème éponyme, elle réalise une bijection de $]0;\alpha[$ dans $\Big]\lim\limits_{x\to0^+}\sigma(x);\lim\ \lim\limits_{x\to\alpha^-}\sigma(x)\Big[$.

L'inégalité : $\forall x\in]0;\alpha[,\ 0<\sigma(x)<x$ garantit d'après le théorème d'encadrement, que $\lim\limits_{x\to0^+}\sigma(x)=0$.

La relation ($\star$) donne : $\displaystyle\lim_{x\to\alpha^-}\sigma(x)=\frac{1}{G_{n-1}(\alpha)}\int_0^\alpha t.g_{n-1}(t)\mathrm{d}t.$

Or $G_{n-1}(\alpha)=F(\alpha)^{n-1}=1$ puisque $X(\Omega)=]0;\alpha[$, et pour la même raison :

$$\int_0^\alpha t.g_{n-1}(t)\mathrm{d}t=\int_{-\infty}^{+\infty}t.g_{n-1}(t)\mathrm{d}t=\mathbb{E}(Y_{n-1}).$$

La fonction σ réalise bien une bijection de $]0;\alpha[$ dans $]0;\beta[$, où $\beta=\mathbb{E}(Y_{n-1})$.

d) On fixe un réel $x \in]0;\alpha[$. Soit $y \in]0;\beta[$, on pose $z = \sigma^{-1}(y)$.

i. Par définition :

$$\begin{aligned}\gamma(x,y) &= (x-y)G_{n-1}\big(\sigma^{-1}(y)\big) = \big(x-\sigma(z)\big)G_{n-1}(z)\\ &= \Big(x - z + \int_0^z \frac{G_{n-1}(t)}{G_{n-1}(z)}\mathrm{d}t\Big).G_{n-1}(z) \qquad \text{d'après } (\star\star)\\ &= (x-z).G_{n-1}(z) + \int_0^z G_{n-1}(t)\mathrm{d}t\end{aligned}$$

ii. Toujours grâce à $(\star\star)$:

$$\gamma\big(x,\sigma(x)\big) = \underbrace{\big(x-\sigma^{-1}(\sigma(x))\big)}_{=x-x=0} G_{n-1}\big(\sigma^{-1}(\sigma(x))\big) + \int_0^{\sigma^{-1}(\sigma(x))} G_{n-1}(t)\mathrm{d}t = \int_0^x G_{n-1}(t)\mathrm{d}t$$

donc :

$$\gamma\big(x,\sigma(x)\big) - \gamma(x,y) = \int_0^x G_{n-1}(t)\mathrm{d}t - (x-z).G_{n-1}(z) - \int_0^z G_{n-1}(t)\mathrm{d}t = (z-x).G_{n-1}(z) - \int_x^z G_{n-1}(t)\mathrm{d}t$$

d'après la relation de Chasles pour les intégrales.

iii. On détermine le signe de l'expression précédente en distinguant deux cas :

- Si $z \geqslant x$: la croissance de G_{n-1} sur $]0;\alpha[$ donne : $\forall t \in [x,z],\ G_{n-1}(t) \leqslant G_{n-1}(z)$, donc d'après l'inégalité de la moyenne :

$$\int_x^z G_{n-1}(t)\mathrm{d}t \leqslant (z-x).G_{n-1}(z) \iff (z-x)G_{n-1}(z) - \int_x^z G_{n-1}(t)\mathrm{d}t \geqslant 0$$

- Si $z \leqslant x$: la croissance de G_{n-1} a lieu cette fois sur l'intervalle $[z;x]$, où l'inégalité de la moyenne donne cette fois la minoration :

$$\int_z^x G_{n-1}(t)\mathrm{d}t \geqslant (x-z).G_{n-1}(z) \iff (z-x).G_{n-1}(z) - \int_x^z G_{n-1}(t)\mathrm{d}t \geqslant 0$$

On en déduit que : pour tout y de $]0;\beta[$, $\quad \gamma\big(x,\sigma(x)\big) - \gamma(x,y) \geqslant 0$, ce qui prouve bien qu'à x fixé, $\gamma(x,y)$ est maximal lorsque $y = \sigma(x)$.

8. Estimation de $\sigma(x)$.

Soit $x \in]0;\alpha[$.

a) On considère la fonction φ_x définie sur $\mathbb{R}^+$ par : $\quad \varphi_x(t) = \begin{cases} t & \text{si } t \leqslant x \\ 0 & \text{sinon}\end{cases}$.

La fonction densité g_{n-1} étant nulle en-dehors de $]0;\alpha[$, et la fonction φ_x étant nulle en-dehors de $[0;x]$ qui est inclus dans $[0;\alpha[$, le théorème de transfert pour les variables à densité garantit l'existence de l'espérance $\mathbb{E}\big(\varphi_x(Y_{n-1})\big)$, qui vaut :

$$\mathbb{E}\big(\varphi_x(Y_{n-1})\big) = \int_0^\alpha \varphi_x(t).g_{n-1}(t)\mathrm{d}t = \int_0^x t.g_{n-1}(t)\mathrm{d}t + \int_x^\alpha 0.g_{n-1}(t)\mathrm{d}t = \int_0^x t.g_{n-1}(t)\mathrm{d}t$$

Et ainsi : $\dfrac{\mathbb{E}\big(\varphi_x(Y_{n-1})\big)}{\mathbb{P}(Y_{n-1} \leqslant x)} = \dfrac{1}{G_{n-1}(x)} \displaystyle\int_0^x t.g_{n-1}(t)\mathrm{d}t = \sigma(x)$.

b) On retrouve ici les deux modes d'estimations de paramètres associés aux probabilités : l'espérance $\mathbb{E}\big(\varphi_x(Y_{n-1})\big)$ est estimée par la moyenne des valeurs prises par un échantillon de simulations indépendantes de la variable aléatoire $\varphi_x(Y_{n-1})$. La probabilité $\mathbb{P}(Y_{n-1} \leqslant x)$, elle, est estimée par la fréquence de réalisation de cet événement lors de simulations sucessives :

```
function s = sigma(x,n)
    E = 0; A = 0;
    for i=1:10000
        X = simulX(n)
        Y = max(X)
        if Y <= x then
            E = E+Y
            A = A+1
        end
    end
endfunction
```

9. Exemples.

   a) Lorsque X suit la loi uniforme sur $]0;\alpha[$, comme on l'a vu plus haut :

   $\forall x \in]0;\alpha[, \quad G_{n-1}(x) = \left(\dfrac{x}{\alpha}\right)^{n-1}$ et $g_{n-1}(x) = \dfrac{n-1}{\alpha}\left(\dfrac{x}{\alpha}\right)^{n-2}$, les deux fonctions étant nulles en-dehors de cet intervalle.

   La formule $(\star)$ donne alors :

$$\forall x \in]0;\alpha[, \quad \sigma(x) = \frac{1}{G_{n-1}(x)}\int_0^x t.g_{n-1}(t)\mathrm{d}t = \left(\frac{\alpha}{x}\right)^{n-1}.\frac{n-1}{\alpha^{n-1}}\int_0^x t^{n-2}\mathrm{d}x = \frac{n-1}{x^{n-1}}.\left[\frac{t^{n-1}}{n-1}\right]_0^x$$

$$\forall x \in]0;\alpha[, \quad \sigma(x) = 1$$

   b) On suppose ici que X suit la loi puissance de paramètres α et λ. Pour tout réel $x \in]0;\alpha[$:

$$\sigma(x) = \frac{1}{G_{n-1}(x)}\int_0^x t.g_{n-1}(t)\mathrm{d}t = \left(\frac{\alpha}{x}\right)^{\lambda.(n-1)}\int_0^x t.\lambda(n-1).\frac{t^{\lambda(n-1)-1}}{\alpha^{\lambda(n-1)}}\mathrm{d}t$$

$$= \frac{\lambda(n-1)}{x^{\lambda(n-1)}}\int_0^x t^{\lambda(n-1)}\mathrm{d}t = \frac{\lambda(n-1)}{x^{\lambda(n-1)}}.\left[\frac{t^{\lambda(n-1)+1}}{\lambda(n-1)+1}\right]_0^x = \frac{\lambda(n-1).x^{\lambda(n-1)+1}}{x^{\lambda(n-1)}\big(\lambda(n-1)+1\big)}$$

$$\forall x \in]0;\alpha[, \quad \sigma(x) = \frac{\lambda(n-1)}{\lambda(n-1)+1}.x$$

   On a donc obtenu une fonction linéaire ; lorsque $n=6$, $\lambda = 0.2$ et $\alpha = 50$, la formule devient :

$$\forall x \in]0;50[, \quad \sigma(x) = \frac{5\times 0.2}{5\times 0.2+1}.x = \frac{x}{2}$$

   ce qui est bien confirmé par le graphique proposé qui a l'allure d'une fonction linéaire telle que $\sigma(50) \approx 25$, donc de coefficient directeur $\dfrac{1}{2}$.

III. Modélisation d'enchères

III.A - Enchère au premier prix

On suppose que chaque acheteur A_k a une valeur privée $x_k = X_k(\omega)$ qui est une réalisation de la variable aléatoire X_k.

On se met à la place de l'acheteur n, et on suppose que les $n-1$ premiers acheteurs appliquent la stratégie σ, c'est-à-dire : pour tout $k \in \{1,\ldots,n-1\}$, l'acheteur k mise $\sigma(X_k)$.

L'acheteur n a une mise privée x_n et choisit une mise y_n.

Onnote E_n l'événement « l'acheteur A_n remporte l'enchère ».

10. Il est clair que l'acheteur n ne peut pas remporter l'enchère si sa mise x_n est inférieure à la mise la plus forte parmi celles des $n-1$ premiers acheteurs ; la fonction σ étant strictement croissante sur son domaine $]0;\alpha[$, on peut écrire :

$$E_n = \left[\max\left(\sigma(X_1),\ldots,\sigma(X_{n-1})\right) \leqslant y_n\right] = \left[\sigma\left(\max(X_1,\ldots,X_{n-1})\right) \leqslant y_n\right] = \left[\sigma(Y_{n-1}) \leqslant y_n\right]$$

Et la bijectivité de σ donne également :

$$E_n = \left[Y_{n-1} \leqslant \sigma^{-1}(y_n)\right]$$

Comme Y_{n-1} est une variable à densité : $\mathbb{P}\left(Y_{n-1} = \sigma^{-1}(y_n)\right) = 0$, donc en effet :

$$\mathbb{P}(E_n) = \mathbb{P}\left(Y_{n-1} < \sigma^{-1}(y_n)\right).$$

On note R_n la variable aléatoire donnant le résultat net de l'enchère pour l'acheteur A_n.

Il n'y a que deux possibilités : soit l'échateur n remporte l'enchère (E_n est réalisé), et dans ce cas le résultat net est égal à $(x_n - y_n)$, soit il ne remporte pas l'enchère, et dans ce cas le résultat net est égal à 0.

La fonction indicatrice $\mathbb{1}_{E_n}$ étant une variable aléatoire définie par :

$$\forall \omega \in \Omega,\ \mathbb{1}_{E_n}(\omega) = \begin{cases} 1 & \text{si } \omega \text{ réalise } E_n \\ 0 & \text{sinon} \end{cases}, \quad \text{on a bien :} \quad R_n = (x_n - y_n)\mathbb{1}_{E_n}$$

Et le passage à l'espérance donne :

$$\mathbb{E}(R_n) = (x_n - y_n)\mathbb{E}\left(\mathbb{1}_{E_n}\right) = (x_n - y_n).\mathbb{P}(E_n) = (x_n - y_n).\mathbb{P}\left(Y_{n-1} < \sigma^{-1}(y_n)\right) = (x_n - y_n).G_{n-1}\left(\sigma^{-1}(y_n)\right)$$

11. On reconnaît dans la relation précédente que : $\mathbb{E}(R_n) = \gamma(x_n, y_n)$.
L'étude menée dans la partie II que le maximumn de $\mathbb{E}(R_n)$ est alors obtenu lorsque $y_n = \sigma(x_n)$: pour maximiser son espérance de résultat, l'acheteur doit choisir pour mise $y_n = \sigma(x_n)$, donc à appliquer lui aussi la stratégie σ.

III.B - Enchère au second prix

On se met à nouveau à la place de l'acheteur n. Soit $m = \max(y_1, \ldots, y_{n-1})$ la meileure offre faite par les acheteurs $A_1, \ldots, A_{n-1}$ (que A_n ne connaît pas).

12. a) Si on suppose que $m \geqslant x_n$, il y n'y a là encore que deux options :
    - Soit $y_n \geqslant m$, auquel cas l'acheteur A_n remporte l'enchère et paie le prix m qui est forcément la deuxième meilleure enchère vu sa définition, pour un résultat net égal à $x_n - m \leqslant 0$.
    - Soit $y_n \leqslant m$ et y_n est au mieux la deuxième meilleure enchère juste après m : dans ce cas l'acheteur A_n perd l'enchère, et son résultat net est nul.

      Ici, c'est notamment le cas lorsque $y_n = x_n$: le résultat net est nul.

    b) Si on suppose que $m < x_n$, on examine encore les deux options possibles :
    - Si $y_n < m$, l'acheteur A_n perd l'enchère et son résultat net est nul.
    - Si $y_n \geqslant m$, alors la mise y_n est supérieure aux $(n-1)$ autres mises, et l'acheteur A_n remporte l'enchère ; il paie alors le prix m qui est la deuxième meilleure offre, pour un résultat net égal à $x_n - m \geqslant 0$ cette fois.

    c) Au vu des deux résultats précédents, on constate donc que : si $x_n \leqslant m$ il faut, pour optimiser le résultat net, avoir choisi $y_n \leqslant m$ aussi, auquel cas le résultat net nul, ne dépend pas du choix de $y_n \leqslant m$; et si $x_n > m$ il faut prendre $y_n > m$ également pour obtenir un résultat net optimal, lequel ne dépend pas alors de la valeur de y_n : le choix $y_n = x_n$ constitue bien, dans tous les cas, une stratégie optimale.

III.C - Équivalence des revenus

On se met maintenant à la place du vendeur.

Les valeurs privées des acheteurs sont données par les variables aléatoires $X_1, \dots, X_n$.

13. Enchère au premier prix.

    On suppose que le vendeur organise une enchère au premier prix, et que les acheteurs adoptent la stratégie d'équilibre σ.

    On note B_n la variable aléatoire donnant le *bénéfice*, ou *revenu*, du vendeur. Il s'agit du montant que paie l'acheteur qui a remporté l'enchère.

    a) Pour tout $\omega \in \Omega$, il existe (au moins) un entier $k \in [\![1;n]\!]$ tel que $Y_n(\omega) = \sigma\big(X_k(\omega)\big)$: k est le numéro de l'acheteur qui a enchéri le plus (il peut y en avoir plusieurs en cas d'ex aequo), et sa mise optimale est alors bien $\sigma\big(X_k(\omega)\big)$ qui correspond au prix d'achat.

    Comme σ est une fonction strictement croissante sur son domaine : l'image du maximum est le maximum des images, c'est-à-dire que :

$$\forall \omega \in \Omega, \quad B_n(\omega) = \max\big(\sigma(X_1(\omega)), \dots, \sigma(X_n)\big) = \sigma\Big(\max(X_1(\omega), \dots, X_n(\omega))\Big) = \sigma\big(Y_n(\omega)\big)$$

    C'est-à-dire qu'on a bien l'égalité de variables aléatoires : $\quad B_n = \sigma(Y_n)$.

    b) Le théorème de transfert, appliqué à la variable Y_n dont une densité est la fonction g_n étudiée en 1.b) au début du sujet, et avec la fonction σ qui est continue, positive et bornée, donne (les intégrales sont convergentes) :

$$\mathbb{E}(B_n) = \int_0^\alpha \sigma(x) g_{n-1}(x)\mathrm{d}x = \int_0^\alpha \sigma(x).n.f(x).\underbrace{\big(F(x)\big)^{n-1}}_{=G_{n-1}(x)}\mathrm{d}x \qquad \text{d'après la question 1.}$$

$$= n\int_0^\alpha \frac{1}{\cancel{G_{n-1}(x)}}\Big(\int_0^x t.g_{n-1}(t)\mathrm{d}t\Big).f(x).\cancel{G_{n-1}(x)}\mathrm{d}x \qquad \text{d'après la relation } (\star)$$

$$\mathbb{E}(B_n) = \int_0^\alpha \Big(\int_0^x t.g_{n-1}(t)\mathrm{d}t\Big).f(x)\mathrm{d}x$$

    c) On réalise directement une intégration par parties dans la grande intégrale précédente, en posant :

$$\begin{array}{lll} u(x) = \displaystyle\int_0^x t.g_{n-1}(t)\mathrm{d}t & \longrightarrow & u'(x) = x.g_{n-1}(x) \\ v'(x) = f(x) & \longrightarrow & v(x) = F(x) \end{array}$$

    Les fonctions u et v sont de classe $\mathcal{C}^1$ et bornées sur $]0;\alpha[$, donc par intégration par parties (faussement impropre en 0) :

$$\mathbb{E}(B_n) = n.\Big[F(x).\int_0^x t.g_{n-1}(t)\mathrm{d}t\Big]_0^\alpha - n\int_0^\alpha x.g_{n-1}(x).F(x)\mathrm{d}x$$

$$= n.\underbrace{F(\alpha)}_{=1}.\int_0^\alpha t.g_{n-1}(t)\mathrm{d}t - 0 - n\int_0^\alpha x.g_{n-1}(x).F(x)\mathrm{d}x$$

$$= n\int_0^\alpha \big(x.g_{n-1}(x) - x.g_{n-1}(x).F(x)\big)\mathrm{d}x \qquad \text{par linéarité de l'intégrale}$$

$$\mathbb{E}(B_n) = n\int_0^\alpha x.[1-F(x)].g_{n-1}(x)\mathrm{d}x \qquad \textit{CQFD}$$

14. Enchère au second prix.

    On suppose que le vendeur organise une enchère au second prix, et que les acheteurs adoptent la stratégie dominante : chacun mise autant que sa valeur privée.

    On note B'_n la variable aléatoire donnant le revenu du vendeur dans cette enchère.

    Dans ce cas, le gain du vendeur correspond à la deuxième plus grande mise (qui correpond aussi au prix payé) parmi les n mises $(X_k)_{1\leqslant k\leqslant n}$: B'_n a bien même loi que la variable Z_n étudiée dans la partie I., et ces deux variables aléatoires ont donc même espérance : $\mathbb{E}(B'_n)=\mathbb{E}(Z_n)$.

15. On peut écrire sous forme intégrale l'espérance $\mathbb{E}(Z_n)$, grâce à sa densité h_n obtenue à la question 2.b) :

$$\begin{aligned}\mathbb{E}(Z_n) &= \int_0^\alpha x.h_n(x)\mathrm{d}x = \int_0^\alpha x.n(n-1).f(x).[1-F(x)].\big[F(x)\big]^{n-2}\mathrm{d}x\\ &= n\int_0^\alpha x[1-F(x)].\underbrace{(n-1).f(x).\big[F(x)\big]^{n-2}}_{=g_{n-1}(x)\ !!}\mathrm{d}x\end{aligned}$$

$$\mathbb{E}(B'_n)=\mathbb{E}(Z_n)=n\int_0^\alpha x[1-F(x)].g_{n-1}(x)\mathrm{d}x=\mathbb{E}(B_n)\qquad \text{d'après ce qui précède.}$$

Ainsi, le revenu moyen pour le vendeur est le même pour les enchères au premier ou au second prix lorsque les acheteurs adoptent tous la stratégie optimale.

★★★ FIN DU SUJET ★★★