Correction du concours blanc n$^{\text{o}}$ 2

Option Économique

MATHÉMATIQUES II

Vendredi 9 juin 2017, de 8h à 12h.

Ce sujet a été posé cette année au concours de l'ESSEC, épreuve de Mathématiques II.

Étudier l'évolution des inégalités dans la répartition des richesses, matérielles ou symboliques, dans une société est un des thèmes majeurs des sciences humaines. Considérons un exemple élémentaire. Le tableau ci-dessous présente le pourcentage d'accès à l'enseignement secondaire en Grande-Bretagne lors de deux périodes pour deux catégories sociales :

	avant 1910	entre 1935 et 1940
Profession libérale	37%	62%
Ouvriers	1%	10%

On propose trois modes de comparaison des inégalités entre les deux classes sociales.

1. En regardant l'augmentation des pourcentages pour les deux classes entre les deux périodes on conclut que l'inégalité a augmenté entre la classe la plus aisée (Profession libérale) et la plus défavorisée (Ouvriers).
2. Si on regarde le taux d'accroissement des pourcentages, comme $\frac{10}{1} > \frac{62}{37}$, on déduit que l'inégalité a diminué.
3. Si on regarde le taux d'accroissement des pourcentages *de ceux qui n'accèdent pas à l'enseignement secondaire*, comme $\frac{90}{99} > \frac{38}{63}$, on déduit que l'inégalité a augmenté puisque le nombre de ceux qui n'ont pas accès à l'enseignement supérieur a proportionnellement plus diminué que celui de ceux qui y ont accès.

Comme on le voit chacune des façons de voir est légitime à sa manière. L'objet du problème est d'introduire des outils afin d'étudier la *concentration* d'une loi de probabilité pour contourner des paradoxes auxquels une analyse trop rapide peut conduire, ou du moins d'en être conscient.

I Indice de Gini

On rappelle qu'une fonction numérique définie sur l'intervalle J de $\mathbb{R}$ est *convexe* sur J si elle vérifie la propriété suivante : $\forall (t_1, t_2) \in J^2,\ \forall \lambda \in [0,1], f(\lambda t_1 + (1-\lambda)t_2) \leq \lambda f(t_1) + (1-\lambda) f(t_2)$.

On rappelle en outre qu'une fonction f est *concave* si $-f$ est convexe.

On désigne par E l'ensemble des applications f définies sur $[0,1]$ à valeurs dans $[0,1]$, continues et convexes sur $[0,1]$, et telles que $f(0) = 0$ et $f(1) = 1$. Pour toute application f de E, on note $\tilde{f}$ l'application associée à f, définie sur $[0,1]$ par $\tilde{f}(t) = t - f(t)$.

On pose enfin $I(f) = 2\displaystyle\int_0^1 \tilde{f}(t)dt = 2\int_0^1 (t - f(t))dt$. $I(f)$ s'appelle l'**indice de Gini** de l'application f.

1. (a) Donnez une interprétation géométrique de la propriété de convexité.

Solution. C'est une question de cours ! La convexité de f se traduit géométriquement par le fait que, sur tout segment inclus dans J, la courbe de la fonction se situe au-dessous de sa corde.

(b) Lorsque f est une fonction de classe C^1 sur $[0,1]$, rappeler la caractérisation de la convexité de f sur $[0,1]$ à l'aide de la dérivée f'.

Solution. Encore une question de cours. Si la fonction f est de classe C^1 sur $[0,1]$, alors elle est convexe sur $[0,1]$ si et seulement si sa dérivée f' est croissante sur $[0,1]$.

2. (a) Justifier que $\tilde{f}$ est concave sur $[0,1]$.

Solution. Attention ! Puisque la fonction n'est pas supposée de classe C^1, on ne peut pas utiliser la caractérisation de la question précédente. On doit donc vérifier directement la définition, c'est-à-dire montrer que la fonction $-\tilde{f}$, définie par $t \mapsto f(t) - t$ est convexe sur $[0,1]$.
Soient $(t_1, t_2) \in [0,1]^2$ et soit $\lambda \in [0,1]$. En notant $t = \lambda t_1 + (1-\lambda)t_2$, la convexité de f montre que l'on a l'inégalité $f(t) \leq \lambda f(t_1) + (1-\lambda) f(t_2)$, ce qui conduit à :

$$f(t) - t \quad = \quad f(t) - \lambda t_1 - (1-\lambda)t_2 \quad \leq \quad \lambda(f(t_1) - t_1) + (1-\lambda)(f(t_2) - t_2).$$

C'est ce qu'il fallait vérifier.

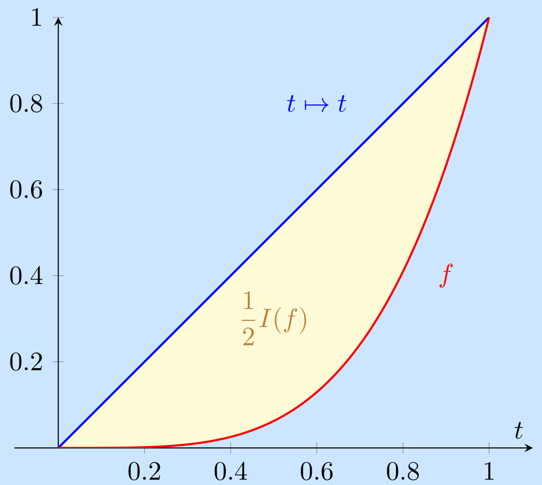
(b) Montrer que $I(f) = 1 - 2\int_0^1 f(t)dt$.

Solution. Les fonctions considérées sont continues sur le segment $[0,1]$, donc les intégrales sont bien définies. Par linéarité de l'intégration, on obtient de plus :

$$I(f) = 2\int_0^1 t\,dt - 2\int_0^1 f(t)\,dt, \quad \text{avec} \quad 2\int_0^1 t\,dt = 2\left[\frac{t^2}{2}\right]_0^1 = 2 \times \frac{1}{2} = 1, \quad \text{d'où le résultat.}$$

(c) Représenter dans un même repère orthonormé les fonctions f et $t \mapsto t$ et donner une interprétation géométrique de $I(f)$.

Solution. La fonction $t \mapsto t$ correspond en fait à la corde de f sur $[0,1]$.



On sait que $\int_0^1 t\,dt$ et $\int_0^1 f(t)\,dt$ mesurent respectivement l'aire sous la courbe de $t \mapsto t$ et l'aire sous la courbe de f. Ainsi, $I(f)$ mesure le double de l'aire entre la courbe de f et sa corde sur $[0,1]$.

3. **Un premier exemple.**
Soit $f : [0,1] \to \mathbb{R}$ telle que $f(t) = t^2$ pour tout $t \in [0,1]$.

(a) Montrer que f est un élément de E.

Solution. La fonction $f : t \mapsto t^2$ est polynomiale, donc en particulier de classe C^1 sur $[0,1]$, et elle admet pour dérivée la fonction $t \mapsto 2t$ qui est croissante sur cet intervalle. Ainsi, f est convexe. De plus $f(0) = 0$ et $f(1) = 1$. Enfin $0 \leq t^2 \leq 1$ pour tout $t \in [0,1]$, donc f est bien un élément de E.

(b) Calculer $I(f)$.

Solution. Une question de calcul facile, à ne surtout pas manquer ! D'après 2.(b) :

$$I(f) = 1 - 2\int_0^1 t^2\,dt = 1 - 2\left[\frac{t^3}{3}\right]_0^1 = 1 - \frac{2}{3} = \boxed{\frac{1}{3}}$$

4. **Propriétés de l'indice de Gini.**

(a) Pour f élément de E, établir que $I(f) \geq 0$.

Solution. Par convexité, la courbe de f est au-dessous de sa corde sur $[0,1]$, qui représente la fonction $t \mapsto t$. On a donc $\tilde{f}(t) = t - f(t) \geq 0$ pour tout $t \in [0,1]$. Puisque les bornes 0 et 1 sont dans l'ordre croissant, on en déduit par propriété de positivité de l'intégration que :

$$\boxed{I(f) = 2\int_0^1 \tilde{f}(t)\,dt \geq 0}$$

(b) Montrer que $I(f) = 0$ si et seulement si $f(t) = t$ pour tout $t \in [0,1]$.

Solution. La fonction f étant continue, il en va de même pour $\tilde{f}$. Elle est de plus positive, donc son intégrale sur $[0,1]$ est nulle si et seulement si : $\forall t \in [0,1],\ \tilde{f}(t) = t - f(t) = 0$.

(c) Montrer que pour tout f élément de E, $I(f) < 1$.

Solution. Soit $f \in E$. D'après la formule 2.(b), il suffit de vérifier que $\displaystyle\int_0^1 f(t)\,dt > 0$.
Or f est continue et positive sur $[0,1]$, donc son intégrale est positive et vaut 0 si et seulement si f est la fonction nulle (positivité de l'intégration). Mais ceci est impossible car $f(1) = 1$.

(d) Pour tout entier $n > 0$, on définit f_n sur $[0,1]$ par $f_n(t) = t^n$.

i. Pour tout entier n strictement positif, calculer $I(f_n)$.

Solution. Encore un calcul facile à ne pas sauter !

$$I(f_n) = 1 - 2\int_0^1 t^n\,dt = 1 - 2\left[\frac{t^{n+1}}{n+1}\right]_0^1 = 1 - \frac{2}{n+1} = \boxed{\frac{n-1}{n+1}}$$

Remarquons que pour $n = 2$, on retrouve bien le résultat de la question 3.

ii. En déduire que pour tout réel A vérifiant $0 \leq A < 1$, il existe f appartenant à E telle que $I(f) > A$.

Solution. D'après la question précédente, on a facilement $\lim_{n\to\infty} I(f_n) = 1$. Pour tout réel $A \in [0,1[$, il existe donc un entier n_0 tel que $I(f_{n_0}) > A$. Il suffit alors de considérer f_{n_0}, qui appartient à bien à E (preuve analogue à la question 3).

5. **Minoration de l'indice de Gini**

(a) Soit f élément de E. Montrer qu'il existe t_0 dans $]0,1[$ tel que $\tilde{f}(t_0) = \max_{t\in[0,1]} \tilde{f}(t)$.

Solution. La fonction $\tilde{f}$ est continue sur le segment $[0,1]$. D'après le théorème des bornes, elle est donc majorée et son maximum est atteint en un certain réel $a \in [0,1]$. Si $0 < a < 1$, il suffit de prendre $t_0 = a$ car $\tilde{f}(a) = \max_{t\in[0,1]} \tilde{f}(t)$.

Il reste donc à traiter l'autre cas : si $a = 0$ ou $a = 1$. Dans ce cas $\tilde{f}(a) = 0$: on sait en effet que $f(1) = 1$ et $f(0) = 0$, donc $\tilde{f}(a) = \tilde{f}(1) = 0$ ou $\tilde{f}(a) = \tilde{f}(0) = 0$. Ainsi $\tilde{f}$ est partout négative puisque son maximum sur $[0,1]$ vaut 0. Mais on a vu à la question 4 qu'elle était partout positive, donc la seule possibilité est que $\tilde{f}$ est nulle sur $[0,1]$. N'importe quel $t_0 \in]0,1[$ convient alors !

(b) Montrer que pour tout t de $[0,t_0]$, $\tilde{f}(t) \geq \tilde{f}(t_0).\dfrac{t}{t_0}$.

Solution. Le plus rapide consiste à remarquer que $t \mapsto \tilde{f}(t_0)\dfrac{t}{t_0}$ est la corde de f sur $[0,t_0]$ et à utiliser la concavité. Une autre méthode (qui revient en fait au même) consite à écrire t sous la forme $t = \lambda t_0 + (1-\lambda)0$, ce qui impose $\lambda = t/t_0$. Lorsque $t \in [0,t_0]$, on a bien $\lambda \in [0,1]$ et, en utilisant le fait que $\tilde{f}(0) = 0$, l'inégalité de concavité de $\tilde{f}$ conduit au résultat :

$$\tilde{f}(t) = \tilde{f}(\lambda t_0 + (1-\lambda)0) \geq \lambda \tilde{f}(t_0) + (1-\lambda)\tilde{f}(0) = \frac{t}{t_0}\tilde{f}(t_0).$$

(c) Montrer que pour tout t de $[t_0,1]$, $\tilde{f}(t) \geq \tilde{f}(t_0).\dfrac{t-1}{t_0-1}$.

Solution. Ici encore, les deux approches sont possibles. Détaillons cette fois la première, qui consiste à remarquer que la fonction $g : t \mapsto \tilde{f}(t_0)\dfrac{t-1}{t_0-1}$ est la corde de f sur $[t_0,1]$. Il s'agit bien d'une fonction affine, donc il suffit de vérifier que :

$$g(t_0) = \tilde{f}(t_0)\frac{t_0-1}{t_0-1} = \tilde{f}(t_0), \qquad \text{et} \qquad g(1) = \tilde{f}(t_0)\frac{1-1}{t_0-1} = 0 = \tilde{f}(1).$$

Par concavité, f est au-dessus de sa corde. Donc pour tout $t \in [t_0,1]$, $\tilde{f}(t) \geq g(t)$.

(d) En déduire que $I(f) \geq \tilde{f}(t_0)$.

Solution. Cette question est en fait assez facile et on pouvait la traiter en admettant les deux précédentes (plus difficiles) ! Par propriété de croissance de l'intégration, on a :

$$\int_0^{t_0} \tilde{f}(t)\,dt \geq \int_0^{t_0} \tilde{f}(t_0)\frac{t}{t_0}\,dt = \frac{\tilde{f}(t_0)}{t_0}\int_0^{t_0} t\,dt = \frac{\tilde{f}(t_0)}{t_0}\frac{t_0^2-0}{2} = \frac{\tilde{f}(t_0)t_0}{2}.$$

Et de même sur l'autre intervalle :

$$\int_{t_0}^1 \tilde{f}(t)\,dt \geq \int_{t_0}^1 \tilde{f}(t_0)\frac{t-1}{t_0-1}\,dt = \frac{\tilde{f}(t_0)}{t_0-1}\int_{t_0}^1 (t-1)\,dt = \frac{\tilde{f}(t_0)}{t_0-1}\frac{0-(t_0-1)^2}{2} = \frac{\tilde{f}(t_0)(1-t_0)}{2}.$$

La relation de Chasles donne enfin :

$$I(f) = 2\int_0^1 \tilde{f}(t)\,dt = 2\int_0^{t_0} \tilde{f}(t)\,dt + 2\int_{t_0}^1 \tilde{f}(t)\,dt \geq \tilde{f}(t_0)t_0 + \tilde{f}(t_0)(1-t_0) = \tilde{f}(t_0).$$

L'indice de Gini donne une indication sur la concentration des richesses d'un pays si l'on suppose que la fonction f rend compte de cette concentration. Par exemple, $f(0,3) = 0,09$ s'interprète par le fait que dans la population classée par ordre de richesse croissante, les premiers 30% de la population possèdent 9% de la richesse totale du pays. Plus l'indice $I(f)$ est grand, plus la répartition des richesses est inégalitaire.

II Le cas à densité

Soit g une densité de probabilité sur $\mathbb{R}$, nulle sur $]-\infty, 0]$, continue et strictement positive sur $]0, +\infty[$. On définit une fonction G sur $\mathbb{R}_+$ par $G(x) = \displaystyle\int_0^x g(\nu)d\nu$ pour $x \in \mathbb{R}_+$. Si g représente la densité de population classée suivant son revenu croissant, $G(x)$ représente la proportion de la population dont le revenu est inférieur à x. On suppose de plus que $\displaystyle\int_0^{+\infty} \nu g(\nu)d\nu$ est convergente et on note m sa valeur qui représente donc la richesse moyenne de la population.

6. (a) Montrer que $m > 0$.

Solution. Par produit, la fonction $\nu \mapsto \nu g(\nu)$ est continue et strictement positive sur $]0, +\infty[$. Son intégrale $m = \displaystyle\int_0^{+\infty} \nu g(\nu)\,d\nu$ est donc strictement positive (positivité de l'intégration).

(b) Montrer que G est une bijection de $[0, +\infty[$ sur $[0, 1[$. On notera G^{-1} son application réciproque.

Solution. Remarquons que G est la fonction de répartition associée à la densité g car :

$$\forall x \in \mathbb{R}, \quad \int_{-\infty}^x g(t)\,dt = \int_{-\infty}^0 0\,dt + \int_0^x g(t)\,dt = 0 + G(x).$$

Elle est donc croissante, continue sur $\mathbb{R}$, et vérifie $\lim\limits_{x\to+\infty} G(x) = 1$. Puisque g est continue sur $]0, +\infty[$, la fonction G est de classe C^1 sur cet intervalle, de dérivée $G' = g$ qui est strictement positive. Ainsi, G est continue et strictement croissante sur $[0, +\infty[$, donc on peut lui appliquer le théorème de la bijection : G est bijective de $[0, +\infty[$ vers l'intervalle image $[G(0), 1[$. Pour conclure, il reste simplement à remarquer que $G(0) = 0$.

(c) Quel est le sens de variation de G^{-1} sur $[0, 1[$?

Solution. Le théorème de la bijection montre aussi que G^{-1} est continue et strictement monotone sur $[0, 1[$, de même sens de variation que G : c'est-à-dire strictement croissante.

7. (a) A l'aide du changement de variable $u = G(\nu)$, établir que pour tout $t \in [0, 1[$,

$$\int_0^t G^{-1}(u)du = \int_0^{G^{-1}(t)} \nu g(\nu)d\nu.$$

Solution. On a vu que G est C^1 sur $[0,1[$ et admet pour dérivée g. Les bornes peuvent s'écrire $t = G(G^{-1}(t))$ et $0 = G(0)$. Enfin G^{-1} est continue, donc la formule de changement de variable conduit à :

$$\int_0^t G^{-1}(u)\,du = \int_0^{G^{-1}(t)} G^{-1}(G(\nu))G'(\nu)\,d\nu = \int_0^{G^{-1}(t)} \nu g(\nu)\,d\nu.$$

Remarque. Ceci n'est pas tout à fait rigoureux car g n'est pas supposée continue en 0. Pour pallier ce manque, il faudrait se borner à un intervalle $[\varepsilon, t]$ avec $\varepsilon > 0$, puis faire tendre ε vers 0. Mais ceci n'est pas dans l'esprit du programme de première année. La correction n'en tient pas compte.

(b) En déduire que la limite à gauche $\lim\limits_{t\to 1^-} \displaystyle\int_0^t G^{-1}(u)du$ existe et donner sa valeur.

Solution. D'après le résultat précédent, il suffit de remarquer que $G^{-1}(t)$ tend vers $+\infty$ pour $t \to 1$ puisque $G(x)$ tend vers 1 lorsque $x \to +\infty$. Par composition des limites, on obtient alors :

$$\boxed{\int_0^t G^{-1}(u)\,du \xrightarrow[t\to 1^-]{} \int_0^{+\infty} \nu g(\nu)\,d\nu = m}$$

8. Soit f la fonction définie sur $[0,1]$ par : $f(t) = \dfrac{1}{m}\displaystyle\int_0^t G^{-1}(u)du$ pour tout $t \in [0,1[$ et $f(1) = 1$.

(a) i. Montrer que f est continue sur $[0,1]$.

Solution. Par définition, f est la primitive s'annulant en 0 de la fonction $\dfrac{1}{m}G^{-1}$, qui est continue sur $[0,1[$. En particulier, f est continue sur $[0,1[$.
Puisque $f(1) = 1$, la question précédente montre que f est aussi continue en 1.

ii. Montrer que f est convexe sur $[0,1[$. **On admettra qu'en fait f est convexe sur $[0,1]$.**

Solution. Comme rappelé ci-dessus, f admet pour dérivée $\dfrac{1}{m}G^{-1}$ qui est continue sur $[0,1[$. Donc f est de classe C^1 et, puisque G^{-1} est croissante, le critère de la question 1 montre que f est convexe sur $[0,1[$.
Remarque. Pour montrer que f est convexe sur $[0,1]$, il resterait à vérifier que sa courbe est située en dessous de ses cordes sur tout intervalle de la forme $[a,1]$ avec $a \in [0,1]$. Il suffit en fait d'utiliser la convexité sur $[a,b]$ pour tout $b < 1$ et de passer à la limite $b \to 1$, ce qui se justifie grâce à la continuité de f sur $[0,1]$. Cette propriété est essentielle !

iii. En déduire que f est un élément de E.

Solution. On a vu que f est continue et convexe sur $[0,1]$. De plus $f(0) = 0$ et $f(1) = 1$.
Il reste à justifier que f est à valeurs dans $[0,1]$. En fait, f est croissante sur $[0,1]$ puisqu'elle est continue sur cet intervalle et que sa dérivée $\dfrac{1}{m}G^{-1}$ sur $]0,1[$ est positive. Ceci entraîne que pour tout $t \in [0,1]$ on a : $0 = f(0) \leq f(t) \leq f(1) = 1$. Autrement dit $f([0,1]) \subset [0,1]$.

(b) Montrer, à l'aide d'une intégration par parties, l'égalité

$$I(f) = -1 + \frac{2}{m}\int_0^{\infty} \nu g(\nu)G(\nu)d\nu.$$

Solution. Voici une question difficile, mais instructive ! Rappelons que $I(f) = 1 - 2\int_0^1 f(t)\,dt$. L'énoncé suggère de calculer l'intégrale en effectuant une intégration par parties, de manière à faire apparaître $f' = \frac{1}{m}G^{-1}$. Mais cette relation n'est vraie que sur $[0,1[$, donc on doit ruser et se borner dans un premier temps à un intervalle $[0,b]$ avec $0 < b < 1$. Sur un tel intervalle, les fonctions $u : t \mapsto t$ et $v : t \mapsto f(t)$ sont C^1, d'où (formule d'intégration par parties) :

$$\int_0^b f(t)\,dt = \Big[t\,f(t)\Big]_0^b - \int_0^b t f'(t)\,dt, \qquad \text{avec} \qquad \int_a^b t f'(t)\,dt = \frac{1}{m}\int_0^b t\,G^{-1}(t)\,dt.$$

Le changement de variable $t = G(\nu)$ déjà justifié dans la question 7 conduit alors à :

$$\int_0^b f(t)\,dt - \Big[t\,f(t)\Big]_0^b = -\frac{1}{m}\int_0^{G^{-1}(b)} G(\nu)\nu g(\nu)\,d\nu.$$

Par continuité, le membre de gauche converge lorsque $b \to 1^-$. Il en est donc de même pour l'intégrale impropre à droite. Puisque $G^{-1}(b)$ tend vers $+\infty$, on obtient par passage à la limite :

$$\int_0^1 f(t)\,dt - f(1) = -\frac{1}{m}\int_0^{+\infty} \nu g(\nu) G(\nu)\,d\nu.$$

Or $f(1) = 1$, donc le résultat demandé en découle directement.

Remarque. Ici encore, on a supposé que g est continue en 0 pour éviter les complications, mais la parade serait la même qu'à la question 7 si on voulait traiter le cas général.

9. Soit λ un réel strictement positif. On suppose dans cette question que g est une densité de la loi exponentielle de paramètre λ.

   (a) Expliciter $G(x)$ pour $x > 0$.

   Solution. Oh, une question de cours ! Pour tout $x > 0$, $G(x) = 1 - e^{-\lambda x}$.

   (b) Expliciter $G^{-1}(u)$ pour $u \in [0,1[$.

   Solution. Soit $u \in [0,1[$. Par définition, $G^{-1}(u)$ est l'unique $x \geq 0$ tel que $G(x) = u$. De plus,

   $$G(x) = u \iff 1 - e^{-\lambda x} = u \iff 1 - u = e^{-\lambda x} \iff \ln(1-u) = -\lambda x.$$

   Puisque λ est non nul, on obtient donc : $\boxed{G^{-1}(u) = -\frac{1}{\lambda}\ln(1-u)}$

   (c) Donner la valeur de m.

   Solution. Encore une question de cours ! $\boxed{m = \frac{1}{\lambda}}$

   (d) Soit $t \in [0,1[$. Montrer que $f(t) = -\int_0^t \ln(1-u)du$.

   Solution. On adapte la définition de la fonction f (question 8) aux questions précédentes :

   $$f(t) = \frac{1}{m}\int_0^t G^{-1}(u)\,du = \frac{1}{\frac{1}{\lambda}}\int_0^t -\frac{1}{\lambda}\ln(1-u)du = -\int_0^t \ln(1-u)\,du. \quad \text{(linéarité de l'intégrale)}$$

(e) En déduire que pour tout t élément de $[0,1[$, on a $f(t) = (1-t)\ln(1-t) + t$.

Solution. Deux solutions pour cette question : effectuer une intégration par parties en dérivant le logarithme ou, plus simplement, vérifier que la fonction proposée $g : t \mapsto (1-t)\ln(1-t) + t$ est bien la primitive de $u \mapsto -\ln(1-u)$ qui s'annule en 0. La vérification de $g(0) = 0$ est immédiate. De plus g est dérivable sur $[0,1[$ par opérations algébriques sur les fonctions dérivables, et :

$$\forall t \in [0,1[, \quad g'(t) = -\ln(1-t) + (1-t)\frac{-1}{1-t} + 1 = -\ln(1-t).$$

(f) Justifier l'existence de la limite à gauche $\lim\limits_{b\to 1^-} \int_0^b (1-t)\ln(1-t)dt$ et la calculer.

Solution. Soit $b > 1$. On calcule d'abord l'intégrale sur $[0,b]$ en intégrant par parties, avec les fonctions $u : t \mapsto -\frac{1}{2}(1-t)^2$ et $v : t \mapsto \ln(1-t)$ qui sont C^1 sur cet intervalle :

$$\begin{aligned}\int_0^b (1-t)\ln(1-t)\,dt &= \left[\frac{-(1-t)^2}{2}\ln(1-t)\right]_0^b - \int_0^b \frac{-(1-t)^2}{2}\frac{-1}{1-t}\,dt \\ &= -\frac{(1-b)^2}{2}\ln(1-b) - \frac{1}{2}\int_0^b (1-t)\,dt \\ &= -\frac{1}{2}(1-b)^2\ln(1-b) + \frac{1}{4}(1-b)^2 - \frac{1}{4}.\end{aligned}$$

Il s'agit maintenat d'établir la convergence pour $b \to 1^-$, qui *a priori*, présente une forme indéterminée. Cependant $\lim\limits_{b\to 1^-}(1-b) = 0$ et $\lim\limits_{x\to 0} x\ln(x) = 0$ par croissances comparées, d'où $\lim\limits_{x\to 0} x^2\ln(x) = 0$.

La composition des limites donne donc finalement : $\boxed{\int_0^b (1-t)\ln(1-t)\,dt \xrightarrow[b\to 1^-]{} -\frac{1}{4}}$

(g) En déduire la valeur de $I(f)$.

Solution. On utilise les résultats précédents et la continuité de f sur $[0,1]$:

$$I(f) = 2\int_0^1 (t - f(t))\,dt = 2\lim_{b\to 1^-}\int_0^b (t-f(t))\,dt = -2\lim_{b\to 1^-}\int_0^b (1-t)\ln(1-t)\,dt = \boxed{\frac{1}{2}}$$

III Application à une population

Une population de N personnes est divisée en deux classes (typiquement hommes et femmes) et en n catégories (par exemple socio-professionnelles), suivant le tableau à double entrée suivant où tous les x_i et y_i pour i dans $[\![1,n]\!]$ sont des entiers naturels. On suppose en outre que pour tout i dans $[\![1,n]\!]$, $x_i \neq 0$.

Classes \ Catégories	c_1	c_2	c_3	...	c_i	...	c_n	Total
I	x_1	x_2	x_3	...	x_i	...	x_n	X
II	y_1	y_2	y_3	...	y_i	...	y_n	Y
Total	n_1	n_2	n_3	...	n_i	...	n_n	N

où on a donc posé $X = \sum\limits_{i=1}^n x_i$, $Y = \sum\limits_{i=1}^n y_i$ et $X + Y = N$. On suppose en outre que $Y > 0$.

Pour i appartenant à $[\![1,n]\!]$, on adopte les notations suivantes :

$$p_i = \frac{n_i}{N},\ q_i = \frac{x_i}{X},\ r_i = \frac{y_i}{Y}.$$

On note aussi $\varepsilon_i = \dfrac{x_i}{n_i}$, et $\varepsilon = \dfrac{X}{N}$, et on suppose que les catégories sont numérotées de telle sorte que

$$\varepsilon_1 \leq \varepsilon_2 \leq ... \leq \varepsilon_n.$$

10. On pose $\Omega = \{c_1, c_2, ..., c_n\}$, ensemble des catégories dans la population.

(a) Montrer que $P = (p_i)_{1\leq i\leq n}$, $Q = (q_i)_{1\leq i\leq n}$ et $R = (r_i)_{1\leq i\leq n}$ sont des distributions de probabilités.

Solution. Notons déjà que $N = \displaystyle\sum_{i=1}^{n} n_i \geq \sum_{i=1}^{n} x_i > 0$ donc toutes les fractions sont bien définies.
On vérifie maintenant les deux conditions pour une loi de probabilité :

- Pour tout $i \in [\![1,n]\!],\ p_i = \dfrac{n_i}{N} \geq 0$ car $n_i \in \mathbb{N}$.
- $\displaystyle\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{n_i}{N} = \frac{1}{N}\sum_{i=1}^{n} n_i = \frac{1}{N}N = 1.$

Ainsi, la famille $P = (p_i)_{1\leq i\leq n}$ définit bien une loi de probabilité.
Le raisonnement est entièrement similaire pour les familles $Q = (q_i)_{1\leq i\leq n}$ et $R = (r_i)_{1\leq i\leq n}$.

(b) Montrer que $\dfrac{q_1}{p_1} \leq .... \leq \dfrac{q_n}{p_n}$ $(*)$

Solution. Remarquons tout d'abord que $\dfrac{q_i}{p_i} = \dfrac{N}{X}\varepsilon_i = \dfrac{\varepsilon_i}{\varepsilon}$ pour tout $i \in [\![1,n]\!]$.
Puisque $\varepsilon_1 \leq \varepsilon_2 \leq ... \leq \varepsilon_n$, le résultat s'en déduit immédiatement en divisant par $\varepsilon > 0$.

(c) Montrer que $\dfrac{r_1}{p_1} \geq .... \geq \dfrac{r_n}{p_n}$

Solution. On procède de même en remarquant que pour tout $i \in [\![1,n-1]\!]$, $\dfrac{r_i}{p_i} = \dfrac{1-\varepsilon_i}{1-\varepsilon}$.

(d) Montrer que pour i appartenant à $[\![1,n]\!]$, $r_i = \dfrac{n_i - x_i}{N - X} = \dfrac{p_i - \varepsilon q_i}{1-\varepsilon}$.

Solution. Soit $i \in [\![1,n]\!]$. Puisque $r_i = \dfrac{y_i}{Y}$ avec $y_i = n_i - x$ et $Y = N - X$, la première égalité est évidente. En divisant numérateur et dénominateur par N, on obtient alors :

$$r_i = \frac{n_i - x_i}{N - X} = \frac{\frac{n_i}{N} - \frac{x_i}{N}}{1 - \frac{X}{N}} = \frac{p_i - \frac{x_i}{N}}{1-\varepsilon}, \qquad \text{avec} \quad \frac{x_i}{N} = \frac{X}{N}\frac{x_i}{X} = \varepsilon q_i.$$

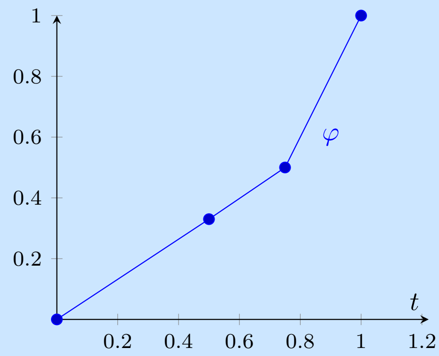
11. Dans un premier temps, nous allons construire une application appartenant à E, qui permet de mesurer les inégalités à l'intérieur de la classe I.

On pose $P_0 = Q_0 = 0$, et pour $i \in [\![1,n]\!]$, $P_i = \displaystyle\sum_{h=1}^{i} p_h$ et $Q_i = \displaystyle\sum_{h=1}^{i} q_h$. On définit alors l'application φ de $[0,1]$ dans $[0,1]$ telle que, pour tout entier $i \in [\![0,n]\!]$, $\varphi(P_i) = Q_i$ et pour tout entier $i \in [\![0,n-1]\!]$, φ est affine sur le segment $[P_i, P_{i+1}]$.

(a) On suppose **dans cette question** $n = 3$.
Représenter dans un repère orthonormé φ lorsque $P = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ et $Q = \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right)$.

Solution. On commence par calculer les (P_i) et les (Q_i) puis on trace le graphe.

i	0	1	2	3
P_i	0	$\frac{1}{2}$	$\frac{3}{4}$	1
Q_i	0	$\frac{1}{3}$	$\frac{1}{2}$	1



(b) Montrer que, dans le plan rapporté à un repère orthonormé, la pente de la droite passant par les points de coordonnées (P_{i-1}, Q_{i-1}) et (P_i, Q_i) est $u_i = \dfrac{q_i}{p_i}$ pour i appartenant à $[\![1, n]\!]$.

Solution. L'accroissement vertical entre les deux points vaut $Q_i - Q_{i-1} = q_i$ tandis que l'accroissement horizontal vaut $P_i - P_{i-1} = p_i$. La pente de la droite est donc $\boxed{\dfrac{Q_i - Q_{i-1}}{P_i - P_{i-1}} = \dfrac{q_i}{p_i}}$.

(c) Montrer que si $i \in [\![0, n-1]\!]$ et $t \in [P_i, P_{i+1}]$, on a $\varphi(t) = u_{i+1}(t - P_i) + Q_i$.

Solution. On sait que, sur l'intervalle $[P_i, P_{i+1}]$, la fonction φ est affine et a pour pente u_{i+1}. Il existe donc un réel b tel que φ soit définie par $t \mapsto u_{i+1}t + b$ sur cet intervalle. Puisque $\varphi(P_i) = Q_i$, on obtient l'équation $Q_i = u_{i+1}P_i + b$ dont l'unique solution est $b = Q_i - u_{i+1}P_i$. Ainsi :

$$\forall t \in [P_i, P_{i+1}], \qquad \boxed{\varphi(t) = u_{i+1}t + Q_i - u_{i+1}P_i = u_{i+1}(t - P_i) + Q_i}$$

(d) **En admettant** que les inégalités $(*)$ de la question 10.(b) permettent d'affirmer que φ est convexe, justifier que φ appartient à E.

Solution. Notons déjà que φ est définie sur $[0, 1]$ car la suite (Q_i) est croissante, à valeurs comprises entre $Q_0 = 0$ et $Q_n = 1$. Ceci découle du fait que Q est une loi de probabilité.

Le même argument montre que (P_i) est à valeurs comprises entre $P_0 = 0$ et $P_n = 1$. Par construction, φ est donc aussi à valeurs dans $[0, 1]$ (elle est même croissante) et elle vérifie :

$$\varphi(0) = \varphi(P_0) = Q_0 = 0, \qquad \varphi(1) = \varphi(P_1) = Q_1 = 1.$$

De plus, φ est continue sur $[0, 1]$, et même affine par morceaux, de par sa construction. En admettant la convexité, tout ceci montre que φ appartient à E.

(e) Pour $i \in [\![0, n-1]\!]$, calculer $\displaystyle\int_{P_i}^{P_{i+1}} \varphi(t)dt$.

Solution. Soit $i \in [\![0, n-1]\!]$. On a vu que φ est donnée par $t \mapsto u_{i+1}(t - P_i) + Q_i$ sur $[P_i, P_{i+1}]$. On a donc par linéarité :

$$\int_{P_i}^{P_{i+1}} \varphi(t)dt = u_{i+1}\int_{P_i}^{P_{i+1}} (t - P_i)\,dt + \int_{P_i}^{P_{i+1}} Q_i\,dt = u_{i+1}\frac{(P_{i+1} - P_i)^2}{2} + Q_i(P_{i+1} - P_i)$$

Or $u_{i+1} = \dfrac{Q_{i+1} - Q_i}{P_{i+1} - P_i}$ donc on peut simplifier ce résultat sous la forme :

$$\int_{P_i}^{P_{i+1}} \varphi(t)\,dt = \frac{1}{2}(Q_{i+1} - Q_i)(P_{i+1} - P_i) + Q_i(P_{i+1} - P_i) = \boxed{\frac{1}{2}(Q_{i+1} + Q_i)(P_{i+1} - P_i)}.$$

Remarque. Il s'agit simplement de l'aire du trapèze sous la courbe de φ entre P_i et P_{i+1}.

(f) Exprimer $I(\varphi)$ sous la forme d'une somme en fonction de $P_0, P_1, ..., P_n, Q_0, ...Q_n$.

Solution. D'après la formule de Chasles, on déduit de ce qui précède que :

$$\int_0^1 \varphi(t)\,dt = \sum_{i=0}^{n-1} \int_{P_i}^{P_{i+1}} \varphi(t)\,dt = \frac{1}{2} \sum_{i=0}^{n-1} (Q_{i+1} + Q_i)(P_{i+1} - P_i).$$

Et donc finalement :

$$I(\varphi) = 1 - 2\int_0^1 \varphi(t)\,dt = \boxed{1 - \sum_{i=0}^{n-1} (Q_{i+1} + Q_i)(P_{i+1} - P_i)}$$

12. Nous allons maintenant étudier l'application correspondante pour la classe II.

On pose $P_0 = R_0 = 0$ et pour $i \in [\![1, n]\!]$, $P_i = \sum_{h=1}^{i} p_h$ et $R_i = \sum_{h=1}^{i} r_h$. De même, on définit pour i élément de $[\![0, n]\!]$, $\Pi_i = 1 - P_{n-i}$. On considère l'application ψ de $[0, 1]$ dans $[0, 1]$ telle que pour tout $i \in [\![0, n]\!]$, $\psi(P_i) = R_i$ et pour tout entier $i \in [\![0, n-1]\!]$, ψ est affine sur le segment $[P_i, P_{i+1}]$.

(a) Montrer que la pente de la droite passant par les points de coordonnées (P_{i-1}, R_{i-1}) et (P_i, R_i) est $v_i = \dfrac{r_i}{p_i}$ pour $i \in [\![1, n]\!]$.

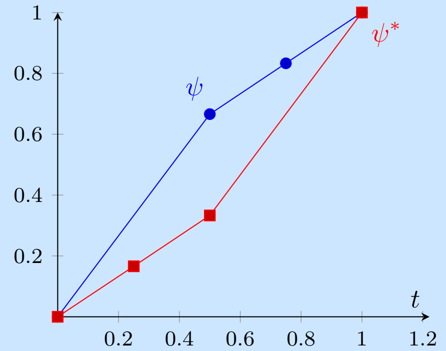
Solution. Même raisonnement que pour la question 11.

(b) On considère l'application ψ^* définie pour tout $t \in [0, 1]$, par $\psi^*(t) = 1 - \psi(1 - t)$.

i. On suppose **dans cette question** $n = 3$.
Représenter dans un même repère orthonormé les courbes représentatives de ψ et ψ^* lorsque $P = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ et $R = \left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$.

Solution. On commence par calculer les (P_i), les (R_i) et les (Π_i) puis on trace les graphes en remarquant que $\psi^*(\Pi_i) = 1 - \psi(1 - \Pi_i) = 1 - \psi(P_{n-i}) = 1 - R_{n-i}$ pour tout $i \in [\![0, n]\!]$.

i	0	1	2	3
P_i	0	$\frac{1}{2}$	$\frac{3}{4}$	1
R_i	0	$\frac{2}{3}$	$\frac{5}{6}$	1
Π_i	0	$\frac{1}{4}$	$\frac{1}{2}$	1



Remarque. La courbe de ψ^* s'obtient par symétrie de celle de ψ par rapport au point central de coordonnées $\left(\frac{1}{2}, \frac{1}{2}\right)$.

ii. Montrer que ψ^* est convexe sur $[0,1]$.

Solution. Comme pour la question 11, on admet que les inégalités 10 (c) entraînent la concavité de la fonction ψ sur $[0,1]$. Montrons maintenat que ψ^* est convexe. Soient $(t_1,t_2) \in [0,1]^2$, $\lambda \in [0,1]$ et $t = \lambda t_1 + (1-\lambda)t_2$. En posant $u = 1-t$, $u_1 = 1-t_1$ et $u_2 = 1-t_2$, on a $u = \lambda u_1 + (1-\lambda)u_2$ car $\lambda + (1-\lambda) = 1$, de sorte que $\psi(u) \geq \lambda\psi(u_1) + (1-\lambda)\psi(u_2)$ par concavité de ψ. Mais alors $\psi^*(t) = 1-\psi(1-t) = 1-\psi(u) \leq 1-\lambda\psi(u_1)-(1-\lambda)\psi(u_2)$, d'où :

$$\psi^*(t) \leq \lambda(1-\psi(u_1)) + (1-\lambda)(1-\psi(u_2)) = \lambda\psi^*(t_1) + (1-\lambda)\psi^*(t_2).$$

C'est ce qu'il fallait démontrer.

iii. Montrer que ψ^* est affine sur $[\Pi_{i-1}, \Pi_i]$ pour $i \in [\![0, n-1]\!]$.

Solution. En notant σ la fonction $t \mapsto 1-t$, on a par définition : $\psi^* = \sigma \circ \psi \circ \sigma$. Or σ est une bijection (affine) de $[\Pi_{i-1}, \Pi_i]$ sur $[P_{n-i}, P_{n-i+1}]$, intervalle sur lequel ψ est affine. Ainsi, sur $[\Pi_{i-1}, \Pi_i]$, la fonction ψ^* est une composition de trois fonctions affines, donc elle est affine.

iv. Montrer que la pente de ψ^* sur $[\Pi_{i-1}, \Pi_i]$ est v_{n-i+1} pour $i \in [\![1, n]\!]$.

Solution. On a vu au (i) que $\psi^*(\Pi_i) = 1-R_{n-i}$ et on obtient de la même manière $\psi^*(\Pi_{i-1}) = R_{n-(i-1)} = 1-R_{n-i+1}$. Ainsi la pente de ψ^* sur cet intervalle vaut-elle :

$$\frac{\psi^*(\Pi_i)-\psi^*(\Pi_{i-1})}{\Pi_i-\Pi_{i-1}} = \frac{(1-R_{n-i})-(1-R_{n-i+1})}{(1-P_{n-i})-(1-P_{n-i+1})} = \frac{R_{n-i+1}-R_{n-i}}{P_{n-i+1}-P_{n-i}} = \frac{r_{n-i+1}}{p_{n-i+1}} = v_{n-i+1}$$

On dit dans cette situation que les fonctions φ et ψ^* sont **adjointes** l'une de l'autre. C'est leur comparaison que Gini a proposé de considérer pour "mesurer les inégalités" entre la population de catégorie I et celle de catégorie II.

Une égalité entre les fonctions adjointes signale notamment l'absence totale d'inégalité sociale. La dernière question précise quelque peu ce point.

13. (a) Montrer que si $\varphi = \psi^*$ alors pour tout i appartenant à $[\![1, n]\!]$,

$$\frac{\varepsilon_i}{\varepsilon} = \frac{1-\varepsilon_{n-i+1}}{1-\varepsilon}.$$

Solution. L'idée est simple : la fonction φ a pour pentes successives $u_1, u_2, \ldots, u_n$ tandis que ψ^* a pour pentes successives $v_n, v_{n-1}, \ldots, v_1$. À condition de pouvoir identifier ces pentes terme à terme, on aura donc pour tout $i \in [\![1, n]\!]$, $u_i = v_{n-i+1}$, ce qui se traduit directement par $\frac{\varepsilon_i}{\varepsilon} = \frac{1-\varepsilon_{n-i+1}}{1-\varepsilon}$ d'après les formules établies aux questions 10 (b) et 10 (c).

Il reste donc à justifier l'identification. Celle-ci ne pose pas de problème lorsque $\varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_n$ car on dispose alors de n pentes distinctes ordonnées. En fait, la situation générale de l'énoncé peut toujours se ramener à ce cas plus simple, quitte à regrouper les catégories par valeur commune du coefficient ε_i (on diminue donc le nombre n). Nous abandonnons les détails à la sagacité du lecteur.

(b) Montrer que si $\varphi = \psi^*$, alors pour tout i appartenant à $[\![1, n]\!]$, $\varepsilon_i + \varepsilon_{n-i+1} = 2\varepsilon$.

Solution. Soit $i \in [\![1, n]\!]$ et soit $j = n-i+1$. D'après la question précédente, on sait que ε_i et ε_j sont reliés par l'égalité $\frac{\varepsilon_i}{\varepsilon} = \frac{1-\varepsilon_j}{1-\varepsilon}$. Mais puisque $n-j+1 = i$, on dispose aussi de $\frac{\varepsilon_j}{\varepsilon} = \frac{1-\varepsilon_i}{1-\varepsilon}$.

Ainsi $\dfrac{\varepsilon_i}{\varepsilon} + \dfrac{\varepsilon_j}{\varepsilon} = \dfrac{1-\varepsilon_j}{1-\varepsilon} + \dfrac{1-\varepsilon_i}{1-\varepsilon}$, d'où $(1-\varepsilon)(\varepsilon_i+\varepsilon_j) = \varepsilon(2-(\varepsilon_i+\varepsilon_j))$ et donc $\varepsilon_i+\varepsilon_j = 2\varepsilon$.

(c) Déduire que si $\varphi = \psi^*$, on a pour tout i appartenant à $[\![1,n]\!]$, $\varepsilon_i(1-2\varepsilon) = \varepsilon(1-2\varepsilon)$

Solution. En reprenant les notations introduites dans la question précédente, on a obtenu l'égalité $\varepsilon_j = 2\varepsilon - \varepsilon_i$, de sorte que $\dfrac{\varepsilon_i}{\varepsilon} = \dfrac{1-\varepsilon_j}{1-\varepsilon}$ peut s'écrire $\dfrac{\varepsilon_i}{\varepsilon} = \dfrac{1-2\varepsilon+\varepsilon_i}{1-\varepsilon}$, ou encore :

$$(1-\varepsilon)\varepsilon_i = \varepsilon(1-2\varepsilon+\varepsilon_i) = \varepsilon(1-2\varepsilon)+\varepsilon\varepsilon_i, \quad \text{puis enfin} \quad (1-2\varepsilon)\varepsilon_i = (1-\varepsilon)\varepsilon_i - \varepsilon\varepsilon_i = \varepsilon(1-2\varepsilon).$$

(d) On suppose que $\varepsilon \neq \frac{1}{2}$. Montrer que si $\varphi = \psi^*$, alors pour tout i appartenant à $[\![1,n]\!]$, $\varepsilon_i = \varepsilon$. Interpréter ce résultat.

Solution. Supposons que $\varphi = \psi^*$. Puisqu'on a aussi supposé $\varepsilon \neq \frac{1}{2}$, on peut diviser par $1-2\varepsilon \neq 0$ le résultat précédent :

$$\forall i \in [\![1,n]\!], \quad \varepsilon_i = \frac{\varepsilon_i(1-2\varepsilon)}{1-2\varepsilon} = \frac{\varepsilon(1-2\varepsilon)}{1-2\varepsilon} = \varepsilon.$$

Lorsque $\varphi = \psi^*$, chaque catégorie présente donc le même coefficient $\varepsilon_i = \varepsilon$, c'est-à-dire la même proportion $\frac{x_i}{n_i} = \varepsilon$ et *a fortiori* la même proportion $\frac{y_i}{n_i} = 1-\varepsilon$. Il y a alors une homogénéité parfaite entre les différentes catégories de la population du point de vue de la répartition dans les classes I et II. Autrement dit, on n'observe aucune inégalité sociale entre la classe I et la classe II.