Calcul de l'Indice de Gini
Exo 1 Question 4 du partiel de Décembre 2006

On reprend les données de cet exercice :

salaire mensuel	effectif de la classe i : n_i
[500, 1500[	50
[1500, 2500[	125
[2500, 5500[	25

On s'intéresse à la répartition **du salaire** (que l'on note x) sur la population **des salariés**. Il nous faut donc calculer :

1. les fréquences cumulées pour la variable de salaire x,
2. les fréquences cumulées pour la masse salariale.

I. Calcul des fréquences cumulées pour la variable x :

1. On calcule les fréquences (ou pourcentage) de chaque classe : pour la classe i, $f_i = n_i/n$, où n est l'effectif total (200 salariés).
2. On calcule les fréquences cumulées : $F(x)$.

On obtient :

salaire mensuel x	effectif de la classe, n_i	fréquence de la classe, f_i	fréquences cumulées $F(x)$
$[500, 1500[$	50	0,25	0,25
$[1500, 2500[$	125	0,625	0,875
$[2500, 5500[$	25	0,125	1

II. Calcul des fréquences cumulées pour la masse salariale :

1. Comme les salaires sont donnés **par classe**, on ne peut pas calculer exactement la masse salariale d'une classe (la somme de tous les salaires des personnes de cette classe). On fait l'approximation suivante :
   - On calcule le **centre de la classe** :
     pour la classe [500, 1500[, le centre (milieu de l'intervalle) est $x_1 = (500 + 1500)/2 = 1000$,
   - On approche la masse salariale de la classe i en faisant comme si les n_i personnes de cette classe gagnaient toutes x_i. La masse salariale de la classe est environ $n_i \times x_i$.
2. On calcule le pourcentage de la masse salariale de chaque classe par rapport à la masse salariale totale (ici $n_1x_1 + n_2x_2 + n_3x_3$). Pour la classe i : $g_i = \frac{n_ix_i}{n_1x_1+n_2x_2+n_3x_3}$.
3. On calcule les fréquences cumulées : $F(nx)$.

On obtient :

salaire mensuel x	effectif de la classe i n_i	Centre de la classe x_i	masse salariale de la classe n_ix_i	pourcentage de la masse salariale g_i	fréquences cumulées de la masse salariale $F(nx)$
[500, 1500[	50	1000	50000	0,125	0,125
[1500, 2500[	125	2000	250000	0,625	0,75
[2500, 5500[	25	4000	100000	0,25	1

III. Courbe de Lorentz

La courbe de Lorentz représente les fréquences cumulées **de la masse salariale** $F(nx)$ en fonction des fréquences cumulées pour la **variable salaire** $F(x)$.

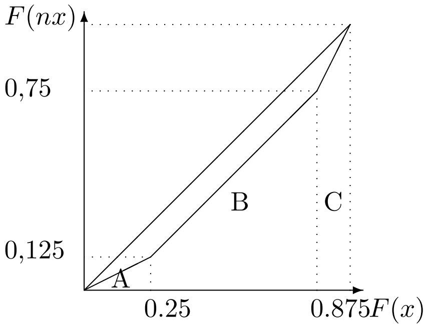

FIG. 1 – *Courbe de Lorentz.*

IV. Calcul de l'indice de Gini :

L'indice de Gini est **2 fois** l'aire entre la courbe de Lorentz et la première bissectrice.

En regardant la figure, on voit que l'aire entre la courbe de Lorentz et la première bissectrice est égale à :

- L'aire du triangle de côté 1 :
$$\frac{1 \times 1}{2} = \frac{1}{2}$$
- **MOINS** l'aire du triangle A :
$$\frac{0,25 \times 0,125}{2} = 0,0156$$
- **MOINS** l'aire du trapèze B. L'aire d'un trapèze étant donnée par :
$$\frac{\text{hauteur(grande base + petite base)}}{2},$$
l'aire du trapèze B est :
$$\frac{(0,875{-}0,25) \times (0,125{+}0,75)}{2} = 0,2734$$
(**Attention aux signes**) car :
  - La hauteur est ici le segment [0.25, 0.875] sur l'axe des abscisses, de longueur 0.875−0.25,
  - La grande base est le segment vertical de longueur 0.75,
  - La petite base est le segment vertical de longueur 0.125.
- **MOINS** l'aire du trapèze C :
$$\frac{(1{-}0,875) \times (0,75{+}1)}{2} = 0,1094.$$

Au final :
$$\text{Ind. Gini} = 2 \times \left(\frac{1}{2} - (0,0156 + 0,2734 + 0,1094)\right) = 0,2032.$$

(On retrouve le résultat du corrigé !)

Commentaire : L'indice de Gini appartient à [0,1] et plus il est proche de 1, plus la répartition des salaires est inégalitaire. Ici, comme 0,2032 est faible, la concentration des salaires est faible : les salaires sont assez bien répartis sur l'ensemble des salariés.