Exercice 4. Salaires des agents de l'état.

D'après l'INSEE, en 1986, la répartition en pourcentage des agents de l'état selon le montant de leur salaire annuel net (en milliers de francs), était :

Hommes		Femmes	
Salaire	%	Salaire	%
30-60	4,5	30-50	1,4
60-70	9,8	50-60	9,2
70-80	19,0	60-70	15,8
80-90	12,6	70-80	21,5
90-100	10,8	80-90	14,5
100-110	11,0	90-100	10,9
110-120	11,2	100-110	10,6
120-140	8,2	110-120	7,8
140-180	8,4	120-140	4,0
180-260	4,0	140-180	3,6
260-400	0,5	180-340	0,7

On admettra, dans chaque classe, une répartition uniforme des salaires.

1°/ Salaire des hommes.

a) Tracer l'histogramme des densités de fréquences.
b) Déterminer le mode.
c) Tracer la fonction de répartition empirique.
d) Déterminer la médiane.
e) Déterminer l'intervalle interquartile.
f) On appelle *rapport des déciles* le quotient du salaire au-dessus duquel se trouvent 10 % des agents, par celui en dessous duquel se trouvent 10 % des agents. Calculer ce rapport.

2°/ Salaires des femmes.

a) Déterminer le mode. Que conclure ?
b) Déterminer le rapport des déciles. Que peut-on conclure sur la comparaison des rémunérations entre les hommes et les femmes ?

3°/ Concentration.

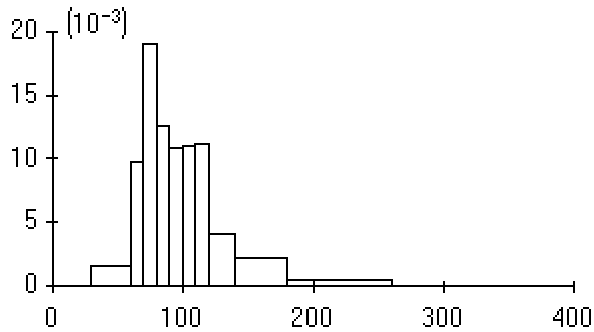
a) Tracer la courbe de concentration des salaires des hommes.
b) En déduire le coefficient de Gini.
c) Sachant que le coefficient de Gini pour les salaires des femmes est donné par 0,165, que peut-on en conclure ?

Solution.

1°/ Salaires des hommes.

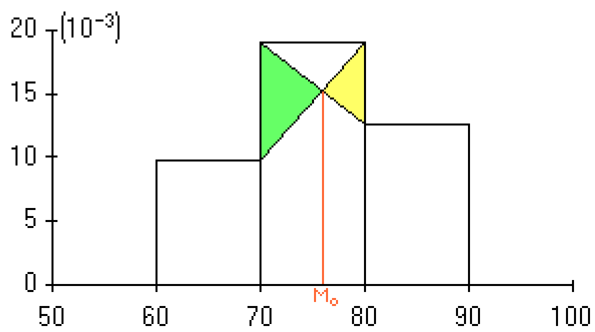
a) Histogramme des densités de fréquences.

Hommes		
Salaire	%	Densité de fréquence
30-60	4,5	$1,5 \times 10^{-3}$
60-70	9,8	$9,8 \times 10^{-3}$
70-80	19,0	$19,0 \times 10^{-3}$
80-90	12,6	$12,6 \times 10^{-3}$
90-100	10,8	$10,8 \times 10^{-3}$
100-110	11,0	$11,0 \times 10^{-3}$
110-120	11,2	$11,2 \times 10^{-3}$
120-140	8,2	$4,1 \times 10^{-3}$
140-180	8,4	$2,1 \times 10^{-3}$
180-260	4,0	$0,5 \times 10^{-3}$
260-400	0,5	$3,6 \times 10^{-5}$



b) Mode.

La classe modale est la classe de plus forte densité de fréquence, c'est la classe [70 ; 80 [.
Le mode se détermine, dans cette classe, en tenant compte des classes adjacentes.



Le mode M_O est donné par :

$$\frac{M_o - 70}{19{,}0 - 9{,}8} = \frac{80 - M_o}{19{,}0 - 12{,}6} = \frac{80 - 70}{2 \times 19{,}0 - 9{,}8 - 12{,}6} = \frac{10}{15{,}6} = \frac{25}{39}$$

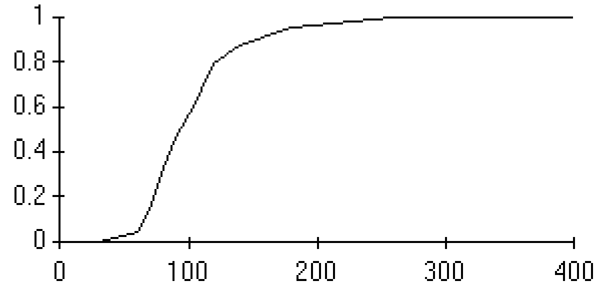
$$M_O = 70 + 9{,}2 \times \frac{25}{39} = 75{,}90$$

$$\boxed{M_O = 75{,}90 \text{ (milliers de francs).}}$$

c) Fonction de répartition empirique.

La fonction de répartition empirique s'obtient par la fréquence cumulée, en fonction de la limite supérieure de classe et en ajoutant 0 pour la limite inférieure de la première classe.

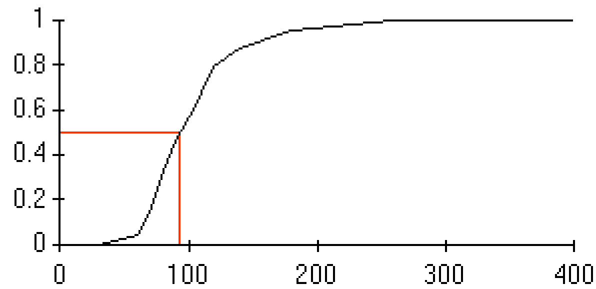
Hommes		
Salaire	%	Fréquence cumulée (%)
30	0,0	0,0
60	4,5	4,5
70	9,8	14,3
80	19,0	33,3
90	12,6	45,9
100	10,8	56,7
110	11,0	67,7
120	11,2	78,9
140	8,2	87,1
180	8,4	95,5
260	4,0	99,5
400	0,5	100,0



d) Médiane.

La médiane se lit sur le graphique de la fonction de répartition empirique : c'est l'abscisse du point d'ordonnée 0,5.
Numériquement, sa valeur s'obtient par interpolation linéaire entre les points (90 ; 0,459) et (100 ; 0,567).

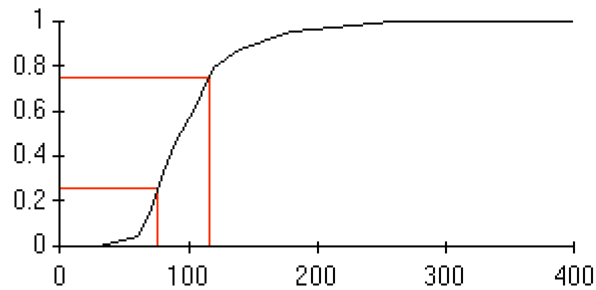
$M_e = 90 + (100 - 90) \times \frac{0{,}500 - 0{,}459}{0{,}567 - 0{,}459} = 90 + 10 \times \frac{41}{108} =$ 93,80



$\boxed{M_e = 93{,}80 \text{ (milliers de francs)}}$

e) Intervalle interquartile.

L'intervalle interquartile est la différence $Q_3 - Q_1$ entre le troisième quartile Q_3, abscisse du point d'ordonnée 0,75 sur le graphique de la fonction de répartition empirique, et le premier quartile Q_1, abscisse du point d'ordonnée 0,25 sur le graphique de la fonction de répartition empirique.
Numériquement, la valeur des quartiles s'obtient :
— Pour Q_3, par interpolation linéaire entre les points (110 ; 0,677) et (120 ; 0,789) :
$Q_3 = 110 + (120 - 110) \times \frac{0{,}750 - 0{,}677}{0{,}789 - 0{,}677} = 110 + 10 \times \frac{73}{112} = 116{,}52$



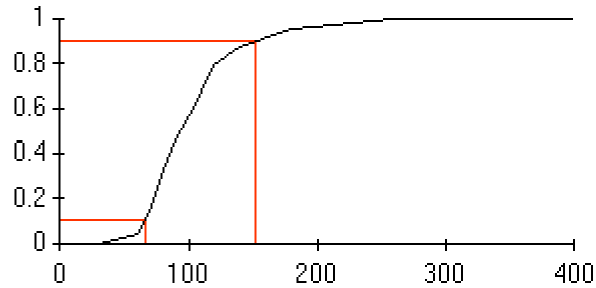
— Pour Q_1, par interpolation linéaire entre les points (70 ; 0,143) et (80 ; 0,333) :
$Q_1 = 70 + (80 - 70) \times \frac{0{,}250 - 0{,}143}{0{,}333 - 0{,}143} = 70 + 10 \times \frac{107}{190} = 75{,}63$

$$Q_3 - Q_1 = 110 + 10 \times \frac{73}{112} - 70 - 10 \times \frac{107}{190} = 40 + 10 \times \frac{943}{10\,640} = 40{,}89$$

$$\boxed{Q_3 - Q_1 = 40{,}89 \text{ (milliers de francs)}}$$

f) Rapport des déciles.

Le rapport des déciles est le rapport $\frac{D_9}{D_1}$ entre le neuvième décile D_9, abscisse du point d'ordonnée 0,90 sur le graphique de la fonction de répartition empirique, et le premier décile D_1, abscisse du point d'ordonnée 0,10 sur le graphique de la fonction de répartition empirique.



Numériquement, la valeur des déciles s'obtient :

— Pour D_1, par interpolation linéaire entre les points (60 ; 0,045) et (70 ; 0,143) :

$D_1 = 60 + (70 - 60) \times \frac{0{,}100 - 0{,}045}{0{,}143 - 0{,}045} = 60 + 10 \times \frac{55}{98} = 65{,}61$

— Pour D_9, par interpolation linéaire entre les points (140 ; 0,871) et (180 ; 0,955) :

$D_9 = 140 + (180 - 140) \times \frac{0{,}900 - 0{,}871}{0{,}955 - 0{,}871} = 140 + 40 \times \frac{29}{84} = 153{,}81$

$$\frac{D_9}{D_1} = \frac{140 + 40 \times \frac{29}{84}}{60 + 10 \times \frac{55}{98}} = \frac{4\,522}{1\,929} = 2{,}34$$

$$\boxed{\frac{D_9}{D_1} = 2{,}34}$$

2°/ Salaires des femmes.

a) Mode.

Pour calculer le mode, il faut d'abord, comme pour les hommes, calculer les **densités de fréquences** de chaque classe.

Femmes		
Salaire	%	Densité de fréquence
30-50	1,4	$0{,}7 \times 10^{-3}$
50-60	9,2	$9{,}3 \times 10^{-3}$
60-70	15,8	$15{,}8 \times 10^{-3}$
70-80	21,5	$21{,}5 \times 10^{-3}$
80-90	14,5	$14{,}5 \times 10^{-3}$
90-100	10,9	$10{,}9 \times 10^{-3}$
100-110	10,6	$10{,}6 \times 10^{-3}$
110-120	7,8	$7{,}8 \times 10^{-3}$
120-140	4,0	$2{,}0 \times 10^{-3}$
140-180	3,6	$0{,}9 \times 10^{-3}$

180-340	0,7	$0{,}04 \times 10^{-3}$

La classe modale est la classe de plus forte densité de fréquence, c'est la classe [70 ; 80 [.
Le mode se détermine, dans cette classe, en tenant compte des classes adjacentes.

Le mode M_O est donné par :

$$\frac{M_O - 70}{21{,}5 - 15{,}8} = \frac{80 - M_O}{21{,}5 - 14{,}5} = \frac{80 - 70}{2 \times 21{,}5 - 15{,}8 - 14{,}5} = \frac{10}{12{,}7} = \frac{100}{127}$$

$$M_O = 70 + 5{,}7 \times \frac{100}{127} = 74{,}49$$

$$\boxed{M_O = 74{,}49 \text{ (milliers de francs).}}$$

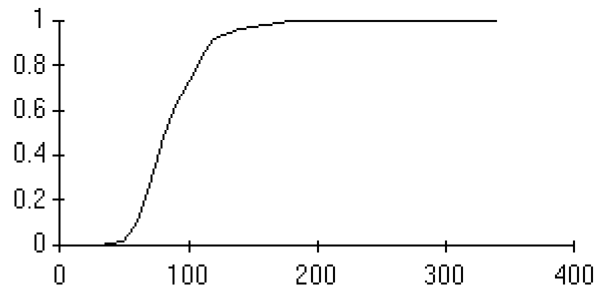
Même si le mode est légèrement inférieur chez les femmes, c'est bien dans la même tranche modale de salaires que chez les hommes que l'on trouve le plus de femmes salariées.

> Les modes du salaire sont pratiquement les mêmes chez les hommes et chez les femmes, aux alentours de 75 milliers de francs.

b) Rapport des déciles.

Pour pouvoir calculer le rapport des déciles $\frac{D_9}{D_1}$ chez les femmes, il faut d'abord calculer la fonction de répartition empirique.

Femmes		
Salaire	%	Fréquence cumulée (%)
30	0,0	0,0
50	1,4	1,4
60	9,2	10,6
70	15,8	26,4
80	21,5	47,9
90	14,5	62,4
100	10,9	73,3
110	10,6	83,9
120	7,8	91,7
140	4,0	95,7
180	3,6	99,3
340	0,7	100,0



Numériquement, la valeur des déciles s'obtient :
— Pour D_1, par interpolation linéaire entre les points (50 ; 0,014) et (60 ; 0,106) :
$D_1 = 50 + (60 - 50) \times \frac{0{,}100 - 0{,}014}{0{,}106 - 0{,}014} = 50 + 10 \times \frac{43}{46} = 59{,}35$
— Pour D_9, par interpolation linéaire entre les points (110 ; 0,839) et (120 ; 0,917) :
$D_9 = 110 + (120 - 110) \times \frac{0{,}900 - 0{,}839}{0{,}917 - 0{,}839} = 110 + 10 \times \frac{61}{78} = 117{,}82$

$$\frac{D_9}{D_1} = \frac{110+10\times\frac{61}{78}}{50+10\times\frac{43}{46}} = \frac{21137}{10647} = 1{,}99$$

$$\boxed{\frac{D_9}{D_1} = 1{,}99}$$

Si l'on compare cette valeur à la valeur de 2,34 obtenue pour les hommes, on voit que les salaires des femmes sont plus resserrés, donc moins dispersés, autour du mode, que ceux des hommes.

En regardant plus en détail la valeur des déciles, on voit que si les 10 % des femmes les moins bien payées gagnent moins de 59,35 milliers de francs, contre 65,61 milliers de francs pour les hommes (soit environ 10 % d'écart de salaire), en revanche, seulement 10 % des femmes gagnent plus de 117,82 milliers de francs, alors que les 10 % des hommes les mieux payés parmi les agents de l'état gagnent plus de 153,81 milliers de francs (soit environ 30 % d'écart de salaire) : c'est de là surtout que provient la différence entre les rapports des déciles des hommes et des femmes.

3°/ Concentration.

a) Concentration des salaires des hommes (courbe de Lorenz).

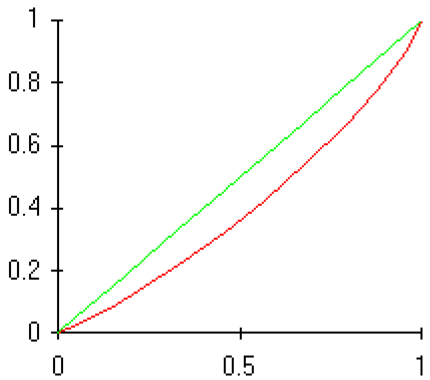
La courbe de concentration des salaires des hommes joint, par des segments de droite, les points ayant pour abscisse la fréquence cumulée d'une classe et, pour ordonnée, la proportion de masse salariale représentée par les salariés dont le salaire est plus petit que la limite supérieure de la classe.
Rappelons que l'interpolation linéaire suppose que la répartition des salaires à l'intérieur d'une classe est uniforme, de sorte que le pourcentage de masse salariale d'une classe est égal au centre de la classe, multiplié par la fréquence de la la classe, et divisé par la somme de ces produits (qui est le salaire moyen) $\frac{x_i f_i}{\sum x_j f_j}$.

Hommes		
Salaire	%	Fréquence cumulée (%)
30-60	4,5	4,5
60-70	9,8	14,3
70-80	19,0	33,3
80-90	12,6	45,9
90-100	10,8	56,7
100-110	11,0	67,7
110-120	11,2	78,9
120-140	8,2	87,1
140-180	8,4	95,5
180-260	4,0	99,5
260-400	0,5	100,0

$\overline{X}$ = å $x_j f_j$ = 45 × 0,045 + 65 × 0,098 + 75 × 0,190 + 85 × 0,126 + 95 × 0,108 + 105 × 0,110 + 115 × 0,112 + 130 × 0,082 + 160 × 0,084 + 220 × 0,040 + 330 × 0,005 = 102,595 kF.

Hommes

Centre de classe x_j	Fréquence f_j	Fréquence cumulée p_j	Fraction de masse salariale $x_j\,f_j$	Fraction de masse salariale cumulée q_j
		0,000		0,0000
45	0,045	0,045	0,0197	0,0197
65	0,098	0,143	0,0621	0,0818
75	0,190	0,333	0,1389	0,2207
85	0,126	0,459	0,1044	0,3251
95	0,108	0,567	0,1000	0,4251
105	0,110	0,677	0,1126	0,5377
115	0,112	0,789	0,1255	0,6632
130	0,082	0,871	0,1039	0,7671
160	0,084	0,955	0,1310	0,8981
220	0,040	0,995	0,0858	0,9839
330	0,005	1,000	0,0161	1,0000



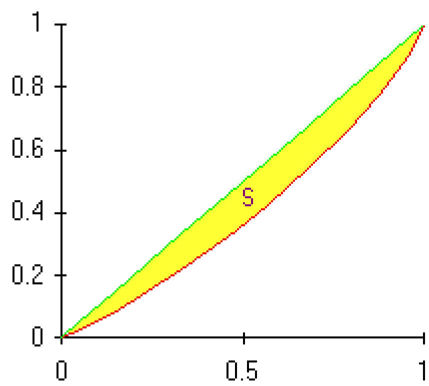
Courbe de Lorenz des salaires des hommes.

b) Indice de Gini.

L'indice de Gini est le double de la surface comprise entre la courbe de Lorenz et la diagonale principale, dans l'intervalle [0 ; 1]. Numériquement, il est donné par la formule :

$$g = 2\,S = 1 - \sum f_{j+1}\,(q_{j+1} + q_j)$$

g = 1 – (0,045 × (0,0197 + 0,0000) + 0,098 × (0,0818 + 0,0197) + 0,190 × (0,2207 + 0,0818) + 0,126 × (0,3251 + 0,2207) + 0,108 × (0,4251 + 0,3251) + 0,110 × (0,5377 + 0,4251) + 0,112 × (0,6632 + 0,5377) + 0,082 × (0,7671 + 0,6632) + 0,084 × (0,8981 + 0,7671) + 0,040 × (0,9839 + 0,8981) + 0,005 × (1,0000 + 0,9839)).



$\boxed{g = 0{,}199.}$

c) Comparaison avec l'indice de Gini des salaires des femmes.

L'indice de Gini des salaires des femmes est 0,165, d'après l'énoncé.
L'indice de Gini des salaires des hommes est 0,199, comme on vient de le calculer.
Celui des hommes est plus élevé que celui des femmes : cela traduit une plus grande inégalité de salaires chez les agents masculins de l'état que chez les agents féminins de l'état.

On avait déjà remarqué une plus grande dispersion des salaires masculins : parmi les hauts salaires, il y a surtout des hommes, alors que, pour les bas salaires, les proportions d'hommes et de femmes sont sensiblement équivalentes.