



Partiel novembre 2004

— ÉNONCÉ —

On étudie la distribution selon la distance de « domicile familial – travail » des salariés d'une entreprise. Après répartition par classes, on obtient le tableau suivant :

distances (km)	effectifs
[1 ; 5[250
[5 ; 10[328
[10 ; 20[165
[20 ; 30[91
[30 ; 50[46

Question 1– Quelle proportion de salariés habite à moins de 20 km de l'entreprise ?

Question 2– Tracer l'histogramme correspondant à cette série.

Question 3– Déterminer la valeur de la médiane.

Question 4– L'écart interquartiles est 10,45 et le troisième quartile est 14,97 (on ne demande pas de calculer ces valeurs). Déterminer le premier quartile.

Question 5– Calculer la moyenne correspondant à cette série. Donner la formule utilisée.

Question 6– Calculer l'écart-type.

— SOLUTION —

Réponse à la question 1– Le nombre de salariés habitant à moins de 20 km de l'entreprise est

$$250 + 328 + 165 = 743.$$

Le nombre total de salariés de l'entreprise est

$$250 + 328 + 165 + 91 + 46 = 880.$$

La proportion de salariés habitant à moins de 20 km de l'entreprise est donc 84,4% puisque

$$\frac{743}{880} = 84,4\%.$$

Réponse à la question 2– Dans l'histogramme de la distribution d'un caractère continu, chaque modalité est représentée par un rectangle dont la longueur est proportionnelle à l'amplitude de la modalité et la surface est proportionnelle à son effectif. La hauteur d'un rectangle est donc proportionnelle au rapport (appelé densité) de l'effectif sur l'amplitude de sa modalité correspondante. On résume les amplitudes et densités des modalités dans le tableau suivant et l'histogramme est représenté sur la figure 1 page suivante.

modalités	[1 ; 5[[5 ; 10[[10 ; 20[[20 ; 30[[30 ; 50[
amplitudes	4	5	10	10	20
densités	62,5	65,6	16,5	9,1	2,3

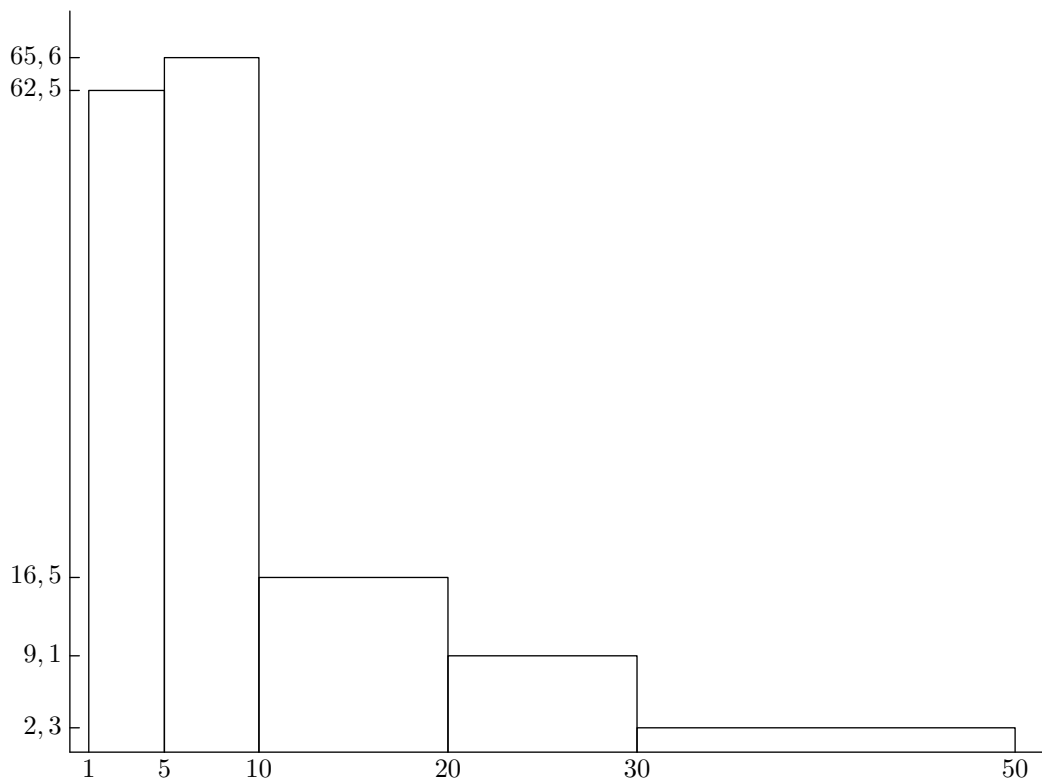


FIG. 1 – Histogramme de la distribution

Réponse à la question 3– La fréquence cumulée de la modalité $[1 ; 5[$ est

$$\frac{250}{880} = 28,4\%$$

et la fréquence cumulée de la modalité $[5 ; 10[$ est

$$\frac{250 + 328}{880} = 65,7\%.$$

La médiane est la valeur en laquelle la fonction de distribution cumulée (extrapolée linéairement entre les valeurs connues) prend la valeur 50%. Elle est donc située dans l'intervalle $[5 ; 10[$ et on la calcule à l'aide (par exemple) du théorème de Thalès¹ (voir la figure 2 page suivante). On a

$$\frac{m - 5}{10 - 5} = \frac{50 - 28,4}{65,7 - 28,4} \quad \text{donc} \quad m - 5 = (10 - 5) \times \frac{50 - 28,4}{65,7 - 28,4}$$

puis

$$m = 5 + (10 - 5) \times \frac{50 - 28,4}{65,7 - 28,4} = 7,9.$$

Ainsi, la médiane est de 7,9 km.

Réponse à la question 4– Si Q_1 et Q_3 sont, respectivement, les premiers et troisièmes quartiles, l'espace interquartiles est $\Delta = Q_3 - Q_1$. Ainsi $Q_1 = Q_3 - \Delta$ et on trouve

$$Q_1 = 14,97 - 10,45 = 4,52.$$

Le premier quartile est 4,52 km.

Réponse à la question 5– Les modalités étant des intervalles, le calcul de la moyenne et de l'écart-type se fait en remplaçant chaque intervalle par son centre. On rappelle que le centre c d'un intervalle $[a, b[$ se calcule par la formule

$$c = \frac{a + b}{2}.$$

modalités	$[1 ; 5[$	$[5 ; 10[$	$[10 ; 20[$	$[20 ; 30[$	$[30 ; 50[$
centres	3	7,5	15	25	40
effectifs	250	328	165	91	46

¹Qui fut mathématicien bien avant que des ambitieux ne donnent son nom à une entreprise.

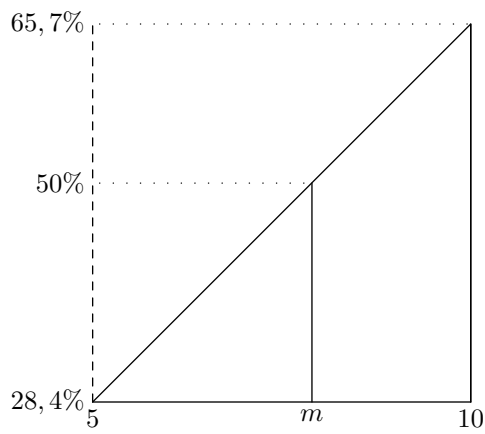


FIG. 2 – Calcul de la médiane à l'aide du théorème de Thalès

Ainsi

$$\text{Moy} = \frac{3 \times 250 + 7,5 \times 328 + 15 \times 165 + 25 \times 91 + 40 \times 46}{880} = 11,136.$$

La distance moyenne entre le domicile et le lieu de travail est 11,1 km.

Réponse à la question 5– La moyenne des carrés est

$$\frac{3^2 \times 250 + 7,5^2 \times 328 + 15^2 \times 165 + 25^2 \times 91 + 40^2 \times 46}{880} = 213,977.$$

La variance est donc

$$\text{Var} = 213,977 - 11,136^2 = 89,97$$

et l'écart-type est

$$\sigma = \sqrt{\text{Var}} = \sqrt{89,97} = 9,5.$$

L'écart-type de la distance entre le domicile et le lieu de travail est 9,5 km.