



MIASS 231
Mathématiques (appliquées aux sciences sociales) 3
© El Hadj Touré, 2022

DIX EXERCICES RÉCAPITULATIFS
Leçon 4. Mesures de tendance centrale
(Solutionnaire)

1) Comparez, en termes de différence et de similitude, les paires de concepts ci-dessous :

a) Distributions de fréquences/pourcentages et mesures de tendance centrale

Les distributions de fréquences/pourcentages et les mesures de tendance centrale relèvent toutes des **statistiques descriptives** : elles permettent de décrire des données d'échantillon relatives à des variables.

Toutefois, alors que les premières aident à connaître l'**ampleur d'un phénomène** à l'aide de **plusieurs résumés statistiques**, les secondes cherchent à établir la **représentativité d'un phénomène** à l'aide d'un **résumé statistique précis**.

b) Médiane et moyenne

La médiane et la moyenne sont des **mesures de tendance centrale**, qui aident à décrire le score typique d'une distribution.

Toutefois, alors que la médiane ne prend en compte que l'**ordre des scores** par leur mise en rang, la moyenne comptabilise la **grandeur des scores** par leur sommation. Par conséquent, si la moyenne est sensible aux **scores extrêmes**, ce n'est pas le cas de la médiane.

2) Répondez aux questions suivantes à propos d'une mesure de tendance centrale

a) Qu'est-ce qu'une mesure de tendance centrale?

Une mesure de tendance centrale est une **statistique descriptive** qui cherche à établir autour de quelle valeur les scores d'une distribution sont agglutinés. Il en est ainsi du mode, de la médiane et de la moyenne.

b) À quoi sert-elle précisément en termes d'étude d'un phénomène social?

Une mesure de tendance centrale sert à décrire la **représentativité** d'un phénomène par le biais du **score typique, le plus commun**.

3) Résolvez les problèmes ci-dessous liés à l'utilisation et l'interprétation des mesures de tendance centrale.

a) Un magasin qui vend des chaussures de femmes veut connaître les deux pointures d'une espadrille les plus représentatives dans la perspective de doubler les commandes pour ces deux pointures. Quelle mesure de tendance centrale est plus appropriée : le mode; la médiane ou la moyenne; la médiane et la moyenne?

Les responsables du magasin ont avantage à choisir les deux pointures les plus fréquemment achetées par les femmes. Contrairement aux pointures médiane et moyenne, les pointures modales sont plus susceptibles de leur assurer une grande clientèle. **On ne porte pas nécessairement une pointure moyenne ou médiane, mais on porte précisément une pointure modale.**

NB : En effet, le mode c'est la valeur à la mode. Lorsque des phénomènes relatifs à la mode sont en jeu, il faut tenir compte du mode en conséquence 😊

b) Le *lock-out* déclenché en septembre 2012 nous offre l'opportunité de réfléchir sur la façon dont les mesures de tendance centrale peuvent faire l'objet d'une instrumentalisation de la part des acteurs impliqués dans le Hockey. Supposons qu'on est dans un processus de négociation des salaires entre la **Ligue nationale de Hockey** (propriétaires) et le **président de l'Association des joueurs**. Pour défendre ses positions,

- lequel des deux camps est plus susceptible d'invoquer le salaire moyen des joueurs
- et lequel va avoir tendance à invoquer le revenu médian des joueurs?

Justifiez votre réponse!

D'un côté, la Ligue nationale de hockey représentée par les propriétaires aura tendance à invoquer et à se targuer du **salaire moyen** des joueurs pour supporter sa position, sachant que le salaire moyen va être tiré par les salaires hautement élevés des vedettes comme Brad Richards, Vincent Lecavalier, Sidney Crosby, Alexander Ovechkin.

De l'autre côté, le président de l'Association des joueurs aura plutôt tendance à invoquer et décrier le **salaire médian** des joueurs pour supporter sa position, étant entendu que le salaire médian n'est pas affecté par les salaires faramineux des joueurs vedettes.

NB : Contrairement à la médiane, la moyenne est affectée par les scores extrêmes. Par conséquent, les inconvénients et avantages de ces mesures de tendance centrale peuvent être instrumentalisés par les utilisateurs des statistiques selon les intérêts et enjeux du moment. Lorsqu'une distribution comporte des cas déviants, à défaut de les exclure pour utiliser la moyenne, vous gagnerez à vous fier plutôt à la médiane. En définitive, il faut bien réfléchir... 😊

4) Selon les données de l'institut de Statistique du Québec, en 2008, les revenus moyen et médian par ménage de 2 personnes sont ainsi présentés pour le Canada et le Québec:

Caractéristiques	Canada	Québec
Revenu moyen \$	73494	62789
Revenu médian \$	59000	55000

a) De combien en pourcentages, le revenu moyen est-il supérieur au revenu médian au Canada en 2008?

$$\text{Taux} = \frac{\text{Valeur d'un élément donné} - \text{Valeur d'un élément de référence}}{\text{Valeur d'un élément de référence}} * 100$$

$$\text{Taux d'augmentation} = ((73494 - 59000)/59000) * 100 = \mathbf{24,6\%}$$

Au Canada, le revenu moyen est supérieur au revenu médian par près de 25% en 2008.

b) De combien en pourcentages, le revenu moyen est-il supérieur au revenu médian au Québec en 2008?

$$\text{Taux} = \frac{\text{Valeur d'un élément donné} - \text{Valeur d'un élément de référence}}{\text{Valeur d'un élément de référence}} * 100$$

$$\text{Taux d'augmentation} = ((62789 - 55000)/55000) * 100 = \mathbf{14,2\%}$$

Au Québec, le revenu moyen est supérieur au revenu médian par plus de 14% en 2008.

c) Comparez les deux pourcentages et avancez une tentative d'explication de la différence calculée.

Au Canada, le revenu moyen est davantage supérieur au revenu médian qu'au Québec, la différence étant de **10,4%** (24,6%-14,2%). Cette différence relative peut s'expliquer par le fait qu'au Canada, il y a plus de salaires élevés dans les ménages, et ces salaires ont tendance à tirer le revenu moyen dans leur sens. Il semble donc y avoir plus de disparités entre les Canadiens qu'entre les Québécois en termes de revenus (les mesures de dispersion, que nous verrons la semaine prochaine, fourniront une réponse plus précise). **En clair, pour le cas canadien, le revenu médian est plus approprié que le revenu moyen pour rendre compte de la tendance centrale de la distribution des salaires.**

NB : Lorsqu'on veut établir la tendance centrale d'une distribution, on a avantage à calculer et comparer la moyenne et la médiane. De surcroît, les taux d'augmentation en % permettent de comparer deux distributions ou plus 😊

5) Pour chacune des situations ci-dessous, dites si vous utiliseriez la moyenne ou la médiane pour rendre compte de la tendance centrale des revenus. Justifiez votre choix.

- a) Un chercheur souhaite connaître le revenu le plus représentatif de la famille typique canadienne pour mieux rendre compte du niveau de vie ou du pouvoir d'achat standard des Canadiens.

Pour mieux rendre compte du pouvoir d'achat standard de la famille typique canadienne, le chercheur aurait intérêt à utiliser le revenu médian. Car, avec la médiane, il s'assurera que les revenus très élevés ou très faibles n'affecteront pas la mesure du revenu le plus représentatif d'un Canadien-type.

- b) Un chef d'entreprise veut connaître le montant le plus représentatif qu'il a dépensé par employé dans le but d'avoir une meilleure idée sur la tendance centrale du coût salarial défrayé.

Le chef d'entreprise gagnerait à utiliser le revenu moyen pour avoir une meilleure idée sur le coût salarial défrayé par employé. Car, avec la moyenne, il prendrait en compte tous les salaires des employés (incluant les cadres et ouvriers), leur faiblesse et leur grandeur.

6) Une étude porte sur le nombre d'heures passées sur Internet/par semaine chez les adolescents. Pour chacun des deux échantillons ci-dessous, calculez la médiane et interprétez-la.

- a) Échantillon 1 (n=7)

Cas 1	Cas 2	Cas 3	Cas 4	Cas 5	Cas 6	Cas 7
24	14	17	10	7	5	22

Pour trouver la médiane, la première étape consiste à arranger les scores (nombre d'heures passées sur Internet/par semaine) du plus petit au plus grand.

5	7	10	14	17	22	24
---	---	----	-----------	----	----	----

Le nombre de cas étant impair (n=7), la médiane correspond au score du quatrième cas dans la mise en ordre $((7+1)/2=4^{\text{e}}$ score). Le nombre d'heures médian est 14. **Au moins 50% des adolescents de l'échantillon passent 14 heures/par semaine ou moins sur Internet.**

- b) Échantillon 2 (n=6)

Cas 1	Cas 2	Cas 3	Cas 4	Cas 5	Cas 6
24	14	17	10	7	5

Pour trouver la médiane, la première étape consiste à arranger les scores du plus petit au plus grand.

5	7	10	14	17	24
---	---	-----------	-----------	----	----

Le nombre de cas étant pair (n=6), la médiane correspond au score point milieu des deux cas centraux dans la mise en ordre $((6+1)/2=3,5^{\text{e}}$ score). Le nombre d'heures

médian est $(10+14)/2=12$. **Au moins 50% des adolescents de l'échantillon passent 12 heures/par semaine ou moins sur Internet.**

NB : Prenez note que la formule $(n+1)/2$ ne donne pas directement la médiane, mais la position de la médiane dans la mise en ordre des cas. La médiane correspond donc à un score séparant une distribution en deux parties égales. Commencez donc toujours par ordonner les cas selon leurs scores 😊

7) À partir du tableau suivant des âges ($n = 70$),

Classes d'âge	Fréquence	Fréquence cumulée
[16-18[6	6
[18-20[33	39
[20-22[20	59
[22-24[11	70
Total (n)	70	

- a) Déterminez le mode, interprétez statistiquement le résultat et caractérisez la distribution

La classe modale est 18-20. **Autrement dit, la plupart des répondants de l'échantillon ($33/70=47\%$) sont âgés de 18 à 20 ans.** La distribution est unimodale.

- b) Quelle hypothèse devez-vous accepter à propos des classes afin de pouvoir estimer la moyenne?

Pour calculer la moyenne, nous avons besoin de prendre en compte le **point milieu des classes**. Par conséquent, l'hypothèse est faite que **les classes sont réellement homogènes**, c'est-à-dire que les âges sont répartis de façon uniforme dans chaque intervalle.

- c) Calculez la moyenne et interprétez statistiquement le résultat

Pour calculer la moyenne, prenons le **point milieu de chaque classe et multiplions-le par sa fréquence !**

Le point milieu s'obtient en additionnant les limites inférieure et supérieure de la classe et en divisant par 2 (par exemple, pour la première classe: $(16+18)/2=17$)

Point milieu X	Fréquence f	Produit X*f
17	6	102
19	33	627
21	20	420
23	11	253
Total (n)	70	1402

Une fois calculée la somme des produits entre les scores et leurs fréquences, divisons le résultat par le nombre de cas (n).

Voici la formule !

$$\bar{X} = \frac{\sum Xf_i}{n}$$

Moyenne = $1402 / 70 = 20,03$.

La moyenne est de 20,03. **Autrement dit, l'âge moyen des répondants de l'échantillon est égal à 20 ans.**

d) Estimez approximativement la médiane à l'aide de l'ogive et interprétez statistiquement le résultat

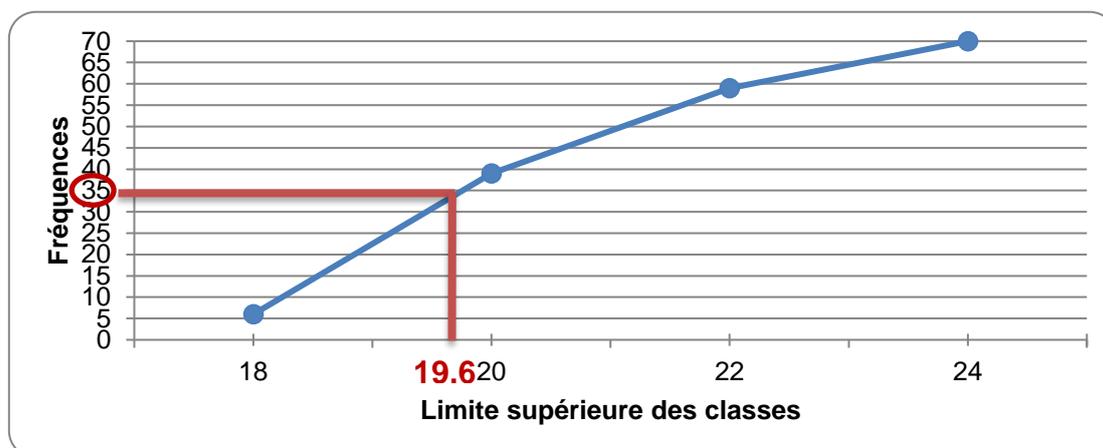
Commençons par diviser la taille de l'échantillon en deux parties égales:

$$n/2 = 70 / 2 = 35$$

La médiane est donc le 35^e score. Pour le retrouver, il suffit de tracer une ligne horizontale au niveau du 35^e score et une ligne verticale à partir de l'intersection pour obtenir la médiane, qui est estimée environ à 19,6.

La médiane est estimée approximativement à 19,6. **Autrement dit, au moins 50% des répondants sont âgés de 19,6 ans ou moins.**

NB : Comme l'ordre des scores bruts est noyé dans les fréquences, l'estimation de la médiane ne peut être qu'approximative.



8) Supposons qu'un étudiant de la classe ait obtenu en mathématiques appliquées aux sciences sociales les notes (sur 20) suivantes :

Quiz: 14

Examen intra: 15

Examen final: 16

Trouvez sa moyenne pour la session si la pondération de chaque évaluation est: Quiz: 20% Examen intra: 40% Examen final: 40%

20% de 14 donne $20\% \cdot 14 = (20 \cdot 14) / 100 = 31,5$

40% de 15 donne $40\% \cdot 15 = (40 \cdot 15) / 100 = 20$

40% de 16 donne $40\% \cdot 16 = (40 \cdot 16) / 100 = 27$

La moyenne pondérée de l'étudiant est égale à 78,5%, soit $31,5+20+27$.

9) Pour chacune des deux situations ci-dessous, calculez la moyenne combinée.

a) Calculez la moyenne combinée sur 10 à un examen pour l'ensemble des trois groupes ci-dessous

Groupe	n	Moyenne	n*Moyenne
1	10	7	70
2	20	6	120
3	30	5	150
Somme	60		340

La moyenne combinée à l'examen pour l'ensemble des trois groupes est égale à $340/60$, soit 5,66.

b) Calculez le salaire moyen des 100 employés d'une entreprise sachant que les 10 cadres administrateurs gagnent en moyenne 65 000\$, les 30 ingénieurs-techniciens 50 000\$ et les 60 ouvriers-mécaniciens 35 000\$.

$$((65000*10) + (50000*30) + (35000*60) / 100 = 42\ 500$$

Le salaire moyen des 100 employés de l'entreprise est égal à 42 500\$.

10) Vous vous intéressez à la tendance centrale de l'espérance de vie en années afin de décrire en quoi le phénomène est représentatif au sein de 10 pays du monde. Vous possédez les informations suivantes :

Individus (cas)	Espréance de vie
Allemagne	81
Argentine	77
Australie	83
Canada	82
Caméroum	59
Chine	77
Nigéria	55
Roumanie	75
Sénégal	68
Uruguay	78

a) Cette distribution comporte-t-elle un mode? Si oui, quelle est l'espérance de vie modale ?

La distribution comporte un mode : 77, soit le score qui revient le plus fréquemment dans la série de scores. En d'autres termes, pour une pluralité de pays, l'espérance de vie est de 77 ans.

b) Calculez la moyenne de l'espérance de vie de cette distribution. Interprétez-la statistiquement.

$$\bar{x} = \frac{81 + 77 + 83 + 82 + 59 + 77 + 55 + 75 + 68 + 78}{10} = 73,5.$$

La moyenne est de 73,5. Autrement dit, l'espérance de vie moyenne est de 73,5

ans au sein des 10 pays.

- c) Calculez la médiane de l'espérance de vie de cette distribution. Interprétez-la statistiquement.

Ici, le nombre de cas est pair. La médiane est la moyenne des scores de rang 5 et 6.

Rang	1	2	3	4	5	6	7	8	9	10
%. financement	55	59	68	75	77	77	78	81	82	83

Ici, la médiane est 77 ($77+77/2$). **Cela signifie que pour au moins 50% des pays, l'espérance de vie est de 77 ans ou moins.**

- d) Dichotomiser l'espérance de vie (EV) selon le score médian comme point de césure : Faible si $EV \leq$ score médian ; Élevée si $EV >$ Score médian. Quel est le pourcentage de pays ayant une espérance de vie faible?

Le pourcentage de pays ayant une espérance de vie faible, c.à.d dont le score est inférieur ou égal à 77 ans, est 60% (6/10).

NB : Vous comprenez donc pourquoi la médiane s'interprète ainsi : « Au moins 50% des cas ont un score donné ou moins ». Il faut bien distinguer le score médian de la proportion de cas inférieure précisément au score médian 😊