

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

DIX EXERCICES RÉCAPITULATIFS
Leçon 3. Fréquences et pourcentages
Solutionnaire

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

a) Fréquence et effectif

Les fréquences et les effectifs réfèrent au décompte du nombre de cas correspondant aux valeurs d'une variable. Elles révèlent des occurrences.

Toutefois, les effectifs comportent une *acception minimaliste* puisque le nombre de cas (unités statistiques) fait surtout référence à des personnes. L'exemple est donné de l'effectif d'une classe (20 étudiants), d'une entreprise (20 employés), etc.

Par contre, les fréquences comportent une *acception maximaliste* en ce sens que le nombre de cas peut-être aussi bien des personnes, des objets ou des évènements. L'exemple est donné de la fréquence d'un type de crime (5 dossiers criminels pour meurtres, 10 pour vols...), d'un type d'élections (10 élections non démocratiques, 20 élections démocratiques dans le monde entre 2010 et 2012...). C'est la raison pour laquelle nous userons du terme « fréquence » dans ce cours.

b) Fréquence absolue et fréquence relative

Les fréquences absolues et relatives sont des fréquences qui révèlent le nombre de fois qu'une valeur apparaît dans un ensemble de données. Elles permettent de connaître l'ampleur ou l'importance d'un phénomène.

Toutefois, les fréquences absolues réfèrent au **décompte originel** du nombre de cas (dont le total constitue l'échantillon) correspondant aux valeurs d'une variable : par exemple, 17 sur 30 personnes sondées croient qu'il faut suivre sa conscience.

Par contre, les fréquences relatives renvoient à la **standardisation** du nombre de cas total à une valeur de référence telle que 100 : par exemple, 57% ($17/30 \cdot 100$) des personnes sondées affirment qu'il faut suivre sa conscience.

c) Diagramme en barres et histogramme

Le diagramme en barres et l'histogramme sont des techniques graphiques qui permettent de représenter des données d'une variable de façon visuelle. Ils ont ceci de commun qu'ils représentent les distributions sous forme de barres.

Toutefois, les barres du diagramme en barres ne se touchent jamais pour illustrer le fait qu'elles représentent surtout les **catégories** d'une variable qualitative. Par contre, les barres de l'histogramme se touchent toujours pour rappeler le **caractère continu** de la variable dont les valeurs sont transformées en classes.

d) Diagramme en barres et diagramme en bâtons linéaires

Le diagramme en barres et le diagramme en bâtons linéaires sont des techniques graphiques qui permettent de représenter des données d'une variable de façon visuelle. Ils ont ceci de commun qu'ils représentent les distributions sous forme de barres séparées les unes des autres.

Toutefois, les barres du diagramme en barres sont plus larges pour illustrer le fait qu'elles représentent surtout les **catégories** d'une variable qualitative. Par contre, les barres du diagramme en bâtons linéaires sont très étroites pour illustrer le **caractère discret** de la variable quantitative dont on veut visualiser les données.

2) Répondez aux questions suivantes en rapport avec une distribution.

- a) Qu'est-ce qu'une distribution de fréquences? À quoi sert-elle précisément en termes d'étude d'un phénomène social?

Une distribution de fréquences met en évidence la façon dont les cas sont distribués par rapport aux valeurs d'une variable.

Elle sert à résumer, décrire les données d'une variable de façon à connaître **l'ampleur ou l'importance d'un phénomène**.

- b) Qu'est-ce qu'une distribution de fréquences (%) cumulatives? Quelle est son utilité en sciences sociales?

Une distribution de fréquences (%) cumulatives réfère à la somme de toutes les fréquences (%) en partant de la valeur (de la variable) la plus faible. Elle donne toujours un total de 100 et requiert une variable ordinale ou quantitative.

Elle sert à répondre à des questions du genre : « **Quel est le nombre de cas (ou %) ayant un score donné ou inférieur?** ». Par exemple, quel est le nombre de personnes (ou %) qui gagnent 50 000\$ ou moins au Canada? La technique graphique qu'est l'ogive permet de trouver facilement une fréquence cumulative.

3) Résolvez les problèmes ci-dessous liés à l'utilisation et l'interprétation des fréquences et pourcentages, ainsi qu'à la gestion des données manquantes.

- a) Un analyste avance que 30% des élèves du secondaire fument la cigarette. Peut-on avoir confiance en cette conclusion si l'on sait que l'analyste n'a sondé qu'un groupe d'amis constitué de 20 élèves? Justifiez!

On ne peut avoir confiance à un sondage effectué auprès de 20 cas seulement. De plus, les élèves constituant un groupe d'amis, il y a de fortes de chances qu'un fumeur entraîne les autres. Une étude plus sérieuse situerait le pourcentage d'élèves fumeurs entre 14% et 18%, au Québec à tout le moins.

Pour avoir confiance aux résultats d'un sondage, l'échantillon doit être représentatif de la population, de par sa taille (un minimum de 30 cas) et de son caractère probabiliste ou aléatoire (sur lequel nous reviendrons dans la leçon 7).

NB : On utilise parfois des pourcentages en arrière-plan duquel se cache un faible nombre de cas; d'où l'importance de toujours mentionner le nombre de cas pour une meilleure interprétation des pourcentages. Contextualisez donc vos pourcentages en mentionnant le nombre de cas sur lequel leur calcul repose, au risque de confondre le lecteur 😊

b) En l'espace d'un an, le taux de chômage d'un pays est passé de 5% à 4,8%, soit un taux de diminution de 4% $((4,8-5/5)*100)$. Peut-on affirmer, à la suite de cette diminution de 4%, qu'il ne reste plus que 1% comme taux de chômage? Justifiez!

On ne peut affirmer que le taux de chômage n'est plus que de 1% à la suite d'une diminution de 4%. En fait, le taux de diminution (-4%) est une différence relative : autrement dit, si l'on standardisait l'écart entre deux valeurs (deux taux de chômage) à 100, que représenterait cet écart par rapport à la valeur de référence. Or, cette différence relative est différente de la différence absolue : $4,8\% - 5\% = -0,2\%$.

NB : Par conséquent, il faut toujours prendre soin de distinguer les mesures absolues des mesures relatives pour apprécier l'ampleur d'une différence 😊

c) Une étude porte sur le racisme dans 50 pays afin d'y mesurer le niveau de tolérance de ses habitants. À une des questions « Accepteriez-vous d'avoir comme voisin une personne d'une autre origine ethnique que la vôtre ? », correspondent trois réponses : 1.Oui; 2.Non; 9.Ne sait pas/Pas de réponse. Le chercheur décide de considérer les « non-réponses » comme des données manquantes et de les exclure systématiquement de son analyse. La décision du chercheur est-elle judicieuse ? Auriez-vous un avis éclairé?

Je ne crois pas que la décision du chercheur soit judicieuse eu égard au problème de recherche. Au fait, lorsqu'une étude porte sur des questions « tabous », comme le racisme, les non-dits comportent souvent une signification. Les personnes n'ayant pas répondu (NSP, NR) ont probablement une attitude intermédiaire ou ne veulent pas assumer leur réponse. Dans ce cas, l'inclusion des valeurs manquantes apporterait une information importante : le fait que le racisme renvoie à une attitude dont il est difficile d'exprimer la réalité face au regard pesant de la société. Surtout que les peuples n'ont pas la même liberté d'expression.

NB : Lorsqu'on doit exclure des données manquantes, il faut beaucoup réfléchir, car les sans-réponses parlent parfois d'elles-mêmes 😊

4) En tant qu'analyste statisticien au ministère de l'éducation, vous devez présenter un rapport technique devant décrire l'évolution du taux de

décrochage scolaire dans les écoles secondaires sénégalaises. Vous disposez des données sur le taux de décrochage annuel de 2000 à 2020. Comment allez-vous y prendre dans la présentation des résultats? Allez-vous privilégier un graphique ou un tableau ou allez-vous recourir à la fois aux deux techniques? Expliquez-vous?

En tant qu'analyste dans un milieu professionnel, je présenterais les résultats en privilégiant un graphique, soit la courbe de tendance. Cette technique graphique permet de décrire l'évolution du taux de décrochage scolaire, de telle sorte que l'on puisse visualiser rapidement les tendances. Or, un tableau de distribution oblige mon chef à aller chercher dans les détails, non sans difficulté, pour dégager les tendances. Par ailleurs, il n'est pas question que je présente l'évolution du taux de décrochage en recourant aux deux techniques pour éviter de doubler l'information.

NB : Si en milieu professionnel, on a tendance à privilégier les techniques graphiques, le patron ayant peu de temps pour comprendre l'essentiel, dans les milieux de recherche, on a tendance plutôt à privilégier les techniques tabulaires, le chercheur préférant souvent chercher dans les détails des chiffres 😊

5) Complétez le tableau ci-dessous en calculant les % et les % cumulés :

Classes d'âge	Fréquence	Pourcentages	Pourcentages cumulés
[16-18[6	8,6	8,6
[18-20[33	47,1	55,7
[20-22[20	28,6	84,3
[22-24[11	15,7	100,0
Total (n)	70	100,0	

a) Quelle est l'intervalle (longueur) des classes?

L'intervalle des classes est égale à 2, soit 18-16.

NB : La variable originale étant continue, ses valeurs ont été arrondies. Et donc, pour la première classe par exemple, [16-18[, on considère qu'elle comprend tous les cas qui sont âgés de 16 à 17,9 ans. Aussi, l'utilisation des crochets facilite le calcul de la longueur des classes : on a juste à prendre la limite sup. MOINS la limite inf. 😊

b) Quel est le pourcentage de cas âgés entre 20 et 22 ans?

Le pourcentage de cas âgés entre 20 et 22 ans est de 28,6.

c) Quel est le pourcentage de cas âgés de 22 ans ou moins?

Le pourcentage de cas âgés de 22 ans ou moins est de 84,3.

d) Quel est le pourcentage de cas âgés d'au plus 22 ans?

Le pourcentage de cas âgés d'au plus 22 ans est de 84,3.

e) En montrant les calculs, retrouvez les fréquences à partir des pourcentages

$$\% = \frac{f}{n} * 100$$

$$\% * n = f * 100$$

$$f = \frac{\% * n}{100}$$

Si	Alors	Donc
		$(8,6*70) / 100 = 6$; $(47,1*70) / 100 = 33$; $(28,6*70) / 100 = 20$; $(15,7*70) / 100 = 11$.

6) Selon la Banque mondiale, avec 15 420 000 habitants, le Sénégal compte 1066 médecins (du public) en 2017.

- a) Quel est le taux de médecins en référence au nombre constant 100 000? Interprétez statistiquement!

Taux = $(1066*100000)/15\ 420\ 000 = 6.9$
 En 2017, le Sénégal comptait 7 médecins par 100000 habitants

- b) Quel est le taux de médecins en référence au nombre constant 10 000? Interprétez statistiquement!

Taux = $(1066*10000)/15\ 420\ 000 = 0.69$
 En 2017, le Sénégal comptait près d'un (.69) médecin par 10000 habitants.

- c) Quel est le ratio d'habitants par médecin? Interprétez statistiquement!

Ratio = $(15\ 420\ 000 / 1066)*1 = 14465,3$
 En 2017, on comptait 14465 habitants pour 1 médecin.

- d) Quel est le ratio d'habitants par 2 médecins? Interprétez statistiquement!

Ratio = $(15\ 420\ 000 / 1066)*2 = 28930,6$
 En 2017, on comptait 28931 habitants pour 2 médecins.

7) Selon l'Institut de la statistique du Québec, le taux de chômage est passé de 8,0% en 2010 à 7,8% en 2011 au Québec. Par contre dans la province de Terre-Neuve-et-Labrador, le taux de chômage est passé de 14,4% en 2010 à 12,7% en 2011.

- a) Quel est le taux de diminution au Québec? Interprétez statistiquement!

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

$$\text{Taux de diminution} = ((7,8 - 8,0)/8,0)*100 = -2,5\%$$

Le taux de diminution est de -2,5%. Autrement dit, en 2011, le taux de chômage a diminué de 2,5% par rapport à 2010 dans la province du Québec.

- c) Quel est le taux de diminution dans la province de Terre-Neuve-et-Labrador ? Interprétez statistiquement

$$\text{Taux de diminution} = ((12,7 - 14,4)/14,4) * 100 = -11,8\%$$

Le taux de diminution est de -11,8%. En 2011, le taux de chômage a diminué de 11,8% par rapport à 2010 dans la province de Terre-Neuve-et-Labrador.

- d) Comparez les deux taux de diminution et dites laquelle des deux situations de chômage est moins dramatique.

Certes, le taux de chômage a plus diminué dans la province de Terre-Neuve-et-Labrador (-11,8%) qu'au Québec (-2,5%). Mais, peut-on dire que la situation est moins dramatique à Terre-Neuve-et-Labrador? N'oublions pas qu'en 2011, le taux de chômage est toujours plus élevé dans la province de Terre-Neuve-et-Labrador (12,4%) qu'au Québec (7,8%). Par conséquent, la situation est toujours moins dramatique au Québec que dans la province de Terre-Neuve-et-Labrador, en dépit de la faiblesse du taux de diminution au Québec.

NB : Lorsqu'on compare deux mesures, il faut toujours prendre soin de bien distinguer les mesures absolues (taux absolus : Qc=7,8%; TL=12,4%) et les mesures relatives (taux relatifs : Qc=-2,5%; TL=-11,8%) ☺

- 8) Selon Perspectives Monde, en 2010, le produit intérieur brut (PIB) des États-Unis et du Canada se chiffraient respectivement à 14 586 et 1 577 milliards de dollars courants. En supposant que le PIB américain soit demeuré stable, de combien en pourcentage le PIB canadien devrait-il s'accroître pour rattraper celui des États-Unis?

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

$$\text{Taux d'augmentation} = ((14586 - 1577)/1577) * 100 = 824,9\%$$

Le PIB canadien devrait s'accroître de 825% pour rattraper le PIB américain. Hum...C'est énorme ça! Autant dire que c'est presque impossible d'ici au moins 10 ans...

NB : Reste à savoir maintenant si le PIB canadien est comparable au PIB américain, vu que la population américaine est près de 10 fois plus importante que la population canadienne. Le revenu par tête d'habitant aurait été plus fiable comme critère de comparaison. Quoi qu'il en soit, un indicateur mesure toujours une réalité infiniment complexe, de telle sorte qu'il faut l'interpréter prudemment, c'est-à-dire avec un esprit critique et non un esprit de critique ☺

- 9) En prenant comme référence l'année 2010, l'indice du prix à la consommation (IPC) de l'abonnement mensuel dans le Réseau de transport

de la Capitale (RTC) était de 105% en 2012.

a) Calculez le coût de l'abonnement mensuel en 2010, sachant qu'il était de 76,75\$ en 2012.

Si

$$\text{IPC} = \frac{\text{Valeur du produit pour une année donnée}}{\text{Valeur du produit pour l'année de référence}} * 100$$

En transposant, on a

$$\text{Valeur pour année référence} = \frac{\text{Valeur du produit pour une année donnée}}{\text{IPC}} * 100$$

$$\text{Coût de l'abonnement en 2010} = (76,75/105)*100 = \mathbf{73,10\$}$$

b) Calculez le coût du même abonnement mensuel en 2011, sachant que l'IPC était de 102,67%.

Si

$$\text{IPC} = \frac{\text{Valeur du produit pour une année donnée}}{\text{Valeur du produit pour l'année de référence}} * 100$$

En transposant, on a

$$\text{Valeur pour année référence} = \frac{\text{Valeur du produit pour une année donnée}}{\text{IPC}} * 100$$

$$\text{Coût de l'abonnement en 2011} = (76,75/102,67)*100 = \mathbf{74,75\$}$$

c) En prenant comme référence le prix de 2011, calculer le taux d'augmentation de l'abonnement mensuel entre 2011 et 2012 et interprétez.

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

$$\text{Taux} = ((76,75-74,75)/74,75)*100 = \mathbf{2,7\%}.$$

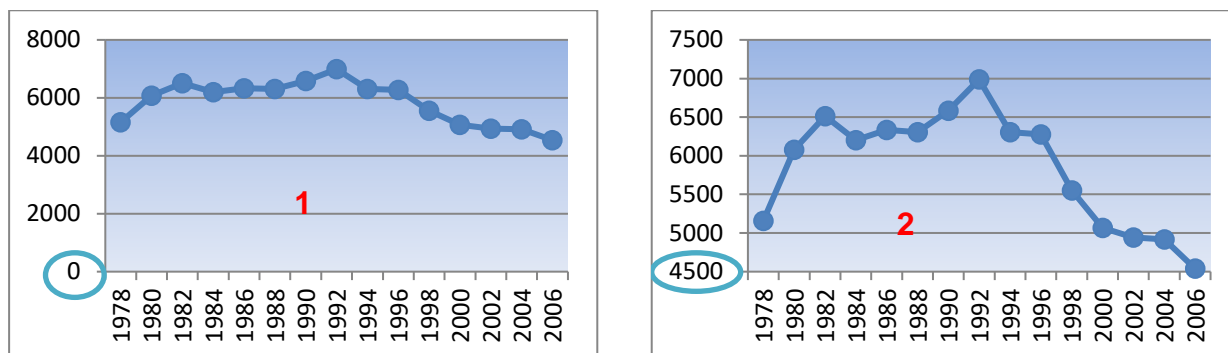
En 2012, l'abonnement mensuel a augmenté de 2,7% par rapport à 2011. Pour 100 dollars consacrés à l'abonnement en 2011, on a dépensé près de 3 dollars de plus en 2012.

NB : Prendre soin de toujours considérer la valeur de référence (la première temporellement) comme la valeur à soustraire et comme le dénominateur. Si vos formules sont incorrectes, les statistiques le seront également 😊

10) Deux catégories d'acteurs se positionnent sur la question de l'évolution

des crimes au Canada : le gouvernement (ses sous-systèmes politique et judiciaire y compris) d'une part et le lobby des armes à feu d'autre part. Les mêmes données sont présentées à l'aide de deux courbes de tendance.

Courbe de tendance des crimes au Canada de 1978 à 2006



a) Y a-t-il une différence entre les deux graphiques ? Expliquez.

La lecture des deux graphiques laisse observer une différence dans l'évolution des crimes. La diminution du nombre de crimes paraît plus importante sur le graphique 2 que sur le graphique 1, notamment à partir de 1992. En fait, l'échelle de l'axe vertical du graphique 1 commence par 0 alors que dans le graphique 2 l'axe horizontal coupe l'axe vertical au niveau de la valeur 4500. Cela suffit pour que les deux courbes de tendance se différencient assez nettement.

b) Lequel des deux camps aura-t-il tendance à présenter le graphique 2 ?

Il est fort probable que le gouvernement supporte ses positions en présentant la courbe de tendance 2, laquelle met en évidence une importante diminution du taux de criminalité, à tout le moins visuellement. Il pourrait ainsi se vanter des effets positifs de ses politiques et programmes de lutte contre la criminalité.

D'un autre côté, le lobby des armes à feu aura surtout tendance à parler d'une diminution négligeable du nombre de crimes. Pour supporter cette prise de position, la présentation de la courbe de tendance 1 ferait l'affaire. L'enjeu serait de convaincre la société, à travers le gouvernement, de la nécessité de libéraliser l'accès aux armes à feu pour permettre aux citoyens d'assurer eux-mêmes leur sécurité.

NB : On se rend compte que les statistiques peuvent être utilisées, voire instrumentalisées selon les intérêts des acteurs et les enjeux du moment. Aussi, si une figure ne ment pas, on peut y faire figurer des mensonges 😊