EXAMEN FINAL DE STATISTIQUES DESCRIPTIVES
L1 ECO - SESSION 1
- Correction -

Exercice 1 :

1) Questions à Choix Multiples (QCM). Cochez la bonne réponse

Classer ces statistiques selon leur nature (indicateur de position ou de dispersion)

	Position	Dispersion
Minimum	X	
Moyenne	X	
Ecart-type		X
Mode	X	
Ecart interquartile		X
Médiane	X	
Premier quartile	X	
Coefficient de variation		X

Considérons un groupe de TD pour lequel la moyenne des notes est égale à 10 et la variance est égale à 9. Trois nouveaux étudiants ayant respectivement 7, 14 et 17 s'inscrivent dans ce groupe. Quelle sera, selon vous, l'évolution des statistiques suivantes (sans calcul) :

	Moyenne	Ecart-type	Médiane	Mode	Etendue	Ecart interquartile
Augmentation	X	X				
Diminution						
identique						
Les données de l'exercice ne permettent pas de conclure			X	X	X	X

2) Après avoir rappelé le principe de la méthode des moindres carrés permettant de trouver la droite d'équation y=ax+b, poser le programme d'optimisation associé et trouver la valeur de b.

La méthode des moindres carrés vise à expliquer un nuage de points par une droite qui lie y à x, telle que la distance entre le nuage de points et droite est minimale. Cette distance matérialise l"'erreur" c'est à dire la différence entre le point réellement observé et le point prédit par la droite. Si la droite passe au milieu des points, cette erreur sera alternativement positive et négative, la somme des erreurs étant par définition nulle. Ainsi, la méthode des moindres carrés consiste à chercher la valeur des paramètres a et b qui minimise la somme des erreurs élevées au carré.

$$Min \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 = \sum_{i=1}^{n}(y_i - \widehat{a}x_i - \widehat{b})^2 = f(a,b)$$

Pour trouver les valeurs $\widehat{a}$ et $\widehat{b}$ au minimum de la fonction, on résoud le système :

$$\begin{cases} \frac{\delta f(a,b)}{\delta \widehat{b}} = 0 \\ \\ \frac{\delta f(a,b)}{\delta \widehat{a}} = 0 \end{cases}$$

$\frac{\delta f(a,b)}{\widehat{\delta b}} = 0$ (dériver par rapport à b : $u^2 \rightarrow 2uu'$)

$$\begin{array}{ll} \frac{\delta f(a,b)}{\widehat{\delta b}} & = (-1) * 2 * \sum_i (y_i - ax_i - b) = 0 \\ \Leftrightarrow & \sum_i (y_i - ax_i - b) = 0 \\ \Leftrightarrow & \sum_i y_i - a \sum_i x_i - nb = 0 \\ \Leftrightarrow & \sum_i y_i = a \sum_i x_i + nb \quad \textit{(on divise par n)} \\ \Leftrightarrow & \overline{y} = a\overline{x} + b \end{array}$$

De la même manière, il est possible de montrer que $a = \frac{cov(x,y)}{V(x)}$.

Exercice 2 :

Le poste " produits pétroliers " comprend différents biens issus de l'hydrocarbure de base appelée pétrole brut. Nous recensons trois principaux produits : essence, gazole et fioul lourd. Le tableau suivant désigne les prix moyens annuels (en dollars par baril) de ces produits ainsi que les quantités consommées (en milliers de barils par jour : b/j) en 2000 et 2005 en France.

	2000		**2005**	
	quantités	**prix**	**quantités**	**prix**
Essence	320	1.025	250	1.500
Gazole	923	0.650	1000	1.062
Fioul	136	0.424	120	0.723

1) Calculer l'indice élémentaire des prix du gazole en 2005 base 100 l'année 2000.

L'indice élémentaire des prix du gazole en 2005 base 100 l'année 2000 est :

$$I_{05/00}(p_g) = 100 * \frac{V_{05}}{V_{00}} = 100 * \frac{1,062}{0,650} = 163,4$$

Ainsi, entre 2000 et 2005, les prix ont augmenté de 63,4%.

2) Calculer, pour le gazole, le taux de croissance annuel moyen des prix entre 2000 et 2005.

Le taux de croissance annuel moyen (TCAM, noté $\overline{g}$) est le taux qui, appliqué chaque année durant cette période (5 années), conduit à une aumgmentation de 100% à l'issue des 8 années.

$$p_{05} = (1 + \overline{g})^5 p_{00} = 1,634 p_{00} \Rightarrow \overline{g} = 1,634^{1/5} - 1 = 0,103$$

Le TCAM est donc égal à 9%.

3) Qu'est ce qu'un indice synthétique de Paasche et de Laspeyres. Comment sont-ils construits pour les prix et pour les quantités ? Justifier.

Voir le cours pour la définition et la construction.

4) Calculer et interpréter l'indice de Laspeyres des prix des produits pétroliers en considérant l'année 2000 comme référence.

$$L_{05/00}(p) = 100 * \frac{\sum_i p_{i05} q_{i00}}{\sum_i p_{i00} q_{i00}} = 100 * \frac{1,5 * 320 + 1,062 * 923 + 0,723 * 136}{1,025 * 320 + 0,650 * 923 + 0,424 * 136} = 158,13$$

Ainsi, si les quantités étaient restées identiques entre 2000 et 2005, la valeur globale (p*q) aurait augmenté de 58,13% : en d'autres termes, l'impact des prix sur la valeur globale conduit à une augmentation de cette dernière de 58,13%.

5) Calculer l'indice élémentaire des valeurs globales en 2005 base 100 année 2000. A partir de ce résultat, que pouvez vous en déduire (sans calcul) à propos de la valeur de l'indice de Laspeyres des quantités (Année 2000 comme référence) ? Justifier votre réponse.

$$VG_{05/00}(p) = 100 * \frac{\sum_i p_{i05}q_{i05}}{\sum_i p_{i00}q_{i00}} = 100 * \frac{1,5*250+1,062*1000+0,723*120}{1,025*320+0,650*923+0,424*136} = 154,6$$

Ainsi, compte tenu des évolutions des prix et des quantités entre 2000 et 2005, la valeur globale a augmenté de 54,6%. L'évolution associée à la seule variation des prix est supérieure à celle observée lorsque l'on considère l'évolution des prix et des quantités, cela suppose donc qu'en considérant seulement l'évolution des quantités nous aurions observé une diminution de la valeur globale. L'indice de Laspeyres des quantités est donc inférieur à 100.

6) Les instances Européennes décideraient le cas échéant de remplacer ce type de gazole par un biocarburant faiblement consommé en France (2000 b/j en 2000 et 2500 b/j) commercialisé à des prix de 2.5$en 2000 et 2.55$ en 2005. Quel sera l'impact sur l'indice ? Pourquoi ? (cet organisme calcule son indice de prix à l'aide de l'indice de Laspeyres)

L'indice de Laspeyres prix va baisser si on tient compte du remplacement du gazole par le biocarburant. En effet, l'indice de Laspeyeres mesure l'impact de la variation de prix, au regard des prix du biocarburant il apparaît que ceux ci ont très peu évolué entre 2000 et 2005.

Si on calcule l'indice de Laspeyres en tenant compte du remplacement du gazole par le biocarburant :

$$L_{05/00}(p) = 100 * \frac{\sum_i p_{i05}q_{i00}}{\sum_i p_{i00}q_{i00}} = 100 * \frac{1,5*320+2,55*2000+0,723*136}{1,025*320+2,5*2000+0,424*136} = 105,43$$

Exercice 3 :

Un potentiel acheteur d'un véhicule automobile se questionne quant au choix du carburant qu'il va privilégier pour sa voiture. Il sait que le cours du pétrole brut influe directement sur les prix à la pompe dans les stations services. Il décide alors de comparer l'impact du cours du prix du pétrole sur le prix de l'essence ainsi que sur celui du diesel. Pour cela il effectue un relevé des prix de chaque produit sur les six premiers mois de l'année 2008.

	Prix du Brut (en dollar par baril)	Prix de l'essence HT (en centimes d'euros par litre)	Prix du diesel HT (en centimes d'euros par litre)
Janvier	85	130	115
Février	95	140	125
Mars	105	130	130
Avril	115	140	135
Mai	125	150	145
Juin	135	155	155

1) En considérant d'une part la relation entre le prix du pétrole brut et le prix de l'essence et d'autre part la relation entre le prix du pétrole brut et le prix du diesel,

et en supposant que le prix du pétrole brut a un impact sur les prix à la pompe des stations services : Tracer sur deux graphiques différents les nuages de points correspondants.

Pour représenter correctement les nuages de points : le prix du pétrôle brut devait être placé en abcisses pour chacun des graphiques et le prix de l'essence (ou du gazole) en ordonnées.

2) Après avoir calculé pour chacune des variables (prix du pétrole brut, prix de l'essence et prix du diesel) la moyenne et la variance, déterminer les deux droites de régression.

Notons x le prix du pétrole brut, y le prix de l'essence et z le prix du diesel.

$$\begin{aligned}
\bar{x} &= \frac{1}{N}\sum_i n_i x_i = \frac{1}{6}(85+\ldots+135) = 110 \\
\bar{y} &= \frac{1}{N}\sum_i n_i y_i = \frac{1}{6}(130+\ldots+155) = 140,83 \\
\bar{z} &= \frac{1}{N}\sum_i n_i z_i = \frac{1}{6}(115+\ldots+155) = 134,17 \\
V(x) &= \frac{1}{N}\sum_i n_i x_i^2 - \bar{x}^2 = \frac{85^2+\ldots+135^2}{6} - 110^2 = 291,67 \\
V(y) &= \frac{1}{N}\sum_i n_i y_i^2 - \bar{y}^2 = \frac{130^2+\ldots+155^2}{6} - 140,83^2 = 87,74 \\
V(z) &= \frac{1}{N}\sum_i n_i z_i^2 - \bar{z}^2 = \frac{115^2+\ldots+155^2}{6} - 134,17^2 = 169,24
\end{aligned}$$

Les droites de régression ont pour équations : $y = ax + b$ avec $a = \frac{cov(x,y)}{V(x)}$ et $b = \bar{y} - a\bar{x}$ et $z = a'x + b'$ avec $a' = \frac{cov(x,z)}{V(x)}$ et $b' = \bar{z} - a'\bar{x}$. Il faut donc au préalable calculer les covariances :

$$cov(x,y) = \frac{1}{N}\sum_i\sum_j x_i y_j - \bar{x}\bar{y} = \frac{85*130+\ldots+135*155}{6} - 110*140,83 = 137,8$$
$$cov(x,z) = \frac{1}{N}\sum_i\sum_j x_i z_j - \bar{x}\bar{z} = \frac{85*115+\ldots+135*155}{6} - 110*134,17 = 220,46$$

Après calcul, les droites de régression ont pour équations : $y = 0,47x + 89,13$ et $z = 0,76x + 50,57$

a. Comparer les deux coefficients directeurs des droites de régression. Que pouvez-vous en conclure quant à l'évolution des prix respectifs de l'essence et du diesel selon l'évolution future du prix du pétrole ?

$a = 0,47 < a' = 0,76$: le coefficient directeur associé au gazole est plus élevé que celui associé à l'essence. Ainsi, si le prix du pétrole brut (x) augmente de 1, le prix de l'essence augmente de 0,47 et celui du gazole de 0,76. Le prix du gazole est donc davantage sensible aux fluctuations du prix du pétrole.

b. Calculer et interpréter pour chacune des deux droites le coefficient de détermination.

Le coefficient de détermination fournit une indication de la qualité de l'ajustement. Il est, par définition, toujours compris entre 0 et 1.

$$R^2(x,y) = \frac{cov(x,y)^2}{V(x)V(y)} = \frac{137,8^2}{291,67*87,74} = 0,74$$

Le modèle explique 74% de la réalité : le fait que le prix de l'essence différe selon les mois peut être expliqué à 74% par le fait que le prix du pétrole brut diffère.

$$R^2(x,z) = \frac{cov(x,z)^2}{V(x)V(z)} = \frac{220,46^2}{291,67 * 169,24} = 0,98$$

Le modèle explique 98% de la réalité : le fait que le prix du gazole différe selon les mois peut être expliqué à 98% par le fait que le prix du pétrole brut diffère.

3) La croissance des cours du pétrole amène l'acheteur à se demander quels seront les prix des deux carburants à la pompe des stations services si le baril du brut atteignait un jour 200 dollars. Calculer ces prix.

Si le prix du pétrole brut (x) est égal à 200, calculons y et z : $y = 0,47 * 200 + 89,13 = 183,13$ et $z = 0,76 * 200 + 50,57 = 202,57$.

4) En pleine réflexion, l'acheteur se dit qu'il sera plus avantageux pour lui de prendre les transports en commun lorsque le prix au litre dépassera les deux euros. Pour quels cours du pétrole brut, devra-t-il s'apprêter à vendre sa voiture selon qu'il ait opté pour une voiture essence ou diesel?

Pour quels prix du pétrole brut, le prix de l'essence (y) et le prix du gazole (z) dépssent-ils les 2 euros, c'est à dire 200 centimes : $200 = 0,47x + 89,13 \Rightarrow x = 235,9$ et $200 = 0,76x + 50,57 \Rightarrow x = 196,6$.