Série 1 - calculatrice - tables

Objectifs: Rappel élémentaire qui consiste à se familiariser (ou confirmer) l'utilisation de la calculatrice en mode statistique. Lecture des tables et interpolation linéaire et logarithmique.

Sur la calculatrice il y a deux touches particulières la touche $\boxed{sx}$, notée également $\boxed{s}$, ainsi que la touche $\boxed{\sigma x}$, notée également $\boxed{\sigma}$ (CASIO fx-991 EX, CASIO fx-85ES PLUS, HP 15C, HP 35s, HP 41, SHARP EL-531-TH, SHARP EL-501T). Dans certains modèles de calculatrice, il s'agit de $\boxed{\sigma_{xn-1}}$ et $\boxed{\sigma_{xn}}$ (TI, par exemple TI-30 ECO RS, et CASIO fx-82 SOLAR II).

Il s'agit de la racine carrée (écart type) de la variance d'échantillon

$$s^2 \triangleq \sum_{k=1}^{N} \frac{(x_k - \bar{x})^2}{N-1}$$

et

$$\sigma^2 \triangleq \sum_{k=1}^{N} \frac{(x_k - \bar{x}_k)^2}{N}$$

qui diffère de la première expression par la division par N au lieu de $N-1$.

Les moyennes ont toujours le dénominateur égal à N

$$\bar{x} \triangleq \sum_{k=1}^{N} \frac{x_k}{N}$$

Exercice 1

Repérer les touches (ou menus) correspondants aux symboles

$$\boxed{\sum X}$$

$$\boxed{\sum Y}$$

$$\boxed{\sum X^2}$$

$$\boxed{\sum Y^2}$$

$$\boxed{\sum XY}$$

$$\boxed{\bar{x}}$$

$$\boxed{\bar{y}}$$

$$\boxed{sx}$$

$$\boxed{\sigma x}$$

et compléter le tableau :

	$\sum X$	$\sum X^2$	$\bar{x}$	s	σ
X					
Y					

Calculer également $\sum XY$ en utilisant les menus ou la mémoire de la calculatrice si la calculatrice le permet, sinon faites le "à la main" à l'aide de la calculatrice tout de même.

$$\sum XY =$$

à partir des données

X	-70	25	-61	19	54	-59	-72	172	42	230	-139	43
Y	208	6	155	-126	171	91	67	-34	-9	-51	110	-121

Que constatez-vous entre les deux séries ?

Exercice 2

Il s'agit d'illustrer différente relations algébriques entre les quantités $\sum X$, $\sum Y$, $\sum XY$, $\sum X^2$, $\sum Y^2$

Vérifier la relation

$$\sum_k x_k^2 - \bar{x} \sum_k x_k = \sum_k (x_k - \bar{x})^2$$

et déduire une façon de calculer la variance d'échantillon s en utilisant uniquement N et les sommes $\sum x_k^2$ et $\sum x_k$. Vérifier vos prévisions à l'aide du tableau obtenu dans l'exercice 1 précédent.

A partir de la relation algébrique bien connue

$$(x + y)^2 = x^2 + 2xy + y^2$$

il s'agit d'améliorer la précision du calcul de s en utilisant la même calculatrice (le même hardware).

Commencer par l'exemple simple suivant qui consiste en l'entrée de chiffres consécutifs dans la calculatrice (remarquer la séquence)

X	12	45	78	13	46	79

qui consiste à prendre la première touche suivie du chiffre correspondant de la seconde colonne de la calculatrice (colonnes du pavé numérique) dans un premier temps, puis de la première touche suivi de la touche de la troisième colonne de la calculatrice. Il n'y a rien de particuler concernant cette succession, mais elle est facile à entrer rapidement dans la calculatrice (contrairement au soin méticuleux nécessaire pour l'exercice 1). Comme pour l'exercice précédent calculer les statistiques afin de remplir le tableau:

$\sum X$	$\sum X^2$	$\bar{x}$	s	s^2

Maintenant décomposer la série en deux avec une série correspondant au premier chiffre et une seconde série correspondant à la succession des seconds chiffres. Ceci donne le tableau:

X	1	4	7	1	4	7
Y	2	5	8	3	6	9

de telle sorte que vous lisez la même suite de nombres si vous lisez le tableau de haut en bas colonne après colonne.

Compléter le tableau

	$\sum X$	$\sum X^2$	$\bar{x}$	s	s^2
X					
Y					

L'identité algébrique $(x+y)^2 = x^2 + 2xy + y^2$ conduit en associant chaque chiffre x_k aux nombres proportionneles à la dixaine et chaque chiffre y_k aux nombres proportionnels à l'unité

$$100 \sum_{k=1}^{N} x_k^2 + 20 \sum_{k=1}^{N} x_k y_k + \sum_{k=1}^{N} y_k^2 = \sum_{k=1}^{N} (10x_k + y_k)^2$$

Vérifier cette relation en remarquant que le membre de droite n'est rien d'autre que le résultat $\sum X^2$ de la première partie de l'exercice.

Voici un exemple plus intéressant. Suivez exactement la même procédure mais appliquée aux échantillons

456 123
789 456
789 123
258 147
369 147
369 258

et la décomposition

X	456	789	789	258	369	369
Y	123	456	123	147	147	258

que l'on décomposera en série de trois chiffres

On demande de comparer $\sum X^2$ entre la première méthode qui consiste à entrer les échantillons sur 6 chiffres, et la seconde qui consiste à décomposer les nombres en deux séries de 3 chiffres.

Exercice 3

C'est un exercice de lecture de table. La compréhension de ce que représente les tables, en particulier les degrés de liberté viendra au fur et à mesure du cours.

Pour la loi normale $\mathcal{N}(0{,}1)$ la table représente la surface grisée de la courbe qui représente la fonction de distribution de la loi normale. Les valeurs corresondant à la première colonne et la première ligne de la table représente la variable Z et donne l'abscisse d'arrivée de la zone grise sur la figure (abscisse Z). La probabilité est donnée par la surface de la zone grise. Lorsque $Z = 0$, on a bien $P(0 < x \leq 0) = 0$, ce que l'on retrouve comme première entrée de la table (première ligne et deuxième colonne). Autre exemple, en considérant la troisième ligne et la quatrième colonne (quatrième colonne de la variable Z, c'est la cinquième colonne correspondant à la quatrième valeur 0.03) on a $Z = 0.2 + 003 = 0.203$, on lit la valeur $P(0 < x \leq 0.203) = 0.09$.

On demande d'obtenir un intervalle $]z_1; z_2[$ qui contient l'abscisse du premier et troisième quartile, en lisant la table directement (le premier quartile représente la région qui contient les 25 % des valeurs les plus basses et le troisième quartile correspond au 75 %).

Exercice 4

Utiliser l'interpolation linéaire pour affiner le résultat de l'exercice précédent.

Exercice 5

Dans la table de la loi du χ^2 de Pearson (que l'on détaillera en leçon 4, pas besoin de comprendre quoi que ce soit à ce stade, ce n'est qu'un exercice de lecture de table), on demande de trouver la valeur de la probabilité correspondant à une valeur de $\alpha = 21.5$ pour un degré de liberté $\nu = 10$. La valeur α correspond à l'abscisse du point de départ de la région grisée de la figure de la table. Utiliser une interpolation linéaire entre les deux valeurs α les plus proches que l'on obtient en lisant la table.

Exercice 6

Même chose que pour l'exercice 6 mais en utilisant une interpolation logarithmique. L'interpolation logarithmique consiste à interpoler entre les deux valeurs les plus proches de $\log \alpha$.

INDICATION: La vraie valeur demandée est calculable avec un ordinateur et donne $\alpha = 0.0178646$. Discuter les résultats obtenus aux exercices 5 et 6 en relation avec cette valeur. Selon vous, quelle est la raison de la supériorité d'une méthode sur une autre ?