

## LSM 3.053

### Informatique - Statistique : TP n°1 et 2

Ce TP couple les notions de statistiques descriptives abordées en cours avec l'utilisation avantageuse du tableur Excel. Le tableur est utilisé pour effectuer les calculs répétitifs (mise à profit des options de COPIER-COUPER-COLLER) et pour ses fonctions graphiques (Histogrammes, Courbes).

Les raccourcis clavier usuels de Windows sont utilisables sous Excel :

Action	Souris	Clavier
Sélection d'une cellule :	Clic gauche	-
Sélection d'une colonne	Clic gauche sur la Lettre	-
Sélection d'une ligne	Clic gauche sur le numéro	-
Sélection de la feuille	Clic sur le coin Haut Gauche	CTRL A
Copier la sélection	Clic droit + Copier	CTRL C
Couper la sélection	Clic droit + Couper	CTRL X
Coller la sélection	Clic droit + Coller	CTRL V
Coller la sélection (spécial)	Clic droit + Collage spécial	-
Répéter une opération Edition/Répéter		CTRL Y
Annuler une opération	Edition/Annuler	CTRL Z

#### RAPPELS SUR LE TABLEUR

Le tableur gère des feuilles de calcul constituées de 16.711.425 cellules arrangées en 255 colonnes (numérotées A à IV) \* 65535 lignes (numérotées 1 à 65535).

#### Cellule ("Cell")

Chaque cellule est identifiée par son adresse, qui peut être stipulée de 4 façons :

- 1) **adresse absolue** : On spécifie directement l'indice de colonne et celui de ligne.
- 2) **adresse relative** : On spécifie le décalage en colonne/ligne par rapport à la cellule de référence.
- 3),4) **adresse semi-relative** : On spécifie de façon relative la ligne et de façon absolue la colonne, ou vice-versa.

Dans le tableur, l'adressage relatif est la règle lorsqu'on fait référence au contenu d'une autre cellule dans un calcul. L'adresse relative est donnée sous la forme (lettre de colonne, n° de ligne) et s'entend comme le décalage amenant de la cellule contenant la formule à la cellule à laquelle on fait référence. La cellule de référence est la cellule active ("ActiveCell") qui est la cellule sélectionnée par clic:

ex. : dans la cellule B4 on tape la formule "=A2 +3"  
 la formule s'interprète en fait comme :  
 "= contenu de la cellule 1 ligne plus haut et 2 colonnes à gauche + 3"

Si on souhaite faire référence de façon absolue à une colonne/ligne, on fait précéder son indice d'un \$.  
 ex.: dans la cellule B4 on tape la formule "=\$A2 +3"  
 la formule s'interprète en fait comme :  
 "= contenu de (la cellule 1 ligne plus haut et en colonne A )+ 3"

Le même type de distinction sur l'adressage se retrouve en VBA (cf. TP suivants).

### **Plage ("Range")**

Une **plage** de cellules est constituée d'un ou plusieurs domaines de cellules contiguës:

- Un domaine de cellule contiguës est sélectionné en "cliquant-tirant" à la souris et est identifié par les coordonnées des cellules situées en haut à gauche et en bas à droite. Ces coordonnées, relatives ou non, sont séparées par ":".  
ex.: "A2:B6" "\$B\$1:\$C\$3"
- Une plage formée de plusieurs domaines contigus est définie par sélection de plusieurs domaines par "cliqué-tiré" successifs, tout en maintenant la touche CTRL enfoncée (sélection multiple). Les plages sélectionnées sont séparées à l'aide de ";".  
ex.: "A2:B3;C2:C4"

### **Contenu d'une cellule :**

Une cellule contient une donnée qui est soit explicite (constante), soit résulte d'une opération (formule). Cette donnée doit être distinguée de ce qui est affiché dans la cellule.

ex. : dans la cellule B4 on entre la constante 34 le résultat affiché est 34

ex. : dans la cellule B5 on entre la formule "= B4" le résultat affiché est 34.

Le contenu d'une cellule appartient à un des types de donnée suivants :

<i>TYPE</i>	<i>MISE EN PAGE</i>
Nombre (Entier /Réal)	calage à droite de la cellule
Chaîne	calage à gauche de la cellule
Booléen	En majuscule, centré

*REMARQUE : Par défaut Excel prend comme séparateur décimal le séparateur système (qui est la virgule si le système est déclaré en Français) et aucun séparateur de milliers. Un chiffre en notation Anglo-Saxonne avec un séparateur décimal de type "point" est donc interprété comme une chaîne de caractères. Pour modifier ces paramètres (importation fréquente de fichiers étrangers), il suffit d'aller dans l'onglet Outils/Options/International et de définir vos propres séparateurs.*

### **Format d'affichage du contenu :**

Le format d'affichage des données peut être personnalisé grâce à l'onglet **Format/Cellule** :

- nombre de décimales	(nombres)	- format scientifique	(nombres)
- unité €, \$, m, s, ...	(nombres)	- fraction	(nombres)
- pourcentage	(nombre)	- data/heure	(nombre)
- police (fonte, taille, couleur)	(tous)	- bordure (épaisseur, couleur)	(tous)
- motif (couleur)	(tous)		

### **Formule :**

Une formule commence nécessairement par un signe "=". Elle fait référence à des variables (contenu d'autres cellules, référencées par leurs adresses) ou des constantes (nombres, e.g.). Un grand nombre de fonctions mathématiques de base (beaucoup plus nombreuses qu'en VBA) sont directement accessibles avec Excel. Des fonctions supplémentaires peuvent être importées grâce à l'onglet Outils/Macros Complémentaires (Solver, Outils de recherche, d'analyse, ...).

Les formules peuvent être entrées directement au clavier si on en connaît déjà la syntaxe, où appelées à partir d'un menu déroulant (Insérer/Fonction/ ), le cas échéant. Il est possible de créer ses propres fonctions, programmées en VBA (voir TP suivants).

Lorsque le résultat est un scalaire, stocké dans une seule cellule, il est affiché dès validation de la formule avec la touche "RETURN"

ex.: dans la cellule B4 on entre la fonction "=Pi()" le résultat affiché est 3,14..:

          dans la cellule B5 on entre: "= sin(B4)\*racine (4)" le résultat affiché est 2

ex.: dans la cellule B5 on entre: "= fonction\_perso(B4)

S'il s'agit d'un calcul matriciel, il faut d'abord sélectionner la plage qui contiendra les cellules résultats, insérer la formule puis la valider avec les touches CTRL+SHIFT+RETURN.

ex.: dans la cellule A1:C2 on entre une matrice  $M_{2,3}$   
 dans la cellule E1:F3 on entre une matrice  $M_{3,2}$   
 dans la cellule A1:B3 on entre la formule : " $=\text{PRODUITMAT}(A1:C2;E1:F3)$ " que l'on valide avec CTRL+Shift+Return (validation d'opérations matricielles).

### Calcul itératif :

Excel offre la possibilité de réaliser des calculs itératifs où une cellule contient une formule lui faisant elle-même référence ("Référence circulaire"). Par défaut, le programme considère que la référence circulaire est fortuite et affiche un message d'alerte. Si la référence circulaire est volontaire, pour obtenir le résultat il est nécessaire de valider l'option de calcul itératif dans le menu déroulant /Outils/Options/Calcul. Le nombre d'itérations et la variation maximale de la variable entre 2 itérations sont ajustables dans ce même menu.

### Exercice 1: Fonctions du Tableur

#### A) Factorielle

Calculez la liste des valeurs de  $n!$ , pour  $0 \leq n \leq 20$ . Les valeurs de  $n$  sont entrées en A2:A21, en utilisant l'option d'incrémement automatique du "Tirer-Copier". Les valeurs  $n!$  définies par une relation de récurrence sont calculées en B2:B21. Le résultat obtenu est vérifié dans les cellules C2:C21 en utilisant la fonction **Fact(n)** du tableur.

#### B) Développement limité de $\cos(x)$

On peut montrer qu'une fonction infiniment dérivable  $f(x)$  peut être approximée pour  $x \rightarrow 0$  par la formule de Taylor - Mac Laurin (développement limité à l'ordre  $n$ ):

$$\lim_{x \rightarrow 0} f(x) = f(0) + \frac{x}{1!} f'(0) + \frac{x^2}{2!} f''(0) + \frac{x^3}{3!} f'''(0) + \dots + \frac{x^n}{n!} f^{(n)}(0)$$

Pour la fonction  $f(x) = \cos(x)$ , on obtient ainsi :

$$\lim_{x \rightarrow 0} \cos(x) = \cos(0) - \frac{x}{1!} \sin(0) - \frac{x^2}{2!} \cos(0) + \frac{x^3}{3!} \sin(0) + \dots + \frac{x^n}{n!} \cos^{(n)}(0)$$

soit :

$$\lim_{x \rightarrow 0} \cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^{2n} \frac{x^{2n}}{2n!} \quad \text{ou} \quad \lim_{x \rightarrow 0} \cos(x) = \sum_k (-1)^{k/2} \frac{x^k}{k!}, \text{ avec } k \text{ pair}$$

Observez la convergence du développement à l'ordre  $n$  de  $\cos(x)$  pour  $0 \leq k \leq 40$ . Les valeurs de  $k$  sont entrées en C2:C21, en utilisant l'option d'incrémement automatique du "Tirer-Copier". L'argument  $x$  de  $\cos(x)$  est inscrit dans la cellule A2 (choisir dans un premier temps  $x = \pi/10$ ). Les valeurs de  $\cos(x)$  évaluées par développement limité sont calculées en D2:D21. Comparer le résultat obtenu à la valeur exacte de  $\cos(x)$  calculée en A3 en utilisant la fonction **cos(x)** du tableur.

**Remarque** : si  $x$  s'éloigne trop de 0, la série limitée diverge et ne permet plus le calcul de  $\cos(x)$ . Dans les calculatrices, les fonctions trigonométriques  $\cos(x)$  et  $\sin(x)$  sont en fait approximées par la somme et la différences des développements en série de Taylor de  $\exp(ix)$  et  $\exp(-ix)$  :

$$\exp(ix) = \sum_n \frac{(ix)^n}{n!} + \theta(n) \quad \text{et} \quad \exp(-ix) = \sum_n \frac{(-ix)^n}{n!} + \theta(n)$$

C) Opérateurs logiques:

Construisez la table de multiplication logique de 2 variables Booléennes A et B stockées en A2:B5, en y faisant figurer les opérations suivantes :

$\bar{A}$ ,  $\bar{B}$ , (A et B), (A ou B), (ni A ni B).

D) Calcul matriciel:

Inscrivez en B2:D4 la matrice  $A = \begin{pmatrix} 1 & 3 & 8 \\ 2 & 4 & 5 \\ 3 & 3 & 6 \end{pmatrix}$ .

Calculez-en le carré  $A^2$  en B7:D9, le déterminant en B12, et l'inverse  $A^{-1}$  en B14:D16.

E) Calcul itératif et séries convergentes

L'option de calcul itératif de Excel peut être utilisée pour calculer des limites de suites et séries (convergentes!). Pour cela, on fixe dans une 1<sup>ère</sup> cellule le premier terme de la suite/série  $u_0$  (Initialisation). La 2<sup>ème</sup> cellule, dans laquelle on effectue le calcul par référence circulaire, contient une commande conditionnelle "**=SI(condition; valeur si VRAI; valeur si FAUX)**". Le test conditionnel porte sur une 3<sup>ème</sup> cellule où on entre la valeur VRAI pour initialiser le calcul. Si ce test est VRAI, on initialise le contenu de la 2<sup>ème</sup> cellule en y recopiant le contenu de la 1<sup>ère</sup> cellule. Si le test est FAUX, Excel calcule la valeur limite de la suite/série par itération à partir d'une formule faisant référence à la 2<sup>ème</sup> cellule elle-même.

Exemple : Calcul du nombre d'or par la représentation en fractions continues.

Le nombre d'or, habituellement désigné par la lettre  $\phi$  de l'alphabet grec en l'honneur de

Phidias, sculpteur et architecte grec du Parthénon, est le nombre :  $\phi = \frac{1+\sqrt{5}}{2} \approx 1,618...$

$\phi$  est l'unique solution strictement positive de l'équation  $\phi^2 = 1 + \frac{1}{\phi}$ , soit  $\phi = 1 + \frac{1}{\phi}$ .

Sur la base de ces 2 équations, on obtient 2 représentations de  $\phi$  à bases de racines carrées ou de fractions continues :

$$\phi = \lim \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \dots}}}} \quad \text{ou} \quad \phi = \lim 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}}$$

Pour calculer  $\phi$  à l'aide de la 1<sup>ère</sup> série, on entre dans le tableur les formules suivantes :

cellule A1: ="initialisation" chaîne de caractères de commentaire

cellule A2: = VRAI ou =FAUX

cellule A4: ="valeur initiale u0" chaîne de caractères de commentaire

cellule A5: =1

cellule A7: ="limite de la série" chaîne de caractères de commentaire

cellule A8: =SI(A2; A5; =RACINE(1+RACINE(A8)))

et on active le mode itératif.

Exercices d'application:

a) Calculez par la même méthode la valeur de  $\phi$  à partir de la représentation fractionnaire.

b) Quelle est la valeur limite de la série :  $mystère = \lim 1 + \frac{1}{1 + 1 + \frac{1}{1 + 1 + \frac{1}{1 + 1 + \dots}}}$  ?

## Exercice 2: Statistiques

### A) Importation des données

Importer les données statistiques à partir d'un fichier séquentiel "donnees.txt" situé sur la zone de partage en utilisant l'option /fichier/ouvrir et en choisissant les options de délimitation appropriées. Les données brutes  $x_i$  (2<sup>ème</sup> colonne) sont rangées en fonction d'un numéro d'ordre  $i$  (1<sup>ère</sup> colonne).

### B) Mise en ordre des données - Histogramme

A l'aide des commandes Tableur MIN (plage), MAX(plage) et NB.SI (condition; plage) construisez les histogrammes des fréquences partielles et cumulées en utilisant 10 classes d'intensité, uniformément distribuées. Tracez ces histogrammes sur 2 feuilles graphiques séparées en imposant que la couleur de batons soit verte.  
Estimez la médiane de la distribution.

### C) Calculs statistiques élémentaires

Calculez en utilisant les fonctions du Tableur la moyenne (MOYENNE(plage)), la variance (VAR(plage) ou VAR.P(plage)), l'écart-type (ECARTYPE(plage) ou ECARTYPEP(plage)), la skewness (COEFFICIENT.ASYMETRIE(plage)) et le kurtosis (KURTOSIS(plage)) de la distribution. Calculez ces mêmes valeurs à partir des formules données en cours de Statistiques.  
Vérifiez l'estimation de la médiane (MEDIANE(plage) trouvée à la question B).  
Quelle Loi Statistique semblent suivre ces données ?

### D) Raccordement de la distribution à une Loi statistique

Calculez les fréquences partielles espérées pour les 10 classes d'intensité choisies à la question B) en assumant que les observations suivent la Loi Statistique trouvée au C).  
Superposez le graphe des valeurs espérées à l'histogramme de la question B).  
Tracer un diagramme "Nuage de points" illustrant les relations entre fréquences observées et espérées.  
Calculer la covariance entre ces valeurs à partir des relations présentées en cours de Statistiques et avec la fonction du Tableur COVARIANCE(plage1; plage2).  
Utilisez le test du  $\chi^2$  (TEST.KHIDEUX(plage\_observations;plage\_valeurs\_espérées) pour valider votre hypothèse de raccordement.