CPH 316

Méthodes de la chimie physique

Analyse de données

Expériences 1 - 5

Table des matières

EXPÉRIENCE UN

Excel

OBJECTIFS PEDAGOGIQUES

Cet exercice a pour objet d'apprendre l'utilisation d'un chiffrier électronique.

Objectifs spécifiques

Après avoir complété cet exercice l'étudiant sera capable de :

a) charger le logiciel Excel
b) créer et ajuster les titres de colonnes pour un tableau électronique
c) entrer des données dans le tableau
d) sauvegarder un fichier Ex. sur une disquette
e) utiliser la commande édition, copier pour positionner des données
f) utiliser la commande édition, copier pour copier des formules mathématiques dans une série de cellules
g) utiliser la commande alignement pour centrer les données et les titres
h) utiliser la commande format, colonne, ajustement automatique ou largeur pour la largeur des colonnes
i) utiliser la commande ? pour le dépannage spécifique vis-à-vis d' une commande
j) tracer des graphes à l'aide de Excel
k) tracer des graphes 3D à l'aide de SigmaPlot et Excel
l) utiliser la commande régression pour faire des calculs de régressions linéaires
m) l'étudiant sera également introduit à l'interprétation des paramètres de sortie de la régression.

Attention : sauvegardez vos fichiers sur le disque de réseau : « Donnees (N:) » dans un dossier portant votre nom; un fichier d'Excel par Expérience. Vous pouvez aussi le sauvegarder sur votre USN drive. N'oubliez pas de sauvegarder votre fichier souvent pendant le labo, utilisez fonction Ctrl S.

Excel

Excel est un chiffrier électronique le plus populaire de MS Office.

Grandeur du chiffrier. La fenêtre que l'on voit à l'écran n'est qu'une fraction du chiffrier. Cette feuille électronique est composée de 256 colonnes (verticales) et de 65536 rangées (horizontales), et de plusieurs pages (onglets).

Calculs automatiques. On peut entrer des formules dans le chiffrier de façon à calculer automatiquement les valeurs dans les cellules. Un changement de valeur dans l'une des cellules change automatiquement tous les calculs.

Édition facile. Il est fort simple de changer une valeur du chiffrier. Toutes les données affectées par ce changement sont automatiquement calculées à nouveau.

Graphe. Des graphes de toutes sortes peuvent être produits à partir des données introduites dans le chiffrier.

Droites de régressions. On peut calculer facilement les droites de régressions à l'aide du chiffrier.

Utilisation de la souris : Dans Excel le bouton de gauche et le bouton de droite sont actifs. Le bouton de gauche sert à activer et à sélectionner des items (données, graphes, etc.). Le bouton de droite active les menus correspondants à l'endroit où repose le pointeur. Par exemple si on clique le bouton droit dans la fenêtre du chiffrier le menu de contrôle des cellules apparaît (police, tracé de ligne, largeur de colonne, etc.).

1. Création d'une page de chiffrier

Dans cette leçon nous allons utiliser Excel pour calculer l'isotherme d'un gaz parfait. Les étapes suivantes seront exécutées :

a) Chargement de Excel
b) Création et ajustement des titres de colonnes dans un tableau
c) Entrée de données dans un tableau
d) Sauvegarde du fichier sur une disquette.
1a) Cliquez démarrer, programmes, office 97, Microsoft Excel afin de démarrer le programme.

Vous devriez maintenant voir un écran semblable à celui de la figure 1.

La **barre de menu** contient les commandes principales. Elle utilise les menus déroulants selon les normes propres à Windows.

La **barre rapide** contient des boutons pour choisir les commandes les plus usuelles. Une brève description suit plus bas.
D'autres barres rapides secondaires peuvent s'ajouter au-dessous de celle-ci.

La **ligne d'entrée** sert à entrer et à éditer les cellules.

La **ligne d'état** donne de l'information sur l'état actuel du programme (par exemple le mot READY indique que Excel est en attente).

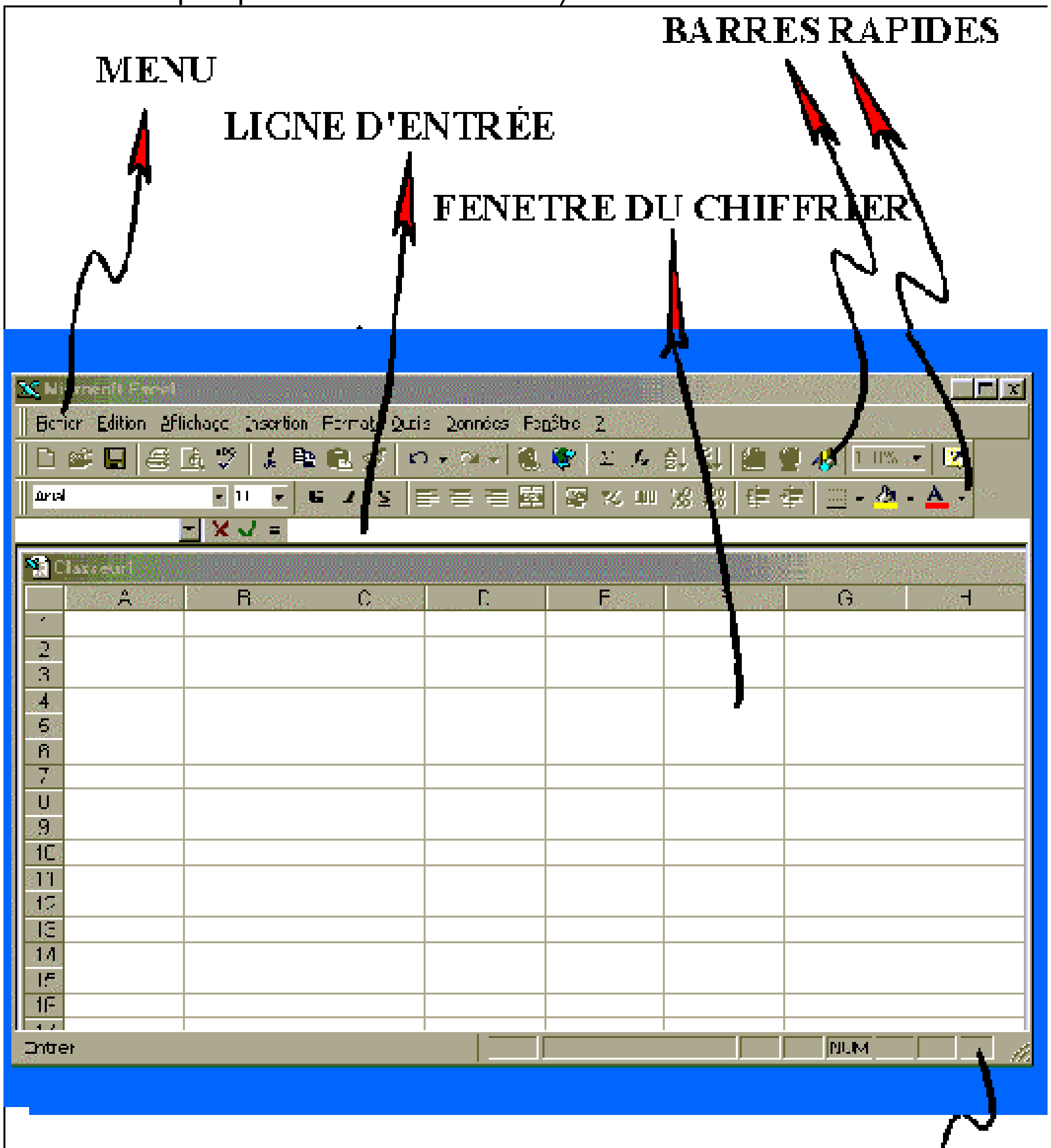


FIGURE 1

La **fenêtre du chiffrier** est l'endroit où vos données apparaissent.

Une cellule est une boîte qui peut contenir des données. La rangée et la colonne indiquent son adresse.

Un sélectionneur est le cadre noir qui montre la cellule active (C3 dans la figure 2).

Un Bloc est un groupe rectangulaire d'une ou plusieurs cellules. Un bloc est identifié par

les coordonnées de la cellule dans le coin extrême gauche en haut et le coin extrême droit en bas le tout séparé par deux points. Dans la figure 2, le block sélectionné est tout noir et les coordonnées sont C3..G16.
Les **onglets** au bas de la page servent à sélectionner les différentes pages (256 pour une fenêtre donnée).

Le **Pointeur de graphe** donne accès rapidement aux graphes ou au chiffrier.

Les barres de déroulement suivent la norme Windows et servent au déplacement à l'intérieur des pages ou entre les pages du chiffrier.
Les commandes **Menu** suivent la norme Windows et contiennent des menus à déroulement facile d'accès.

 Les boutons **couper, copier, coller** transportent, déplacent et insèrent des données et des objets dans la pile de Windows.

La seconde barre rapide contient plusieurs boutons qui permettent de choisir le type de police, la grosseur du caractère, l'imprimerie, etc... Il est également possible d'ajouter d'autres barres rapides à l'aide du menu **Outils**. Cliquez à quelques endroits de ce menu et vous verrez apparaître d'autres barres rapides. Pour les faire disparaître-il suffit de cliquer sur le X rouge à l'extrême droite de la barre.
1b) Dans la leçon qui suit vous serez introduit aux termes et procédures suivantes :
Sélectionneur de cellule. Le rectangle qui apparaît dans la fenêtre du chiffrier, indique la cellule sélectionnée.
Libellé (label). Un libellé est tout texte que vous entrez dans une cellule. Un libellé peut commencer avec une lettre quelconque ou une marque de ponctuation autre que . / + - $ (@ #.
Entrée des libellés. Le libellé apparaît à la ligne d'entrée à mesure que vous tapez. Si vous faites une erreur utilisez "backspace" et retapez. Pour écrire le libellé dans la cellule active utilisez la touche return.
Dans le but d'atteindre une certaine familiarité dans l'utilisation de Excel un calcul simple sera effectué.

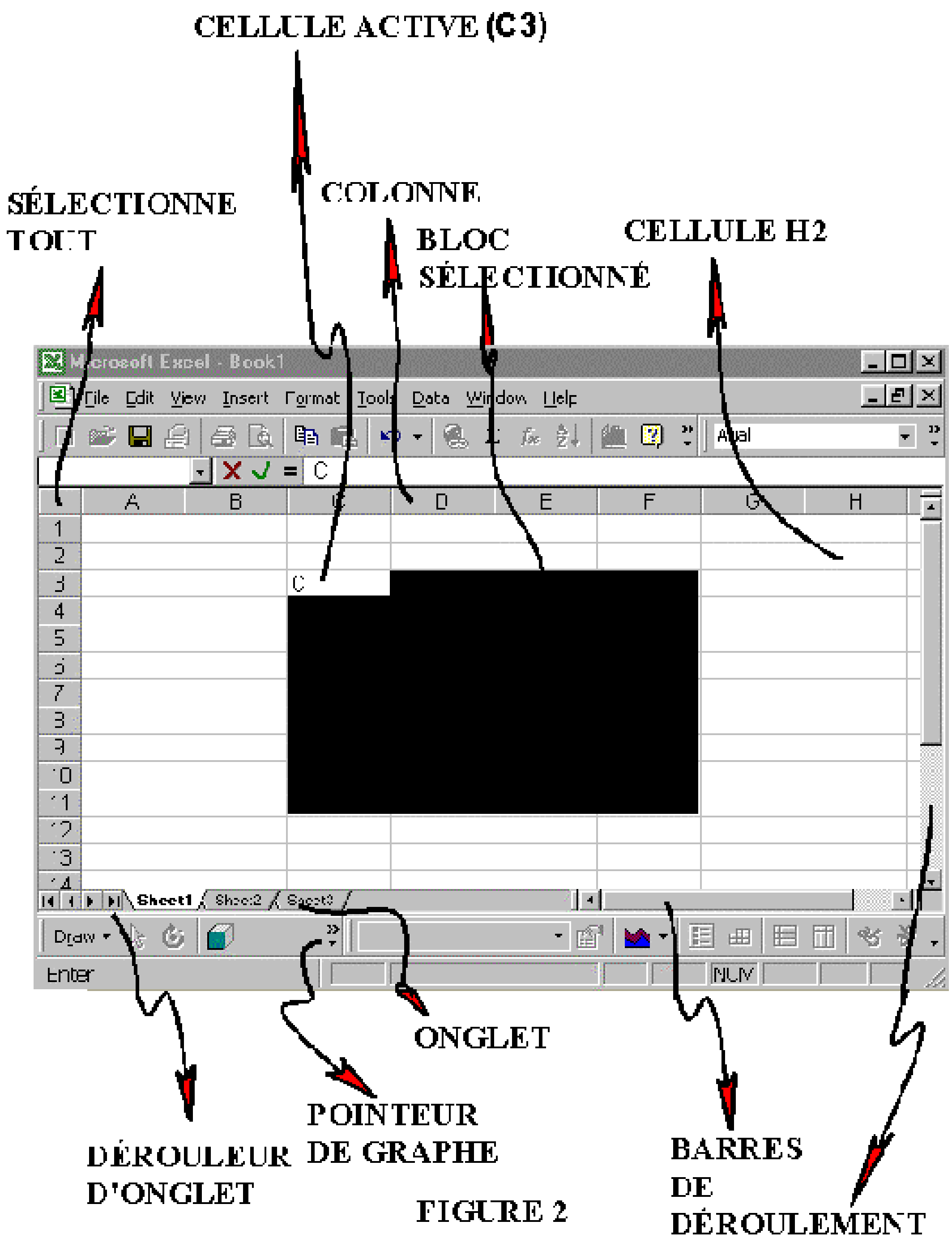

CELLULE ACTIVE (C3)
SÉLECTIONNE TOUT
COLONNE
BLOC SÉLECTIONNÉ
CELLULE H2
Microsoft Excel - Book1
File Edit View Insert Format Tools Data Window Help
Arial
A B C D E F G H
Sheet1 Sheet2 Sheet3
Draw
Enter
NUM
ONGLET
POINTEUR DE GRAPHE
DÉROULEUR D'ONGLET
BARRES DE DÉROULEMENT

FIGURE 2

Exercice 1

Dans ce qui suit nous allons calculer une isotherme de gaz parfait à une température de 300 K.
Pour une mole de gaz l'équation des gaz s'écrit

PV = RT

Nous allons utiliser Excel pour calculer le volume de gaz à différente Pression pour une température constante donnée. Avant de faire ces calculs nous allons apprendre comment écrire un tableau comme celui de la TABLE 1.

TABLE 1

R=	**8.31451**	
PRESSION / Pa	**TEMPERATURE / K**	**VOLUME / m^3**
50000	**300**	**0.04989**
60000		**0.04157**
70000		**0.03563**
80000		**0.03118**
90000		**0.02772**
100000		**0.02494**
110000		**0.02268**
120000		**0.02079**
130000		**0.01919**
140000		**0.01782**
150000		**0.01663**

Création de titres de colonnes

Dans ce qui suit entrez signifie dactylographiez.
↵ signifie enfoncez la touche "entrée" ou "return"

Placez le curseur à la cellule A1
Entrez 8.31451 ↵
Entrez PRESSION/Pa ↵

Déplacez le curseur en B2
Entrez TEMPERATURE/K ↵

Déplacez le curseur en C2
Entrez VOLUME/m^3 ↵ (pour changer 3 en index : choisissez 3, Format, Cellule, Index)

Nous allons maintenant entrer les pressions dans la colonne A.

Déplacez le curseur en A3
entrez

50000 ↵ (pas d'espaces, pas de virgules)
60000 ↵
70000 ↵
80000 ↵
...
150000 ↵
(comme la 1ère colonne de la TABLE 1).

Il y a une autre méthode automatique d'entrer une colonne des chiffres, qui augmentent par une constante :

entrez 50000 ↵, restez dans la même cellule
ÉditionRecopier Série :
Série en colonne
Type : linéaire
Valeur du pas : 10000
Dernière valeur : 150000

IMPORTANT : n'entrez pas les chiffres qui apparaissent dans la 3ème colonne, ces données seront calculées plus loin en utilisant Excel

Afin de produire les lignes autour des données, nous allons utiliser une 1ère commande d'Excel.

Placez la flèche de la souris sur la case A1 et cliquez le bouton gauche de la souris. La cellule s'illumine aussitôt. Gardez le doigt sur le bouton et déplacez la flèche vers la droite. Vous constatez que les cellules s'illuminent tour à tour. Illuminez toutes les cellules autour desquelles vous voulez tracer un trait.
Cliquez sur le bouton de **droite** de la souris
Une fenêtre apparaît alors
Cliquez sur Format de cellule, bordure
Cliquez Contour
Cliquez le type de ligne (simple, double, épais)

Cliquez OK
Les cellules devraient maintenant ressembler au tableau 1 (moins les chiffres de la colonne trois).
L'alignement des chiffres et du texte des colonnes est peut être différent. Nous allons laisser le tableau tel quel pour l'instant. Nous verrons comment utiliser la commande "**Alignement**" plus tard.

Sauvegarde d'un fichier
1d) Nous allons maintenant sauvegarder la page sur une disquette.

Cliquez **Ficher** sur la barre du menu
Cliquez **Enregistre sous**
Cliquez **Votre dossier sur le disque : DONNEE**

Entrez le nom du fichier dans **FILE NAME** (par exemple GAZ)
Cliquez **Enregistrer**

Maintenant pour être certain que le fichier est bien sauvegardé
Cliquez **Fichier**
Cliquez **Ouvrir**
Toutes les données disparaissent de la fenêtre.
Maintenant nous allons retourner chercher le fichier sur la disquette
Cliquez **Fichier**
Cliquez **Ouvrir**
Cliquez **A:** dans la fenêtre
Sélectionnez GAZ dans la fenêtre
Cliquez **Ouvrir**

Après quelques secondes l'écran devrait contenir le tableau que vous avez créé à l'origine. Si tel n'est pas le cas, demandez de l'aide immédiatement. Vous avez probablement oublié une étape.

2. Dans cette section nous allons

a) Utiliser les commandes "Édition, copier" pour calculer le volume d'une mole d'un gaz parfait correspondant aux pressions indiquées dans la colonne des données.
b) Nous utiliserons également les commandes "Format, cellules, Alignement" pour centrer les données et les titres de colonnes.
c) Nous utiliserons la commande "Format, colonne, ajustement, ou largeur" pour ajuster la largeur d'une colonne (plus grand ou plus petit).

2a) Calcul avec Excel

L'utilisation la plus évidente de la commande "copier" est la copie d'une colonne ou d'un bloc de donnée ailleurs dans la page par exemple, si on fait l'opération suivante :
Sélectionnez un bloc de donné à l'aide de la souris
Relâchez le bouton de la souris
Amenez le pointeur vers le centre du bloc à déplacer.
Enfoncez à nouveau le bouton gauche et gardez-le enfoncé jusqu'à ce que la main apparaisse (au lieu du pointeur).
Gardez le doigt sur le bouton et déplacez le bloc de données à l'endroit voulu.
Relâchez le bouton pour que les données apparaissent à leur nouvelle position.
La commande "**Copier**" du menu "**Édition**" peut être utilisée pour effectuer la même opération.
Nous allons maintenant utiliser la commande "**Copier**" du menu "**Édition**" pour copier des équations dans les cellules.

Déplacez le curseur à la cellule C3. Vous êtes alors dans la colonne des volumes dans la rangée vis-à-vis la pression de 50000 Pa.
Cliquez sur la cellule C3.

Nous allons maintenant entrer dans cette cellule la formule RT/P

Entrez : =+$B$1*$B$3/A3 (Entrée)

Dans la cellule C3 apparaît alors 0.04988865 soit la valeur du volume d'une mole d'un gaz parfait à une pression de 50000 Pa et à une température de 300°K.
La signification des symboles + $ * / sera expliquée en peu plus loin.
Nous allons maintenant copier cette formule dans toutes les cellules de la colonne C (celles qui correspondent à une Pression donnée).

Cliquez **Édition**
Cliquez **Copier**

Dans la fenêtre de conversation sélectionner les cellules C5..C15

Cliquez **Coller**

La même série de nombres qui apparaît dans la colonne 3 de la table 1 devrait alors apparaître sur l'écran. Si vous placez le curseur sur l'une des cellules de la colonne C vous pouvez voir le contenu de celle-ci sur la ligne d'entrée.
Vous constaterez que la formule qui apparaît sur la ligne d'entrée change, en déplaçant le curseur d'une cellule à l'autre,

i.e.	+$B$1*$B$3/A4	Cellule C4
	+$B$1*$B$3/A5	Cellule C5
	+$B$1*$B$3/A6	Cellule C6

on voit donc que la commande "Copier" a reproduit la formule dans chacune des cellules en gardant la cellule A1 constante, remarquez cependant la division se fait par A4, A5, A6, . Nous allons maintenant expliquer la signification des symboles + - $ /.

Le **+** ou = au début est nécessaire pour indiquer que ce qui suit est une opération mathématique. Remarquez que lorsque l'on copie une formule, l'adresse varie à chaque cellule, i.e. A5, A6, A7...
Si on veut qu'une valeur demeure constante lors de l'opération "Copier" on précède la rangée ou la colonne ou les deux qui doivent demeurer constant par le signe $.
Le **$.** Ici le $B$1 représente la valeur 8.314510... et doit demeurer constante dans toutes les cellules, c'est pourquoi on écrit $B$1. On peut changer automatiquement B1 en $B$1 en se plaçant juste après B1 et pressant une **touche F4**.
Le ***** c'est le symbole pour la multiplication de deux nombres.
Le **/** la division, ^ pour exponentiel.

La commande Enregistrer :
Il est bon d'utiliser cette commande de temps à autre durant une session de travail afin d'éviter la perte d'un travail.

Cliquez Fichier
Cliquez Enregistrer

Le fichier se sauvegarde sous le nom et dans le répertoire indiqué en "Enregistre sous".

2b) Les commandes Format, cellule, alignement servent à justifier les données et le texte dans les colonnes. Il faut éviter de centrer les données à l'aide d'espaces spécialement pour les nombres, car ceux-ci deviennent alors interprétés différemment et une suite de problèmes sérieux s'ensuit. Donc pour centrer ou aligner les colonnes on utilise les commandes

" Format, cellule, alignement ".

Sélectionnez le bloc A1..C15 à l'aide de la souris.

Cliquez le bouton **centre** des commandes rapides.

2c) Largeur de la colonne

Sélectionnez le bloc de donnée à l'aide de la souris. "Format, colonne, ajustement, ou largeur" pour ajuster la largeur d'une colonne (plus grande ou plus petite).

3. Graphes

Nous allons maintenant utiliser le menu Insertion, graphique pour produire un graphe du volume en fonction de la pression

a) Cliquez **Insertion**

Cliquez **Graphique**

Une fenêtre de dialogue apparaît

Type de graphique

b) Cliquez Nuages des points, suivant.

La fenêtre apparaît et vous pouvez sélectionner les données de l'axe des X ,Y.

Plage des données: A3:C12 (Par défaut a:l'axe X)

Séries en : Colonne
c) Cliquez Suivant

Entrez le titre du graphique (**GAZ IDEAL**)

Entrez le titre de l'axe X: Pression/Pa
Entrez le titre de l'axe Y: Volume/m^3

Le graphe de la FIGURE 3 devrait apparaître à l'écran.

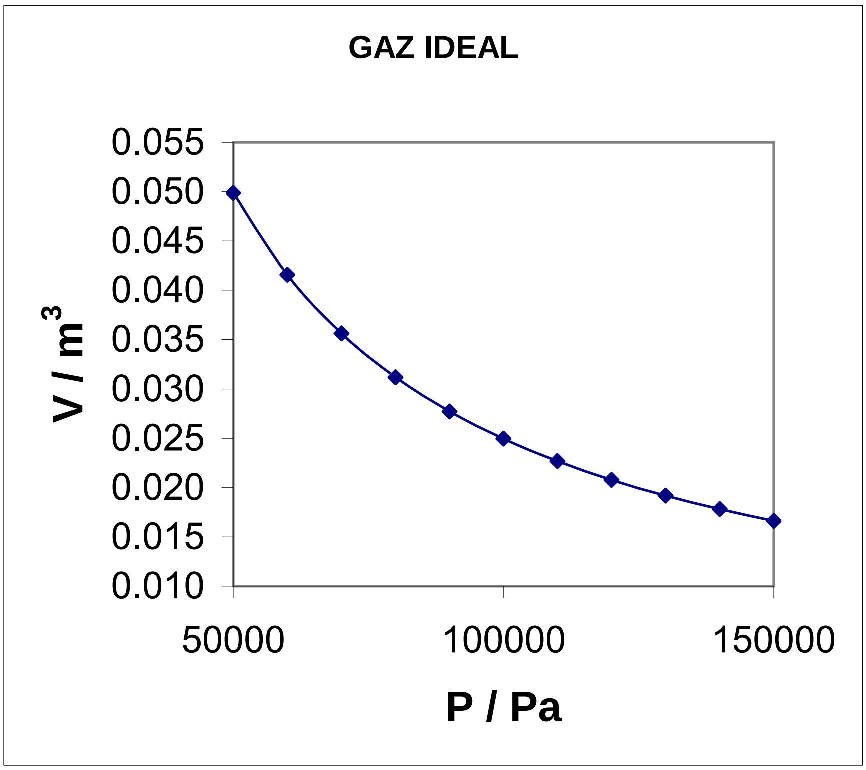


FIGURE 3

Vous pouvez ajuster tous les paramètres du graphe en cliquant le bouton de droite à l'endroit approprié (Axe des X, des Y, etc...).

Graphique 3D en utilisant Sigma Plot

Exercice 2

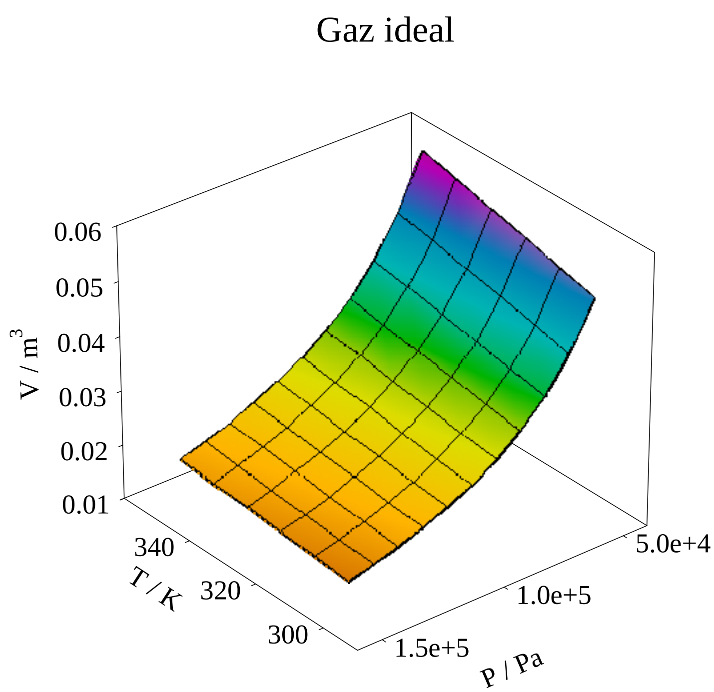
Préparer un tableau des valeurs de volumes dans la colonne C, V=RT/P ou (8.31451*B1/A1) en fonction de T(B1) et de P(A1). Exemple :

Continuez le Tableau jusqu'à la température de 350 K. Copier (Ctrl C) ce tableau dans la mémoire presse-papiers d'ordinateur.
Démarrer le programme Sigma Plot
Recopier le tableau dans la cellule 1,1 de Sigma Plot (sans titres)
Graph
Create Graph
3D Mesh Plot
Next
XYZ Triplet
Next
X : Column 2 (Température T)
Y : Column 1 (Pression P)
Z : Column 3 (Volume V)
Finish
Cliquer sur la souris à droite
Graph Properites
Graph
Rotation
Jouer avec horizontal et vertical view, OK.

50000	300	0.049887
60000	300	0.041573
70000	300	0.035634
80000	300	0.031179
90000	300	0.027715
100000	300	0.024944
110000	300	0.022676
120000	300	0.020786
130000	300	0.019187
140000	300	0.017817
150000	300	0.016629
50000	310	0.051550
60000	310	0.042958
70000	310	0.036821
80000	310	0.032219
90000	310	0.028639
100000	310	0.025775
110000	310	0.023432
120000	310	0.021479
130000	310	0.019827
140000	310	0.018411
150000	310	0.017183
50000	320	0.053213
60000	320	0.044344
70000	320	0.038009
80000	320	0.033258
90000	320	0.029563
100000	320	0.026606
110000	320	0.024188
120000	320	0.022172

130000	320	0.020466
140000	320	0.019005
150000	320	0.017738
50000	330	0.054876
60000	330	0.045730
70000	330	0.039197
80000	330	0.034297
90000	330	0.030487
100000	330	0.027438
110000	330	0.024944
120000	330	0.022865
130000	330	0.021106
140000	330	0.019598
150000	330	0.018292

On ne doit pas copier les titres des colonnes dans le SigmaPlot.



Ajoutez les titrez aux colonnes

Cliquer Y data, changer le titre en T/K, choisir la police au 18 points
Cliquer X data, changer le titre en P/Pa, choisir la police au 18 points
Cliquer Z data, changer le titre en V/m^3, choisir la police au 18 points
Cliquer 3D Graph, changer le titre en Gaz Idéal, choisir la police au 24 points
Cliquer le bouton droit dans le Graph
Cliquer Graph Properties
Plots Mesh Colir : Rainbow, Transition : Gradient
Enregistrer le Graphique.

Exercice 3
La moyenne

L'Excel permet de calculer une moyenne l'écart-type, l'écart-type de la moyenne, etc. Dans cet exemple on va générer une série des chiffres aléatoires et calculer la moyenne, l'écart-type et les intervalles de confiance.

Statistiques descriptives

1) génération des nombres aléatoires avec
Pressez :

Outils	
Utilitaire d'analyse	
Génération de nombres aléatoires	
Nombre de variables	1
Nombre d'échantillons générés	20
Distribution	Normale
Moyenne	4.95
Écart-type	0.1
Plage de sortie :	A1
OK	

2) statistiques descriptives des valeurs générées

Outils	
Utilitaire d'analyse	
Statistiques descriptives	
Plage d'entrée	a1..a20
Groupées par	Colonnes
Plage de sortie	b1
Rapport détaillé	
Niveau de confiance pour la moyenne	95%
Kième maximum	1
Kième minimum	1

x_i	Statistique descriptive			
4.969976784	*Colonne1*			
4.872231683				
5.024425731	Moyenne	4.955828122		
5.127647354	Erreur-type	0.025430154	=C7/20^0.5	de la moyenne
5.119835022	Médiane	4.951401378		
5.17331331	Mode	#N/A		
4.781641236	Écart-type	0.113727107		
4.976581876	Variance de l'échantillon	0.012933855		
5.109502253	Kurstosis (Coefficient d'aplatissement)	-0.584378401		
4.891329935	Coefficient d'asymétrie	0.324021929		
4.930979584	Plage	0.391672074	MAX-MIN	
4.830956767	Minimum	4.781641236		
4.815308911	Maximum	5.17331331		
4.90223705	Somme	99.11656244		
4.922649295	Nombre d'échantillons	20		
4.788206878	Maximum(1)	5.17331331		
4.943207513	Minimum(1)	4.781641236		
4.959595243	Niveau de confiance(95.0%)	0.053225941	5.00905406	=C3+C18
5.013485305			4.90260218	=C3-C18
4.963450705			intervalle de confiance	

LOI.STUDENT.INVERSE(0.05,19)= t	2.093024705
C4*B23=xm*t	0.053225941

Exercice 4

La Méthode de Moindres Carrés

Excel est également utilisé dans le calcul des régressions linéaires (moindres carrés). Pour la compréhension de ce qui va suivre il est suffisant de savoir que la régression est une méthode mathématique qui permet de calculer, à partir de points x et y observés, la meilleure valeur de la pente et de l'ordonnée à l'origine pour la droite de régression (la droite qui passe le plus près possible du plus grand nombre possible de points x, y).
Dans le but d'étudier l'utilisation de l'Excel dans ces calculs, la température d'ébullition de l'eau pure est donnée à différentes pressions dans la table 3.

Table 3

TEMP / C	PRESSION / mm Hg	TEMP / K	Ln P	1 / T K^{-1}
10	9.21	283.15	2.22029	0.00353
20	17.54	293.15	2.86448	0.00341
30	31.82	303.15	3.46010	0.00330
40	55.32	313.15	4.01313	0.00319
50	92.51	323.15	4.52732	0.00309
60	149.38	333.15	5.00649	0.00300
70	233.70	343.15	5.45404	0.00291
80	355.10	353.15	5.87240	0.00283
90	525.76	363.15	6.26484	0.00275
100	760.00	373.15	6.63332	0.00268

La température d'ébullition est donnée dans la première colonne et la pression correspondante est donnée dans la seconde colonne. Il est bien connu que ces données sont décrites par l'équation de Clausius-Clapeyron

$$\ln P = -\frac{\Delta H_v}{RT_v} + a \qquad (1)$$

où T_V est la température d'ébullition de l'eau en K, et P est la pression correspondante, R est la constante des gaz et ΔH_V est l'enthalpie de vaporisation de l'eau pure, a est une constante.

Si on fait

$y = \ln P$ $\qquad X = 1/T_v$ $\qquad$ et $\qquad b = -\frac{\Delta H_v}{R}$

Alors l'équation (1) s'écrit

$y = bx + a$

On voit que si on détermine b à partir d'un moindre carré on peut calculer ΔH_V car

$\Delta H_V = - bR$

Si l'écran contient des données faire file close
Entrez dans la première colonne les données de la TABLE 3 ainsi que celles de la deuxième colonne. La température est donnée en C et on doit la convertir en K pour que l'équation (1) s'applique.

Placez le curseur à la case C3
Entrez=+A3+273.15

La température en Kelvin apparaît à la case C3.
Utilisez la commande "Édition". Copier pour calculer les autres températures en Kelvin.
Nous allons maintenant calculer la valeur ln P = y.
Placez le curseur en D3 et cliquez sur la ligne d'entrée.
Entrez =+ln (B3).
Faite entrée sur le clavier.
Utilisez la commande "Édition", Copier pour reproduire cette formule pour toutes les températures.
Il reste à calculer 1/T.
Placez le curseur en E3
Entrez =+1/C3
la valeur de 1/T apparaît en E3

Utilisez "Édition", Copier pour calculer toutes les autres valeurs de 1/T. Pour Copier une formule on peut aussi cliquer deux fois dans le point en bas à droite d'une cellule.

Le tableau contient maintenant les valeurs montrées à la table 3. Vous pouvez utiliser les commandes **"Format, cellule, alignement" et "Bordure"** pour améliorer l'apparence du tableau.

Les colonnes nécessaires pour faire la régression sont la quatrième, ln P = y et la cinquième 1/T = x.
Cliquez Outils
Cliquez Utilitaires d'analyse
Cliquez Régression linéaire
Entrez E3.E12 dans la variable indépendante
Entrez D3.D12 dans la variable dépendante
Entrez A17 dans la plage de sortie
Cliquez OK

Le résultat de la régression apparaît dans la partie de l'écran que vous avez choisie (Table 4).

RAPPORT DÉTAILLÉ

Statistiques de la régression	
Coefficient de détermination multiple	0.999941817
Coefficient de détermination R^2	0.999883638
Coefficient de détermination R^2	0.999869092
Erreur-type	0.016942091
Observations	10

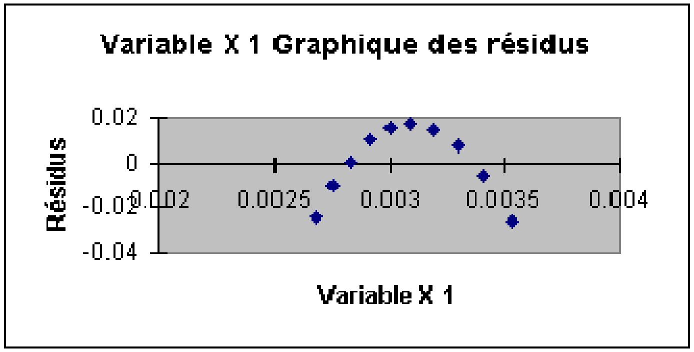
ANALYSE DE VARIANCE

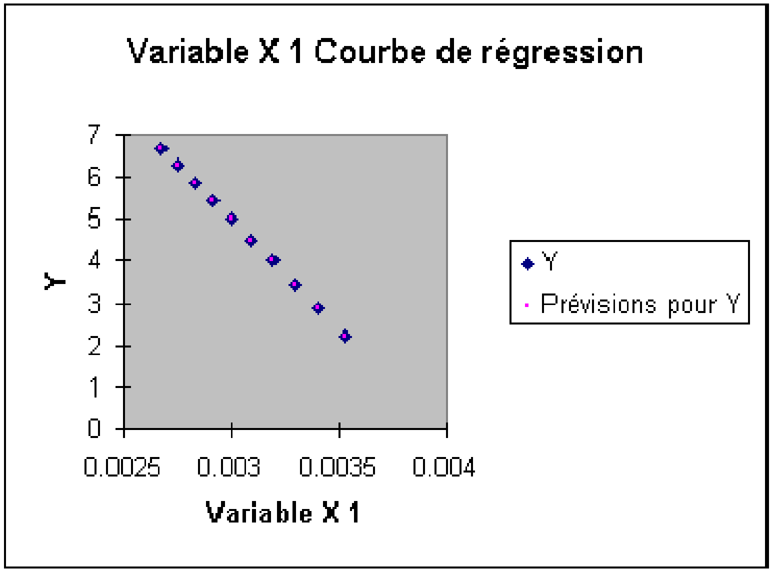
	Degré de liberté	*Somme des carrés*	*Moyenne des carrés*	*F*	*Valeur critique de F*
Régression	1	19.73151617	19.73151617	68742.6737	5.01338E-17
Résidus	8	0.002296276	0.000287034		
Total	9	19.73381244			

	Coefficients	*Erreur-type*	*Statistique t*	*Probabilité*	*Limite inférieure pour seuil de confiance = 95%*	*Limite supérieure pour seuil de confiance = 95%*	*Limite inférieure pour seuil de confiance = 95.0%*	*Limite supérieure pour seuil de confiance = 95.0%*
Constante	20.53593817	0.060895974	337.2298183	6.6942E-18	20.39551171	20.67636463	20.39551171	20.67636463
Variable X 1	-5178.769682	19.75210506	-262.1882411	5.0134E-17	-5224.318147	-5133.221217	-5224.318147	-5133.221217

ANALYSE DES RÉSIDUS

Observation	*Prévisions pour Y*	*Résidus*	*Résidus normalisés*
1	2.246092919	-0.025803069	-1.615401951
2	2.870000313	-0.005516326	-0.34534976
3	3.452746077	0.007348945	0.460080965
4	3.998273527	0.01486098	0.930372101
5	4.510037867	0.01727888	1.081744821
6	4.991079449	0.015413946	0.964990584
7	5.44408422	0.009954022	0.62317184
8	5.8714339	0.00096554	0.060447665
9	6.275247868	-0.010403034	-0.651282252
10	6.657418317	-0.024099884	-1.508774012





Comme exercice supplémentaire vous pouvez preparer le graphe de la figure 5.

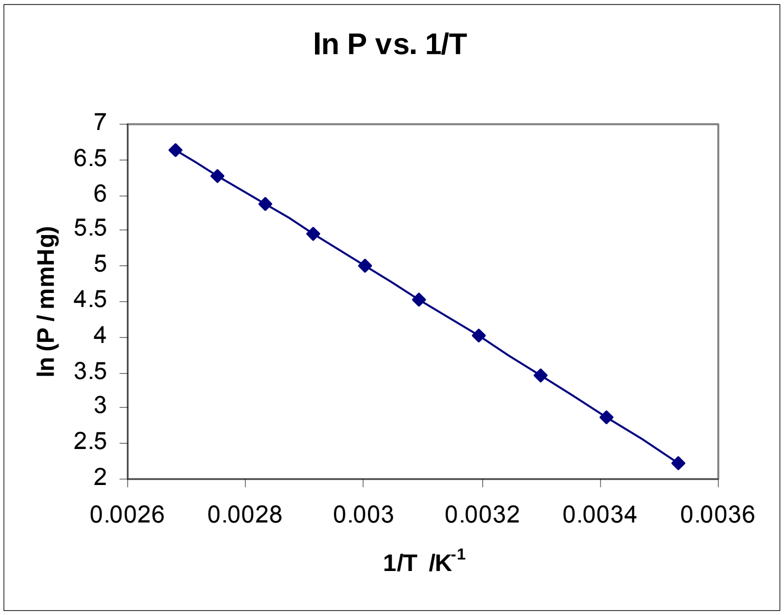


FIGURE 5

Exercice 5

Générez 10, 20, 50 et 100 chiffres aléatoires possédants une distribution gaussienne, avec la moyenne 0 et l'écart type 1. Déterminez la valeur moyenne et l'écart type de ces séries.

Outils
Utilitaire d'analyse
Génération de nombres aléatoires

Nombre de variables	1
Nombre d'échantillons générés	20
Distribution	Normale
Moyenne	0
Écart-type	1
Plage de sortie :	A1

OK

Calculez $\overline{x}$
Allez à la place où vous voulez entrer vos résultats
Insérer, Fonction" ou l'icône: f_x
Catégorie de fonctions: statistiques
Nom de fonction: MOYENNE"
Entrez une liste des cellules avec une souris ou dans le format, p. ex.: "a1:a20"

Calculez *s*:
f_x, Catégorie de fonctions: statistiques
Nom de fonction: ECARTYPE
Calculez l'écart-type de la moyenne : = $s/\sqrt{N}$, [*s*/ RACINE(20)]
Donnez les résultats dans un tableau et les encadrez.

Exercice 6

Générez une courbe *y* = 1.2 + 0.7**x* entre *x* = -5 et 9, (Δx = 2). Ajoutez un bruit gaussien aux valeurs de *y* (*s* = 0.5) (moyenne 0, écart-type 0.5). Déterminez les paramètres de la régression linéaire : la pente, l'ordonnée à l'origine est ses écarts-types, les limites de confiance, le coefficient de corrélation. Comparez des données expérimentales (générées) avec la ligne de régression sur un graphique avec les points expérimentaux, une ligne calculé en utilisant la régression linéaire.

But : comparaison des valeurs (y) obtenues par la régression linéaire avec des valeurs (y) obtenues par simulation (simulation des valeurs expérimentales de y avec bruit gaussienne).

- Créer une colonne des valeurs de «x» de -5 à +9 par intervalle de 2 (Δx = 2)
- Dans une autre colonne calculer une droite théorique selon la relation suivante : y = 1.2 + 0.7• x ($y_{calculé}$)
- Créer un bruit gaussien (chiffres aléatoires) avec une moyenne = 0 et l'écart type = 0.5 (colonne 3 = y_{bruit})
- Additionner les valeurs de «y_{bruit}» aux valeurs de «y» de la deuxième colonne; les valeurs ainsi obtenues («$y_{expérimental}$») simulent les valeurs expérimentales de «y», vous aller voir sur le graphique qu'ils sont plus éparpillés que les valeurs de la deuxième colonne
- Faire la régression linéaire sur les points expérimentaux simulés.
- Mettre dans un tableau les paramètres de la régression tel : ordonnée à l'origine, pente, écart type et intervalles de confiance
- Comparer l'ordonnée et la pente avec les paramètres de la droite théorique y = 1.2 + 0.7• x ($y_{calculé}$).
- Mettre sur le même graphique : les points expérimentaux (en points), la droite théorique (en ligne) et la ligne de la régression linéaire (ligne pointillée).

Exercice 7

Générez des valeurs X-Y de –5 á 5 dans l'Excel, selon l'exemple :

-5	-5
-5	-4
-5	-3
-5	-2
-5	-1
-5	0
-5	1
-5	2
-5	3
-5	4
-5	5
-4	-5
-4	-4
-4	-3
-4	-2
-4	-1
-4	0
-4	1
-4	2
-4	3
-4	4
-4	5
-3	-5
-3	-4
-3	-3
-3	-2

etc. jusqu'à +5.

Calculez des fonctions dans la zone de x et y {-5,5} :

a)

$$z = \exp\left[-\left(\frac{x}{3}\right)^2 - \left(\frac{y}{3}\right)^2\right]$$

utilisez doubles parenthèses EXP(-((A2/3)^2)-((B2/3)^2))

b)

$$z = \exp\left[-\left(\frac{x}{3}\right)^2 - \left(\frac{2y}{3}\right)^2\right]$$

c)

$$z = \frac{1}{1 + x^2 + (2 * y)^2}$$

d)
$z = \sin(x+y)$

Exécutez les graphiques dans le Sigma Plot. Exécutez les calcules dans l'Excel (vérifiez les calcules). Placez chaque graphique sur une autre feuille graphique dans le SigmaPlot.

EXPÉRIENCE DEUX

STATISTIQUES

OBJECTIFS PEDAGOGIQUES

Objectif général
Cette expérience a pour but de compléter l'étude du logiciel Excel. Comme autre objectif, cette expérience introduit l'étude des paramètres statistiques, soit la moyenne, l'écart-type et l'écart-type de la moyenne. Ces paramètres sont étudiés en fonction du nombre de lectures. L'intégration numérique de la fonction de Gauss sera également étudiée ainsi que les intervalles de confiance. La régression linéaire.

Objectifs spécifiques
Dans cette expérience l'étudiant
a) fera la différence entre les concepts de population et d'échantillon;
b) définira le concept de mesure aléatoire;
c) déduira les propriétés de la moyenne en fonction du nombre de mesures;
d) déduira les propriétés de l'écart-type en fonction du nombre de mesures;
e) déduira les propriétés de l'écart-type de la moyenne en fonction du nombre de lectures;
f) calculera l'intervalle de confiance 95% de la moyenne et d'une lecture;
g) tracera le graphe de la moyenne, de l'écart-type et de l'écart-type de la moyenne en fonction du nombre de lectures;
h) tracera le graphe de la distribution de Gauss;
i) analysera l'évolution de la précision d'une mesure en fonction du nombre de lectures, intégrera la fonction de Gauss numériquement pour divers intervalles.
j) effectuera la régression linéaire

CHAPITRES À LIRE : Chapitre 4, Chapitre 5, Chapitre 6

POPULATION ET FONCTION DE DISTRIBUTION

BUT: Calculer la moyenne, l'écart type et l'écart type de la moyenne en fonction du nombre de lectures (valeurs cumulatives). Également on veut tracer la courbe gaussienne qui correspond à une moyenne et un écart type donné et faire l'intégration numérique de la fonction de Gauss pour divers intervalles. Effectuer la régression linéaire.

THEORIE

On suppose que tout ensemble de points expérimentaux est un échantillon aléatoire tiré d'une population hypothétique infinie distribuée suivant une loi normale de Gauss donnée par:

$$P_G(x,\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad (1)$$

FONCTION DE GAUSS

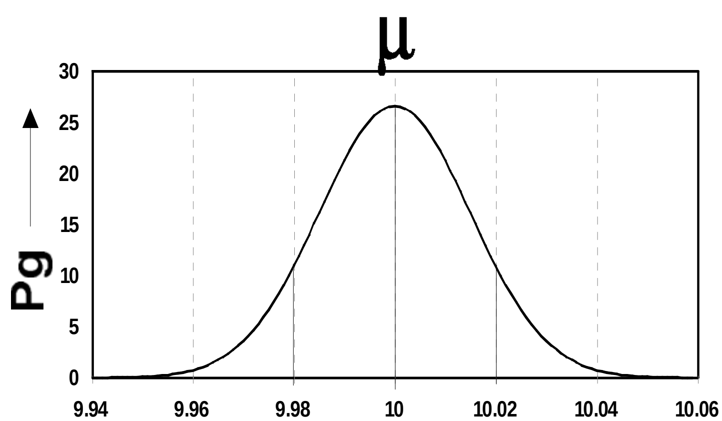


FIGURE 1

où x est la valeur d'une observation aléatoire, μ est la moyenne de la population et σ est l'écart type. La Figure 1 montre un exemple de graphe de P_G en fonction de x. Pour déterminer μ et σ, il faudrait en théorie un nombre infini de mesures, en pratique, si le nombre de lectures est d'au moins 20 à 30, on peut estimer de façon très satisfaisante μ et σ. Ces 20 ou 30 lectures forment ce que l'on appelle l'échantillon. La moyenne $\bar{x}$ de l'échantillon ainsi que l'écart type sont les meilleurs estimés de μ et σ. On a donc

$$\mu=\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i \qquad (2)$$

$$\sigma \cong s = \left[\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2 \right]^{1/2} \tag{3}$$

On pourrait démontrer que l'écart type de la moyenne σ_μ est donné par:

$$\sigma_\mu = s_\mu = \frac{s}{\sqrt{N}} \tag{4}$$

où N est le nombre total de mesures.

PROBABILITE INTEGRALE ET INTERVALLE DE CONFIANCE

$P_G(x,\mu,\sigma)$ n'est qu'une fonction mathématique qui définit la distribution des points autour de la "vraie valeur" μ. Ce qui nous intéresse de connaître, c'est la probabilité qu'une mesure aléatoire sera située dans les limites $\mu \pm z\sigma$ (z = 1, 2, 3, etc.); en terme mathématique, on veut évaluer l'intégrale suivante:

$$A_G = \int_{\mu-z\sigma}^{\mu+z\sigma} P_G(x,\mu,\sigma)\,\mathrm{d}x \qquad \text{la probabilité intégrale} \tag{5}$$

i.e. on veut connaître la surface sous la courbe comprise entre $\mu - z\sigma$ et $\mu + z\sigma$, (FIGURE 2). Cependant, la fonction de Gauss n'est pas intégrable analytiquement et on doit utiliser des méthodes de calculs numériques pour obtenir la valeur de A_G.

PROBABILITÉ INTÉGRALE

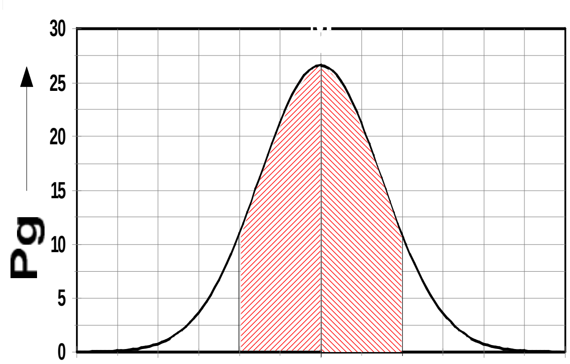


FIGURE 2

Comme nous avons vu auparavant il y a une distribution pour chaque paire de μ et σ. Il serait évidemment impossible de construire une table pour chaque paire, aussi nous allons en sélectionner une, construire une table de surfaces, et utiliser cette table avec les formules de conversion appropriées pour calculer les probabilités pour n'importe laquelle variable distribuée normalement. Cette distribution est la distribution normale standard (normalisée), et elle est définie comme la distribution normale de moyenne 0 et de variance 1.

Si on définit $z = \frac{|x-\mu|}{\sigma}$ et σ = 1 on peut écrire P_G en terme de z de la façon suivante:

$$P_G(z,0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

On dit alors que $P_G(z,0,1)$, est une distribution normale standard de moyenne 0 et d'écart type 1 (figure 5).Et ainsi on écrit

$$A_G(z, 0, 1) = \int_{-z}^{+z} P_G(u, 0, 1)\, du \qquad (6)$$

COURBE STANDARD NORMALE

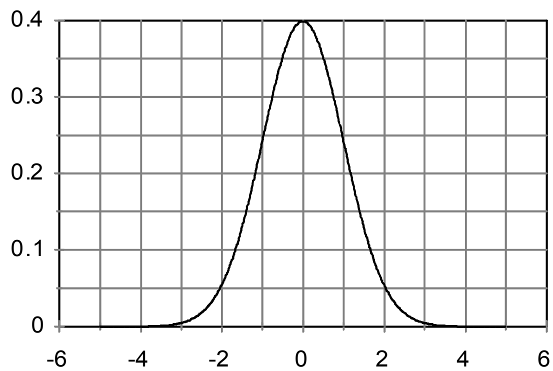


FIGURE 5

Les valeurs de $A_G(z, 0, 1)$ sont tabulées dans la table 1 de l'appendice pour différentes valeurs de z.

Exercice 1

But : Calculer la moyenne, l'écart-type et l'écart-type de la moyenne en fonction de nombre des échantillons. Placer les résultats dans des graphiques.

Plusieurs expériences classiques sont décrites dans la littérature, ayant comme objectifs l'étude de la distribution des points autour de la moyenne. Par exemple on pourrait vous demander de couper quelques centaines de pailles à la même longueur de mesurer la longueur de celles-ci à l'aide d'un micromètre et de porter les résultats en graphique.

D'autres expériences consistent à peser le volume d'eau contenu dans une pipette jaugée à vingt ou trente reprises et à l'aide des données on tente d'établir la courbe de Gauss.

Dans cette expérience nous allons utiliser les commandes Outils, Utilitaire d'analyse, Génération de nombres aléatoires, dans le but de générer 100 points (mesures) que nous utiliserons pour étudier la courbe de distribution Gaussienne.

Pour générer le nombre de points (mesures) :

Placer le curseur à la ligne A1. Entrez 1, ↓.
Rester dans A1, Cliquer
Édition, recopier, série : Série en : Colonne
Type : linéaire
Valeur du pas : 1
Dernière valeur : 100
Vous devriez maintenant avoir une liste de nombre de mesures dans la colonne A1-A100 (1-100).
Pour générer le nombre aléatoire du volume moyen d'une pipette jaugée de 10 mL.
Placer le curseur à la ligne B1.
Outils, utilitaire d'analyse, génération de nombres aléatoires, OK.
Une nouvelle fenêtre apparaît
Nombre de variables: 1
Nombre d'échantillons générés :100
Distribution: normale
Moyenne: 10.00
Écart-type: 0.015
Plage de sortie: B1.

Le programme calcule alors une série de valeurs distribuées normalement autour de 10.00 (ou de la moyenne choisie) en cliquant OK.

Notez que ces valeurs possèdent toutes les caractéristiques de mesures effectivement faites dans le laboratoire. Vous pourriez d'ailleurs comparer les valeurs générées par le programme avec celles que vous avez mesurées au laboratoire de chimie analytique quand vous avez étalonné la pipette de 10 ml.

CALCULS

A) CALCULEZ LA MOYENNE, L'ECART TYPE ET L'ECART TYPE DE LA MOYENNE APRES 1,2,3, ETC ..., DE VOS 100 MESURES (GÉNÉRÉE) :

$(\overline{x}_1,...,\overline{x}_{100})$ et (s_2, ..., s_{100}) aussi $(s_{2\overline{x}},...,s_{100\overline{x}})$

(i.e. moyenne et écart type cumulatif) en utilisant les fonctions d'Excel : moyenne, ecartype, ecartype / $\sqrt{N}$.

Portez $\overline{x}_1 \rightarrow \overline{x}_{100}$, $s_2 \rightarrow s_{100}$, $s_{2\overline{x}} \rightarrow s_{100\overline{x}}$ en graphiques en fonction de N, le nombre de mesures.

Ces paramètres cumulatifs n'ont aucune utilisation dans le domaine des statistiques. Nous voulons simplement, en calculant la moyenne après une mesure, après deux mesures, etc. être capable d'établir un graphe de $\overline{x}$, s et $s_{\overline{x}}$ en fonction de

N le nombre de mesures, ce qui nous permettra d'étudier le comportement de ces paramètres à mesure que N augmente.
Vous devriez maintenant avoir une liste de nombre dans la colonne B1-B100, ainsi que dans la colonne A1 - A100.
Les trois colonnes suivantes contiendront la moyenne cumulative, l'écart type cumulatif ainsi que l'écart type de la moyenne cumulative, et seront calculées à l'aide de la commande Édition, copier.

Calcul des moyennes cumulatives

1. Placez le curseur en C1
2. Entrez =Moyenne(B$1:B1), ↓

Rester dans C1, Cliquer sur le bouton à gauche 2 fois, on voit apparaître =Moyenne(B$1:B1) et une croix noire au côté droit de la cellule C1. Cliquer 2 fois sur cette croix.

Après cette opération les moyennes cumulatives apparaissent dans la colonne C.
Vérifiez que les cellules contiennent bien des formules (exemple =Moyenne(B$1:B20) à la ligne d'entrée et non pas des nombres.

Calcul des écarts-types cumulatifs

1. Placez le curseur en D2 (pourquoi pas D1 ?)
2. Entrez =ecartype(B$1:B2)
3. Utilisez Édition, « copier » pour copier cette formule de D2 à D100 ou, rester dans D2, Cliquer sur le bouton à gauche 2 fois, on voit apparaître = ecartype(B$1:B2) et une croix noire au côté droit de la cellule D2. Cliquer 2 fois sur cette croix.

Calcul des écarts-types cumulatifs de la moyenne

1. Placez le curseur en E2
2. Entrez +D2/racine(A2)
3. Copiez cette formule de E2 à E100

Graphe des moyennes cumulatives

1. Utilisez l'option Insertion, graphique afin de produire le graphe de la moyenne cumulative en fonction de N
2. Comme 1st séries, entrez les données de la colonne C
3. Comme l’axe des X, entrez les valeurs de la colonne A

N'oubliez pas le titre du graphe ainsi que les titres pour l'axe des y et des x.

Graphe des écarts-types cumulatifs

1. Utilisez l'option Insertion, graphique pour produire le graphe de l'écart type cumulatif

en fonction de N.
2. Entrez dans Y la colonne D. (N'entrez pas D1)
3. Entrez dans X la colonne A.
4. Faites imprimer le graphe sur papier.

Graphe des écarts-types de la moyenne cumulative

Répétez la même chose que ci-haut en entrant la colonne E comme valeurs dans les séries. (N'entrez pas E1)

N.B.: Dans tous ces graphes n'oubliez pas d'utiliser l'option XY comme graph type.
Si le graphe ne sort pas ou vous apparaît étrange, vérifiez que dans l'option X-axis et Y-axis l'option échelle est à automatique.
Vous pourrez utiliser l'option manuelle après que vous aviez visionné vos graphes dans le mode automatique.
Utilisez la touche F1 pour apprendre la signification des différentes options du menu.

Si vous n'avez pas sauvegardé votre fichier, il serait bon de le faire maintenant!

B) CALCULEZ LE MEILLEUR ESTIME DE LA MOYENNE ET DE L'ECART TYPE DE VOS CENT MESURES.

Calcul de la moyenne

Vous n'utiliserez sans doute plus jamais la façon particulière avec laquelle nous avons employé les fonctions Moyenne, écart-type dans la partie a). Nous avons utilisé cette procédure dans un but bien précis, soit l'analyse du comportement de la moyenne et de l'écart type en fonction du nombre de mesures.
Nous allons maintenant utiliser les fonctions Moyenne, écart-type d'une façon plus réelle, soit pour calculer la moyenne et l'écart-type d'un certain nombre de mesures.

1. Placez le curseur en F1
2. Entrez =Moyenne(B1:B100)
3. Notez que la moyenne de cent valeurs apparaît alors en F1

Calcul de l'écart type

1. Placez le curseur en F2
2. Entrez =ecartype(B1:B100)

L'écart type apparaît en F2 (la valeur en F1 est la même qu'en C100; la valeur en F2 est la même qu'en D100, Pourquoi ?
La moyenne et l'écart type peuvent aussi être calculés en utilisant un programme Statistique descriptive dans l'Utilitaire d'analyse.

Exercice 2

But : Gênerez la fonction de Gauss en utilisant la formule et la fonction d'Excel, distribution normalisée, intégration numérique (méthode des rectangles et de trapèzes), fonction A_G.

A) Graphe de la fonction de Gauss P_G.
La fonction de Gauss de vos mesures est décrite par la fonction

$$P_G(x,\bar{x},s) = \frac{1}{s\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{(x-\bar{x})^2}{s^2}\right] \tag{7}$$

ou

$$P_G(x,\bar{x},s) = y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

où $\bar{x}$ et s sont la moyenne et l'écart type de vos mesures.

Pour la distribution normalisée $\mu = 0$ et $\sigma = 1$, $z = \dfrac{x-\mu}{\sigma}$:

$$= P_G(z,0,1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{9}$$

Nous allons maintenant faire le graphe de l'équation (9), à l'aide d'Excel.

a) Générez la série des chiffres de -5 à 5, Δz =0.01.
Dans la colonne A entrez les valeurs z, dans la case A1, taper –5. Générer une colonne de A1 à A1001 avec des valeurs de -5 à +5 par incréments (steps) de 0.01.
Rester dans A1, Cliquer :
Édition, recopier, série : Série en : Colonne
Type : linéaire
Valeur du pas : 0.01
Dernière valeur : 5

b) calculez les valeurs de $P_G(z,0,1)$; la valeur de π dans l'Excel est Pi().
Dans la colonne B entrez dans la case B1, =1/racine(2*Pi())*exp(-0.5*A1^2).
Copier cette formule de B1 à B1001.
Dans le même Tableau calculez P_G dans la colonne C en utilisant une fonction d'Excel : LOI.NORMALE(A1; 0; 1; FAUX). Comparez les résultats.
Présentez les résultats dans un Tableau 1 (voir plus tard). Il est recommandé que le tableau porte les noms de colonnes dans le premier rangé (voir les pages suivantes).

c) Préparez un graph de P_G en fonction de *z*.
Générez un graphe XY en utilisant le contenu de A1 à A1001 comme axe des x et B1 à B1001 comme 1st séries.

Donnez un titre à ce graphe ainsi qu'aux axes X et Y.
Vous devriez voir apparaître la fameuse cloche inversée.

B) Intervalles de confiance et intégrale de $P_G(z,0,1)$.

a) Intégration-méthode des rectangles
En utilisant la courbe de Gauss, la probabilité d'obtenir un résultat à l'intérieur d'un certain intervalle autour de la moyenne est reliée à la surface sous la courbe comprise à l'intérieur de l'intervalle. Premièrement, calculons la surface comprise sous la courbe dans un intervalle de z compris entre -5 et +5 (ou si on veut dans un intervalle -5σ à +5σ). Dans le tracé du graphe nous avons divisé l'intervalle en incréments de 0.01. Si on veut trouver la surface sous la courbe nous allons diviser cette surface en 1000 rectangles, chacun ayant une largeur de 0.01 et une hauteur Y, donnée par la valeur de $P_G(z,0,1)$ listée dans la colonne B.

$$\int_{x_1}^{x_N} f(x)dx = y_1\Delta x + y_2\Delta x + ... + y_{N-1}\Delta x$$

$$= \Delta x \sum_{i=1}^{N-1} y_i$$

Parce que la largeur des rectangles est constante (0.01), la surface de chacun rectangle est donnée par 0.01 Y. La surface totale sous la courbe est donnée par 0.01 Σy_i, qui est obtenue simplement en utilisant la fonction =0.01*Somme(B1:B1001), et en plaçant la réponse dans une cellule non utilisée (F1 par exemple). Cette valeur sera très près de 1 (0.999999...) comme requis par la nature de la fonction de Gauss. Cette méthode d'intégration est le plus simple. On peut améliorer les calculs en utilisant la méthode de trapèzes.

b) Intégration-méthode des trapèzes
La surface entre chaque deux points est donnée comme une aire d'un trapèze : $(y_i + y_{i+1})/2*\Delta x$. L'aire totale est la somme des trapèzes :

$$\int_{x1}^{x_N} f(x)dx = \left[\frac{y_1+y_2}{2} + \frac{y_2+y_3}{2} + ... + \frac{y_{N-1}+y_N}{2}\right]\Delta x$$

$$= \left[\frac{y_1}{2} + \sum_{i=2}^{N-1} y_i + \frac{y_N}{2}\right]\Delta x$$

Pour le premier trapèze on calcule l'air de surface directement en utilisant une formule : +(b1+b2)/2*0.01, et à partir de troisième point, on peut utiliser la formule générale :
+(b1/2+somme(b2 :b2)+b3/2)*0.01 et recopier cette formule jusqu'à la fin.

Un exemple d'intégration numérique est démontré dans l'e fichier d'Excel ci-dessus.

Exemple des calculs d'intégration.

	A	B	C	D
1	x	y	Int. rectangles	Int. trapezes
2	-1.41	0.14764	0	0
3	-1.4	0.14973	=0.01*SUM($B$2:B2)	=0.01*(B1+B2)/2
4	-1.39	0.15183	=0.01*SUM($B$2:B3)	=+($B$2/2+SUM($B$3:B3)+B4/2)*0.01
5	-1.38	0.15395	=0.01*SUM($B$2:B4)	=+($B$2/2+SUM($B$3:B4)+B5/2)*0.01
6	-1.37	0.15608	=0.01*SUM($B$2:B5)	=+($B$2/2+SUM($B$3:B5)+B6/2)*0.01
7	-1.36	0.15822	0.007592248	0.007645179
8	-1.35	0.16038	0.009174495	0.00923822
9	-1.34	0.16256	0.010778329	0.010852911
10	-1.33	0.16474	0.012403879	0.012489385
11	-1.32	0.16694	0.014051276	0.014147769
12	-1.31	0.16915	0.015720647	0.015828188
13	-1.3	0.17137	0.017412114	0.017530765
14	-1.29	0.1736	0.0191258	0.019255619
15	-1.28	0.17585	0.020861823	0.021002867
16	-1.27	0.1781	0.022620297	0.022772624
17	-1.26	0.18037	0.024401336	0.024564999
18	-1.25	0.18265	0.026205047	0.0263801
19	-1.24	0.18494	0.028031538	0.028218032
20	-1.23	0.18724	0.029880911	0.030078895
21	-1.22	0.18954	0.031753265	0.031962788
22	-1.21	0.19186	0.033648697	0.033869805
23	-1.2	0.19419	=0.01*SUM($B$2:B22)	=+($B$2/2+SUM($B$3:B22)+B23/2)*0.01

c) Calculez les intégrales pour $z = \pm 1, \pm 2, \pm 3$ en utilisant les valeurs de P_G déjà calculées en utilisant la méthode des trapèzes

Pour trouver la surface des courbes entre ± 1, ± 2 et ± 3 identifiez simplement les intervalles appropriées dans la colonne A (la surface de -1 à +1 égale 2*surface de 0 à 1σ). Puis trouvez les surfaces, en multipliant les valeurs (par exemple) B400/2+Somme(B401:B600)+b601/2, Somme B300/2+(B301:B700)+B701/2, et B200/2+Somme(B201 :B800)+B801/2 par 0.01. Ces valeurs, quand elles sont multipliées par 100, représentent la fraction en pourcentage de la surface sous la courbe à l'intérieur des limites désignées. Elles sont exactement équivalentes à la probabilité en % de retrouver une valeur de x dans l'intervalle ± 1σ, ± 2σ, ± 3σ (environ 68, 95 et 99.7%). Vous pouvez les comparer avec les valeurs calculées en utilisant une fonction d'Excel LOI.NORMALE pour vérifier que vos valeurs pour A_G sont très près des valeurs tabulées. Rapportez vos valeurs calculées pour les différents intervalles et comparez avec les valeurs de la table 1, pour les mêmes intervalles.

d) Calculez la fonction A_G en fonction de z.

Calculez les valeurs de A_G en intégrant P_G de 0 vers z positifs; multipliez les résultats par 2 (voir Remarque plus bas).

Présentez un Tableau : z, $\int_{-5}^{z} P_G(z,0,1)dz$ calculées en utilisant (i) la méthode des rectangles, (ii) des trapèzes et (iii) les valeurs $\int_{-\infty}^{z} P_G(z,0,1)dz$ calculées en utilisant l'Excel: LOI.NORMALE(z,0,1,VRAI). Comparez les résultats dans un graphique (trois lignes, sans symboles). Comparez aussi les valeurs de A_G calculées par l'intégration et en utilisant la méthode des trapèzes et l'Excel (calculez l'erreur absolue entre ces valeurs, voir Tableau).

Présentez les résultats dans le tableau (Tableau 1):

x	P_G formule	P_G d'Excel	$\int_{-5}^{z} P_G(z,0,1)dz$ méthode des rectangles	$\int_{-5}^{z} P_G(z,0,1)dz$ méthode des trapèzes	Excel $\int_{-\infty}^{z} P_G(z,0,1)dz$
-5	1.48672E-06	1.48672E-06			2.87105E-07
-4.99	1.56287E-06	1.56287E-06	1.48672E-08	1.52479E-08	3.02369E-07
-4.98	1.64275E-06	1.64275E-06	3.04959E-08	3.1276E-08	3.18415E-07
-4.97	1.72654E-06	1.72654E-06	4.69234E-08	4.81225E-08	3.35279E-07
…					
5					

A partir de x = 0, à droite de la Table ci-dessus :

x	A_G trapèzes	A_G d'Excel	Err($A_G$trap-$A_G$calc)
0	0	0	0
0.01	0.007978646	0.007978758	-1.11635E-07
0.02	0.015956494	0.015956708	-2.1322E-07
0.03	0.023932747	0.023933053	-3.05565E-07

Remarque : Dans l'Excel on peut calculer directement la fonction A_G(z) comme :

A_G(z)=LOI.NORMALE(z,0,1,VRAI)-LOI.NORMALE(-z,0,1,VRAI)

où A_G(z) est défini comme :

$$A_G(z) \equiv A_G(z,0,1) = \int_{-z}^{z} P_G(u,0,1)du$$

Ou

A_G(z)=LOI.NORMALE.STANDARD(z)-LOI.NORMALE.STANDARD(-z)

Dans les calcules de A_G tenez compte de la relation :

$$A_G(z) = \int_{-z}^{z} P_G(u,0,1)du = 2\int_{0}^{z} P_G(u,0,1)du$$

C) Calculez l'intégrale $\int_{-2}^{-1} P_G(z,0,1)dz$ en utilisant la méthode des trapèzes et la fonction d'Excel.

Exercice 3

But : calculez la moyenne et l'écart-type d'une série des données.
Générez une série de 20 et de 100 données possédant une valeur moyenne 5 (en utilisant le générateur des chiffres aléatoires) et l'écart type 0.2 (distribution normale). Calculez la valeur moyenne et l'écart type de ces séries en utilisant « Statistiques descriptives ».

Exercice 4

But : Comparez lez courbes de Gauss pour les valeurs de σ différentes.

Tracez les courbes de Gauss avec la moyenne de 0 et l'écarte type de 0.2, 0.5, 1, 2 pour les valeurs x de -5 à 5, pas 0.01.

Exercice 5

LA MOYENNE PONDÉRÉE

But : Calculez la moyenne pondérée à partir des données suivantes :

x_i	s_i
3.5	0.2
3.8	0.3
3.3	0.1
3.6	0.2
3.4	0.3

Vérifiez si la distribution est normale selon le test kchi-deux. Utilisez les formules :

$$\overline{x} = \frac{\sum w_i x_i}{\sum w_i}, \qquad s_{\overline{x}}^2 = \frac{1}{\left(\sum w_i\right)} = \left(\sum \frac{1}{s_{x_i}^2}\right)^{-1}, \qquad \chi^2 = \sum_{i=1}^{N} \frac{(x_i - \overline{x})^2}{s_{x_i}^2} = \sum_{i=1}^{N} w_i (x_i - \overline{x})^2$$

Exercice 6

RÉGRESSION LINÉAIRE PONDÉRÉE

Calculez une régression linéaire pondérée à partir des données suivantes :

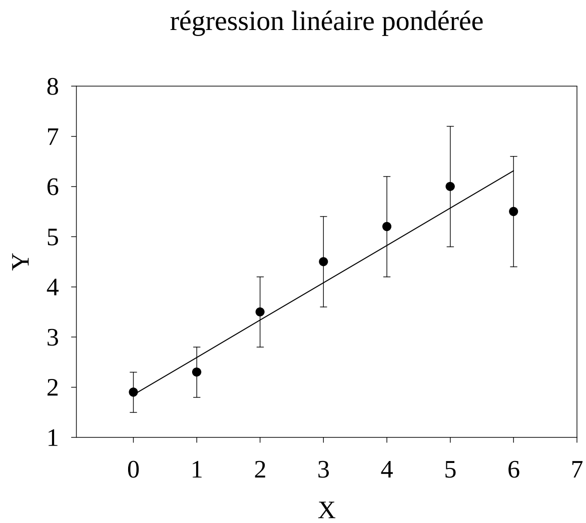
x_i	y_i	s_i
0	1.9	0.4
1	2.3	0.5
2	3.5	0.7
3	4.5	0.9
4	5.2	1.0
5	6.0	1.2
6	5.5	1.1

Pour cette exercice changez le virgule décimal en point (le programme utilise la notation avec le point) dans les Options régionales de Windows.
Préparez les données dans un format texte (ASCII) en utilisant Excel (recopiez les valeurs dans un Bloc - Note ou Notepad en anglais). Sauvegardez ce fichier dans un format texte :

x_1, y_1, s_1
x_2, y_2, s_2
x_3, y_3, s_3
…
x_N, y_N, s_N

Utilisez le programme POLFIT (c'est un programme de DOS, n'utilisez pas de souris). Sauvegardez les résultats.

Présentez un graphique $y = f(x)$ avec : points – expérimentaux avec des erreurs (l'icône des points et puis Error bars dans le SigmaPlot) et la ligne – calculée.
Graph Create Graph Scatter Plot Simple Error Bars XY Pairs
Choisissez les colonnes en cliquant sur le numéro de colonne.



DISCUSSION

a) Commentez le graphe des valeurs <u>cumulatives</u> de $\overline{x}$, *s* et $s_{\overline{x}}$.

b) L'écart type vous apprend-t-il quelque chose sur la précision de vos résultats?

c) D'après le graphe de P_G, pouvez-vous déduire certaines propriétés de cette fonction?

Si s était plus petit, quelle apparence prendrait cette courbe?

Si $\overline{x}$ est plus petit, qu'arrive-t-il à la courbe?

d) En général, si pour une mesure donnée, l'intervalle de confiance 95% est deux fois plus grand, la mesure est-elle plus précise ou moins précise?

e) En utilisant les mêmes instruments, comment peut-on augmenter la précision pour la moyenne par un facteur deux? (i.e. deux fois plus grande précision).

IMPORTANT

POUR LE RAPPORT DE CETTE EXPÉRIENCE, FAITES
UNE BRÈVE INTRODUCTION (BUT ET THÉORIE).

- IL N'Y A PAS DE SECTION SUR L'APPAREILLAGE NI LES MANIPULATIONS
- PAS DE TABLEAU DE LECTURES
- PAS D'EXEMPLE DE CALCULS NI D'ANALYSE D'ERREUR
- RÉPONDEZ AUX QUESTIONS a), b), c), d), e) DE LA DISCUSSION

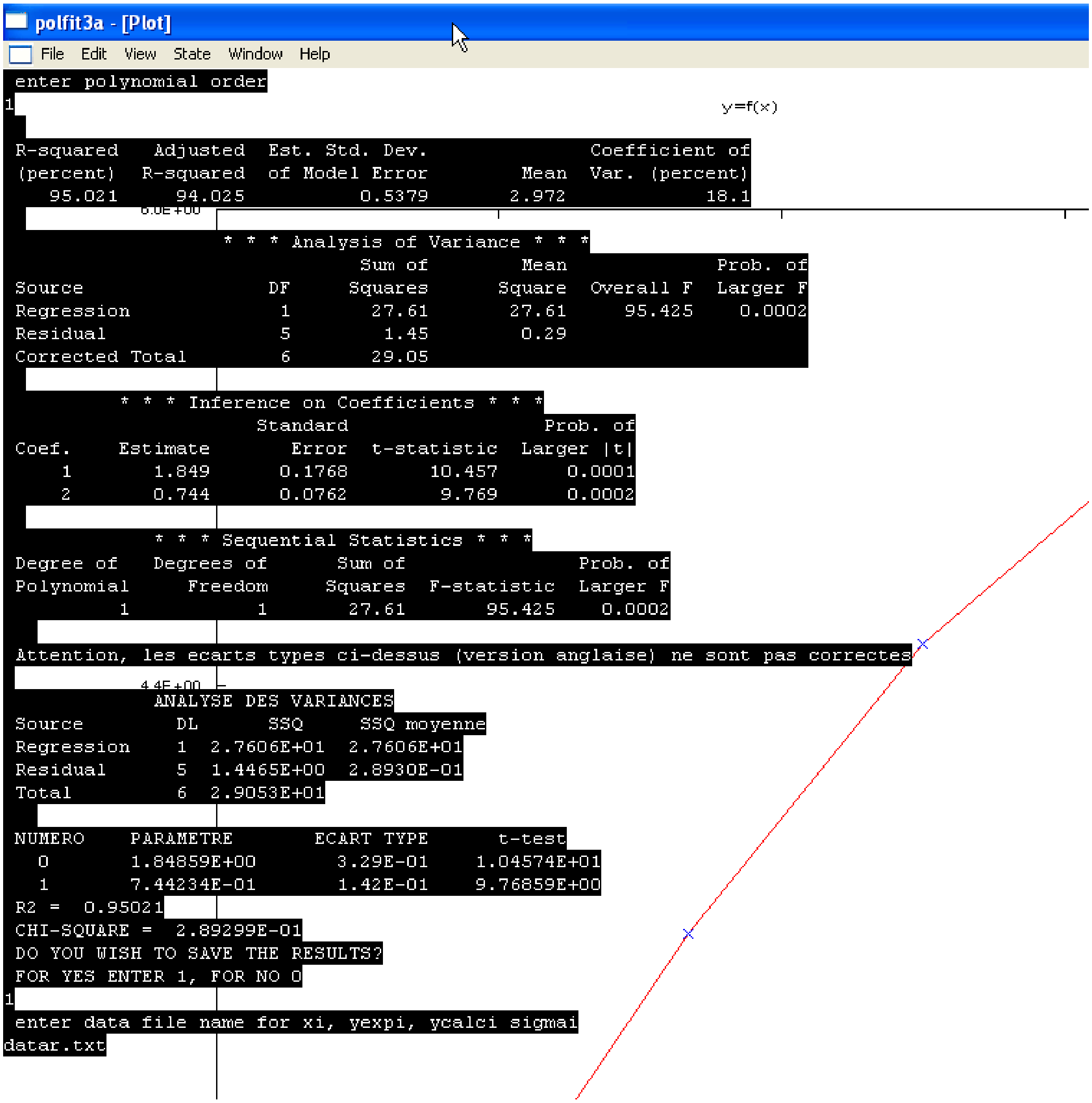
Utilisation de Polfit.exe

polfit3a - [Graphic1]

File Edit View State Window Help

```
 for statistic weights wi=1 enter 1
 for known experimental values of si enter 2
 for statistic weights proportional wi~1/si^2  enter 3
2

 program reads data xi, yi, si
 enter data file name in the format: xi, yi, sigmai
data.txt
    0.00000E+00   1.90000E+00   4.00000E-01
    1.00000E+00   2.30000E+00   5.00000E-01
    2.00000E+00   3.50000E+00   7.00000E-01
    3.00000E+00   4.50000E+00   9.00000E-01
    4.00000E+00   5.20000E+00   1.00000E+00
    5.00000E+00   6.00000E+00   1.20000E+00
    6.00000E+00   5.50000E+00   1.10000E+00
Fortran Pause - Enter command<CR> or <CR> to continue.
```



Attention : recopiez les résultats dans l'Excel. Les fonctions de choisit et de copier dans le programme POLFIT sont différentes que dans le Windows. Utilisez les commandes dans Edit.

EXPÉRIENCE TROIS

TESTS STATISTIQUES

I. Tests sur les moyennes

Objectifs

Cette expérience a pour but de montrer l'utilisation des différents tests statistiques connus sous le nom des tests de signification. Ces tests sont utilisés pour évaluer des résultats expérimentaux; par exemple ils permettent de comparer deux séries des données, deux moyennes, d'évaluer l'importance des paramètres de la régression, etc.

À lire : notes de cours, Chapitres 6, 7 et 8.

1. Comparaison de deux moyennes expérimentales (test t)

Ce test répond à la question suivante :
Est-ce que les deux moyennes (obtenues par deux méthodes différentes ou encore à partir de deux séries de mesures indépendantes) sont statistiquement identiques?
Dans ce cas nous avons deux moyennes : $\bar{x}_1$ et $\bar{x}_2$:
On utilisera les hypothèses suivantes :

- hypothèse nulle, $H_0 \rightarrow \mu_1 = \mu_2$; c'est-à-dire deux moyennes sont statistiquement identiques.
- hypothèse alternative, $H_1 \rightarrow \mu_1 \neq \mu_2$, les moyennes sont différentes.

Ils existent deux tests :
1) Test d'égalité des espérances: deux observations de variances identiques
2) Test d'égalité des espérances: deux observations de variances différentes
Si les deux séries de mesures ont les mêmes variances on utilise le test 1 est quand les variances sont différentes on utilise le test 2.

A) Test d'égalité des espérances: deux observations de variances identiques

Dans ce cas on calcule :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \qquad \text{où} \qquad s^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

et le nombre de dégrées de liberté $N_1 + N_2 - 2$. Il faut comparer la valeur de *t* calculée avec celle des tables (Excel) $t(\alpha,\ N_1 + N_2 - 2)$. Si $t_{exp} < t_{table}$ on garde H_0.

B) Test d'égalité des espérances: deux observations de variances différentes
Pour décider si l'hypothèse nulle est vraie lorsque les deux échantillons proviennes de populations avec des variances différentes (s_1 et s_2) il faut calculer le « *t* » selon :

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

pour le nombre des degrés de liberté D :

$$D = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\left(\frac{s_1^2}{N_1}\right)^2 \frac{1}{N_1+1} + \left(\frac{s_2^2}{N_2}\right)^2 \frac{1}{N_2+1}} - 2$$

en comparant la valeur de *t* calculée avec celle des tables $t(\alpha,D)$.
Pour calculer la valeur de « t » on utilise la fonction d'Excel :
$t(\alpha,D)$ = LOI.STUDENT.INVERSE(α,D).

2) Comparaison de deux variances (test *F*)

Ce test répond à la question suivante :
Les deux séries des données ont-elles le même écart-type?
Ceci nous dit si les deux séries de mesures ont la même précision.
Pour tester si la différence entre les deux variances est significative, c'est-à-dire, pour tester

$H_0 : \sigma_1^2 = \sigma_2^2$
contre $H_1 : \sigma_1^2 \neq \sigma_2^2$
Il faut calculer la valeur de F comme suit :

$$F = \frac{s_2^2}{s_1^2} \geq 1$$

où 1 et 2 sont placés de telle façon que F, soit toujours ≥ 1. Les nombres de degrés de liberté au numérateur et au dénominateur sont respectivement, n_1-1 et n_2-1. La valeur calculée doit être comparée avec celle des tables pour $F(\alpha,D_1,D_2)$, où $D_1 = n_1-1$, $D_2 = n_2$ sont les nombres des dégrées de liberté de numérateur et de dénominateur, respectivement. Pour calculer la valeur de *F* on utilise la fonction d'Excel : $F(\alpha,D_1,D_2)$ = INVERSE.LOI.F(α, D_1,D_2).
Le test assume que les échantillons proviennent des populations de distribution normale.

Exercice 1

Nous avons deux séries des donnéessuivantes de $x_{1,i}$ et $x_{2,i}$:

x_1	x_2
5.04	5.61
5.21	5.53
5.29	5.48
5.2	4.67
4.68	6.09
5.4	5.1
4.46	6.02
	5.83
	6.03

Vérifiez si les deux séries des données possèdent-elles des variances statistiquement identiques (ou non) au niveau de confiance de 95% (pour α=0.05).

H_0 : $s_1^2 = s_2^2$

H_1 : $s_1^2 \neq s_2^2$

- allez dans « outils » $\rightarrow$ « utilitaires d'analyse » $\rightarrow$ « test d'égalité des variances (F test) ».

Les résultats suivants sont obtenus :

Test d'égalité des variances (F-Test)

	Variable 1	*Variable 2*	
Moyenne	5.04	5.595555556	
Variance	0.118766667	0.223552778	
Observations	7	9	
Degré de liberté	6	8	
F	0.531269027		F_1=1/F=1.882286
P(F<=f) unilatéral	0.228290243		
Valeur critique pour F (unilatéral)	0.241149323		F_1=1/F=4.146808

La valeur obtenue de F est plus petite que 1. Il faut donc choisir comme variable 1 celle qui a la variance plus grande, c.-à-d. la deuxième colonne.

Test d'égalité des variances (F-Test)

	Variable 1	Variable 2	
Moyenne	5.595555556	5.04	
Variance	0.223552778	0.118766667	
Observations	9	7	
Degré de liberté	8	6	
F	1.882285527		
P(F<=f) unilatéral	0.228290243		
Valeur critique pour F (unilatéral)	4.146812671		=F(8,6,0.05)

Conclusion :

$F_{expérimental}$ = 1.8823 < F(0.05,8,6)$_{table}$ = 4.1468 donc on peut dire qu'au niveau de confiance de 95% il n'a pas des différences statistiques entre les deux séries de données, on garde H_0.

Exercice 2

Vous allez analyser les données de l'Exercice 1.

Question :

Les deux séries des données possèdent-elles les moyennes statistiquement identiques au niveau de confiance de 95%?

H_0 : $x_1 = x_2$

H_1 : $x_1 \neq x_2$

- Allez dans « Outils » -- « Utilitaire d'analyse » -- « test d'égalités des espérances » --« deux observations de variances égales ».
- Introduisez les ranges des valeurs et effectuez l'analyse pour α=0.05.

Les résultats :

Test d'égalité des espérances: deux observations de variances égales

	Variable 1	Variable 2
Moyenne	5.04	5.595555556
Variance	0.118766667	0.223552778
Observations	7	9
Variance pondérée	0.178644444	
Différence hypothétique des moyennes	0	
Degré de liberté	**14**	
Statistique t	**-2.608212794**	
P(T<=t) unilatéral	0.010322303	
Valeur critique de t (unilatéral)	1.76130925	
P(T<=t) bilatéral	0.020644605	
Valeur critique de t (bilatéral)	**2.144788596**	

| t expérimental | = 2.61 > $t(14,0.05)_{table}$ = 2.14
Conclusion :
Au niveau de confiance de 95% les deux moyennes sont différentes, on rejette H_0.

Exercice 3

Déterminez si les deux séries des données ont statistiquement la même moyenne et le même écart-type.

x_1	x_2
3.9	4.6
4.2	4.4
3.7	4.8
4.6	4.7
4.1	4.6
4.5	4.7
	4.4

II. Tests sur la régression linéaire

Dans le cas de la régression linéaire nous supposons que les données expérimentales peuvent être modélisées par une équation d'ordre 1 : $y = b_0 + b_1 x$.
Cependant, on peut aussi proposer deux autres modèles, soit :
1) $y = b_0$ (les données peuvent être modélisées par une moyenne et la pente n'est pas importante);
2) $y = b_1 x$ (l'ordonnée à l'origine n'est pas importante).
Cela veut dire, qu'il faut trouver si un terme additionnel est vraiment nécessaire dans l'équation de la régression. Pour tester cela **on peut utiliser le test « t » ou le test « F ».**

Théorie : test pour *b_1* :
Nos données de y_i sont-elles bien décrites par une moyenne ($y = b_0$) ou faut-il faire une régression linéaire ($y = b_0 + b_1 x$)?
Il faut tester deux hypothèses :
$H_0 : \rightarrow \; b_1 = 0$ donc $y = b_0$
contre $H_1 : \rightarrow \; b_1 \neq 0$ donc $y = b_0 + b_1 x$

Test *t*
On compare $t_{\text{experimental}} = \frac{b_1}{s_{b_1}}$ avec la valeur des tableaux $t(\alpha,N\text{-}2)$. Si $t_{exp} < t(\alpha,N\text{-}2)_{table}$ on garde l'hypothèse H_0.

Test *F*
On teste deux hypothèses :

$H_0 \rightarrow\ b_1 = 0$ donc $y = b_0 = \bar{y}$ $\qquad S_1^2 = \sum (y_i - \bar{y})^2$

contre $H_1 \rightarrow\ b_1 \neq 0$ donc $y = b_0 + b_1 x$ $\qquad S_2^2 = \sum (y_i - \hat{y}_i)^2$

- déterminez la somme des carrées S^2 et calculez la valeur de F
- S_1^2 pour $y = b_0$ et S_2^2 pour $y = b_0 + b_1 x$

$$F = \frac{\dfrac{S_1^2 - S_2^2}{1}}{\dfrac{S_2^2}{N-2}} = \frac{S_1^2 - S_2^2}{s_y^2}$$

Les sommes des carrés se trouvent dans la table de l'analyse de variance :

Source	Degré de liberté	Somme des carrés	Moyenne des carrés	Test F
Régression	1	$\sum(\hat{y}_i - \bar{y})^2$	$MS_R = \sum(\hat{y}_i - \bar{y})^2 /1$	$F = \dfrac{MS_R}{s_y^2}$
Résidus	N - 2	$SS = \sum(y_i - \hat{y}_i)^2$	$s_y^2 = \dfrac{SS}{N-2}$	
Total par rapport de la moyenne	N - 1	$\sum(y_i - \bar{y})^2$		

$$\sum_{i=1}^{N}(y_i - \bar{y})^2 = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2$$

On compare F_{calc} avec $F(\alpha,1,N\text{-}2)_{table}$. Si la valeur de $F_{calc} > F_{table}$ on accepte l'hypothèse H_1, c.à.d. que le b_1 est important car on observe une amélioration importante de l'approximation et la diminution de la somme des carrées est statistiquement importante.

Théorie : Test pour b_0

En chimie analytique et en chimie physique on cherche une corrélation linéaire entre les paramètres mesurés, par exemple une courbe de calibrage. En chimie physique l'ordonné à l'origine a souvent une signification physique. Cependant, on n'est pas sur si sa valeur peut être déterminée par la régression. Pour s'assurer, que la valeur de b_0 est statistiquement important on doit faire le test « t » pour b_0.
Une fois encore on va tester deux hypothèses :

$H_0 : \rightarrow\ b_0 = 0$ donc $y = b_1 x$ $\qquad S_1^2 = \sum \left(y_i - \hat{y}_{i,H_0}\right)^2$

Contre $H_1 : \rightarrow\ b_0 \neq 0$ donc $y = b_0 + b_1 x$ $\qquad S_2^2 = \sum \left(y_i - \hat{y}_{i,H_1}\right)^2$

$$t = \frac{b_0}{s_{b_0}}$$

Si $t_{exp} < t\,(\alpha, N\text{-}2)$ on rejette H_1 et il n'y a pas de raisons de garder le paramètre b_0.

On peut aussi utiliser le test F pour vérifier l'importance de paramètre b_0 en utilisant les paramètres S_1^2 et S_1^2 définie ci-dessus.

Exercice 4

x	y
-2.0	0.27
-1.8	0.19
-1.6	0.36
-1.4	0.49
-1.2	0.50
-1.0	0.57
-0.8	0.20
-0.6	0.42
-0.4	0.57
-0.2	0.37
0.0	0.43

Déterminer si le terme b_1 est important dans la régression. Utilisés les tests t et F.

$H_0 \rightarrow b_1 = 0$ donc $y = b_0 = \bar{y}$ $\qquad S_1^2 = \sum (y_i - \bar{y})^2$

$H_1 \rightarrow b_1 \neq 0$ donc $y = b_0 + b_1 x$ $\qquad S_2^2 = \sum (y_i - \hat{y}_i)^2$

- allez dans « outils » → « utilitaire d'analyse » → « régression linéaire »
- marquez les données x et y, niveau de confiance de 95%, demandez les résidus et les graphiques.

Tracez aussi les graphiques de :

1) y expérimental (points) et y calculé (ligne droite) en fonction de x,
2) les résidus en fonction de x.

Quel modèle : régression linéaire ou la moyenne explique le mieux les valeurs de y?
Donnez les résultats de ce modèle.

RAPPORT DÉTAILLÉ

Statistiques de la régression		*commentaires*
Coefficient de détermination multiple	0.383635718	
Coefficient de détermination R^2	0.147176364	**=R^2**=SS/total=0.026582727/0.180618182
=Coefficient de détermination R^2	0.052418182	
Erreur-type	0.130824503	
Observations	11	

ANALYSE DE VARIANCE

	Degré de liberté	*Somme des carrés*	*Moyenne des carrés*	*F*	*Valeur critique de F*
Régression	1	**0.026582727**	0.026582727	**1.55317843**	0.24412875
Résidus	9	0.154035455	0.017115051		
Total	10	**0.180618182**			

	Coefficients	*Erreur-type*	*Statistique t*	*Probabilité*	*Limite inférieure pour seuil de confiance = 95%*	*Limite supérieure pour seuil de confiance = 95%*	*Limite inférieure pour seuil de confiance = 95.0%*	*Limite supérieure pour seuil de confiance = 95.0%*
Constante	0.475	0.073794972	6.436752902	0.000119987	0.30806405	0.641936	0.308064	0.641936
Variable X 1	**0.077727273**	**0.062368135**	**1.246265794**	0.244128754	-0.0633594	0.218814	-0.06336	0.218814

ANALYSE DES RÉSIDUS

Observation	*Prévisions pour Y*	*Résidus*	*Résidus normalisés*
1	0.319545455	-0.049545455	-0.3992027
2	0.335090909	-0.145090909	-1.1690413
3	0.350636364	0.009363636	0.075445648
4	0.366181818	0.123818182	0.997640509
5	0.381727273	0.118272727	0.952959106
6	0.397272727	0.172727273	1.391715835
7	0.412818182	-0.212818182	-1.7147404
8	0.428363636	-0.008363636	-0.06738835
9	0.443909091	0.126090909	1.015952559
10	0.459454545	-0.089454545	-0.72076231
11	0.475	-0.045	-0.3625786

Test *t* (la valeur de *t* est calculée par l'Excel) *t* = 1.246265794
Ou peut être calculé comme :
$t_{expérimental} = b_1/s_{b1}$ = 0.077727273/0.062368135 = 1.246265794
t(0.05,9) = LOI.STUDENT.INVERSE(0.05,9) = 2.262158887
$t_{expérimental} < t(0.05,9)_{table}$, donc on rejet l'hypothèse H_1 : il n'y a pas des raisons de garder le paramètre b_1 (il n'est pas statistiquement important), les données peuvent être bien expliquées par une moyenne.

Test *F*

L'analyse des variances (ANOVA) nous donne **F = 1.55317843**.
La valeur de F(0.05,1,9) = INVERSE.LOI.F(0.05,1,9) = 5.117357205

Conclusion :
$F_{exp} < F(0.05,1,9)$ donc on rejette l'hypothèse H_1 et on conclue que le paramètre b_1 n'est pas important.

Les données sont donc décrites par la moyenne. Répétez les calculs en utilisant : Statistiques descriptives.

Moyenne	**0.397**
Erreur-type	**0.041**
Médiane	0.42
Mode	0.57
Écart-type	**0.13**
Variance de l'échantillon	0.018062
Kurstosis (Coefficient d'aplatissement)	-1.02408
Coefficient d'asymétrie	-0.32632
Plage	0.38
Minimum	0.19
Maximum	0.57
Somme	4.37
Nombre d'échantillons	11
Niveau de confiance(95.0%)	**0.090**

Exercice 5

Nous avons une série des données suivantes :

x	y
-2.0	-0.93
-1.8	-0.93
-1.6	-0.68
-1.4	-0.47
-1.2	-0.38
-1.0	-0.23
-0.8	-0.52
-0.6	-0.22
-0.4	0.01
-0.2	-0.11
0.0	0.03

- exécutez la régression linéaire et vérifiez l'importance du paramètre b_0
- Utilisez les tests t et F.
- Donnez les paramètres de la régression, écarts-types et les limites de confiance.

Exercice 6

Vous avez les données suivantes :

x	Y
1	3.04
1.2	3.18
1.4	3.83
1.6	4.38
1.8	4.70
2	5.15
2.2	4.85

- déterminez l'équation de la régression, les limites de confiance des paramètres de la régression et les intervalles de confiance;
- utilisez les tests « t » et « F » pour déterminer l'importance des paramètres b_0 et b_1 de la régression;
- préparer un graphique de y(expérimental, points), y(calculé, ligne de régression), les limites de confiance de $\hat{y}_i$ ($y_{calculé}$) et les limites de confiance de y_i ($y_{expérimentale}$) vs. x comme lignes continues en utilisant les équations suivantes :

$$\hat{y}_i \pm t_{0.05,N-k}\ s_{\hat{y}_i} \qquad \hat{y}_i \pm t_{0.05,N-k}\ s_{\Delta}$$

$$s_{\hat{y}_i} = s_y \sqrt{\frac{1}{N} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \qquad s_{\Delta} = s_y \sqrt{1 + \frac{1}{N} + \frac{(x_i - \bar{x})^2}{S_{xx}}}$$

Répétez le dernier graphique en utilisant SigmaPlot, qui prépare ce graphique automatiquement dans la régression linéaire.

Exercice 7

Vous avez une série des données comme suit :

x	Y
0.0	-0.9
0.5	-0.6
1.0	0.2
1.5	0.4
2.0	2.0
2.5	1.5
3.0	1.8
3.5	2.5
4.0	2.9
4.5	3.7
5.0	4.0

- montrez les valeurs de $y_{expérimentale} = f(x)$ sur un graphique;
  - calculez les paramètres de la régression linéaire

- Préparez le graphique $y_{expérimentale}$, $\hat{y}_i$ calculé, les limites de confiance de $\hat{y}_i$ ($y_{calculé}$) et les limites de confiance de y_i ($y_{expérimentale}$) vs. x.
- préparez le graphique des résidus $y_i - \hat{y}_i$ en fonction de x;
- vérifiez s'il faut rejeter un point (si oui, lequel?);
- recalculez la régression après le rejet (s'il y a lieu)
- rapportez les paramètres de la régression, les écarts types et les limites de confiance de ces paramètres
- Préparez le graphique $y_{expérimentale}$, $\hat{y}_i$ calculé, les limites de confiance de $\hat{y}_i$ ($y_{calculé}$) et les limites de confiance de y_i ($y_{expérimentale}$) vs. x pour les nouvelles données.

EXPÉRIENCE QUATRE

OBJECTIFS

Dans cette expérience on va explorer les méthodes statistiques avancées, la transformé de Fourier et les calculs en utilisant Maple; à lire Chapitres 6, 8, 9, 10, 11.

Exercice 1. Régression polynomiale

Considérons un exemple de la fonction y=f(x) :

x	y
0	0.4
1	0.7
2	1.3
3	1.7
4	2.1
5	2.5
6	2.4
7	2.9
8	3.3
9	3.3
10	3.6
11	3.7
12	3.9
13	4.1
14	4.3
15	4.3

Ces données représentent, par exemple, une courbe de l'absorbance en fonction de la concentration. Avec laquelle fonction est-ce qu'on peut décrire ces données (linéaire, parabolique)? Au début il faut faire un graphique de y = f(x). Est-ce que cette courbe est linéaire?

Pour une régression polynomiale avec les poids statistiques w = 1 on peut utiliser : a) Excel, b) programme POLFIT, c) SigmaPlot (pour les poids statistiques différentes il faut utiliser le programme POLFIT).

I) Excel

Comme premier modèle on peut choisir un modèle linéaire : $y = b_0 + b_1 x$

1a) Déterminez les paramètres de la régression linéaire et déterminez s'ils sont statistiquement importants en utilisant les tests t et F.

1b) Déterminez les paramètres de la régression parabolique : $y = b_0 + b_1 x + b_2 x^2$

- Pour effectuer cela préparez 3 colonnes contenants x, x^2, y (l'ordre est important; allez dans « outils » → « utilitaire d'analyse » → « régression linéaire »
- marquez les données *y* et deux colonnes *x* et x^2 comme donnés *x*, niveau de confiance de 95%, demandez les résidus et les graphiques.

Préparez un graphique contenant les données x, y, et les lignes : droite et parabole.

Vérifiez l'importance des paramètres en utilisant le test t et F.

II) Sigma Plot
Comme premier modèle on peut choisir un modèle linéaire : $y = b_0 + b_1 x$
1a) Déterminez les paramètres de la régression linéaire et déterminez s'ils sont statistiquement importants.
Dans le SigmaPlot entrez les données *x* et *y*. Utilisez :
Statistics
Regression Wizard
Equation Category : Polynomial
Equation Name : Linear
Next
x : Column 1
y : Column 2
Finish
Une nouvelle page « Report » est créé et dans la feuille Data 1 il y 2 paramètres de la régression, y calculée et y expérimentale – y calculée.

Préparez un graphique x, $y_{\text{expérimentale}}$, y_{calc}. On voit, que la courbe est légèrement non linéaire.
On peut aussi tester un modèle parabolique : $y = b_0 + b_1 x + b_2 x^2$.
Utilisez :
Statistics
Regression Wizard
Equation Category : Polynomial
Equation Name : Quadratic
Next
x : Column 1
y : Column 2
Finish

Une nouvelle page « Report » est créé et dans la feuille Data 1 il y 3 paramètres de la régression, y calculée, y expérimentale – y calculée.
Préparez un nouveau graphique.

Pour tester si l'addition d'un nouveau paramètre b_2 est importante utilisez un test F pour un paramètre additionnel :

$$F = \frac{s_1^2}{s_2^2} = \frac{S_1^2 - S_2^2}{\dfrac{S_2^2}{N-3}} = \frac{S_1^2 - S_2^2}{s_y^2}$$

où :

$$s_1^2 = \frac{S_1^2 - S_2^2}{(N-2)-(N-3)} = \frac{S_1^2 - S_2^2}{1} = S_1^2 - S_2^2$$

$$s_2^2 = \frac{S_2^2}{N-3}$$

S_1^2 est la somme de carrées $\sum\left(\hat{y}_i - y_i\right)^2$, calculée pour le modèle : $y = b_0 + b_1 x$, N – 2 dégrées de liberté

S_2^2 est calculée pour le modèle : $y = b_0 + b_1 x + b_2 x^2$, N – 3 dégrées de liberté.

La valeur de S_2^2 s'est trouve dans le Report, Analyse of Variances, Residual SS.
Testez l'importance du b_2.
Vérifiez importance des paramètres de la régression en utilisant le test *t*. Utilisez l'équation correcte.
Préparez un graphique : y expérimentale (points), y calculée pour une ligne droite (ligne) et y calculée pour une parabole (ligne) en fonction de x.
Sigma Plot permet aussi d'utiliser plusieurs autres fonctions d'approximation.

Transformée de Fourier rapide (FFT)

La transformée de Fourier d'une série des données f(0), f(1), f(2),..f(N-1) est définie comme :

$$F(u) = \frac{1}{N}\sum_{i=0}^{N-1} f(i)\exp\left(-\frac{j2\pi ui}{N}\right)$$

où la fréquence f_u est donnée par $f_u = u/(N\Delta t)$ pour les points de u = 0 à u = N/2 et u sont les chiffres : 0, 1, 2, …N-1.
Par exemple, si N = 16 et Δt = 0.05 s, T = N Δt = 16*0.05 s = 0.8 s, la fréquence fondamentale est f_1 = 1/0.8 s^{-1} = 1.25 Hz et les fréquences harmoniques sont f_2 = 2/0.8 Hz = 2.5 Hz, f_3 = 3.75 Hz, et 5, 6.25, 7.5, 8.75, 10 Hz. La fréquence f_8 = 10 Hz est la plus grande pour laquelle on a l'information, elle s'appelle la fréquence de Nyquist.

Le nombre des points expérimentaux doit être égal à 2^n, où n est un nombre entier.
L'Excel permet d'utiliser la FFT dans les cas simples seulement. Cependant, les résultats sont donnés comme un texte et on ne peut pas préparer les graphiques avec ces résultats.

Exemple de calculs des fréquences

Fonction : $f = \cos(2\pi t/0.016 + \pi/3)$ $\qquad \nu = 1/0.016 = 12.5\ \mathrm{s}^{-1}$

$$\nu_u = \frac{u}{N\Delta t} = \frac{u}{T} \qquad u = 0.....\frac{N}{2}$$

i	t/s	f	u	ν_u /Hz	F'	F''
0	0.00	0.50	0	0	0	0
1	0.02	-0.87	1	6.25	0	0
2	0.04	-0.50	2	**12.5**	**0.25**	**0.433**
3	0.06	0.87	3	18.75	0	0
4	0.08	0.50	4	25	0	0
5	0.1	-0.87	5	31.25	0	0
6	0.12	-0.5	6	37.5	0	0
7	0.14	0.87	7	43.75	0.25	-0.43

Exercice 2. FFT Expériences avec Excel

Le programme d'Excel ne divise pas la fonction *F*(*u*) par *N*!
Au début générez une fonction $\cos(2\pi t\, f_0)$, avec $t = i\Delta t$ et $f_0 = 0.25$ Hz ($T = 4$ s), $\cos(2\pi t\, f_0) = \cos(2\pi i*0.25)$:

t / s	i	Cos(πi/2)
0	0	1
1	1	0
2	2	-1
3	3	0

Pour effectuer la transformée de Fourier exécutez :
Outils
Utilitaire d'analyse
Transformée de Fourier
Entrez les valeurs et calculez la FFT. **Le programme FFT accepte une seule colonne des données; la colonne de temps n'est pas utilisée! Les données doivent utiliser la séparation décimale le point « . ».**
A partir de ces résultats calculez la transformée inverse et comparez avec les données de départ.
Répétez la même opération pour la fonction $\sin(2\pi t\, f_0)$.

Exercice 3. Utilisation des programmes FFTF et FFTB

En utilisant l'Excel calculez la FFT de la fonction suivante:

t / s	F
0	2
0.5	3
1.0	4
1.5	4

Pour la transformée de plus grand nombre des points, on doit utiliser les programmes FFTF.EXE pour une transformée directe et FFTB.EXE pour une transformée inverse. Les programmes acceptent jusqu'à 32768 points.
Pour exécuter ces programmes il faut préparer un fichier DOS contenant seulement une colonne des chiffres à transformer.
Pour sauver les données il faut les copier dans une autre feuille comme valeurs (Coller spécial, Valeurs) et sauver sous le format *.txt (Enregistrer sous, Texte (séparateur : tabulation).
Exécutez la FFT directe : FFTF.EXE, entrez le nom de fichier *.txt à transformer et le nom de fichier des résultats. Pour entrer les résultats dans l'Excel il faut utiliser : Données, Données externes, Importer le fichier texte.
Exécutez la FFT inverse : FFTB.EXE, en utilisant le fichier des résultats et comparez avec les résultats de l'Excel (l'Excel ne divise pas des résultats par *N*).

Exercice 4.

Générez des nombres de 0 à 63 (notre *t* temps en secondes). Calculez dans les trois colonnes les trois fonctions suivantes :
$\cos(2\pi t/64)$; $\cos(2\pi*4t/64)$; $\cos(2\pi*8t/64)$.
Calculez la somme de ces 3 fonctions.
Générez les graphiques de ces 3 fonctions et de la somme en fonction du temps.
Calculez la FFT de cette fonction (FFTF.exe) et les fréquences correspondantes.
Est-ce qu'on peut déterminer les fréquences des fonctions originales? Quelle sont ces fréquences présentes dans les données (dans les équations)?

Exercice 5.

Générez des nombres de 0 à 63 (notre *t*). Calculez la fonction suivante : $\cos(8\pi t/56)$.
Préparez le graphique de cette fonction.
Calculez la FFT de cette fonction (FFTF). Est-ce qu'on peut déterminer la fréquence de la fonction originale? Expliquez.

Exercice 6.

Générez 512 points de 0 à 51.1 s, Δt = 0.1 s. Calculez une fonction sin(2*PI()*A1/2), qui correspond à la fréquence 0.5 Hz ou T = 2 s.
Préparez trois fichiers des données contenants 32, 128 et 521 points. Exécutez la FFT (FFTF.EXE) et mettez les résultats dans un fichier d'Excel. Pour chaque série des données, dans une feuille de calculs différente, préparez deux graphiques : 1) la fonction sin calculée en fonction du temps, 2) de la transformée de Fourier réelle, imaginaire et 3) l'amplitude, $\sqrt{\mathrm{Re}[F(u)]^2 + \mathrm{Im}[F(u)]^2}$, en fonction de la fréquence. Trouvez entre lesquelles valeurs de fréquence se trouve le maximum de l'amplitude.
Ex. Pour 32 points, les plus grandes valeurs de la transformée se trouvent entre les points No 2 et 3. Dans ce cas, la fréquence d'échantillonnage est f_0 =1/3.2 = 0.3125 s^{-1}. Cela signifie, que notre fréquence se trouve entre (2-1)*f_0 = 0.3125 s^{-1} et (3-1)*f_0 = 0.625 s^{-1} (le premier point c'est la fréquence = 0).

Calculez la fréquence à partir d'autres graphiques/tableaux.
Expliquez les résultats.

Exercice 7. Méthode de Savitzky-Golay

Cet exemple illustre l'utilisation de la méthode de lissage d'une courbe avec le bruit aléatoire.

Générez une fonction $y = \exp\left[-(x/3)^2\right]$ de –5 à 5, intervalle 0.1 (attention : utilisez les doubles parenthèses, $\exp\left(-\left((x/3)^2\right)\right)$, ou une fonction Puissance).

Dans une autre colonne générez un bruit gaussien y_{bruit} avec μ = 0, *s* = 1.

Ajoutez le bruit à la fonction générée : $y_{nouveau} = y_{calc} + 0.02\, y_{bruit}$.

Utilisez la méthode de Savitzky-Golay pour 9 et 15 points (valeurs des coefficients sont dans le Tableau dans les notes de cours).

Pour comparer les courbes : calculée, avant et après lissage (séparément pour deux filtres à 9 et à 15 points) et préparez les graphiques suivant :

a) $y_{nouveau}$ et y_{lisse} (Savitzky-Golay 9 points) en fonction de *x*

b) $y_{nouveau}$ et y_{lisse} (Savitzky-Golay 15 points) en fonction de *x*

Attention : $y_{nouveau}$ - points, y_{lisse} - lignes.

c) $y_{nouveau}$ - y_{lisse} (Savitzky-Golay 9 points) et $y_{nouveau}$ - y_{lisse} (Savitzky-Golay 15 points) en fonction de *x* (lignes)

EXPÉRIENCE CINQ

MAPLE – applications en chimie

OBJECTIFS

Dans cette expérience on va apprendre les calculs en utilisant Maple,

Exercice 1.

1.a. Déterminez l'écart type de $y = \frac{a-b}{a+b}$ pour a = 4, b = 2, s_a = 0.1, s_b = 0.2.

```
> restart;
> y:=(a-b)/(a+b);
```

$$y := \frac{a-b}{a+b}$$

derivé partielle dy/da

```
> diff(y,a);
```

$$\frac{1}{a+b} - \frac{a-b}{(a+b)^2}$$

```
> dyda:=simplify(%);
```

$$dyda := 2\frac{b}{(a+b)^2}$$

```
> dydb:=simplify(diff(y,b));
```

$$dydb := -2\frac{a}{(a+b)^2}$$

```
> err:=sqrt((sa*dyda)^2+(sb*dydb)^2);
```

$$err := 2\sqrt{\frac{sa^2\,b^2}{(a+b)^4} + \frac{sb^2\,a^2}{(a+b)^4}}$$

```
> simplify(%);
```

$$2\sqrt{\frac{sa^2\,b^2 + sb^2\,a^2}{(a+b)^4}}$$

```
> sa:=0.1;
```

$$sa := .1$$

```
> sb:=0.2;
```

$$sb := .2$$

```
> a:=4;
```

$$a := 4$$

```
> b:=2;
```

$$b := 2$$

```
> eval(err);
```

$$.04581228472$$

```
>
```

1.b. Déterminez l'écart type de $y = \dfrac{a+b}{c-d}$ pour a = 3, b = 2, c = 4, d = 1, s_a = 0.1, s_b = 0.2, s_c = 0.05, s_d = 0.15.

```
> restart;
> y:=(a+b)/(c-d);
```

$$y := \frac{a+b}{c-d}$$

```
> dyda:=diff(y,a);
```

$$dyda := \frac{1}{c-d}$$

```
> dydb:=diff(y,b);
```

$$dydb := \frac{1}{c-d}$$

```
> dydc:=diff(y,c);
```

$$dydc := -\frac{a+b}{(c-d)^2}$$

```
> dydd:=diff(y,d);
```

$$dydd := \frac{a+b}{(c-d)^2}$$

```
> std:=sqrt((sa*dyda)^2+(sb*dydb)^2+(sc*dydc)^2+(sd*dydd)^2);
```

$$std := \sqrt{\frac{sa^2}{(c-d)^2} + \frac{sb^2}{(c-d)^2} + \frac{sc^2\,(a+b)^2}{(c-d)^4} + \frac{sd^2\,(a+b)^2}{(c-d)^4}}$$

```
> sa:=0.1;
```

$$sa := .1$$

```
> sb:=0.2;
```

$$sb := .2$$

```
> sc:=0.05;
```
$$sc := .05$$
```
> sd:=0.15;
```
$$sd := .15$$
```
> a:=3;
```
$$a := 3$$
```
> b:=2;
```
$$b := 2$$
```
> c:=4;
```
$$c := 4$$
```
> d:=1;
```
$$d := 1$$
```
> eval(std);
```
$$.1152024520$$
```
>
```

1.c. Déterminez l'écart type de $y = ab^c$; a = 2.1, b = 3.1, c = 1.5, s_a = 0.1, s_b = 0.15, s_c = 0.2.

```
> restart;
> y:=a*b^c;
```
$$y := a\, b^c$$
```
> dyda:=diff(y,a);
```
$$dyda := b^c$$
```
> dydb:=diff(y,b);
```
$$dydb := \frac{a\, b^c\, c}{b}$$
```
> dydb:=simplify(%);
```
$$dydb := a\, b^{(c-1)}\, c$$
```
> dydc:=diff(y,c);
>
```
$$dydc := a\, b^c \ln(b)$$
```
> std:=sqrt((sa*dyda)^2+(sb*dydb)^2+(sc*dydc)^2);
```
$$std := \sqrt{sa^2 \left(b^c\right)^2 + sb^2\, a^2 \left(b^{(c-1)}\right)^2 c^2 + sc^2\, a^2 \left(b^c\right)^2 \ln(b)^2}$$
```
> a:=2.1;
```
$$a := 2.1$$
```
> b:=3.1;
```

$$b := 3.1$$

```
> c:=1.5;
```

$$c := 1.5$$

```
> sa:=0.1;
```

$$sa := .1$$

```
> sb:=0.15;
```

$$sb := .15$$

```
> sc:=0.2;
```

$$sc := .2$$

```
> std;
>
```

$$2.777939088$$

```
>
```

1.d. Déterminez l'écart type de $y = (a + b^2)\log(a + b^2)$; a=2.7, s_a = 0.2, b=1.9, s_b=0.25

```
> restart;
> y:=(a+b^2)*log10(a+b^2);
```

$$y := \frac{(a + b^2)\ln(a + b^2)}{\ln(10)}$$

```
> dyda:=diff(y,a);
```

$$dyda := \frac{\ln(a + b^2)}{\ln(10)} + \frac{1}{\ln(10)}$$

```
> dydb:=diff(y,b);
```

$$dydb := 2\,\frac{b\ln(a + b^2)}{\ln(10)} + \frac{2\,b}{\ln(10)}$$

```
> err:=sqrt((dyda*sa)^2+(dydb*sb)^2);
```

$$err := \sqrt{\left(\frac{\ln(a + b^2)}{\ln(10)} + \frac{1}{\ln(10)}\right)^2 sa^2 + \left(2\,\frac{b\ln(a + b^2)}{\ln(10)} + \frac{2\,b}{\ln(10)}\right)^2 sb^2}$$

```
> simplify(%);
```

$$\frac{\sqrt{(\ln(a + b^2) + 1)^2\,(sa^2 + 4\,b^2\,sb^2)}}{\ln(2) + \ln(5)}$$

```
> a:=2.7;
```

$$a := 2.7$$

```
> b:=1.9:
> sa:=0.2;
```

$$sa := .2$$

```
> sb:=0.25;
```

$$sb := .25$$

\>
\> **evalf(err);**

$$1.198311693$$

\> **evalf(y);**

$$5.048185257$$

\>

1.e. Déterminez l'écart-type de *y*, $y = 3^{2D}\dfrac{C^2+B}{C+\sqrt{B}}$ **; B=5, C=4.2, D=1.7, s_B = 0.3, s_C=0.4, s_D=0.1.**

2. Intégration et évaluation des fonctions :

2.a. Exemple de l'intégration

Intégrez une fonction $y = \mathrm{e}^{-x^2}$ de 0 à 1.

```
> restart;
> y:=exp(-x^2);
```

$$y := \mathbf{e}^{(-x^2)}$$

"int" = integral de y, de x=0 a x=1.

```
> iy:=int(y,x=0..1);
```

$$iy := \frac{1}{2}\,\mathrm{erf}(1)\sqrt{\pi}$$

```
> evalf(iy);
```

$$.7468241330$$

2.b. Intégrez la fonction P_G de 0 à 1, de –1 à 1, etc.

```
> Pg:=1/sqrt(2*Pi)*exp(-z^2/2);
```

$$Pg := \frac{1}{2}\,\frac{\sqrt{2}\,\mathbf{e}^{(-1/2\,z^2)}}{\sqrt{\pi}}$$

Integration de y de 0 à 1

```
> Ag:=int(Pg,z=-x..x);
```

$$Ag := \mathrm{erf}\left(\frac{1}{2}\sqrt{2}\,x\right)$$

```
> Ag1:=int(Pg,z=-1..1);
```

$$Ag1 := \operatorname{erf}\left(\frac{1}{2}\sqrt{2}\right)$$

```
> evalf(Ag1);
```

$$.6826894920$$

```
>
> iy:=int(Pg,z=-2..2);
```

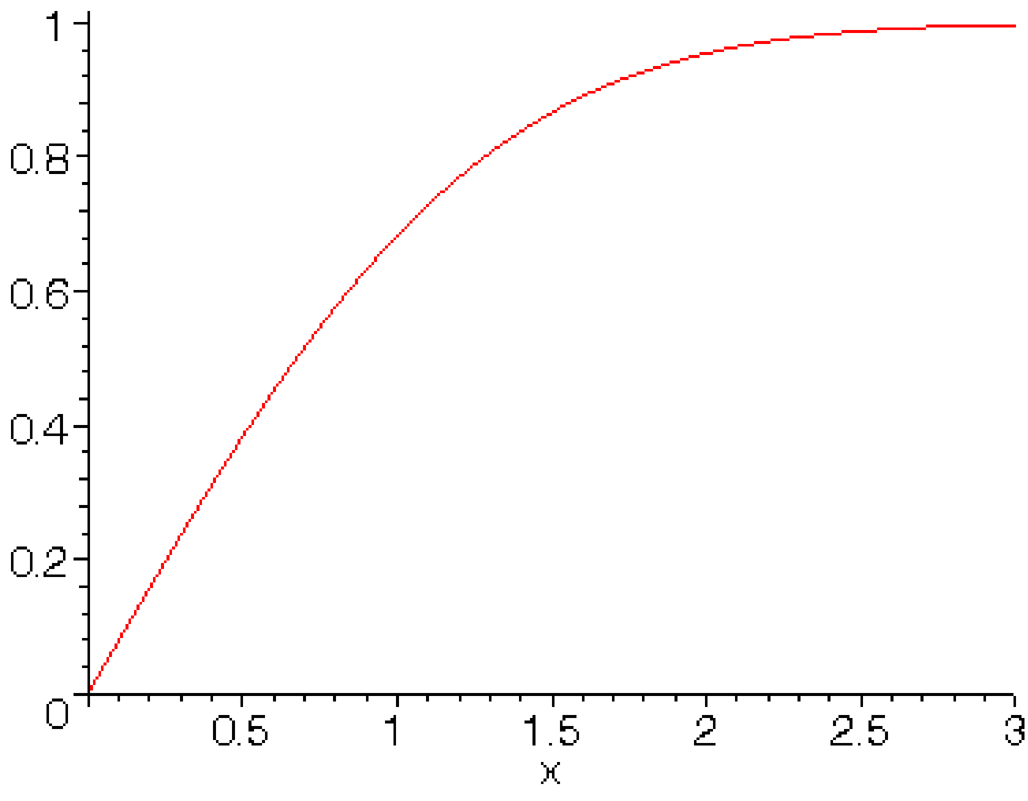
$$iy := \operatorname{erf}(\sqrt{2})$$

```
> evalf(iy);
```

$$.9544997360$$

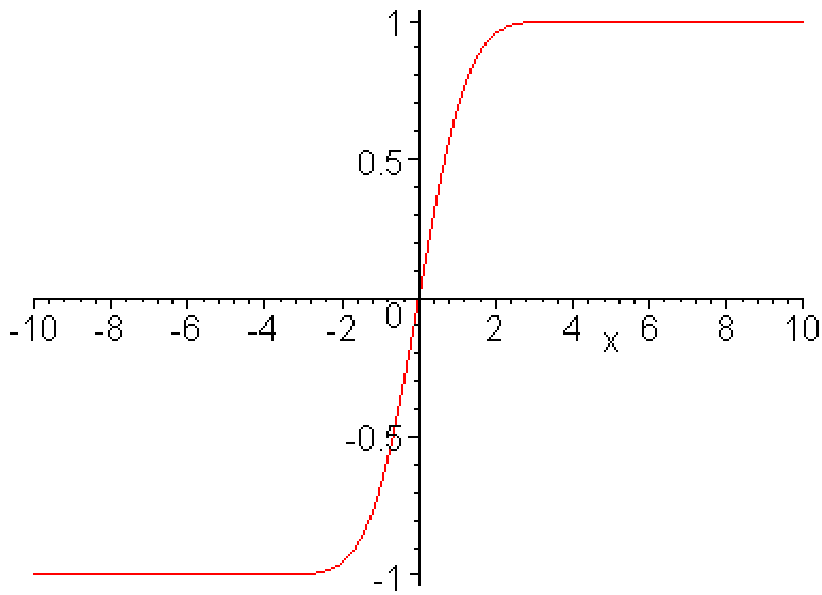
2.c. Graphique de la fonction f(x) = A_G

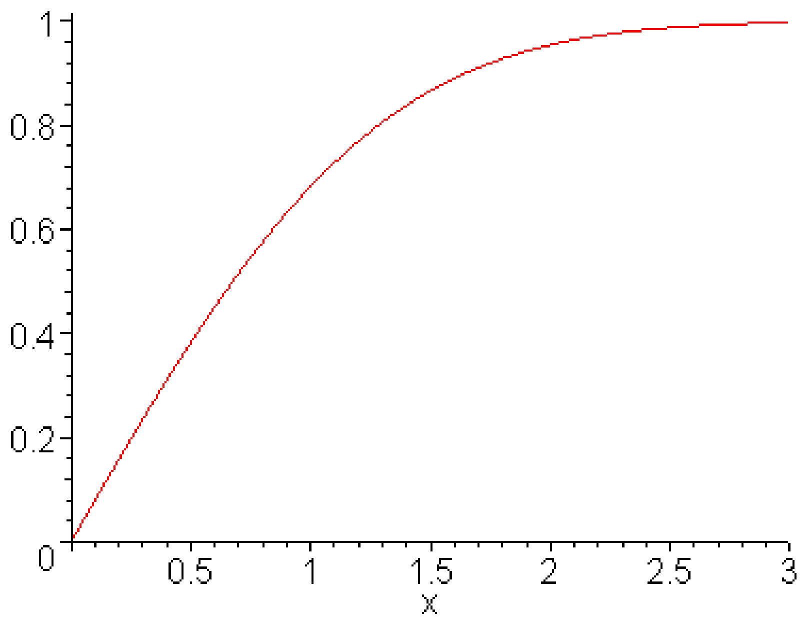
```
> plot(Ag,x=0..3);
```



Une autre méthode d'avoir un graphique est de cliquer à droit sur la formule et choisir une option graphique. Vous pouvez formater ce graphique en choisissant des options des axes. Cela va produire des résultats suivants :

```
> smartplot(Ag);
```





2.d. Calculez les valeurs de la fonction erf(x).
La fonction erf(x) est équivalente à l'intégrale de la fonction de Gauss normalisée :

$$\operatorname{erf}(x) = \frac{2}{\pi}\int_0^x e^{-t^2}\,dt$$

```
> evalf(erf(0.5));
```

$$.5204998778$$

```
> evalf(erf(1.));
```

$$.8427007929$$

```
> evalf(erf(2.));
```

$$.9953222650$$

```
> > f:=x->Ag;
```

$$f := x \to Ag$$

```
> Ag;
```

$$\mathrm{erf}\left(\frac{1}{2}\sqrt{2}\,x\right)$$

```
> f(x);
```

$$\mathrm{erf}\left(\frac{1}{2}\sqrt{2}\,x\right)$$

```
> eval(f(x),x=2);
```

$$\mathrm{erf}(\sqrt{2})$$

```
> evalf(%);
```

$$.9544997360$$

2.e. Intégrez la fonction : $y = \dfrac{x}{(1+x)^2}$ **et évaluez l'intégrale de x=0 à x = 1.**

```
> restart;
> y:=x/(1+x)^2;
```

$$y := \frac{x}{(1+x)^2}$$

```
> iy:=int(y,x);
```

$$iy := \frac{1}{1+x} + \ln(1+x)$$

```
> aa:=int(y,x=0..1);
```

$$aa := -\frac{1}{2} + \ln(2)$$

```
> evalf(aa);
```

$$.1931471806$$

2.f. Intégrez la fonction de Gauss normalisée de -2 à -1.

3. Solution des équations

3.a. Solution d'une équation non linéaire

Calculez le pH du phénol 5 10^{-5} M, K_a = 10^{-10} (acide très faible). Comparez avec une solution approximative (en négligeant la dissociation de l'eau). Quelle est le taux de dissociation du phénol?

Pour une réaction de dissociation dans le milieu aqueux :

$$HA + H_2O = H_3O^+ + A^- \quad (1)$$

on peut écrire les équilibres suivants :

$$K_a = \frac{[H_3O^+][A^-]}{[HA]} \quad (2)$$

$$[H^+][OH^-] = k_w \quad (3)$$

et les balances : de charges

$$[H^+] = [OH^-] + [A^-] \quad (4)$$

et de masses

$$[HA] + [A^-] = C_a \quad (5)$$

```
> restart;
> Ka:=1e-10;
```

$$Ka := .1\ 10^{-9}$$

```
> kw:=1e-14;
```

$$kw := .1\ 10^{-13}$$

```
> C:=5e-5;
```

$$C := .00005$$

```
> eq1:=Ka=CH*CA/CHA;
```

$$eq1 := .1\ 10^{-9} = \frac{CH\ CA}{CHA}$$

```
> eq2:=CH*COH=kw;
```

$$eq2 := CH\ COH = .1\ 10^{-13}$$

```
> eq3:=CH=COH+CA;
```

$$eq3 := CH = COH + CA$$

```
> eq4:=C=CHA+CA;
```

$$eq4 := .00005 = CHA + CA$$

```
>
sol:=solve({eq1,eq2,eq3,eq4,CH>0,CA>0,COH>0,CHA>0},{CH,CA,COH,CHA}
);
```

$$sol := \{ CA = .4079706676\ 10^{-7}, COH = .8166076391\ 10^{-7}, CHA = .00004995920293, CH = .1224578307\ 10^{-6} \}$$

solution alternative sans restrictions donne 3 séries des racines:

```
> sol1:=solve({eq1,eq2,eq3,eq4},{CH,CA,COH,CHA});
```

$$sol1 := \{ COH = -.8163854169\ 10^{-7}, CHA = .00005004085262, CA = -.4085262233\ 10^{-7}, CH = -.1224911640\ 10^{-6} \}, \{ CHA = .00004995920293, CA = .4079706676\ 10^{-7}, COH = .8166076391\ 10^{-7}, CH = .1224578307\ 10^{-6} \}, \{ COH = -.0001500000222, CHA = -.00009999995556, CA = .0001499999556, CH = -.6666665679\ 10^{-10} \}$$

```
> CH:=.1224578307e-6;
```

$$CH := .1224578307\ 10^{-6}$$

```
> pH:=-log10(CH);
```

$$pH := 6.912013438$$

On peut aussi réarranger les équations; l'éqn (5) nous donne:

$$[\mathrm{HA}] = C_a - [\mathrm{A}^-] \qquad (6)$$

En combinant avec les éqns. (3) et (4) nous obtenons :

$$[\mathrm{H}^+] = \frac{k_w}{[\mathrm{H}^+]} + [\mathrm{A}^-] \qquad (7)$$

Les équations (6) et (7) forment us système de deux éqns. avec 2 inconnues: $[\mathrm{H}^+]$ et $[\mathrm{A}^-]$

$$\begin{cases} [\mathrm{H}^+]^2 - [\mathrm{H}^+][\mathrm{A}^-] - k_w = 0 \\ K_a = \dfrac{[\mathrm{H}^+][\mathrm{A}^-]}{C_a - [\mathrm{A}^-]} \end{cases} \qquad (8)$$

qui est équivalent à une équation de 3ème ordre.
Voici la solution dans le Maple:

```
> restart;
> Ka:=1e-10;
```

$$Ka := .1\ 10^{-9}$$

```
> kw:=1e-14;
```

$$kw := .1\ 10^{-13}$$

```
> C:=5e-5;
```

$$C := .00005$$

```
> sol:=solve({CH^2-CH*CA-kw=0,Ka=(CH*CA)/(C-CA)},{CH,CA});
```

$$sol := \{CH = -.1224911640\ 10^{-6}, CA = -.4085262233\ 10^{-7}\},$$
$$\{CH = .1224578307\ 10^{-6}, CA = .4079706676\ 10^{-7}\},$$
$$\{CH = -.6666665679\ 10^{-10}, CA = .0001499999556\}$$

Il y a 3 solutions, mais seulement les concentrations positives ont un sens physique.

```
>sol:=solve({CH^2-CH*CA-kw=0,Ka=(CH*CA)/(C-
CA),CH>0,CA>0},{CH,CA});
```

$$sol := \{CH = .1224578307\ 10^{-6}, CA = .4079706676\ 10^{-7}\}$$

```
> CH:=.1224578307e-6;
```

$$CH := .1224578307\ 10^{-6}$$

```
> pH:=-log10(CH);
```

$$pH := 6.912013438$$

```
> evalf(pH);
```

$$6.912013438$$

pH = 6.91

Par contre, la solution approximative, en négligeant la dissociation de l'eau est erronée :

$$K_a = \frac{[H^+][A^-]}{[HA]} = \frac{[H^+]^2}{C_a - [H^+]}$$

```
> restart;
> Ka:=1e-10;
```

$$Ka := .1\ 10^{-9}$$

```
> kw:=1e-14;
```

$$kw := .1\ 10^{-13}$$

```
> Ca:=5e-5;
```

$$Ca := .00005$$

```
> sol:=solve({CH^2+CH*Ka-Ka*Ca=0},{CH});
```

$$sol := \{CH = .706606958010^{-7}\}, \{CH = -.707606958010^{-7}\}$$

```
sol:=solve({CH^2+CH*Ka-Ka*Ca=0,CH>0},{CH});
```

$$sol := \{CH = .706606958010^{-7}\}$$

```
> CH:=.7066069580e-7;
```

$$CH := .706606958010^{-7}$$

```
> pH:=-log10(CH);
```

$$pH := 7.150822090$$

pH=7.15!

3.b. Exercice

Calculez pH de deux solutions : 2 10^{-5} et 9 10^{-5} M du phénol dans l'eau.
Comparez avec les solutions approximatives.

3.c. Exercice

Calculez le pH de la solution 2 10^{-5} M CH_3COOK. pK_a de l'acide acétique égale 4.76.
Comparez avec la solution approximative.
Écrivez les équations chimiques qui décrivent le système d'équilibres chimiques (deux équations d'équilibre, une de balance des masses et une de balance des charges).
(Rép. pH = 7.17).

4. Déterminez les coefficients de la méthode de Savitzky-Golay pour 7 points

Voir les notes de cours.

5. Déterminez les coefficients de la méthode de Savitzky-Golay pour 9 points.

Pour installer l'Utilitaire d'analyse
Ouvrir:
Outils
Macro complémentaire
Utilitaire d'analyse

Fonction statistiques d'Excel

$P_G(x,\mu,\sigma)$ = LOI.NORMALE(x, μ, σ,FAUX) $= \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2 \right]$

$P_G(z)$ =LOI.NORMALE(z,0,1,FAUX) = $\dfrac{1}{\sqrt{2\pi}} e^{-z^2/2}$

English : NORMDIST(x,μ,σ,FALSE)

LOI.NORMALE(z,0,1,VRAI) = $\int_{-\infty}^{z} P_G(u,0,1)du$

LOI.NORMALE(x,μ,σ,VRAI) = $\int_{-\infty}^{x} P_G(u,\mu,\sigma)du$

English : NORMDIST(x,μ,σ,TRUE)

$A_G(z)$ = LOI.NORMALE(z,0,1,VRAI) -LOI.NORMALE(-z,0,1,VRAI)

$= \int_{-z}^{z} P_G\left(z,0,1\right)dz = \dfrac{1}{\sqrt{2\pi}} \int_{-z}^{z} \mathrm{e}^{-\frac{z^2}{2}} \mathrm{d}z$

$A_G(x)$ = LOI.NORMALE(x,μ,σ,VRAI) -LOI.NORMALE(-x,μ,σ,VRAI)

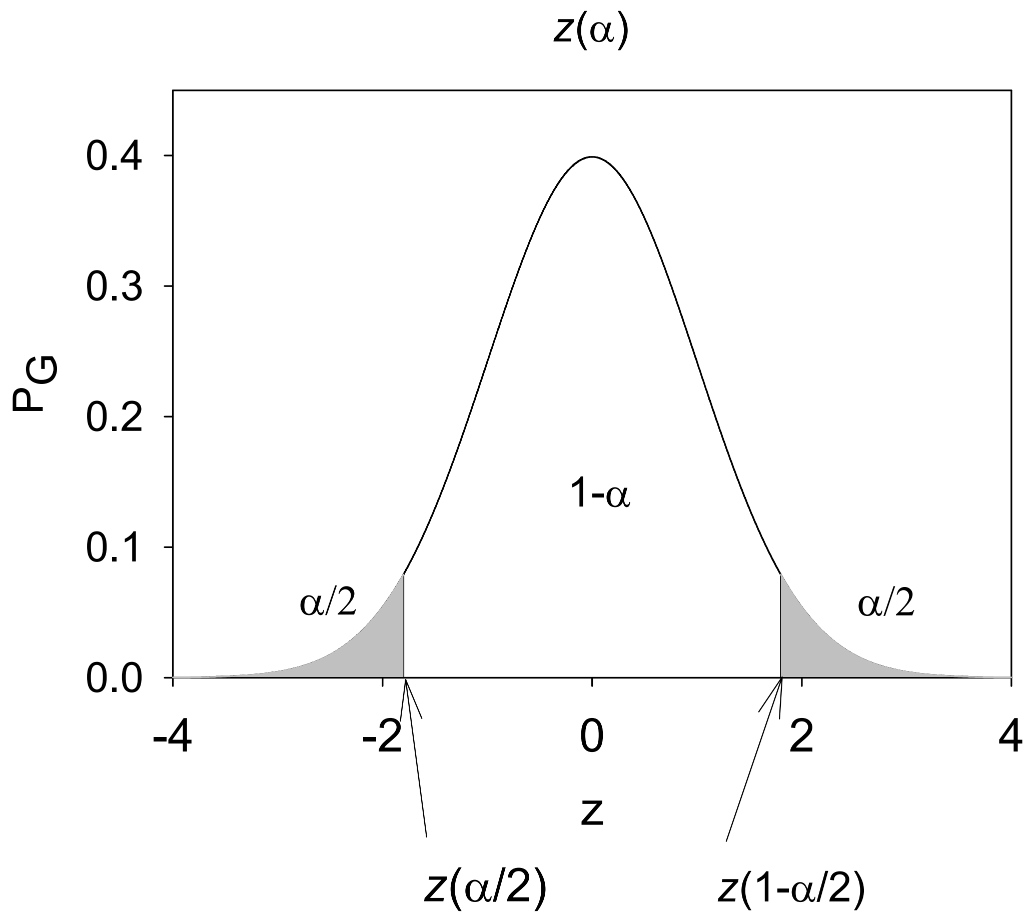
LOI.NORMALE.STANDARD.INVERSE(α) = z(α) $\int_{-\infty}^{z} P_G(z,0,1)dz = \alpha$

English : NORMSINV(α)

LOI.NORMALE.INVERSE(α,μ,σ) = *x*(α) $\int_{-\infty}^{x} P_G(x,\mu,\sigma)dz = \alpha$

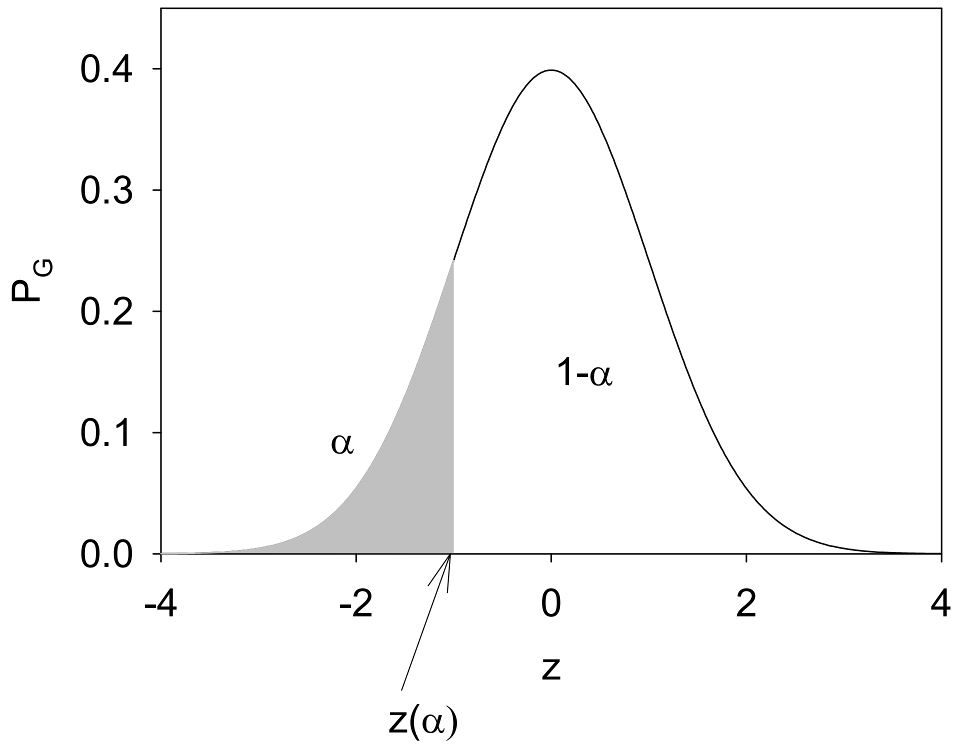
English : NORMINV(α,μ,σ)

z(α) = LOI.NORMALE.STANDARD.INVERSE(1-α/2) voir la figure

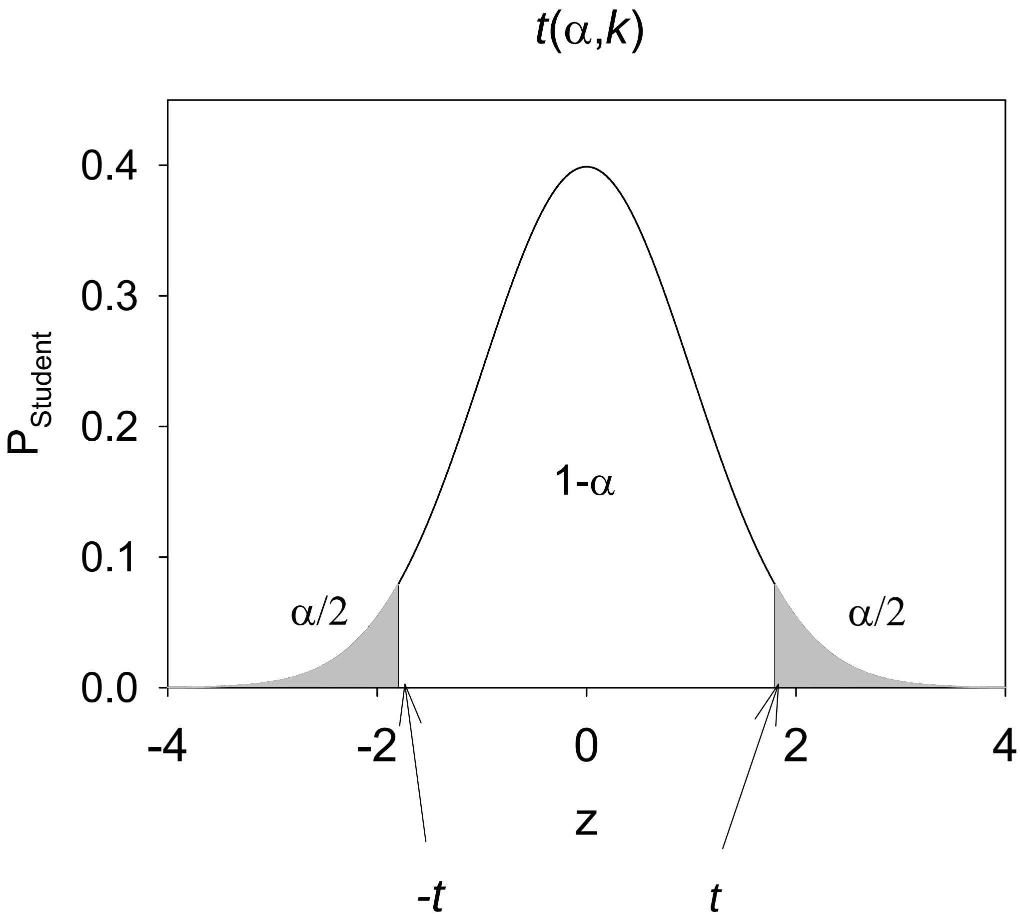


LOI.NORMALE.STANDARD.INVERSE(α) = z(α) $\int_{-\infty}^{z} P_G(z,0,1)dz = \alpha$

English : NORMSINV(α)



$t(\alpha,k)$ = LOI.STUDENT.INVERSE(α, k)   English : TINV(α, k)



Français	English
$F(\alpha, D_1, D_2)$ = INVERSE.LOI.F(α, D_1, D_2)	FINV(α, D_1, D_2)
χ^2 = KHIDEUX.INVERSE(α, k)	CHIINV(α, k)
ECARTYPE; p. ex. ECARTYPE (A1:A10)	STDEV
SOMME	SUM
MOYENNE	AVERAGE

Fonctions complexes

Insérer fonction f_x, choisir Scientifiques
(les cellules A1, A2… sont montrées à titre d'exemple

COMPLEXE.REEL(A1) extraire la partie réelle d'un nombre complexe (texte)

COMPLEXE.IMAGINAIRE(A1) extraire la partie imaginaire d'un nombre complexe (texte)

Exemple :
A1= 3-4i
COMPLEXE.REEL(A1) = 3
COMPLEXE.IMAGINAIRE(A1) = -4

COMPLEXE(A1,A2) changer en nombre complexe deux nombres réels, A1 partie réelle, A2 partie imaginaire

Exemple :
B1 = 3 B2 = -4
COMPLEXE(B1,B2) = 3 -4i

COMPLEXE.MODULE(A1) calculer un module d'un nombre complexe (texte)
A1= 3-4i
COMPLEXE.MODULE(A1) = 5

COMPLEXE.ARGUMENT(A1) calculer un angle en radians d'un nombre complexe (texte)
A1= 3-4i
COMPLEXE.ARGUMENT(A1) = -0.9273 rad = -53.13^o