Diplôme Universitaire de pharmacocinétique de Toulouse

*** Année 2007 ***

Module d'analyse des données

P.L. Toutain, A. Bousquet-Mélou

UMR 181 de Physiopathologie et Toxicologie Expérimentales INRA/ENVT Ecole Nationale Vétérinaire de Toulouse

Avant propos

Ce document correspond au cours donné dans le module d'analyse des données en 2007; pour être profitable, il exige l'acquisition d'un ensemble de connaissances (mathématiques, statistiques, physiologiques) présentées dans les modules précédents.

Pour des raisons pédagogiques, la réalisation des exercices sera faite avec l'aide du logiciel Excel dont une connaissance de base est nécessaire.

On se contentera ici de rappeler comment réaliser une régression linéaire ; de même, il est demandé d'utiliser le point et non la virgule comme séparateur de décimales (à configurer dans les paramètres régionaux de votre ordinateur).

La régression linéaire dans Excel

Excel peur réaliser des régressions linéaires selon différentes approches.

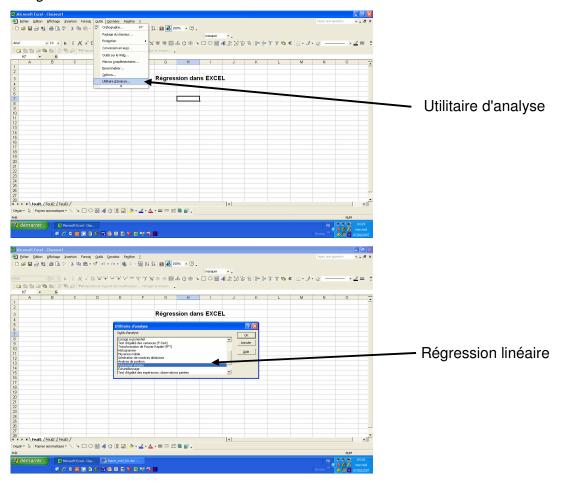
- (1) en faisant appel à une macro "regression"
- (2) en faisant appel à la fonction "DROITEREG"
- (3) en utilisation les "courbes de tendance"

Pour des raisons pédagogiques, la première approche sera préférée.

1. L'utilisation de la macro "régression"

La procédure pour faire appel à la macro "regression" est la suivante : Dans Outil>utilitaire d'analyse>régression (fig.1).

Figure 1: procédure pour accéder à « l'utilitaire d'analyse » & à la macro « régression »



Au préalable, il conviendra de vérifier que les macros complémentaires "utilitaire d'analyse" et "utilitaire d'analyse VBA" sont accessibles. Suivant la façon dont Excel a été installé, on a accès à ces utilitaires d'analyse ou pas. Pour le vérifier faire (fig. 2):

Outil>macro complémentaires et vérifier que les cases "utilitaire d'analyse" et "utilitaire d'analyse – VBA" ont bien été cochées. Cela donne accès à un ensemble de macros statistiques et mathématiques.

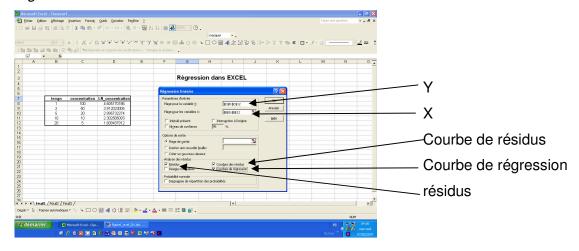
| Companies | Comp

Figure 2: vérification que les utilitaires d'analyse sont installés et accessibles

Si les utilitaires d'analyse n'apparaissent pas dans la liste des macros complémentaires, il faudra relancer le programme d'installation d'Excel (ou celui de Microsoft Office) en choisissant l'option personnalisée qui permettra de copier sur votre disque dur les fichiers nécessaires à l'utilitaire d'analyse.

La macro "regression" (fig.3) donne l'ensemble des informations nécessaires à la réalisation des exercices du cours y compris une présentation des résidus.

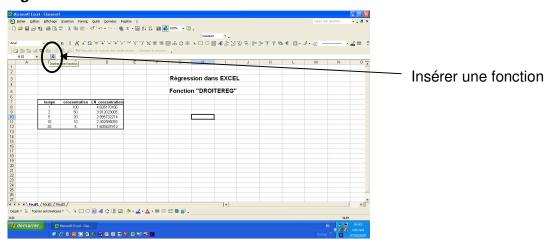
Figure 3 : Boîte de dialogue de la macro « régression » d'EXCEL ; il conviendra de sélectionner les vecteurs à analyser (exemple : temps pour les X et les LOG népériens des concentrations) ; on cochera les cases pour obtenir la courbe de régression et les résidus.



2. Utilisation de la fonction "DROITEREG"

La fonction "DROITEREG" (ou LINEST pour la version anglaise) est accessible en cliquant sur la fonction *fx* située sur le côté de la barre de calculs (fig.4).

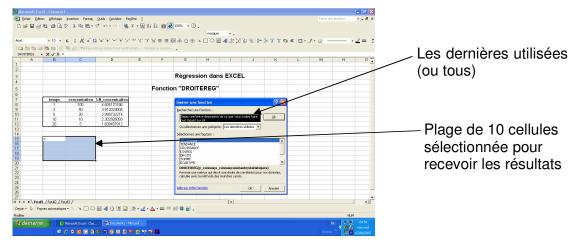
Figure 4: accès à la fonction « DROITEREG »



DROITEREG est une formule matricielle qui s'applique à une plage de cellules. Cette fonction s'utilise selon la procédure suivante (fig.5) :

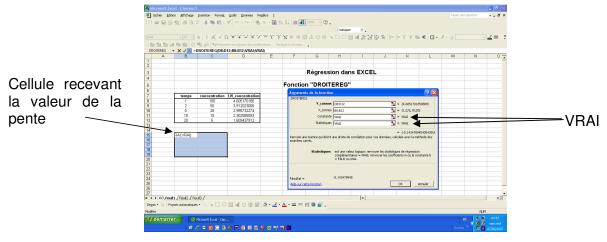
(1) définir à l'aide de la souris une plage de 2 colonnes par 5 lignes c'est-à-dire 10 cellules destinées à recevoir les résultats

Figure 5 : procédure pour utiliser « DROITEREG » avec la sélection des 10 cellules destinées à recevoir les résultats



- (2) accéder à la fonction DROITEREG en cliquant sur fx (insérer une fonction) située à gauche de la barre de calculs. Eventuellement, il faudra changer la catégorie offerte par défaut (fig.5) en sélectionnant "tous". Si vous avez déjà utilisé la fonction DROITEREG, elle s'affichera dans la catégorie "les dernières utilisées".
- (3) Après avoir cliqué sur "DROITEREG" il faudra documenter la boîte de dialogue avec la plage d'entrée des "Y_connus" (concentrations observées ou leurs logarithmes) et la plage des "X_connus" (temps). Dans les cellules constante et statistique, il faudra écrire "vrai" pour qu'Excel calcule l'ordonnée à l'origine et différentes statistiques (fig.6). Puis valider par OK.

Figure 6 : Boîte de dialogue de la fonction « DROITEREG »



(4) Cela donne une première valeur dans la cellule à gauche du pavé des 10 cellules précédemment sélectionnées (il s'agit de la pente). Pour avoir les autres résultats, il faut taper F2 (qui permet d'accéder à la saisie matricielle) puis faire CTRL+Shift+Entrée; cela affiche les résultats dans les 10 cellules (fig.7).

Figure 7.1 : les résultats des calculs sont donnés dans les 10 cellules précédemment sélectionnées

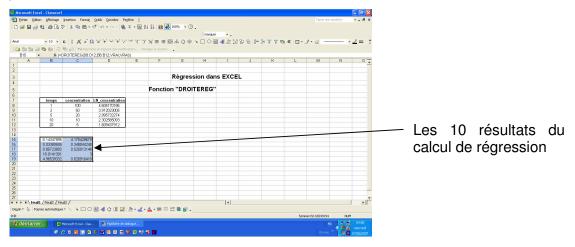


Figure 7.2 : Pavé de cellules des résultats donnés par la fonction DROITEREG

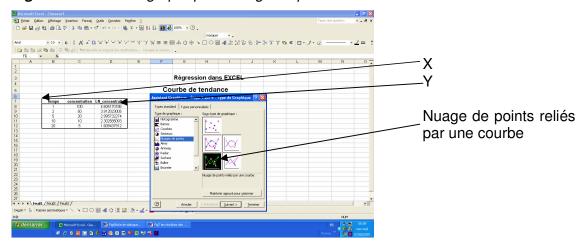
Pente	Ordonnée à l'origine
Se (pente)	Se (ordonnée à l'origine)
R ²	se _y
F	df (degré de liberté)
SCM (ss rég)	SCR (ss resid)

Avec Se, les erreurs types des coefficients, R² le coefficient de détermination (c'est-à-dire le carré du coefficient de corrélation) SCM, la somme des carrés du modèle et SCR la somme des écarts des résidus, et se_y, l'erreur type pour la valeur y estimée (cliquer sur Aide pour plus d'informations).

3. Utilisation de "Tendance"

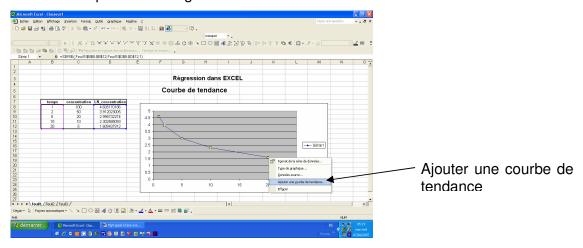
Une méthode rapide pour faire une régression consiste à passer par l'assistant graphique et à sélectionner le type de graphique "nuage de points" (fig.8).

Figure 8 : accès au graphique « nuage de points »



Sélectionnez vos données (X et Y), faites votre graphique, puis avec le click droit de la souris sur le nuage de points sélectionnez "ajouter une courbe de tendance" (fig.9) dans le menu contextuel.

Figure 9 : appel à faire une « courbe de tendance » avec le click droit de la souris sur l'un des points de la figure



Dans l'onglet "type" sélectionner "type linéaire" (fig 10) puis dans "options" cocher les cases "afficher l'équation sur le graphique" et "afficher le coefficient de détermination" (fig.11) et valider par OK. On peut ensuite éditer le pavé de résultats.

Figure 10 : sélection du type de régression « linéaire »

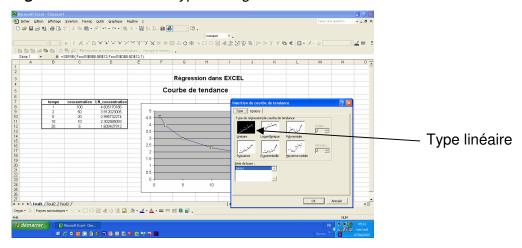
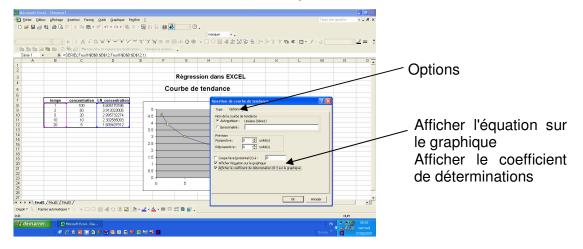


Figure 11: boîte de dialogue de « option » pour cocher « afficher le coefficient de détermination » et « afficher la droite de régression »



Cette méthode peut être utilisée pour confirmer ce que nous ferons pas à pas avec la macro "regression".

Sur la figure 9 on constate que l'ajustement est de mauvaise qualité. On peut vouloir ne faire passer la courbe de tendance que par les 4 derniers points. Pour cela, avec le click droit de la souris, sélectionner "données sources" (fig 12).

Figure 12 : Activation avec le clic droit de la souris « données source » pour sélectionner uniquement les 4 derniers données du vecteur « LN_concentration »

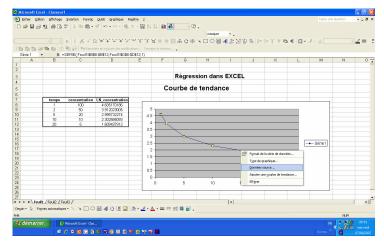
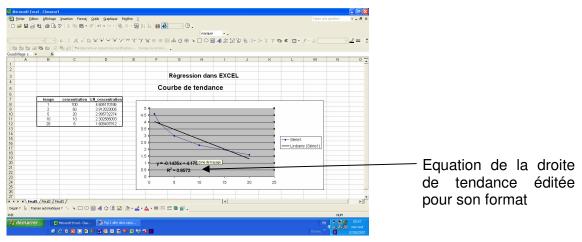
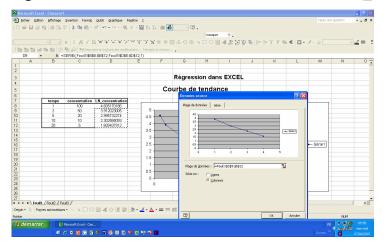


Figure 13 : Courbe de tendance montrant un mauvais ajustement entre le vecteur « temps » et le vecteur logarithme népérien des concentrations (LN_concentration)



On peut alors sélectionner la nouvelle plage de données d'intérêt (les 4 dernières valeurs de Ln-concentration). On ajoute la nouvelle courbe de tendance et on obtient un bien meilleur ajustement (fig.14).

Figure 14 : sélection des 4 dernières données de LN_concentration



et ajustement des 4 dernières données de LN_concentration

