

Nous travaillons sous Excel durant cette séance.

Nous utiliserons les supports suivants :

REF 1 - [http://chirouble.univ-lyon2.fr/~ricco/cours/cours/diapo\\_analyse\\_de\\_correlation.pdf](http://chirouble.univ-lyon2.fr/~ricco/cours/cours/diapo_analyse_de_correlation.pdf)

REF 2 - <http://tutoriels-data-mining.blogspot.fr/2017/04/probabilites-et-quantiles-sous-excel-r.html>

REF 3 - [http://chirouble.univ-lyon2.fr/~ricco/cours/cours/Regression\\_Lineaire\\_Simple.pdf](http://chirouble.univ-lyon2.fr/~ricco/cours/cours/Regression_Lineaire_Simple.pdf)

Les résultats en gris sont fournis à titre indicatif pour que vous puissiez étalonner vos calculs.

## Régression simple

Nous travaillons sur le fichier « **cars\_acceleration\_regression\_simple - sujet.xlsx** ». Nous souhaitons expliquer l'accélération des véhicules (Y : temps effectué pour parcourir une certaine distance) à partir de leur puissance (**horsepower** : X).

1. **Formules directes.** A partir des résultats obtenus en cours (REF 3, pages 6 et 7), calculer les estimations de la pente ( $a^{\wedge}$ ) et de la constante ( $b^{\wedge}$ ) de la régression linéaire simple c.-à-d. former  $(X - \bar{X})$ ,  $(Y - \bar{Y})$ ,  $X^2$ , etc. (Réponse :  $a^{\wedge} = -0.04639...$ ,  $b^{\wedge} = 20.27567...$ )
2. **Courbe de tendance.**
  - a. Construire un nuage de points (X) horsepower, (Y) acceleration.
  - b. Dans le graphique, ajouter une courbe de tendance (clic droit sur un des points dans le graphique) linéaire en faisant afficher le  $R^2$  et l'équation (Réponse :  $R^2 = 0.4678...$ )
  - c. Ajoutez la colonne  $Y^{\wedge}$  (prédiction d'accélération) (REF 3, page 30) dans votre feuille de calcul. Calculez alors la SCR (Réponse : SCR = 850.3038...).
  - d. Compléter la feuille de manière à obtenir le SCT et calculer alors le  $R^2$ . On devrait obtenir la même valeur.
3. **DROITEREG. Régression simple.**

Remarque : Pour utiliser la fonction DROITEREG, qui est une fonction matricielle complétant directement un bloc de cellules, il faut valider par l'appui simultané sur les touches CTRL + MAJ + ENTREE). Voir entre autres les tutoriels en ligne : <http://www.emse.fr/~pbreuil/capmes/droiterег.htm> ; <https://support.office.com/fr-fr/article/DROITEREG-DROITEREG-fonction-84d7d0d9-6e50-4101-977a-fa7abf772b6d?ui=fr-FR&rs=fr-CA&ad=CA> ; [https://wiki.documentfoundation.org/FR/Calc: fonction\\_DROITEREG](https://wiki.documentfoundation.org/FR/Calc: fonction_DROITEREG)).

- a. Sur les mêmes données, réalisez la régression avec DROITEREG. Repérez les différentes informations fournies par DROITEREG (REF 3, page 28).
- b. Construisez le tableau d'analyse de variance à partir des résultats de DROITEREG (REF 3, pages 26 et 27).
- c. La régression est-elle globalement significative à 5 % ? (REF 3, pages 26 et 27) (pour le calcul des quantiles INVERSE.LOI.F.DROITE et des p-values de la loi de Fisher LOI.F.DROITE, voir REF 2, pages 18 et 19) (Réponse : F-calculé = 200.44...)
- d. La pente de la régression est-elle significative à 5 % ? (REF 3, pages 24 et 25) (Réponse : t-calculé = -14.1577...)
- e. Vous remarquez une relation entre les valeurs du F-calculé et du t-calculé des deux tests précédents ? (Oui, vous devriez...)
- f. Calculez la colonne des valeurs estimées de l'endogène ( $\hat{y}_i$ ) ; puis la colonne des résidus estimés ( $\hat{\varepsilon}_i = y_i - \hat{y}_i$ )
- g. Construisez le graphique des résidus vs. exogène ( $x_i, \hat{\varepsilon}_i$ ). Avez-vous un commentaire à faire ? Idem, construisez le graphique ( $y_i, \hat{\varepsilon}_i$ ), commentaire ? (Remarque : Nous constatons – on devrait – un sérieux problème dans le second graphique. Normalement, nous devons y apporter une réponse pour pouvoir continuer. Pour simplifier dans cette phase de découverte de la pratique de la régression simple sous Excel, nous passons outre et nous enchaînons avec l'exercice suivant).

#### 4. Prédiction.

- a. Effectuer la prédiction ponctuelle sur la deuxième feuille "à prédire" (REF 3, page 30) (Réponse :  $\hat{y}_* = 17.5849$ )
- b. Compléter la prédiction avec les intervalles de projection à 95% (REF 3, pages 31 à 33) (Réponse :  $\hat{\sigma}_{\hat{\varepsilon}_*} = 1.93633$  ; intervalle de confiance à 95% = [13.7695, 21.4003])
- c. Vérifier que votre intervalle contient bien la bonne valeur de Y (cf. la feuille "vraie valeur") (Réponse : oui)