Régression linéaire : Deux fonctions personnalisées pour Excel

Henry P. AUBERT

hpa@free.fr

I Présentation

Ces fonctions sont illustrées et leur code disponible dans le classeur Excel Fonctions Régression linéaire.xls. Elle ont été créées par Henry AUBERT en 2005.

Ce code est un logiciel "libre", avec son code lisible et modifiable, dans l'esprit des logiciels ouverts. La seule condition mise à toute réutilisation publique de cette macro est l'obligation de citer son origine et indiquer le nom de l'auteur et de celui des éventuels correcteurs.

Visualiser et modifier(1) le code comme indiqué dans Comment créer des macros en visual basic.

Transférer le code du classeur Fonctions Régression linéaire.xls dans un autre classeur, ou dans votre classeur de macros personnelles Perso.xls, comme indiqué dans Comment transférer les macros d'un classeur Excel à un autre

Les auteurs seraient reconnaissants à toute personne qui améliorerait le fonctionnement de cette fonction, de leur faire partager cette amélioration.

II Contexte d'utilisation

On fait l'hypothèse que la variable d'intérêt numérique **Y** *aléatoire* dépend *linéairement* d'une variable numérique **X** dite indépendante ou explicative, selon la relation :

$$\mathbf{Y} = \alpha + \beta.\mathbf{X} + \varepsilon$$

Pour valider cette hypothèse, on dispose d'un échantillon indépendant de *n* observations.

On voudrait calculer, à partir de ces données, l'*estimation*, dans la population, du coefficient de corrélation linéaire $\hat{\rho} = \sqrt{1 - \frac{n-1}{n-2}.(1-\mathbf{r}^2)}$ à partir de celui calculé sur les données, $\mathbf{r^2}$.

On cherche ensuite un intervalle de confiance bilatéral de :

- La valeur moyenne $\bar{y}_x = \hat{\alpha} + \hat{\beta}.x_i$ attendue pour une valeur x non observée. Pour cela il faut calculer l'estimation de son écart-type, $\hat{\sigma}_{\hat{y}_x} = \sqrt{\frac{(1-\mathbf{r}^2)\cdot\sigma_y^2}{(n-2)\cdot\sigma_x^2}} \cdot \sqrt{\sigma_x^2 + (x-\bar{x})^2}$.
- La prévision $\hat{y}_i$ attendue pour une valeur x non observée Pour cela il faut calculer l'estimation de son écart-type, $\hat{\sigma}_{\hat{y}_x} = \sqrt{\frac{(1-\mathbf{r}^2)\cdot\sigma_y^2}{(n-2)\cdot\sigma_x^2}} \cdot \sqrt{(n+1)\cdot\sigma_x^2 + (x-\bar{x})^2}$.

(1) En cas d'amélioration ou de correction d'une erreur, l'auteur vous serait reconnaissant de bien vouloir lui en faire part.

III Description des fonctions

III.1 CoefficientCorréllationEstimé(Y_connus ; X_connus)

Ses arguments sont :

- `Y_connus :` La référence ou le nom d'une plage de même nombre de cellules que la plage `X_connus`, et contenant obligtoirement des *nombres*..
- `X_connus :` La référence ou le nom d'une plage de même nombre de cellules que la plage `Y_connus`, et contenant obligtoirement des *nombres*.

III.2 EcartypePrévision(Y_connus ; X_connus ; Val_X ; Prévision)

Ses arguments sont :

- `Y_connus :` La référence ou le nom d'une plage de même nombre de cellules que la plage `X_connus`, et contenant obligtoirement des *nombres*..
- `X_connus :` La référence ou le nom d'une plage de même nombre de cellules que la plage `Y_connus`, et contenant obligtoirement des *nombres*.
- `Val_X :` La valeur, ou la référence ou le nom d'une plage contenant la valuer de la variable explicative pour laquelle on cherche à calculer l'écart-type de la prévision.
- `Prévision :` Un booléen = `Vrai` (par défaut) pour l'écart-type de la *prévision*, = `Faux` pour l'écart-type de la `moyenne`.