

# BOXCOXTRANS

## Syntaxe

```
BOXCOXTRANS X;  
  LAMBDA LI LS DL;  
  XSUPPL XS;  
  LEVEL ALPHA;  
  OUT LAMBD LOGLIK RNS;  
  OUTOPT L L1 L2 CORR LV LVLIM;  
  XTRANS XT;  
  XSTRANS XST;  
  DETAILS;  
  NOCPRECIS;  
  NOXSUPPR;  
  NOGL;  
  NOGRNS;  
  NOPRINT.
```

## Sous-commandes

LAMBDA : définition des valeurs de lambda à considérer

XSUPPL : données supplémentaires à transformer

LEVEL : définition du degré de confiance

OUT : enregistrement des informations pour les différentes valeurs de lambda

OUTOPT : enregistrement des informations pour la valeur de lambda correspondant au maximum de vraisemblance

XTRANS : enregistrement des données initiales transformées

XSTRANS : enregistrement des données supplémentaires transformées

DETAILS : impression d'informations détaillées pour chaque valeur de lambda

NOCPRECIS : suppression de la deuxième étape lors du calcul de la valeur de lambda qui maximise la fonction de vraisemblance

NOXSUPPR : pas d'impression des données supplémentaires transformées

NOGL : pas de graphique de la fonction de vraisemblance

NOGRNS : pas de graphique de l'évolution de la corrélation entre les données transformées en fonction de lambda et les scores normaux

NOPRINT : suppression des impressions.

## **Résumé**

La transformation de BOX et COX est la suivante :

$$y(\lambda) = \begin{cases} x^\lambda & \text{si } \lambda \neq 0 \\ \log_e x & \text{si } \lambda = 0, \end{cases}$$

la valeur de  $\lambda$  étant estimée par la méthode du maximum de vraisemblance.

Dans la macro, cette estimation se fait en deux temps. La vraisemblance est d'abord calculée pour les valeurs de  $\lambda$  variant de LI à LS par pas de DL. Soit L0, la valeur qui, parmi les valeurs testées, maximise la vraisemblance. Une nouvelle recherche est ensuite réalisée, en faisant varier  $\lambda$  de L0 – DL à L0 + DL, par pas de DL/10. La valeur finalement retenue pour  $\lambda$  est celle qui maximise la vraisemblance pour ce nouvel ensemble de valeurs  $\lambda$  testées. La sous-commande NOCPRECIS supprime cette seconde étape dans la recherche de la valeur optimale de  $\lambda$ .

En présence de données négatives ou nulles, la macro ajoute une constante aux données initiales :

$$x(\text{corrigé}) = x + |x_{\min}| + 0.001\hat{\sigma}_x,$$

$x_{\min}$  et  $\hat{\sigma}_x$  étant le minimum et l'écart-type de la série à transformer.

La macro donne également, sous forme graphique, l'évolution de la vraisemblance et de la corrélation entre les données transformées et les scores normaux correspondants en fonction de  $\lambda$ .

La macro réalise la transformation des données de départ, ainsi que d'éventuelles données supplémentaires.

Différentes options permettent de contrôler le volume des sorties.

## **Macro utilisée**

BCT : voir BCT.doc

ITLIN : voir ITLIN.doc

## **Arguments d'entrée**

X : colonne contenant la variable pour laquelle la valeur de lambda est déterminée

LI LS et DL : valeur inférieure, valeur supérieure et pas de variation de lambda. Par défaut LI = - 2, LS = 2 et DL = 0.1

XS : colonne contenant les données supplémentaires à transformer avec la valeur de lambda déterminée sur les données  $x_i$

ALPHA : constante permettant de fixer le degré de confiance, égal à  $1 - \text{ALPHA}$  utilisé pour déterminer les limites de confiance de la valeur estimée de lambda

### **Arguments de sortie**

LAMBD LOGLIK et RNS : colonnes contenant, pour les différentes valeurs de lambda, le logarithme de la vraisemblance et la corrélation de la variable transformée avec les scores normaux

L L1 L2 CORR LV et LVLIM: constantes contenant la valeur estimée de lambda et ses limites de confiance, la corrélation de la variable transformée avec les scores normaux, la valeur du logarithme de vraisemblance  $L(\hat{\lambda})$  et la valeur limite du logarithme de vraisemblance :  $L(\hat{\lambda}) - \frac{1}{2} \chi^2_{1-\alpha}$

XT et XST : variables  $X$  et  $XS$  transformées