Économétrie II

Ch. 2. Inférence Classique & Bootstrap
L3 Économétrie – L3 MASS

Pr. Philippe Polomé, U. Lyon 2

Année 2015-2016

Table des matières

Ch. 2. Inférence Classique & Bootstrap

- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Contexte

- MRL $y = X\beta + \epsilon$ et hypothèses
- Un des effets secondaires des ruptures d'hypothèse est que l'**inférence** est invalidée
  - Tests t, F, ...
- Dans ce chapitre
  - Rappel d'inférence
  - Introduction au bootstrap

Principe de l'inférence

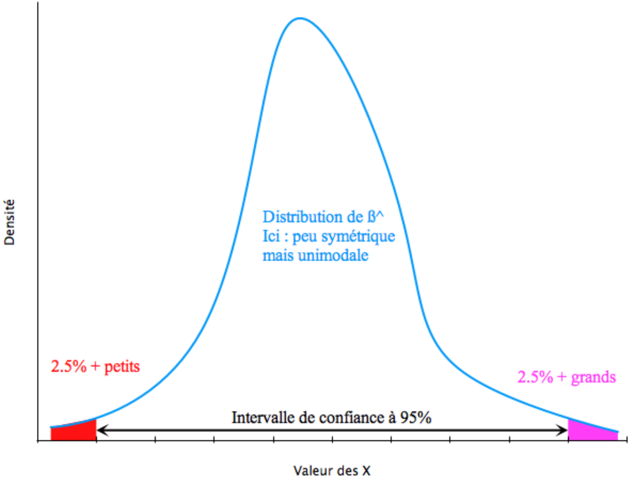
- On a une statistique $\hat{\beta}$ estimation d'un paramètre inconnu β
  - On peut calculer la variabilité statistique de $\hat{\beta}$: sa **distribution**
  - Nombreuses manières
- On veut savoir si on peut dire que $\beta =$ un certain chiffre b
  - $H_0 : \beta - b = 0$ en tenant compte du bruit statistique
  - On se donne des marges d'erreur acceptables : 5% de rejeter H_0 quand elle est vraie
  - Seuil le plus communément utilisé mais arbitraire
- Ces marges définissent un **intervalle** dans la variabilité statistique de $\hat{\beta}$
  - Pas forcément symétrique
- Si $b \in$ intervalle, alors $\neg R\ H_0$
  - On $R\ H_0$ 5% du temps lorsque H_0 est en fait vraie
  - D'où le nom "intervalle de confiance à 95%"

Types d'erreur

	H_0 vraie	H_0 fausse
R	Erreur type I – Prob. $\alpha = 5\%$	Correct
$\neg R$	Correct	Erreur type II – Prob. β

- La probabilité β de l'erreur de type II dépend de H_0, $1-\beta$ est la puissance du test.

Graphiquement



Supposons que la vraie valeur du paramètre soit effectivement b, quelle est la probabilité qu'on ait observé $\hat{\beta}$?

- Si cette probabilité est <5%, alors on décide que H_0 était fausse : $\beta \neq b$

Table des matières

Ch. 2. Inférence Classique & Bootstrap
- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Inférence classique (économétrie I)

- En inférence classique, la plupart des tests
  - t, F, ...
  - reposent sur l'hypothèse de **normalité** du terme d'erreur
  - au moins pour de petits échantillons
- Procédure analytique de test
  - On formule une hypothèse, p.e. $H_0 : \beta_i = 0$
  - On calcule une **statistique du test** dont on peut savoir quelle distribution elle a si H_0 est vraie et $\epsilon \sim n()$
- La normalité permet de déduire la distribution de la statistique de test quand H_0 est vraie

Exemple de la t de Student

- $$\frac{\hat{\beta}_i}{ecart\ type\left(\hat{\beta}_i\right)} \sim t_{n-k} \text{ si}$$
  - $H_0 : \beta_i = 0$ vraie
  - et si $\epsilon \sim n\left(\right)$
  - et si les hypothèses du MRL sont vraies
- La distribution de Student est **tabulée** : il existe des tables avec ses valeurs
  - livres d'économétrie & tableurs
  - Gretl menu Outils $\rightarrow$ Tables statistiques
- On **compare** la valeur calculée de la statistique du test aux valeurs tabulées
  - Si la statistique de test est dans les 5% extrêmes de la t de Student, on décide que c'est H_0 qui est fausse

Exemple de table : t de Student

Critical Values of the *t* Distribution

		Significance Level				
1-Tailed:		.10	.05	.025	.01	.005
2-Tailed:		.20	.10	.05	.02	.01
Degrees of	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898

Matrice de variance-covariance des coefficients

- La plupart des statistiques de test sur les coefficients sont construites à partir de la matrice de variance-covariance (conditionnelle à X) des coefficients
- Plus les estimations seront précises, plus on devrait rejeter facilement (toutes autres choses égales)
- Si les hypothèses MRL sont satisfaites :
  $V\left(\hat{\beta}_{MCO}|X\right) = \sigma^2 \left(X'X\right)^{-1}$
  - Où $\sigma^2 = var\left(\epsilon_i\right) < \infty$ est la variance des erreurs
  - On écrit souvent $V\left(\hat{\beta}\right)$

Estimation des variances

- La variance σ^2 des erreurs ϵ est inconnue
- L'estimateur MCO de cette variance est $\hat{\sigma}^2 = \dfrac{\hat{\epsilon}'\hat{\epsilon}}{n-k}$ où
  - $\hat{\epsilon} = Y - \hat{Y} = Y - X\hat{\beta}$ résidu MCO
  - k nombre de régresseurs (y compris l'intercept)
- Cet estimateur est non-biaisé
- L'estimateur MCO de la matrice de variance-covariance (conditionnelle à X) des coefficients est

$$\widehat{V\left(\hat{\beta}|X\right)} = \hat{\sigma}^2 \left(X'X\right)^{-1} = \frac{\hat{\epsilon}'\hat{\epsilon}}{n-k} \left(X'X\right)^{-1}$$

Table des matières

Ch. 2. Inférence Classique & Bootstrap
- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Exemple

- Pour clarifier ces notions : détail d'un exemple

1. Charger des données de Gretl dans un tableur
2. Estimateur MCO
3. Calcul des résidus
4. Estimation de la matrice de var-cov
5. Calcul des t-stats

Données

- Dans Gretl, charger le jeu de données bwages des données de Verbeek
  - Source gretl.sourceforge.net/gretl_data.html
- 1472 observations de salaires (wage) horaires bruts de 1994 en Belgique
  - Educ, Experience, Genre
- Exportation classique vers tableur
  - File –> Export Data –> CSV (Comma Separated Values)
- Ouvrir le fichier à partir tableur
  - Sélect. col.
  - Menu Données –> Convertir –> valeur délimitées –> virgule
  - Sauver bwages.ods

Modélisation

- Modèle économétrique Wages = F(Experience, Education, Sex)
  - Certainement autres régresseurs pertinents (secteur...), mais pas données
- Transformation des données
  - ln(wages) (au lieu de wages directement) : limite l'hétéroscédasticité (on verra + loin)
  - ln(1+Experience) Experience = 0 : ln(0) impossible
- A priori modèle linéaire

$$ln\left(wages\right) = \beta_0 + \beta_1 ln\left(1 + exp\right) + \beta_2 "Educ" + \beta_3 sex + \epsilon$$

- $\partial ln\left(wage\right) / \partial ln\left(exp\right) = \beta_1$ élasticité du salaire à l'expérience

Régresseurs qualitatifs

- Male $\in \{0,1\}$: variable binaire/dichotomique (aussi dummy/fictive/indicatrice)
  - 1 = homme, donc 0 = ...
- Educ $\in \{1,\ldots,5\}$: variable catégorique
  - 1 = peu d'éducation (diplôme) ... 5 = beaucoup
  - Les chiffres ne sont qu'un code, leur différence ne veut rien dire
  - On transforme en 4 dichotomiques : Deduc1 = 1 si educ = 1, 0 sinon etc...
- Male et Educ sont des variables **qualitatives**
- Dans bwages.ods : création de données

Modèle

- Proposition de modèle :

$$\ln(w) = \beta_0 + \sum_{i=1}^{4} \beta_i dE_i + \beta_5 \ln(1 + exp) + \beta_6 Male + \epsilon$$

- 3 remarques
  - i = 1 à 4 : La 5ème catégorie est **implicite**, les autres coefficients s'interprètent en référence à elle
    - p.e. si $\beta_2 = -1$, alors avoir un niveau d'éducation de la catégorie 2 implique que le ln du salaire sera de 1 inférieur comparativement à la catégorie 5 toute autre chose égale
  - $\beta_5 = \partial \ln(w) / \partial \ln(1 + exp)$ = élasticité du salaire p/r à l'expérience
  - $\beta_6 > 0 \implies$ à éducation et expérience égales, les hommes gagnent plus que les femmes

$\widehat{V\left(\hat{\beta}|X\right)}$

- Une fois les coefficients estimés $\hat{\beta}$
  - calculer les résidus $\hat{\epsilon} = y - X\hat{\beta}$
- $\widehat{V\left(\hat{\beta}|X\right)} = \hat{\sigma^2}\left(X'X\right)^{-1} = \frac{\hat{\epsilon}'\hat{\epsilon}}{n-k}\left(X'X\right)^{-1}$
  - Calcul dans bwages.ods

La huitième hypothèse du MRL : normalité

- Hypothèse de normalité des erreurs : $\epsilon \sim n\left(0, \sigma^2\right)$
- $\Rightarrow \hat{\beta} \sim n\left(\beta, \sigma^2 \left(X'X\right)^{-1}\right)$
- $\Rightarrow t = \dfrac{\hat{\beta}_i}{\sqrt{diag_i\left(\widehat{V\left(\hat{\beta}\right)}\right)}} \sim t_{n-k}$ est vrai
  - Si H0 $\beta_i = 0$ est vraie
  - Si l'hyp de normalité et toutes les autres hyp du MRL sont vraies

Principe du test de significativité classique

- Pour $n - k$ suffisamment grand (une centaine), le percentile 0,975 de t_{n-k} vaut 1,96
- Test de significativité

Si le t du coefficient β_i de la variable x_i

soit le coefficient estimé divisé par son écart-type estimé

vaut moins de 1,96 en valeur absolue

alors x_i n'est pas significative ($\neq 0$)

"règle du 2"

- Pour des valeurs petites de $n - k$, on compare avec le percentile tabulé t_{n-k}

Illustration graphique

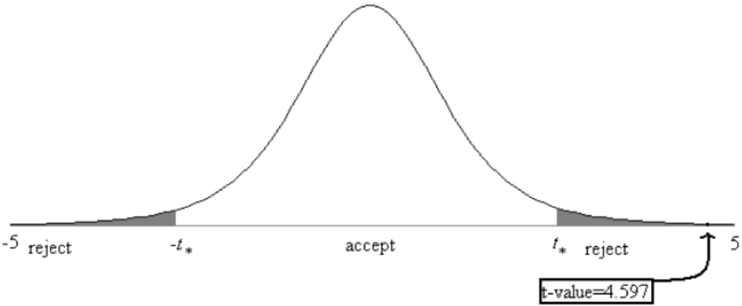


Figure 2 Two-sided t-test of H_0: $\beta = 0$ against the alternative hypothesis H_1: $\beta \neq 0$.

- Calcul des t-stat dans bwages.ods

P-valeur

- De manière équivalente, on peut aussi calculer à partir de quel α (risque de rejeter une hypothèse vraie / zone grise) le coefficient est significatif
  - Plus on prend α petit, plus l'intervalle de confiance est grand, plus il est probable qu'il contienne le zéro, moins il est probable que le coefficient soit significatif
- Lorsque la p-valeur est inférieure à 5%, on dit que le coefficient est significatif (à 5%)
  - Percentile 0,975 [intervalle 95%, $\alpha = 5\%$] de t_{n-k} vaut 1,96
  - Percentile 0,95 [intervalle 90%, $\alpha = 10\%$] de t_{n-k} vaut 1,65
  - Percentile 0,995 [intervalle 99%, $\alpha = 1\%$] de t_{n-k} vaut 2,56
- Calcul des p-valeurs dans bwages.ods
- Sortie Gretl pour vérifier l'équivalence
  - Création des dichotomiques : clic droit $\rightarrow$ "dummify"

Autres tests d'égalité de coefficients

- Au moyen de procédures similaires, on peut tester
  - si le coefficient estimé est significativement différent d'un certain chiffre b
  - si plusieurs coefficients sont significativement différents les uns des autres ou d'un même chiffre
  - si des combinaisons linéaires de coefficients valent un certain chiffre
- Tests dit en F ou de Wald dès que plus d'un coefficient
  - Tous les logiciels fournissent la p-valeur
  - Application particulière standard : H_0 : tous les coefficients (sauf l'intercept) sont simultanément nuls
    - voir sortie dans Gretl
- Voir Test ou Analyse après estimation dans Gretl
  - Postestimation dans Stata (menu Statistique)
  - p.e. b[2]=b[3] à écrire b[2]-b[3]=0

Limites de l'inférence classique

- Test de normalité des résidus
  - Que faire si non-normaux ?
- Intervalle de confiance pour des statistiques plus complexes
  - Combinaison non-linéaire de paramètres
  - Surplus du consommateur (intégrale sous une droite de demande)
- Plusieurs techniques existent
  - Bootstrap : plus versatile, vraisemblablement plus robuste

Table des matières

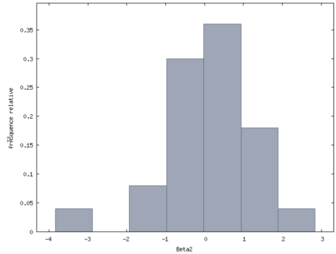
Ch. 2. Inférence Classique & Bootstrap
- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Échantillon aléatoire

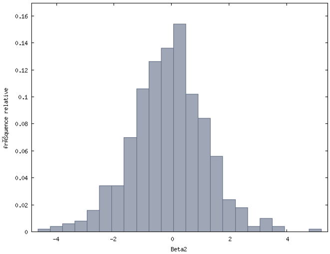
- On a vu que chaque échantillon est un tirage d'une population
  - L'échantillon peut être "aléatoire simple" ou plus complexe
- Les coefficients estimés $\hat{\beta}$ sont des nombres aléatoires
  - Chaque échantillon e de la même population produit un $\hat{\beta}_e$ différent
- Imaginons qu'on ait plusieurs échantillons, donc plusieurs $\hat{\beta}_e$
  - La **distribution empirique** de $\hat{\beta}$ est l'histogramme des $\hat{\beta}_e$
  - Quand le nombre d'échantillons $\rightarrow \infty$, la distribution empirique **converge** vers la fonction de densité de $\hat{\beta}$

Distribution empirique

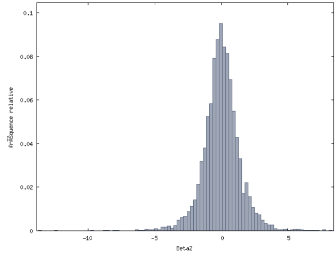
50 échantillons



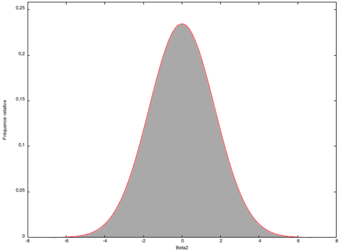
500 échantillons



5000 échantillons



∞ échantillons



Intervalle de confiance empirique

- Si on connaissait la fonction de densité de $\hat{\beta}$
  - On pourrait définir un intervalle de confiance 95% entre les quantiles 2.5% et 97.5% de la distribution
- Avec un nombre E fini d'échantillons :
  - **L'intervalle de confiance empirique** à 95% d'un élément $\hat{\beta}_k$ de $\hat{\beta}$ est formé par les quantiles 2.5% et 97.5% de la distribution empirique
  - Soit $\hat{\beta}_{ke}$ la valeur estimée de β_k dans l'échantillon e
  - On a E valeurs estimées : $\hat{\beta}_{k1}, \ldots, \hat{\beta}_{kE}$
    - Disons 1000 pour simplifier
  - On les ordonne de la plus petite à la plus grande
  - La 25º est le quantile 2.5%
- Pour d'autres quantités, élasticité η p.e.
  - on calcule la quantité d'intérêt $\hat{\eta}_e$ pour chaque e
  - et on prend les quantiles de leur distribution empirique

Re-tirage avec remplacement

- En réalité, on a rarement plus d'un échantillon
- Expérience de Monte-Carlo
  - À partir de chaque échantillon artificiel m, on pouvait calculer $\hat{\beta}_m$ à chaque fois
- Avec un échantillon réel, on **ne peut** en tirer un nouveau, mais
  - Supposons que l'échantillon n'est pas atypique par rapport à la population d'où il est tiré
    - Principe de **médiocrité**
    - $\neq$ représentativité
  - Alors, si on avait pu tirer un autre échantillon, on aurait observé au moins une partie des mêmes chiffres
- Si on tire des obs. de l'échantillon observé, on peut considérer qu'il s'agit d'un autre échantillon de la même population
  - Pour garder un échantillon de même taille : tirer avec **remplacement**
- Le re-tirage avec remplacement constitue le **bootstrap**

Bootstrap

- Par exemple, soit l'échantillon {1,-1,2,3}
  - Un nouvel échantillon bootstrap de celui-là pourrait être {1,1,-1,3}
  - Un autre {2,2,3,3}
  - Tous ces échantillons sont équiprobables
- Donc avec le bootstrap, on est dans un situation semblable au **Monte-Carlo**
  - L'échantillon original est vu comme la population
  - On a plusieurs échantillons e issus de cette population
  - On peut alors calculer $\hat{\beta}_e$ pour chacun d'eux
  - En répétant, on obtient une **distribution empirique** des $\hat{\beta}$
    - Combien de fois ? la littérature suggère que 1000 couvrirait la plupart des cas

Intervalle de confiance bootstrap

- La distribution empirique représente toutes les valeurs observées de $\hat{\beta}$
  - En retirant les 2.5% plus petites valeurs et les 2.5% plus grandes, on obtient l'intervalle à 95%
    - Donc : Si $H_0 : \beta_i = 0$: si $0 \in$ IC, alors H_0 n'est pas invalidée
    - Mais si $0 \notin$ IC, 0 est considéré comme une valeur improbable de β_i et on rejette H_0
  - En calculant l'écart-type de la distribution empirique de $\hat{\beta}_j$, on obtient l'écart-type de $\hat{\beta}_j$
- Exemple tableur bootstrap.ods
- De même, dans chaque échantillon bootstrap, on peut calculer des fonctions des estimations et des données
  - Par exemple, une élasticité $\hat{\eta}_e = H\left(Y_e, X_e, \hat{\beta}_e\right)$
  - L'intervalle de confiance pour l'élasticité est obtenu comme pour $\hat{\beta}$

Types de bootstrap

- Le bootstrap décrit est le bootstrap **non-paramétrique** ou bootstrap **par paires**
  - on ré-échantillonne des paires (Y, X) de l'échantillon original
- $\exists$ techniques de bootstrap où on va vouloir trouver juste de nouveaux y : y^e
  - Boostrap **paramétrique** : on suppose p.e. que $y \sim n\left(X\hat{\theta}, \Sigma\right)$ et on tire des y^e d'une telle normale
  - Bootstrap **résiduel** : On ré-échantillonne les résidus $y_i^e = X_i\hat{\beta} + \hat{\epsilon}_i^e$
- Le non-paramétrique est conceptuellement plus simple
  - Mais les autres bootstraps permettent d'obtenir de meilleures approximations (aux intervalles de confiance ou autres) *pour autant que leurs hypothèses soient satisfaites*

Bootstrap par bloc

- Lorsque les observations ne sont pas IID, les techniques précédentes ne sont plus valables
  - parce qu'elles **détruisent** la corrélation entre observations
  - L'échantillon bootstrap ne peut être considéré comme un nouvel échantillon
  - Bootstrap est plus utile pour coupes transversales que pour séries chronologiques
- **Bootstrap par bloc**
  - On coupe l'échantillon en r blocs de longueur l de telle sorte que $rl \simeq N$ (pas nécessaire exactement car peut ne pas tomber juste)
  - On ré-échantillonne sur ces blocs seulement (pas par observation)
    - Les blocs deviennent donc indépendants, mais la corrélation est préservée à l'intérieur
- Application en particulier en panel

Table des matières

Ch. 2. Inférence Classique & Bootstrap
- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Bootstrap dans Gretl

- Avant les ordinateurs modernes, le bootstrap était impossible
- Après estimation, dans le menu "Analyse", la commande "Bootstrap" permet de calculer
  - Un IC pour chaque coefficient et pour chaque t
  - Par bootstrap résiduel
  - Exemple airq.gdt
- Gretl a une commande spéciale **resample**
  - genr xr = resample(x) crée une nouvelle série XR par ré-échantillonnage de X où X peut être une matrice ou un vecteur
    - Si Xest une matrice, resample est le bootstrap par paire
    - Illustration avec les données airq.gdt (Verbeek) : observations pour 30 "standard metropolitan statistical areas" = villes (SMSAs) en Californie en 1972

Programmation d'un bootstrap par paire dans Gretl

```
# data exemple : airq.gdt (Verbeek) a charger avant bootstrap
ols airq 0 vala rain coas dens medi # MCO sur donnees de base

matrix A = {airq, vala, rain, coas, dens, medi}

scalar replics = 1000 # nbr replications bootstrap
bM = zeros(6,replics) # pour stocker les valeurs des betas estimes
qua = zeros(1,replics) # pour stocker la ∑ des coeff de vala, rain, dens et medi
i=0 # initialiser compteurs
loop replics --quiet
    # Generer echantillon bootstrap B a partir de A
    B=resample(A) # les lignes de B sont tirees au sort avec remplacement
    series airqR=B[,1] # creer une serie pour utilisation dans OLS
    series valaR=B[,2]
    series rainR=B[,3]
    series coasR=B[,4]
    series densR=B[,5]
    series mediR=B[,6]

    ols airqR 0 valaR rainR coasR densR mediR
    i=i+1
    bM[,i]=$coeff
    qua[1,i]=$coeff(valaR)+$coeff(rainR)+$coeff(densR)+$coeff(mediR)
endloop
```

(pour ceux que ça intéresse, hors examen)

Table des matières

Ch. 2. Inférence Classique & Bootstrap
- Motivation
- Inférence classique
- Inférence classique : exemple
- Inférence Bootstrap
- Bootstrap : Exemple
- Comparaison bootstrap / classique

Comparaison Bootstrap / Classique

Bootstrap	Classique
IID & médiocrité	IID & normalité
$H_0 : \beta_i = 0$	
Distribution empirique des $\hat{\beta}_i^s$	Stat de test $t = \hat{\beta}_i / \sqrt{diag_i\left(\widehat{V\left(\hat{\beta}\right)}\right)} \sim t_{n-k}$ si H_0 est vraie
RH_0 si $0 \notin \left[\hat{\beta}_{i,.025}^s, \hat{\beta}_{i,.975}^s\right]$	RH_0 si $t \notin [t_{n-k,.025}, t_{n-k,.975}]$

Cas asymptotique $n \to \infty$

- Lorsque les observations sont très nombreuses, plusieurs centaines au moins
  - Les coefficients estimés sont (presque toujours) asymptotiquement $\sim n()$
  - L'inférence classique ne dépend donc de l'hypothèse de normalité des erreurs que pour les petits échantillons
- L'avantage du bootstrap est moindre lorsque $n \to \infty$
  - Bootstrap reste valable en présence de certaines ruptures des hypothèses du MRL, alors que l'inférence classique non

Devoir #2 : bootstrap

- Prendre un échantillon réel de Gretl
- Estimer un modèle MCO
- Calculez les intervalles de confiance empiriques des $\hat{\beta}$
    - Comparer avec les intervalles de confiance classiques (sorties standards) : est-ce que vous obtenez les mêmes conclusions ?
- Estimer les t-stats par bootstrap en calculant l'écart-type de la distribution empirique des $\hat{\beta}$
    - Comparez avec les t-stats classiques : est-ce qu'elles sont proches ? Aboutissent-elles aux mêmes conclusions ?
        - En général, ça dépendra de votre échantillon
        - Il n'y a pas de réponse universelle