Séance 7 – Régression linéaire et test d'hypothèses

Pierre Pora

Introduction

Rappel des séances précédentes

- Ce que l'on regarde dans la **population** toute entière : on écrit quelque chose comme $Y = X'\beta + \epsilon$ avec $\mathcal{C}(X, \epsilon) = 0$
  - $X'\beta$ est la **meilleure approximation** de Y par une fonction linéaire de X
  - β s'interprète comme des **différences de moyennes** entre des sous-populations définies par les valeurs de X
  - Sous des hypothèses très fortes, et à rediscuter, β a une interprétation causale

Rappel des séances précédentes

- Ni β ni ϵ ne sont connus !
- On n'a pas accès à la population prise toute entière, mais seulement à un petit **échantillon aléatoire**
  - Avec ce petit échantillon on sait construire un **estimateur** $\hat{\beta}$ qui nous donne une approximation de β
    - C'est une v.a. qui dépend du tirage de l'échantillon !
    - Quand la taille d'échantillon devient très grande (par rapport à 1), $\hat{\beta}$ devient très proche de β
  - On sait quantifier les **fluctuations aléatoires** de $\hat{\beta}$ autour de β dans ce régime asymptotique
    - $\sqrt{n}\left(\hat{\beta} - \beta\right)$ s'identifie à une gaussienne centrée multivariée dont on connaît la matrice de variance-covariance
    - **Matrice de variance-covariance robuste à l'hétéroscédasticité**, éventuellement prise en compte des données groupées lorsque c'est nécessaire

L'objet de cette séance

- Comment utiliser ces résultats pour **tester des hypothèses** ?
  - Cela demande de formaliser (un peu) ce qu'est une hypothèse
  - Ensuite on ne fait jamais autre chose que de bricoler autour de la **normalité asymptotique** de $\sqrt{n}\left(\hat{\beta} - \beta\right)$
    - Il y a tout un bestiaire de lois de probabilité dérivées de la loi normale qui est utile pour cela

Le déroulé de la séance

- Ce que l'on pourrait faire
  - Partir de ce que l'on sait sur la distribution asymptotique de $\hat{\beta}$
  - Faire toute la **botanique des tests classiques** en explicitant chaque fois la loi de probabilité que l'on utilise etc.
- C'est bien si on veut juste savoir recoder presque à la main chaque test que R permet d'automatiser
- Mais pas très amusant, ni très utile quand on va vouloir discuter sérieusement de ce que tout cela veut dire
- Et pas très éclairant sur le **concept de test statistique** de façon générale

Le déroulé de la séance

- Ce que je propose
  - Passer un peu plus de temps sur le **concept même de test**
    - Sans aller dans un formalisme mathématique trop exigeant
  - Expliciter sur un exemple le test le plus utilisé
    - Mettre les autres en TD
  - Introduire une petite discussion sur l'intérêt et surtout les **difficultés posées par l'usage des tests statistiques en pratique**
    - L'usage maladroit des tests statistiques soulève des questions sérieuses quant à la façon dont on peut combiner les résultats d'articles différents !

Concept de test statistique

Qu'est-ce qu'un test statistique ?

- Jusque là, on s'intéressait à **une population donnée**
- Et on cherchait à connaître les valeurs d'un vecteur β qui synthétise des **différences de moyennes conditionnelles** d'une variable dépendante Y en fonction d'un lot de variables dépendantes X
- Pour cela, on s'intéressait à un **échantillon de cette population**, de taille petite devant celle de la population mais grande devant 1
- On connaît les (X_i, Y_i) des individus de cet échantillon
- Lorsque la taille d'échantillon devient suffisamment grande, on peut se donner une idée raisonnable de β en regardant $\hat{\beta}$, et on sait aussi estimer de combien on se trompe vraisemblablement en confondant β et $\hat{\beta}$

Qu'est-ce qu'un test statistique ?

- Maintenant, il faut imaginer que notre population d'intérêt est prise dans un **super-ensemble de populations possibles** pour lesquelles on peut chaque fois définir β et ϵ
  - Qui ne dépendent que de la distribution jointe de X et Y dans chacune de ces populations possibles
- On veut dire quelque chose de la **validité d'une théorie économique dans la population effective à laquelle on s'intéresse**, prise dans l'ensemble des popuations possibles
  - Pour certaines de ces populations possibles, cette théorie est vraie, pour d'autres, elle est fausse
- Et on garde notre processus d'échantillonage qui nous permet de connaître les (X_i, Y_i) pour un échantillon fini tiré de la population

Qu'est-ce qu'un test statistique ?

- On s'intéresse à une théorie économique qui dépend des valeurs de β
  - La vraie valeur des coefficients, **prise dans chacune des populations possibles**
  - C'est un vecteur de $\mathbb{R}^d$
- Schématiquement, cette théorie est
  - Fausse si β appartient à un certain sous-ensemble $\mathbf{B}_0 \subset \mathbb{R}^d$
  - Vraie si β appartient à un certain sous-ensemble $\mathbf{B}_1 \subset \mathbb{R}^d$
  - Avec $\mathbb{R}^d = \mathbf{B}_0 \cup \mathbf{B}_1$ et $\mathbf{B}_0 \cap \mathbf{B}_1 = \emptyset$

Qu'est-ce qu'un test statistique ?

- Tester statistiquement cette théorie à partir de l'échantillon dont on dispose, ce n'est finalement pas autre chose que de se doter d'**une règle qui part de** $(X_1, Y_1), \ldots, (X_n, Y_n)$ et qui renvoie ou bien `"La théorie est fausse (pour la population étudiée)"` ou bien `"La théorie est vraie (pour la population étudiée)"`
  - **C'est donc une v.a.**
  - Selon la façon dont on veut présenter les choses, elle prend ses valeurs dans :
    - $\{\{\mathbf{B}_0\}, \{\mathbf{B}_1\}\}$
    - Ou bien plus simplement dans $\{0, 1\}$

Deux types de risque

- Deux possibilités pour notre théorie
  - Être vraie ou fausse pour la population étudiée

- Deux possibilités pour notre test
  - Renvoyer `"La théorie est vraie (pour la population étudiée)"` ou `"La théorie est fausse (pour la population étudiée)"`

  - Attention : ces valeurs affichent qu'elles parlent de la population, mais **elles ne sont calculées qu'à partir de l'échantillon !**

- Il y a donc **4 possibilités en tout !**

Deux types de risque

	"La théorie est fausse"	"La théorie est vraie"
La théorie est fausse		Erreur de première espèce
La théorie est vraie	Erreur de seconde espèce	

Deux types de risque

- Ces deux erreurs permettent de définir **deux types de risques**
  - Le risque de déclarer la théorie vraie à partir de l'échantillon alors qu'elle est fausse dans la population $\rightarrow$ **risque de première espèce**
  - Le risque de déclarer la théorie fausse à partir de l'échantillon alors qu'elle est vraie dans la population $\rightarrow$ **risque de seconde espèce**
- On ne peut pas minimiser les deux risques en même temps !!
  - Un test qui a un risque de première espèce nul : renvoie tout le temps `"La théorie est fausse"`
  - Un test qui a un risque de seconde espèce nul : renvoie tout le temps `"La théorie est vraie"`
  - Ce sont deux tests absolument sans intérêt !

Deux types de risque

- **On ne traite pas les deux risques de façon symétrique**
  - Justification historique : souvent l'un est plus intéressant ou plus pressant que l'autre et c'est celui que l'on veut maîtriser
  - Cela s'applique-t-il vraiment au cas des théories économiques ?
    - C'est une bonne question que l'on discutera en fin de séance
- Risque de première espèce : $\sup_{\beta \in \mathbf{B}_0} \mathbb{P}_\beta (T = 1)$
  - où $\mathbb{P}_\beta(T = 1)$ correspond à la probabilité que le test renvoie `"vraie"`, pour le processus d'échantillonnage que l'on met en place, dans une population potentielle caractérisée par une valeur de β
  - Quand $\mathbf{B}_0 = \{\beta_0\}$ avec $\beta_0 \in \mathbb{R}^d$ (par exemple $\beta_0 = 0$) c'est tout bonnement $\mathbb{P}_{\beta_0}(T = 1)$

Niveau / taille d'un test

- Niveau (de significativité) / taille d'un test : un test est de niveau $\alpha \in]0, 1[$ si le risque de première espèce est inférieur ou égal à α
  - En d'autre termes : un test est de niveau α si, dans le cas où l'on fait l'échantillonnage dans une population pour laquelle la théorie est fausse, la probabilité que l'on déclare la théorie vraie à partir de l'échantillon est au plus α
  - **C'est le risque sur lequel on se dote de la contrainte la plus forte**, on sait que quoi qu'il arrive on ne sera jamais pire que α
- Remarque : en réalité, comme notre approche se fonde sur le comportement asymptotique de $\hat{\beta}$, il faut plutôt parler de test asymptotique de niveau α

Puissance d'un test

- Pour une population potentielle caractérisée par β, dans laquelle on fait l'échantillonnage, la **puissance d'un test** est la probabilité que le test renvoie `"La théorie est vraie"`
  - C'est une fonction qui dépend de β

- Pour une population potentielle caractérisée par β dans $\mathbf{B}_1$ (donc quand la théorie est effectivement vraie), un test de niveau α est plus puissant qu'un autre s'il renvoie plus souvent `"La théorie est vraie"`
  - Il peut exister des tests uniformément plus puissants = plus puissants que les autres pour tout $\mathbf{B}_1$ mais ce n'est pas toujours le cas

- Schématiquement, on cherche à considérer des tests aussi puissants que possibles pour un certain niveau α choisi à l'avance
  - Sans la contrainte sur le niveau, il suffirait de toujours renvoyer `"La théorie est vraie"` pour être le plus puissant

Test de Student pour la régression linéaire par les moindres carrés ordinaires

Les hypothèses que l'on veut confronter

- On se demande si un certain coefficient β_k est nul (hypothèse nulle, $\mathbf{B}_0 = \{0\}$) ou pas ($\mathbf{B}_1 = \mathbb{R} - \{0\}$)

Comment procéder ?

- On sait qu'asymptotiquement, β s'identifie à un **vecteur gaussien dont on sait estimer la matrice de variance-covariance**
- Avec un peu de travail, on peut montrer que $\frac{\hat{\beta_k}-\beta_k}{\hat{\sigma_k}}$ suit asymptotiquement une loi de Student à $n-d$ degrés de liberté
  - Quand $n \gg d$ on peut l'assimiler à une loi normale centrée réduite
- En particulier, sous l'hypothèse nulle $\beta_k = 0$, $T = \frac{\hat{\beta_k}}{\hat{\sigma_k}}$ suit asymptotiquement une **loi de Student** à $n-d$ degrés de liberté, assimilable à une loi normale centrée réduite
  - Généralisation facile à n'importe quelle hypothèse nulle de la forme $\beta_k = \beta_k^0$ en considérant $\frac{\hat{\beta_k}-\beta_k^0}{\hat{\sigma_k}}$

Comment procéder ?

- On connaît les **quantiles** de cette loi
  - On définit un test qui renvoie
    - 0 / `"La théorie est fausse"` si T est compris entre les quantiles d'ordre $\frac{\alpha}{2}$ et $1 - \frac{\alpha}{2}$ de cette loi
    - et 1 / `"La théorie est vraie"` sinon
- Sous l'hypothèse nulle, compte-tenu de ce qui précède, on a bien $\mathbb{P}_0(T = 1) = \alpha$
  - C'est donc un test de niveau α
- Remarque : ce n'est pas le seul test de la nullité de β_k de niveau α fondé sur la statistique de Student $\frac{\hat{\beta}_k}{\hat{\sigma_k}}$ possible !
  - Par exemple :
    - 0 / `"La théorie est fausse"` si T est inférieur au quantile d'ordre α
    - et 1 / `"La théorie est vraie"` sinon

Niveaux usuels

- Seuils usuels : 0.1, 0.05, 0.01...
  - Il n'y a **pas de véritable bonne raison** de se concentrer sur ces niveaux en dehors de l'habitude
  - Important : un test de niveau 5% se trompe une fois sur vingt sous l'hypothèse nulle !!
- Pour $\alpha = 0.05$ les deux quantiles sont approximativement -1.96 et 1.96
  - Quand on est dans le régime asymptotique où on assimile la loi de Student à la loi normale centrée réduite

Une petite vérification

```r
library(AER)

data("CPS1985")

reg <-
  lm(wage ~ education + experience,
     data = CPS1985)
```

Une petite vérification

```r
test_student <- coeftest(reg,
                         vcov. = vcovHC)
test_student
```

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept) -4.904482   1.267195 -3.8703 0.0001222 ***
education    0.925965   0.088813 10.4260 < 2.2e-16 ***
experience   0.105132   0.018117  5.8030 1.121e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Une petite vérification

```r
ecarts_types <-
  sqrt(diag(vcovHC(reg)))

student <-
  reg$coefficients / ecarts_types

student
```

```
(Intercept)   education  experience
  -3.870347   10.425977    5.802955
```

```r
all.equal(as.numeric(student),
          as.numeric(test_student[,"t value"]))
```

```
[1] TRUE
```

Seuil critique

- Quand on utilise une famille de tests que l'on peut indicer par $\alpha \in]0, 1[$, tels que le test indicé par α soit de taille α
  - Par exemple : tous les tests de Student de la nullité de β_k de taille α en faisant varier α
- Le **seuil critique / p-valeur** est la valeur α^* telle que
  - On rejette l'hypothèse nulle pour toutes les valeurs de α plus petite que α^*
  - On ne la rejette pas pour toutes les valeurs plus grandes

Seuil critique

- Dans le cas du test de Student avec la construction détaillée auparavant : il faut juste réinverser les quantiles !
  - Et donc regarder la **fonction de répartition**
- On a dit qu'un rejettait si on se trouvait en dehors de $\left[q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right]$
  - Donc le seuil critique est à $|T| = q_{1-\frac{\alpha^*}{2}}$
    - En utilisant $q_{\frac{\alpha}{2}} = q_{1-\frac{\alpha}{2}}$
  - Autrement dit $\alpha^* = 2\left\{1 - \Phi(|T|)\right\}$
    - Φ étant la fonction de répartition de la loi normale centrée réduite

Seuil critique

- C'est (en général) de cela qu'on parle quand on entend ou qu'on lit **"statistiquement significatif au seuil de"**
  - Ce que cela veut dire exactement : quand on considère le test de Student pour l'échantillon que l'on considère, on rejette l'hypothèse nulle $\beta_k = 0$

- Ce que l'on ne devrait **jamais** dire
  - "Le coefficient (ou pire : l'effet) est significatif à"
    - Le coefficient est un réel fixe dont la valeur inconnue dépend de la population
    - Le test et les p-valeurs sont des v.a. qui dépendent de l'échantillon
    - Et en particulier de la taille d'échantillon
    - Les propriétés du coefficient ne peuvent pas dépendre de l'échantillon !!

Une petite vérification

```
p_valeurs_approx_norm <-
  2 * (1 - pnorm(abs(student)))
p_valeurs_approx_norm
```

```
 (Intercept)    education   experience
1.086807e-04 0.000000e+00 6.515646e-09
```

```
p_valeurs_st <-
  2 * (1 - LaplacesDemon::pst(abs(student),
                              nu = nrow(CPS1985) - 3))
p_valeurs_st
```

```
[1] 1.222074e-04 0.000000e+00 1.121104e-08
```

```
all.equal(
  as.numeric(p_valeurs_st),
  as.numeric(test_student[,"Pr(>|t|)"])
)
```

```
[1] TRUE
```

Test de Student avec R

- Maintenant on comprend tout ce que renvoie `coeftest`

```r
coeftest(reg,
         vcov. = vcovHC)
```

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept) -4.904482   1.267195 -3.8703 0.0001222 ***
education    0.925965   0.088813 10.4260 < 2.2e-16 ***
experience   0.105132   0.018117  5.8030 1.121e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Tests statistiques et construction d’intervalles de confiance

Intervalle de confiance : définition

- Un **intervalle de confiance** de niveau $1-\alpha$, avec $\alpha \in]0,1[$ pour un coefficient β_k est la donnée de deux v.a. réelles $a(\mathbf{X},\mathbf{Y})$ et $b(\mathbf{X},\mathbf{Y})$ telles que
  - $a(\mathbf{X},\mathbf{Y}) < b(\mathbf{X},\mathbf{Y})$ pour toutes les valeurs possibles de $\mathbf{X}$ et $\mathbf{Y}$
  - $\mathbb{P}_{\beta_k}(\beta_k \in [a(\mathbf{X},\mathbf{Y}), b(\mathbf{X},\mathbf{Y})]) \geq 1-\alpha$ pour toutes les valeurs possibles de β_k
    - Autrement dit : pour chaque population potentielle caractérisée par β_k, la probabilité que β_k se trouve dans cet intervalle est au moins $1-\alpha$
- Remarque : en réalité ici on traitera des intervalles de confiance asymptotiques
- Généralisation : une **région de confiance** c'est la même chose mais avec un ensemble aléatoire qui est une partie de l'espace des paramètres

Construire un intervalle de confiance à partir d'un ensemble de tests

- Pour toutes les valeurs possibles de β_k^0, on a un test de Student de l'hypothèse nulle $\beta_k = \beta_k^0$ contre l'alternative $\beta_k \neq \beta_k^0$ de niveau α
  - Ce test a la forme $1 - \mathbb{1}\left\{\frac{\hat{\beta_k} - \beta_k^0}{\hat{\sigma_k}} \in \left[q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right]\right\}$
- Un intervalle de confiance de niveau $1 - \alpha$ est l'ensemble des β_k^0 tels qu'on ne rejette pas l'hypothèse nulle $\beta_k = \beta_k^0$ au niveau α

Pourquoi ça marche ?

- Pour tout β_k^0, sous l'hypothèse $\beta_k = \beta_k^0$, la probabilité que cet ensemble ne contienne pas β_k est la probabilité que l'on rejette l'hypothèse nulle
- Elle est donc au plus égale à α
- La probabilité que β_k appartienne à cet ensemble est donc au moins $1 - \alpha$

A quoi ça ressemble ?

- On se met à rejeter l'hypothèse quand $\left|\frac{\hat{\beta}_k - \beta_k^0}{\hat{\sigma_k}}\right| \geq q_{1-\frac{\alpha}{2}}$
  - Donc les bornes de l'intervalle de confiance ont la forme $\left[\hat{\beta}_k - \hat{\sigma}_k \, q_{1-\frac{\alpha}{2}}, \hat{\beta}_k + \hat{\sigma}_k \, q_{1-\frac{\alpha}{2}}\right]$
- Remarque : ce n'est pas le seul intervalle de confiance possible qui satisfait la définition générale d'être un intervalle de confiance de niveau $1-\alpha$!
  - C'est juste celui qui se déduit de cette série de tests
- Remarque : inversement, on peut déduire un ensemble de tests de niveau α d'une construction d'intervalle de confiance de niveau $1-\alpha$

En pratique avec R

```
coefci(reg,
       vcov. = vcovHC)
```

```
                 2.5 %     97.5 %
(Intercept) -7.3938120 -2.4151526
education    0.7514962  1.1004330
experience   0.0695420  0.1407212
```

Une petite vérification

```r
confint_approx_norm <-
  cbind(reg$coefficients - ecarts_types * qnorm(0.975),
        reg$coefficients + ecarts_types * qnorm(0.975))
confint_approx_norm
```

```
                   [,1]       [,2]
(Intercept) -7.38813805 -2.4208266
education    0.75189389  1.1000353
experience   0.06962312  0.1406401
```

Une petite vérification

```
confint_st <-
  cbind(reg$coefficients -
          ecarts_types *
          LaplacesDemon::qst(0.975,
                             nu = nrow(CPS1985) - 3),
        reg$coefficients +
          ecarts_types *
          LaplacesDemon::qst(0.975,
                             nu = nrow(CPS1985) - 3))
confint_st
```

```
                  [,1]       [,2]
(Intercept) -7.3938120 -2.4151526
education    0.7514962  1.1004330
experience   0.0695420  0.1407212
```

Une petite vérification

```r
all.equal(
  as.numeric(confint_st),
  as.numeric(coefci(reg,
         vcov. = vcovHC))
)
```

```
[1] TRUE
```

Que faut-il (ne pas) faire des tests statistiques ?

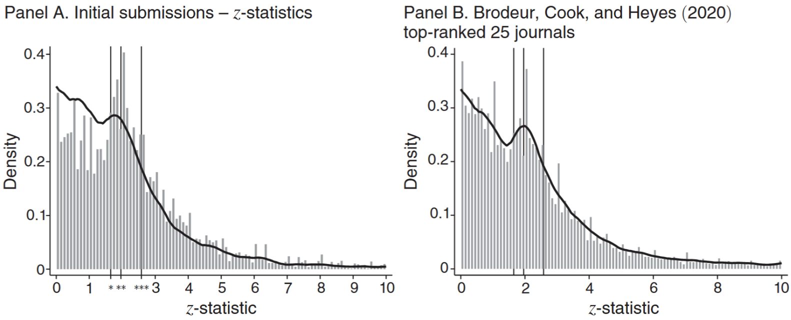
Les dangers des tests statistiques

- **Un test statistique ne devrait pas conduire à publier ou ne pas publier un résultat !**
  - Quand on ne rejette pas l'hypothèse nulle, c'est parfois tout bonnement... que la théorie que l'on voulait tester n'est pas vraie dans la population étudiée
    - Et d'autant plus qu'on regarde un test de niveau faible
  - Inversement, **rejeter l'hypothèse nulle a des chances de se produire régulièrement**
    - Pour un test à 5%, dans une population dans laquelle l'hypothèse nulle est vraie, on rejette l'hypothèse nulle une fois sur vingt !
    - Se concentrer sur les cas où l'on rejette peut conduire à donner beaucoup d'importance à des fluctuations statistiques attendues, et à perdre de l'information précieuse

La distribution empirique des statistiques de test fait apparaître des distorsions préoccupantes

- Brodeur, Abel, Scott Carrell, David Figlio, et Lester Lusher. 2023. "Unpacking P-hacking and Publication Bias." *American Economic Review* 113 (11) : 2974–3002.

La distribution empirique des statistiques de test fait apparaître des distortions préoccupantes



Y a-t-il quelque chose à apprendre de la significativité ?

- Jusque là on a fait comme si on regardait une population parmi toutes les populations possibles telles qu'elles sont caractérisées par les valeurs de β
- Mais **toutes les valeurs de β ne sont pas aussi plausibles** les unes que les autres !
  - Certaines sont *ex-ante* très improbables
  - Et d'autres semblent nettement plus plausibles au regard de l'information que l'on a déjà accumulée
    - Notamment en regardant la littérature antérieure

Y a-t-il quelque chose à apprendre de la significativité ?

- Pour formaliser cette idée, on peut se doter d'un **cadre bayésien** où l'on dote l'espace dans lequel vit β d'une mesure de probabilité
  - Représente le degré de crédibilité que l'on accorde *a priori* à chaque valeur possible de β
- Tout le travail que l'on a fait permet de savoir quantifier la probabilité que notre estimation à partir de l'échantillon renvoie une certaine valeur, sachant que la vraie valeur est fixée $\mathbb{P}(\hat{\beta} \mid \beta)$

Y a-t-il quelque chose à apprendre de la significativité ?

- **Formule de Bayes** : on peut obtenir la probabilité révisée *a posteriori* que β prenne une certaine valeur, sachant que l'on a observé une certaine valeur de $\hat{\beta}$: $\mathbb{P}(\beta \mid \hat{\beta}) = \frac{\mathbb{P}(\hat{\beta}|\beta)\mathbb{P}(\beta)}{\mathbb{P}(\hat{\beta})}$
- Schématiquement, on a acquis beaucoup d'information si on change beaucoup entre la probabilité *a priori* $\mathbb{P}(\beta)$ et la probabilité *a posteriori* $\mathbb{P}(\beta \mid \hat{\beta})$
- Au vu de ce qui précède : on gagne beaucoup d'information en observant une nouvelle valeur $\hat{\beta}$ si c'est une valeur *a priori* improbable $\rightarrow \mathbb{P}(\hat{\beta})$ faible

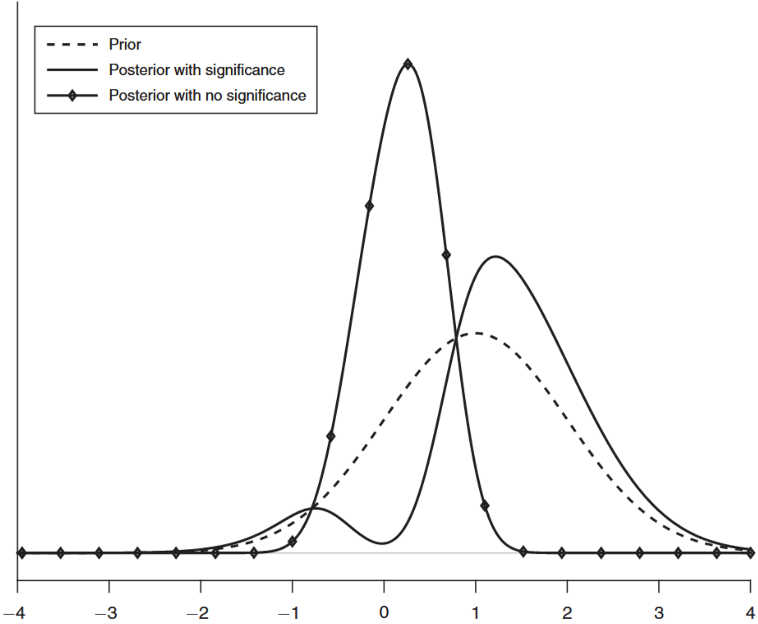
Y a-t-il quelque chose à apprendre de la significativité ?

- Abadie, Alberto. 2020. “Statistical Nonsignificance in Empirical Economics.” *American Economic Review : Insights* 2 (2) : 193–208.

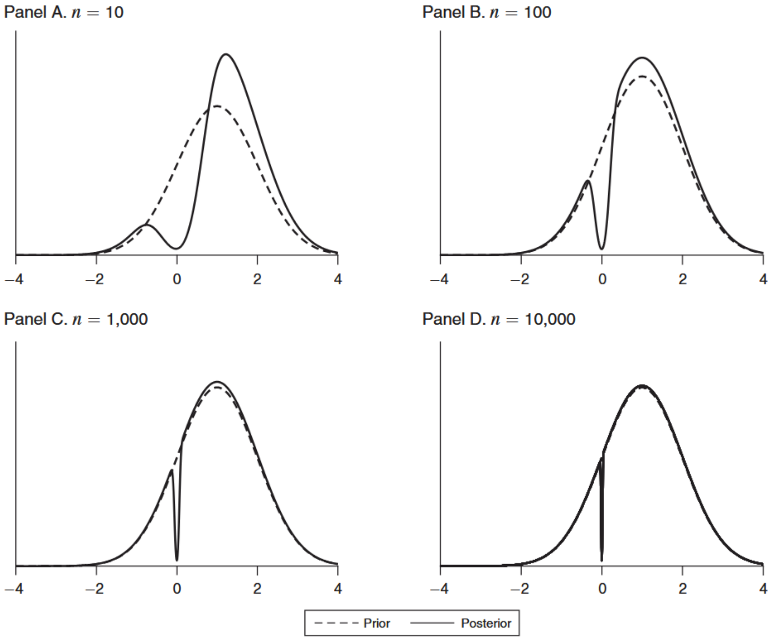
Y a-t-il quelque chose à apprendre de la significativité ?

- Adaptation de ce cadre pour répondre à la question : apprend-on vraiment plus de choses quand on apprend qu'un coefficient diffère statistiquement de 0 à un certain niveau, que s'il ne diffère pas ?
- La réponse est négative pour la plupart des applications en économie !
  - On peut rarement supposer *a priori* qu'un coefficient doit être très proche de 0
  - C'est pourtant dans ce cas que le rejet de l'hypothèse nulle est le plus intéressant...
  - Dans le cas général, on apprend juste que le coeffiicent ne se trouve vraisemblablement pas dans une petite région autour de 0
    - Et cette région est d'autant plus petite que l'on fait une estimation précise / que l'on travaille sur un gros échantillon

Y a-t-il quelque chose à apprendre de la significativité ?



Y a-t-il quelque chose à apprendre de la significativité ?



Que faut-il en tirer ?

- Est-ce à dire que les tests sont à jeter ? Non !
- C'est juste qu'**il y a beaucoup moins d'information dans le rejet / non-rejet d'un test de nullité que dans la valeur du coefficient et dans l'incertitude sur ce coefficient**
- Cette incertitude peut-être assez bien représentée par la construction d'intervalles de confiance
- Comme construire un intervalle de confiance c'est *in fine* la même chose que de construire un test sur les coefficients, il faut bien continuer à faire des tests...
- L'enjeu est surtout de **bien réfléchir à l'information que l'on reporte et sur laquelle on veut insister !**