ÉCONOMETRIE THEORIQUE

(M1 MBFA – Année 2023-2024)

A. *QUESTIONS A CHOIX MULTIPLES*

(Une seule réponse est correcte !)

Chapitre 1 : L'estimation par moindres carrés ordinaires (MCO)

1. L'élément ε dans un modèle économétrique est habituellement appelé :
    a) le terme d'erreur.
    b) un paramètre.
    c) une hypothèse.
    d) une variable dépendante.

2. L'élément β dans un modèle économétrique est habituellement appelé :
    a) le terme d'erreur.
    b) un paramètre.
    c) une hypothèse.
    d) une variable dépendante.

3. Les paramètres d'un modèle économétrique :
    a) comprennent tous les facteurs non observés affectant la variable étudiée.
    b) mesurent l'importance des variables explicatives du modèle.
    c) effectuent des prédictions de la variable dépendante du modèle.
    d) décrivent la force de la relation entre la variable étudiée et les facteurs qui l'affectent.

4. Les paramètres d'un modèle économétrique sont :
    a) aléatoires
    b) fixes.
    c) dépendants des erreurs.
    d) fonction des variables explicatives.

5. Le modèle de régression linéaire multiple
    a) comprend une seule variable explicative
    b) est linéaire dans les variables explicatives
    c) ne comprend pas de constante
    d) est linéaire dans les paramètres

6. L'erreur d'un modèle économétrique
    a) mesure l'effet observé des variables explicatives.
    b) comprend les erreurs de mesure sur les variables explicatives.
    c) incorpore les effets des variables explicatives inobservables sur la variable dépendante.
    d) mesure la variation des paramètres du modèle.

7. Dans le modèle de régression multiple : $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, β_0 est
    a) la variable dépendante
    b) la variable indépendante
    c) l'ordonnée à l'origine
    d) le paramètre de pente

8. Une variable dépendante est également connue sous le nom de ______.
    a) variable explicative
    b) variable de contrôle
    c) variable prédictive
    d) variable de réponse

9. On a une absence de multicolinéarité parfaite si :
    a) $rang(\boldsymbol{X}) = K$
    b) $rang(\boldsymbol{X}) = N$
    c) $N = K$
    d) $N > K$

10. On a une stricte exogénéité des variables explicatives si
    a) $E(\varepsilon_i) = 0$
    b) $E(\varepsilon_i|\boldsymbol{X}) = 0$
    c) $E(\boldsymbol{X}|\varepsilon_i) = 0$
    d) $E(\varepsilon_i|\boldsymbol{y}) = 0$

11. Le critère des moindres carrés
    a) minimise la somme des carrés des erreurs.
    b) maximise la somme des carrés des erreurs.
    c) minimise le carré de la somme des erreurs.
    d) minimise la somme des erreurs.

12. Dans l'estimation par moindres carrés, la matrice $\boldsymbol{X}'\boldsymbol{X}$:

   a) est singulière.
   b) est rectangulaire.
   c) est de rang-plein.
   d) est nulle.

13. La matrice $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ est

   a) inversible.
   b) orthogonale à la variable dépendante.
   c) orthogonale au terme d'erreur.
   d) symétrique et idempotente.

14. Le levier d'une observation mesure :

   a) la taille de l'observation.
   b) la variance de l'observation.
   c) l'influence d'une observation dans une régression.
   d) la force d'une observation.

15. La somme des leviers dans une régression est égale… :

   a) au nombre de variables explicatives.
   b) au nombre d'observations.
   c) au degré de liberté de la régression.
   d) à 1.

Chapitre 2 : Propriétés en petits échantillons de l'estimateur des MCO

16. L'estimateur des moindres carrés ordinaire de $\boldsymbol{\beta}$ est :
    a) un estimateur linéaire.
    b) un estimateur quadratique.
    c) une fonction convexe des erreurs.
    d) la vraie valeur de $\boldsymbol{\beta}$.

17. Sous les hypothèses H1 à H3, l'estimateur MCO est :
    a) de variance minimale.
    b) linéaire dans les variables explicatives.
    c) indépendant de l'erreur.
    d) sans biais.

18. L'estimateur des MCO $\widehat{\boldsymbol{\beta}}$ est sans biais
    a) sous les hypothèses H1 et H2.
    b) sous les hypothèses H1 à H3.
    c) sous les hypothèses H1 à H4.
    d) sous les hypothèses H1 à H5.

19. Sous les hypothèses H1 à H3, si on oublie une variable explicative pertinente dans une régression :
    a) $\widehat{\boldsymbol{\beta}}$ reste sans biais.
    b) $\widehat{\boldsymbol{\beta}}$ ne peut pas être estimé.
    c) le paramètre de cette variable est nul.
    d) $\widehat{\boldsymbol{\beta}}$ devient biaisé.

20. Sous les hypothèses H1 à H4, la variance conditionnelle de l'estimateur des MCO est :
    a) $V\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \sigma^2(\boldsymbol{X'X})^{-1}$
    b) $V\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \sigma^2(\boldsymbol{X'X})$
    c) $V\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \sigma^2(\boldsymbol{XX'})$
    d) $V\left(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right) = \sigma^2(\boldsymbol{XX'})^{-1}$

21. Sous les hypothèses H1 à H5, l'estimateur des MCO de la variance conditionnelle de $\boldsymbol{\varepsilon}$:
    a) égal à σ^2.
    b) suit une loi normale.
    c) suit une loi t de Student.
    d) suit une loi du Khi-deux.

22. Sous les hypothèses H1 à H4, la variance conditionnelle d'un paramètre estimé par MCO est d'autant plus grande :
    a) que la variance de l'erreur est faible.
    b) ne dépend pas de la variance de l'erreur.
    c) que le nombre d'observation est faible.
    d) que les degrés de liberté de la régression sont grands.

23. Sous les hypothèses H1 à H5, l'estimateur des MCO de $\boldsymbol{\beta}$:
    a) suit une loi normale.
    b) suit une loi t de Student.
    c) suit une loi du Khi-deux.
    d) suit une loi F de Fisher.

24. On a une erreur de type 1 ou de première espèce
    a) si on accepte l'hypothèse nulle alors qu'elle est fausse.
    b) si on rejette l'hypothèse nulle alors qu'elle est fausse.
    c) si on accepte l'hypothèse nulle alors qu'elle est vraie.
    d) si on rejette l'hypothèse nulle alors qu'elle est vraie.

25. On a une erreur de type 2 ou de seconde espèce
    a) si on accepte l'hypothèse nulle alors qu'elle est fausse.
    b) si on rejette l'hypothèse nulle alors qu'elle est fausse.
    c) si on accepte l'hypothèse nulle alors qu'elle est vraie.
    d) si on rejette l'hypothèse nulle alors qu'elle est vraie.

26. Le niveau d'un test:
    a) mesure le nombre d'observations pour effectuer le test.
    b) est la probabilité de rejeter H_0 alors que celle-ci est vraie
    c) est la probabilité d'accepterH_0 alors que celle-ci est vraie.
    d) est toujours égal à 5%.

27. La puissance d'un test:
    a) est la probabilité de rejeter H_0 alors que celle-ci est fausse.
    b) est la probabilité de rejeter H_0 alors que celle-ci est vraie.
    c) est la probabilité de l'erreur de $2^{ème}$ espèce.
    d) est toujours égal à 1 – le niveau du test.

28. Si la contrainte est correcte, et sous les hypothèses de Gauss-Markov, l'estimateur des moindres carrés contraint :
    a) est plus précis que l'estimateur des MCO.
    b) est identique à l'estimateur des MCO.
    c) a une variance plus grande que l'estimateur des MCO.
    d) est plus petit que l'estimateur des MCO.

29. La statistique t est le ratio
    a) du paramètre estimé à l'écart-type de la régression
    b) du paramètre estimé à son écart-type estimé.
    c) de l'écart-type estimé du paramètre à l'écart-type de la régression.
    d) de l'écart-type estimé du paramètre au paramètre estimé.

30. La statistique t permet
    a) de mesurer l'effet de la variable considérée.
    b) de donner l'écart-type du paramètre estimé.
    c) de tester la significativité de la variable considérée.
    d) de tester la multicolinéarité.

31. La statistique t se compare à
    a) suit une loi normale
    b) suit une loi t de Student
    c) suit une loi du Khi-deux
    d) suit une loi F de Fisher

32. Un intervalle de confiance à 95% :
    a) est plus grand qu'un intervalle de confiance à 90 %.
    b) est plus grand qu'un intervalle de confiance à 99 %.
    c) dépend de la valeur du vrai paramètre.
    d) est toujours plus grand que le paramètre estimé.

33. Un intervalle de confiance d'un paramètre :
    a) est centré sur le vrai paramètre.
    b) est symétrique autour du paramètre estimé.
    c) est normalement distribué.
    d) est distribué selon une loi t de Student.

34. Un paramètre estimé est dit « significatif » :
    a) si on accepte l'hypothèse nulle du test t.
    b) si la probabilité critique du test t est supérieur au niveau de test.
    c) si la statistique t de Student est inférieure au seuil critique.
    d) si on rejette l'hypothèse nulle du test t.

35. Le test F de significativité conjointe :
    a) est distribué selon une loi $F(N-K\,,K-1)$.
    b) dépend de la somme des carrés totaux de la variable dépendante.
    c) teste tous les paramètres de la régression.
    d) teste l'ensemble des paramètres de pente de la régression.

Chapitre 3 : Les propriétés asymptotiques de l'estimateur des MCO

36. La convergence d'un estimateur :
    a) implique une absence de biais.
    b) montre l'efficacité asymptotique.
    c) est une bonne propriété de l'estimateur.
    d) implique le théorème de Gauss-Markov.

37. Si $\Pr\left(\lim_{N\to\infty} Z_N = \alpha\right) = 1$, la séquence de variable aléatoire :
    a) converge en distribution.
    b) converge en moyenne quadratique.
    c) converge presque sûrement.
    d) converge en probabilité.

38. Soit une séquence de variables aléatoires Z_N qui converge en probabilité vers son espérance $E(Z_N) = \mu$, le carré de cette variable aléatoire :
    a) est d'espérance : $E\left({Z_N}^2\right) = \mu^2$
    b) converge en probabilité vers μ^2
    c) converge en probabilité vers zéro
    d) ne converge pas vers une valeur précise.

39. Un théorème central-limite établit :
    a) la convergence d'une séquence de variable aléatoire
    b) la convergence de la moyenne d'une variable aléatoire vers son espérance
    c) la distribution asymptotique de la moyenne d'une variable aléatoire
    d) la limite de la moyenne de la variable aléatoire

40. Un estimateur $\hat{\beta}$ est convergent si :
    a) $plim\ \hat{\beta} = 0$
    b) $plim\ \hat{\beta} = \beta$
    c) $\lim_{N\to\infty} \hat{\beta} = \beta$
    d) $plim\ \beta = \hat{\beta}$

41. Si le modèle est correct, si $plim\ (\boldsymbol{X}'\boldsymbol{X}/N) = Q > 0$ et $plim\ (\boldsymbol{X}'\boldsymbol{\varepsilon}/N) = 0$, alors l'estimateur des *MCO* :
    a) est sans biais.
    b) est de variance minimale.
    c) est normalement distribué.
    d) est convergent.

42. Un processus stochastique dont l'espérance et la variance de chaque variable aléatoire sont finies, et les covariances les variables aléatoires sont nulles, est un processus :
    a) bruit blanc.
    b) fortement stationnaire.
    c) faiblement stationnaire.
    d) de marche aléatoire

43. Un bruit blanc est un processus :
    a) fortement stationnaire.
    b) faiblement stationnaire.
    c) de marche aléatoire.
    d) muet.

44. Une marche aléatoire est :
    a) une séquence cumulée de constante.
    b) une séquence cumulée de variables aléatoires.
    c) une séquence cumulée de variables aléatoires *i.i.d*.
    d) une séquence cumulée de variables aléatoires normales.

45. Si $E(x_t \varepsilon_s) = 0, pour\ tout\ t\ et\ s$ alors la variable explicative x est :
    a) strictement stationnaire.
    b) faiblement stationnaire.
    c) strictement exogène.
    d) faiblement exogène.

46. Lequel des tests suivants aide à la détection de l'hétéroscédasticité ?
    a) Le test de Breusch – Pagan.
    b) Le test de Breusch – Godfrey.
    c) Le test de Durbin – Watson.
    d) Le test de Farrar - Glauber.

47. Le test de Breusch – Pagan est distribué, sous l'hypothèse nulle, selon
    a) une loi normale.
    b) une loi F de Fisher
    c) une loi du Khi-deux
    d) une loi t de Student

48. Le test de White est distribué, sous l'hypothèse nulle, selon
    a) une loi normale.
    b) une loi F de Fisher
    c) une loi du Khi-deux
    d) une loi W de White

49. Le test de Wooldridge repose sur le R^2 :
    a) de la régression.
    b) d’une régression auxiliaire sur les valeurs calculées de la variable dépendante.
    c) d’une régression auxiliaire sur les valeurs calculées et leurs carrés de la variable dépendante.
    d) d’une régression auxiliaire sur les variables explicatives, leurs carrés et leurs produits croisés uniques.

50. Qu'allez-vous conclure d'un modèle de régression si le test de Breusch – Pagan donne une petite probabilité critique ?
    a) Les erreurs sont homoscédastiques.
    b) Les erreurs sont hétéroscédastiques.
    c) Le modèle contient des variables indicatrices.
    d) Le modèle omet certains facteurs explicatifs importants.

Chapitre 4 : L'estimateur de la Méthode des Moments Généralisée (GMM)

51. On a un problème de convergence de l'estimateur MCO si :
    e) les erreurs sont autocorrélées.
    f) les erreurs sont indépendantes.
    g) la variable dépendante est mesurée avec erreurs.
    h) une variable explicative est mesurée avec erreurs.

52. L'omission d'une variable explicative pertinente corrélée avec les autres variables explicatives implique :
    a) une multicolinéarité.
    b) un biais dans l'estimateur *MCO*.
    c) une perte d'efficacité de l'estimateur *MCO*.
    d) une autocorrélation dans les erreurs.

53. S'il y a des erreurs de mesures non corrélées avec la variable explicative du modèle de régression simple, le paramètre estimé sera :
    a) plus grand que sa vraie valeur.
    b) plus petit que sa vraie valeur.
    c) plus proche de zéro que sa vraie valeur.
    d) plus proche de l'infini en valeur absolue que sa vraie valeur.

54. La méthode des variables instrumentales a été premièrement proposée par :
    a) Henri Theil.
    b) Denis Sargan.
    c) Philip Wright.
    d) Jerry Hausman.

55. Un instrument valide est :
    a) une variable corrélée avec l'erreur mais non corrélée avec la variable explicative.
    b) une variable non corrélée avec l'erreur et non corrélée avec la variable explicative.
    c) une variable corrélée avec l'erreur et corrélée avec la variable explicative.
    d) une variable non corrélée avec l'erreur et corrélée avec la variable explicative.

56. Si il y a moins d'instruments que de variables explicatives dans l'équation à estimer, on dit que l'équation est :
   a) juste identifiée.
   b) sous-identifiée.
   c) sur-identifiée.
   d) non indentifée.

57. Pour une matrice de pondération $\widehat{\boldsymbol{W}}$, l'estimateur de la méthode des moments généralisée est :
   a) $\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}}) = \left(\boldsymbol{S}'_{xz}\widehat{\boldsymbol{W}}\boldsymbol{S}_{xz}\right)^{-1}\boldsymbol{S}'_{xz}\widehat{\boldsymbol{W}}\boldsymbol{S}_{xy}$.
   b) $\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}}) = \boldsymbol{S}'_{xz}\widehat{\boldsymbol{W}}\boldsymbol{S}_{xy}$.
   c) $\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}}) = \left(\boldsymbol{S}'_{xz}\widehat{\boldsymbol{W}}^{-1}\boldsymbol{S}_{xz}\right)^{-1}\boldsymbol{S}'_{xz}\widehat{\boldsymbol{W}}^{-1}\boldsymbol{S}_{xy}$.
   d) $\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}}) = \left(\boldsymbol{S}_{xz}\widehat{\boldsymbol{W}}^{-1}\boldsymbol{S}'_{xz}\right)^{-1}\boldsymbol{S}_{xz}\widehat{\boldsymbol{W}}^{-1}\boldsymbol{S}_{xy}$.

58. L'estimateur des doubles moindres carrés revient à :
   a) régresser la variable dépendante sur les variables explicatives.
   b) régresser la variable explicative endogène sur les variables exogènes.
   c) régresser la variable dépendante sur les variables explicatives projetées sur les variables exogènes.
   d) régresser la variable dépendante sur les variables explicatives exogènes.

59. Sous les hypothèses $\text{plim}(\boldsymbol{X}'\boldsymbol{\varepsilon}/N) = 0$ et $\text{plim}(\boldsymbol{X}'\boldsymbol{Z}/N) = \boldsymbol{\Sigma}_{xz} \neq 0$, l'estimateur des doubles moindres carrés est :
   a) endogène.
   b) convergent.
   c) sans biais.
   d) non convergent.

60. Le choix optimal de la matrice de pondération $\widehat{\boldsymbol{W}}$ pour l'estimateur de la méthode des moments généralisée est :
   a) $\widehat{\boldsymbol{W}} = \boldsymbol{S}_{xx}$ avec $\boldsymbol{S}_{xx} = E(\boldsymbol{x}_i\boldsymbol{x}'_i)$.
   b) $\widehat{\boldsymbol{W}} = \boldsymbol{S}_{xx}^{-1}$ avec $\boldsymbol{S}_{xx} = E(\boldsymbol{x}_i\boldsymbol{x}'_i)$.
   c) $\widehat{\boldsymbol{W}} = \boldsymbol{S}^{-1}$ avec $\boldsymbol{S} = E(\boldsymbol{g}_i\boldsymbol{g}'_i)$.
   d) $\widehat{\boldsymbol{W}} = \boldsymbol{S}$ avec $\boldsymbol{S} = E(\boldsymbol{g}_i\boldsymbol{g}'_i)$.

61. Le test de Hansen permet de tester la validité des instruments :
   a) dans le cas juste identifié.
   b) dans le cas sur-identifié.
   c) dans le cas sous-identifié.
   d) dans tous les cas.

62. Sous l'hypothèse nulle de validité des instruments, le test de Hansen est distribué selon une loi :
   a) $\chi^2(N-L)$.
   b) $\chi^2(N-K)$.
   c) $\chi^2(K-L)$.
   d) $\chi^2(K)$.

63. Le test de Hausman permet de tester :
   e) l'hétéroscédasticité.
   f) la convergence des estimateurs.
   g) l'endogénéité des régresseurs.
   h) La multicolinéarité des régresseurs.

64. Dans le test de Hausman, la matrice de variance-covariance de la différence des deux estimateurs s'écrit :
   a) $V(\hat{\beta}_{MCO}-\hat{\beta}_{VI})=V(\hat{\beta}_{VI})-V(\hat{\beta}_{MCO})$.
   b) $V(\hat{\beta}_{MCO}-\hat{\beta}_{VI})=V(\hat{\beta}_{MCO})-V(\hat{\beta}_{VI})$.
   c) $V(\hat{\beta}_{MCO}-\hat{\beta}_{VI})=V(\hat{\beta}_{MCO})+V(\hat{\beta}_{VI})$.
   d) $V(\hat{\beta}_{MCO}-\hat{\beta}_{VI})=V(\hat{\beta}_{VI})+V(\hat{\beta}_{MCO})+Cov(\hat{\beta}_{MCO},\hat{\beta}_{VI})$.

65. Un instrument faible :
   a) entraîne une faible convergence de l'estimateur par variables instrumentales
   b) rend l'estimateur par variables instrumentales moins efficace.
   c) implique une faiblesse du biais de l'estimateur par variables instrumentales
   d) affaiblit la corrélation entre la variable dépendante et la variable explicative endogène.

Chapitre 5 : La méthode du maximum de vraisemblance (MV)

66. La fonction de log-vraisemblance d'un modèle de régression linéaire normal pour des observations indépendantes est :
    a) $\ell = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - \frac{N}{2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$
    b) $\ell = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - (\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/\sigma^2)$
    c) $\ell = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\sigma^2 - \log(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/(2\sigma^2))$
    d) $\ell = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - (\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/(2\sigma^2))$

67. Le score est :
    a) la valeur de la log-vraisemblance maximisée.
    b) la valeur de la vraisemblance maximisée.
    c) le vecteur du gradient de la log-vraisemblance.
    d) la matrice Hessienne de la log-vraisemblance.

68. La matrice d'information de *Fisher* est :
    a) le produit extérieur du vecteur du gradient.
    b) le produit scalaire du vecteur du gradient.
    c) l'espérance du produit extérieur du vecteur du gradient.
    d) l'information contenue dans les variables du modèle.

69. La borne inférieure de *Rao – Cramèr* établit :
    a) le minimum de la fonction de log-vraisemblance.
    b) le minimum de la variance asymptotique d'un estimateur convergent et asymptotique-ment normal.
    c) le minimum de l'estimateur convergent des paramètres.
    d) le minimum de la matrice d'information de *Fisher*

70. L'estimateur du maximum de vraisemblance n'est pas convergent si :
    a) les erreurs ne sont pas normalement distribuées.
    b) le modèle n'est pas linéaire dans les paramètres.
    c) la matrice d'information de Fisher n'atteint pas la borne inférieure de *Cramèr – Rao*.
    d) le modèle n'est pas correctement spécifié.

71. Le test du rapport des vraisemblances nécessite :
    a) l'estimation du modèle contraint uniquement.
    b) l'estimation du modèle non contraint uniquement.
    c) l'estimation des modèles contraint et non contraint.
    d) L'estimation d'aucun de ces modèles.

72. L'estimateur *MV* de σ^2 d'un modèle de régression linéaire normal est :

a) sans biais et convergent
b) sans biais et non convergent
c) biaisé et convergent
d) biaisé et non convergent

73. Dans le modèle *LOGIT*, le résidu *LOGIT* est :

a) nul.
b) orthogonal aux variables explicatives.
c) de variance minimale.
d) orthogonal aux paramètres estimés.

74. La log-vraisemblance d'un modèle *PROBIT* est :

a) localement convexe.
b) globalement convexe.
c) localement concave.
d) globalement concave.

75. Dans un modèle *LOGIT* ou *PROBIT*, les effets marginaux indiquent :

a) l'effet d'une variable explicative sur la probabilité de choix.
b) l'effet d'une variable explicative sur la log-vraisemblance.
c) l'effet d'une variable explicative sur l'estimateur du *MV*.
d) l'effet d'une variable explicative sur la variable dépendante.

B. QUESTIONS VRAI OU FAUX

Chapitre 1 : L'estimation par moindres carrés ordinaires (MCO)

1. Le critère des moindres carrés consiste à minimiser la somme des carrés des résidus.

Vrai Faux

2. L'estimation d'une constante est nécessaire pour que la moyenne des résidus soit nulle.

Vrai Faux

3. La covariance de l'échantillon entre un régresseur et les résidus des moindres carrés ordinaires (MCO) dépend du signe du paramètre estimé pour ce régresseur.

Vrai Faux

4. Les résidus sont distribués symétriquement autour de zéro.

Vrai Faux

5. Dans une régression multiple, il y a $N - K$ degrés de liberté dans les résidus de moindres carrés ordinaires.

Vrai Faux

6. Le R^2 est le rapport de la variation expliquée par rapport à la variation totale.

Vrai Faux

7. Le $\overline{R^2}$ est toujours supérieur au R^2.

Vrai Faux

8. Le $\overline{R^2}$ peut être négatif.

Vrai Faux

9. Le R^2 est le coefficient de corrélation entre la variable dépendante et la variable dépendante calculée sur le plan de régression.

Vrai Faux

10. La matrice de projection $\boldsymbol{M}$ est orthogonale à la matrice de projection $\boldsymbol{P}$.

Vrai Faux

Chapitre 2 : Propriétés en petits échantillons de l'estimateur des MCO

11. L'estimateur des moindres carrés ordinaire de $\boldsymbol{\beta}$ est toujours sans biais.

Vrai Faux

12. L'hypothèse d'homoscédasticité n'est pas nécessaire pour que $\hat{\boldsymbol{\beta}}$ soit sans biais

Vrai Faux

13. L'hypothèse de stricte exogénéité est suffisante pour que $\hat{\boldsymbol{\beta}}$ soit sans biais

Vrai Faux

14. L'omission d'une variable pertinente dans une régression entraîne toujours l'apparition d'un biais.

Vrai Faux

15. La variance inconditionnelle de $\hat{\boldsymbol{\beta}}$ est égale à sa variance conditionnelle :

Vrai Faux

16. Si une variable explicative est le produit de deux autres variables explicatives, il y a multicolinéarité parfaite.

Vrai Faux

17. Si la somme des variables explicatives est égale à zéro pour toutes les observations, il y a multicolinéarité parfaite.

Vrai Faux

18. On doit prendre le R^2 du modèle pour calculer le VIF.

Vrai Faux

19. Une très forte corrélation entre deux variables explicatives affecte la propriété d'absence de biais de l'estimateur des MCO.

Vrai Faux

20. L'hypothèse de normalité des erreurs est nécessaire pour que $\hat{\boldsymbol{\beta}}$ soit « BLUE »

Vrai Faux

21. L’hypothèse de normalité des erreurs implique la normalité de l’estimateur de la variance de l’erreur $\widehat{\sigma^2}$.

Vrai Faux

22. La loi t de Student est identique à la loi normale si les degrés de liberté sont supérieurs à 30 :

Vrai Faux

23. Plus la probabilité critique est grande, plus vraisemblable sera l’hypothèse nulle.

Vrai Faux

24. On peut calculer la statistique F de significativité globale à partir du R^2 de la régression.

Vrai Faux

25. On peut calculer la statistique F de significativité globale à partir du $\overline{R^2}$ de la régression.

Vrai Faux

26. Plus le R^2 de la régression est grand, plus la statistique F de significativité globale est grande.

Vrai Faux

27. La loi $F(1\ ,N-K)$ est équivalente au carré d’une loi $t(N-K)$

Vrai Faux

28. La loi $F(J\ ,N-K)$ est équivalente à la loi $F(N-K\ ,J)$.

Vrai Faux

29. Il suffit d’estimer le modèle contraint pour obtenir une statistique F pour tester les contraintes.

Vrai Faux

30. La statistique de test de Wald est identique à la statistique de test F lorsqu’il n’y a qu’une restriction à tester.

Vrai Faux

Chapitre 3 : Les propriétés asymptotiques de l'estimateur des MCO

31. Si une séquence de variables aléatoires converge en probabilité, alors elle converge en moyenne quadratique.

Vrai Faux

32. Si une variable aléatoire converge en distribution, alors elle converge en probabilité.

Vrai Faux

33. Toute fonction continue d'un processus stationnaire est aussi stationnaire.

Vrai Faux

34. Une réalisation est une observation particulière d'un processus stochastique.

Vrai Faux

35. Un processus stochastique stationnaire a une espérance et une variance constante.

Vrai Faux

36. On peut estimer de manière convergente les deux premiers moments d'un processus stochastique ergodique avec une seule réalisation suffisamment longue de ce processus.

Vrai Faux

37. La moyenne d'un processus stochastique stationnaire et ergodique est un estimateur convergent de l'espérance des variables aléatoires de ce processus.

Vrai Faux

38. Une marche aléatoire est une martingale.

Vrai Faux

39. Si le processus des variables dépendantes et explicatives est stationnaire et ergodique et si les régresseurs sont endogènes, l'estimateur des MCO est convergent.

Vrai Faux

40. L'estimateur des MCO est normalement distribué si le processus de la variable dépendante est une marche aléatoire.

Vrai Faux

Chapitre 4 : L'estimateur de la Méthode des Moments Généralisée (GMM)

41. Plus la variance de l'erreur de mesure est importante, plus le biais est important.

Vrai Faux

42. Plus le ratio signal / bruit est important, plus l'estimateur MCO sera proche de zéro.

Vrai Faux

43. Si $E(\boldsymbol{x}_t \varepsilon_t) \neq \boldsymbol{0}$ pour tout $t = 1,2, \dots; T$, alors l'estimateur des *MCO* n'est pas convergent.

Vrai Faux

44. On peut avoir certaines variables explicatives du modèle comme variables instrumentales.

Vrai Faux

45. Un instrument est une variable explicative du modèle.

Vrai Faux

46. Les variables explicatives du modèle ne peuvent pas être des variables instrumentales.

Vrai Faux

47. Si le nombre d'instruments est supérieurs au nombre de variables explicatives endogènes, on utilise une méthode des doubles moindres carrés.

Vrai Faux

48. L'estimateur des doubles moindres carrés est équivalent à l'estimateur des moindres carrés quasi-généralisés en deux étapes.

Vrai Faux

49. Dans le test de Hausman, les deux estimateurs que l'on compare sont convergent sous l'hypothèse nulle.

Vrai Faux

50. Un désavantage de l'estimateur des VI est d'avoir une moins bonne précision que celui des *MCO*.

Vrai Faux

Chapitre 5 : La méthode du maximum de vraisemblance (MV)

51. La maximisation de la fonction de vraisemblance ou de la fonction de log-vraisemblance donne les mêmes paramètres estimés.

Vrai Faux

52. La fonction de vraisemblance est toujours positive.

Vrai Faux

53. L'estimateur MV $\widetilde{\sigma^2}$ d'un modèle de régression linéaire normal est plus petit que l'estimateur des MCO $\widehat{\sigma^2}$.

Vrai Faux

54. L'estimateur MV $\widetilde{\sigma^2}$ d'un modèle de régression linéaire normal est biaisé et convergent.

Vrai Faux

55. Dans un modèle linéaire normal, les résidus de l'estimation par MCO sont identiques aux résidus de l'estimation par MV.

Vrai Faux

56. Dans un modèle linéaire de probabilité, la variance de l'erreur dépend des probabilités de chaque choix.

Vrai Faux

57. Le modèle Probit est plus précis que le modèle Logit pour estimer le choix entre deux alternatives.

Vrai Faux

58. Dans un modèle linéaire normal, les résidus de l'estimation par MCO sont identiques aux résidus de l'estimation par MV.

Vrai Faux

59. On estime les paramètres d'un modèle Logit à un facteur de proportionnalité près.

Vrai Faux

60. Le test du multiplicateur de Lagrange requiert l'estimation du modèle non contraint.

Vrai Faux

C. EXERCICES THEORIQUES

Introduction : Exercices de Calcul Matriciel

1. En supposant deux matrices carrées $\boldsymbol{A}$ et $\boldsymbol{B}$ de même dimension, démontrez la propriété de circularité de la trace : $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$.

2. Soit $\boldsymbol{A}$ une matrice de dimension $r \times c$, démontrez que la trace de $\boldsymbol{A'A}$ est la somme des carrés de tous les éléments de $\boldsymbol{A}$: $tr(\boldsymbol{A'A}) = \sum_{i=1}^{c}\sum_{j=1}^{r} a_{i,j}^2$.

3. Expliquez pourquoi chacune des propositions suivantes est vraie (on suppose que tous les produits sont conformes) :
   a) $(\boldsymbol{A} + \boldsymbol{A'})\boldsymbol{B} = \boldsymbol{AB} + \boldsymbol{A'B}$
   b) $tr(\boldsymbol{B'B}) = tr(\boldsymbol{BB'})$
   c) $\boldsymbol{x'B'Bx} = (\boldsymbol{Bx})'(\boldsymbol{Bx}) = tr(\boldsymbol{B'Bxx'})$

   Vérifiez chaque proposition avec : $\boldsymbol{A} = \begin{bmatrix} -1 & 2 \\ 7 & 3 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 6 & 8 & -1 & 0 \\ 2 & 3 & 1 & 4 \end{bmatrix}$ et $\boldsymbol{x}' = [1 \quad 1 \;\; 2 \quad 3]$

4. Prouvez que la matrice $\boldsymbol{A}$ est toujours symétrique dans une forme quadratique $\boldsymbol{x'Ax}$. Montrez que les formes quadratique $\boldsymbol{x'Ax}$ et $\boldsymbol{x'Bx}$ sont identiques pour :

$$\boldsymbol{A} = \begin{bmatrix} 2 & 2.5 & 6 \\ 2.5 & 7 & 3 \\ 6 & 3 & 9 \end{bmatrix} \quad \text{et} \quad \boldsymbol{B} = \begin{bmatrix} 2 & 3 & 8 \\ 2 & 7 & -5 \\ 4 & 11 & 9 \end{bmatrix}$$

5. Démontrez que pour les vecteurs 3×1 $\boldsymbol{x}$ et $\boldsymbol{y}$, et pour la matrice :

$$\boldsymbol{B} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix}$$

   a) $\boldsymbol{B'B} = \boldsymbol{BB'} = \boldsymbol{I}$. Comment appelle-ton alors la matrice $\boldsymbol{B}$ **?**
   b) $\boldsymbol{x'B'Bx} = (\boldsymbol{Bx})'\boldsymbol{Bx} = \boldsymbol{x}'(\boldsymbol{B'B})\boldsymbol{x} = x_1^2 + x_2^2 + x_3^2$
   c) $\boldsymbol{x'B'By} = x_1y_1 + x_2y_2 + x_3y_3$

6. Démontrez que les produits de deux matrices orthogonales sont othogonaux.

7. On a 3 variables aléatoires x_1, x_2 et x_3 qui ont les caractéristiques suivantes :
    - Leurs moyennes sont respectivement : 5 , 10 et 8.
    - Leurs variances sont respectivement : 6 , 14 et 1.
    - La covariance entre x_1 et x_2 est 3, la covariance entre x_1 et x_3 est 1, et la covariance entre x_2 et x_3 est 0.
    - On a 3 combinaisons linéaires de ces 3 variables aléatoires :

$$\begin{cases} y_1 = x_1 + 3x_2 - 2x_3 \\ y_2 = \quad 7x_1 - 4x_2 + x_3 \\ y_3 = -2x_1 - x_2 + 4x_3 \end{cases}$$

    a) Ecrivez le vecteur des moyennes et la matrice de variance-covariance du vecteur $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.
    b) Ecrivez la matrice $\boldsymbol{T}$ qui représente la transformation linéaire de $\boldsymbol{x}$ en $\boldsymbol{y}$.
    c) Utilisez cette matrice $\boldsymbol{T}$ pour dérivez le vecteur des moyennes et la matrice de variance-covariance du vecteur aléatoire $\boldsymbol{y}$.
    d) Calculez la matrice de corrélation de $\boldsymbol{x}$.

8. Dérivez le déterminant de $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Montrez que ce déterminant est égal au déterminant de $\boldsymbol{A'}$. Calculez par exemple le déterminant de $\boldsymbol{A} = \begin{bmatrix} 7 & 3 \\ 4 & 6 \end{bmatrix}$.

9. Dérivez le déterminant de $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ par la méthode directe et par celle des mineurs. Calculez par exemple le déterminant de $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$.

10. Quel est le déterminant
    a) de la transposée d'une matrice ?
    b) d'une matrice triangulaire ?
    c) de l'inverse d'une matrice ?
    d) du produit de deux matrices ?

e) du carré d'une matrice ?
f) d'une matrice orthogonale ?
g) d'une matrice idempotente ?

[voir démonstrations dans « *Matrix Algebra for Applied Economics* », par Shayle R. Searle et Lois Schertz Willett, John Wiley & Sons, 2001.]

11. Dérivez l'inverse de la matrice $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Calculez par exemple l'inverse de $\boldsymbol{A} = \begin{bmatrix} 7 & 3 \\ 4 & 6 \end{bmatrix}$. Vérifiez votre résultat.

12. Si on a deux matrices symétriques $\boldsymbol{A}$ et $\boldsymbol{B}$, démontrez que $[(\boldsymbol{AB})']^{-1} = \boldsymbol{A}^{-1}\boldsymbol{B}^{-1}$.

13. Démontrez que la transposée et l'inverse d'une matrice orthogonale sont identiques.

14. Quels sont les valeurs propres et les vecteurs propres de la matrice $\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 9 & 1 \end{bmatrix}$? Montrez que la trace et le déterminant de $\boldsymbol{A}$ sont respectivement la somme et le produit des valeurs propres ?

15. Démontrez que les valeurs propres d'un matrice idempotente sont soit 1, soit 0. Quelle est la matrice idempotente dont toutes les valeurs propres sont unitaires ? Quel est le déterminant et la trace d'une matrice idempotente ?

Chapitre 1 : L'estimation par moindres carrés ordinaires (MCO)

16. Prouvez que $\boldsymbol{X'X}$ est définie positive si $\boldsymbol{X}$ est de rang plein colonne.
    Astuce : on doit montrer que **c'**$\boldsymbol{X'Xc} > \mathbf{0}$ pour $\boldsymbol{c} \neq \mathbf{0}$.

17. Vérifiez que $\boldsymbol{X'X}/N = \frac{1}{N}\sum_i \boldsymbol{x}_i\boldsymbol{x}_i' = \boldsymbol{S_{xx}}$ et $\boldsymbol{X'y}/N = \frac{1}{N}\sum_i \boldsymbol{x}_i y_i = \boldsymbol{S_{xy}}$.

18. Dans une régression simple : $K = 2$ et $x_{i,1} = 1$: $y_i = \beta_1 + \beta_2 x_{i,2} + \varepsilon_i$, montrez que
$$\boldsymbol{S_{xx}} = \begin{bmatrix} 1 & \overline{x_2} \\ \overline{x_2} & \frac{1}{N}\sum_i x_{i,2}^2 \end{bmatrix} \quad \text{et} \quad \boldsymbol{S_{xy}} = \begin{bmatrix} \bar{y} \\ \frac{1}{N}\sum_i x_{i,2}y_i \end{bmatrix}$$
    L'estimateur MCO des paramètres sera alors :
$$\widehat{\beta_2} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)(y_i - \bar{y})}{\frac{1}{N}\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)^2} = \frac{Cov(x,y)}{V(x)} \quad \text{et} \quad \widehat{\beta_1} = \bar{y} - \widehat{\beta_2}\overline{x_2}$$
    Montrez également que !
$$\widehat{\beta_2} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)(y_i - \bar{y})}{\frac{1}{N}\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)^2} = \frac{\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)y_i}{\sum_{i=1}^{N}\left(x_{i,2} - \overline{x_2}\right)x_{i,2}}$$
    et
$$\widehat{\beta_2} = \frac{\sum_{i=1}^{N} x_{i,2}y_i - N\overline{x_2}\bar{y}}{\sum_{i=1}^{N} x_{i,2}^2 - N\overline{x_2}^2}$$

19. Pour les matrices de projections $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'}$ et $\boldsymbol{M} = \boldsymbol{I}_N - \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'}$, démontrez
    a) $\boldsymbol{P}$ et $\boldsymbol{M}$ sont symétriques et idempotentes
    b) $\boldsymbol{PX} = \boldsymbol{X}$
    c) $\boldsymbol{MX} = \mathbf{0}$

20. Montrez que
    a) $\hat{\boldsymbol{y}} = \boldsymbol{Py}$ et $\boldsymbol{e} = \boldsymbol{My} = \boldsymbol{M\varepsilon}$.
    b) $\boldsymbol{SCR} = \boldsymbol{e'e} = \boldsymbol{\varepsilon' M \varepsilon}$.

21. Dans une régression linéaire multiple de $\boldsymbol{y}$ sur $\boldsymbol{X}$ (contenant une constante), si les variables explicatives sont non-corrélées entre-elles, montrez que les paramètres de pente sont identiques aux paramètres de régressions simples (avec constante) de $\boldsymbol{y}$ sur chacun des $\boldsymbol{x}_k$ $(k = 2,3, \ldots, K)$ séparément.
(Théorème 3.3 dans Greene [2018])

22. Démontrez que la corrélation entre la valeur calculée $\hat{\boldsymbol{y}}$ et le résidus $\boldsymbol{e}$ dans une régression par MCO est nulle. Interprétez ce résultat.

23. On a une régression linéaire multiple de $\boldsymbol{y}$ sur $\boldsymbol{Z}$: $\boldsymbol{y} = \boldsymbol{Z\delta} + \boldsymbol{\varepsilon}$ où $\boldsymbol{Z} = \boldsymbol{XP}$ avec $\boldsymbol{P}$ une matrice non-singulière de dimension $K \times K$ qui transforme les variables $\boldsymbol{X}$.

    i) Montrez que l'estimateur des MCO sera : $\widehat{\boldsymbol{\delta}} = \boldsymbol{P}^{-1}\widehat{\boldsymbol{\beta}}$ où $\widehat{\boldsymbol{\beta}}$ est l'estimateur des MCO de la régression de $\boldsymbol{y}$ sur $\boldsymbol{X}$: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$.

    ii) Montrez que le R^2 est identique dans les deux régressions précédentes.

    iii) Interprétez les résultats que vous avez obtenus.

    (Théorème 3.8 dans Greene [2018])

24. Est-ce que le changement des unités de mesure de la variable dépendante change le coefficient de détermination R^2 d'une régression linéaire ? Pourquoi ?
De même, est-ce que le changement des unités de mesure d'un régresseur change le coefficient de détermination R^2 d'une régression linéaire ? Pourquoi ?

25. Etablissez la relation entre le R^2 et le R^2_{nc} :
$$R^2 = 1 - \left(1 + \frac{N\bar{y}^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}\right)(1 - R^2_{nc})$$

26. Montrez que
$$R^2_{nc} = \frac{y'Py}{y'y}$$

27. Comparer l'estimateur de β dans la régression : $y = \alpha + \beta x + \varepsilon$, la régression : $y = \beta(x - \bar{x}) + \varepsilon$ et la régression : $(y - \bar{y}) = \beta(x - \bar{x}) + \varepsilon$ (une barre au-dessus d'une variable indique la moyenne cette variable) ?

28. Dans une régression par MCO, démontrez que :
    a. le vecteur des résidus est orthogonal aux variables explicatives ?
    b. la moyenne des résidus est nulle s'il y a une constante dans le modèle ?
    c. la somme des carrés totaux est égale à la somme des carrés expliqués par la régression plus la somme des carrés des résidus : $SCT = SCE + SCR$?
    d. le vecteur des résidus est orthogonal au vecteur de la variable dépendante prédite ?
    e. la moyenne de la variable dépendante est égale à la moyenne de la variable dépendante prédite ?
    f. le R^2 est le carré du coefficient de corrélation entre la variable dépendante observée et la variable dépendante prédite ?

29. Démontrez le Théorème de Frisch – Waugh en utilisant les règles d'inversion sur les matrices partitionnées : soit une matrice carrée symétrique partitionnée en 4 blocs :

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{C}' & \boldsymbol{B} \end{pmatrix}$$

Si on suppose que $\boldsymbol{A}^{-1}$ existe, l'inverse de $\boldsymbol{D}$ sera :

$$\boldsymbol{D}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1}(\boldsymbol{I} + \boldsymbol{C}(\boldsymbol{B} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1}) & -\boldsymbol{A}^{-1}\boldsymbol{C}(\boldsymbol{B} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{C})^{-1} \\ -(\boldsymbol{B} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1} & (\boldsymbol{B} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{C})^{-1} \end{bmatrix}$$

30. Soit la matrice de projection $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ qui permet d'obtenir le résidu d'une régression par MCO, et soit la matrice de projection de la régression partielle : $\boldsymbol{M_1} = \boldsymbol{I} - \boldsymbol{X_1}(\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'$ avec $\boldsymbol{X_1} \subset \boldsymbol{X}$, un sous-ensemble de variables explicatives du modèle. Que vaut le produit de ces deux matrices : $\boldsymbol{M_1 M}$? Interprétez le résultat.

31. Démontrez que la somme des leviers d'une régression est égal au nombre de variables explicatives.

$$h_i = \boldsymbol{x}_i'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_i \quad \rightarrow \quad \sum_{i=1}^{N} h_i = K$$

32. Dans le numéro de Décembre 1969 de l'*American Economic Review*, Nathaniel Leff reporte dans son article « *Dependency Rates and Savings Rates* » (p.886-896) les régressions linéaires multiples suivantes basée sur une coupe instantanée de 74 pays en 1964 :

$$\ln\frac{S}{Y} = 7.3439 + 0.1596 \ln\frac{Y}{N} + 0.0254 \ln G - 1.3520 \ln D_1 - 0.3990 \ln D_2$$

$$\ln\frac{S}{N} = 2.7851 + 1.1486 \ln\frac{Y}{N} + 0.0265 \ln G - 1.3438 \ln D_1 - 0.3966 \ln D_2$$

avec S/Y : le ratio de l'épargne domestique au PIB, S/N : l'épargne par tête, Y/N : le PIB par tête, G : la croissance du revenu par tête, D_1 : le pourcentage de la population âgée de moins de 15 ans, et D_2 : le pourcentage de la population âgée de plus de 65 ans.

i) Selon vous, pourquoi n'a-t-il pas introduit dans les régressions, le log de la part de la population âgée de 15 à 65 ans ?

ii) Est-ce que ces régressions vous semblent correctes ? Pourquoi ?

(Voir l'article d'Arthur Goldberger : « *Dependency Rates and Savings Rates : Further Comment* », *American Economic Review*, Vol.63, Issue 1, March 1973, p.232-233)

Chapitre 2 : Propriétés en petits échantillons de l'estimateur des MCO

33. Soit deux variables aléatoires x et y avec :

$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad et \quad V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

   a) Montrez que ρ est la corrélation entre x et y ?
   b) Quelle est l'espérance et la variance de $3x$?
   c) Quelle est l'espérance et la variance de $x - y$?
   d) Quelle est la covariance et la corrélation entre $(x + y)$ et $(x - y)$?

34. Précisez la notion d'exogénéité stricte (forte) ? Soit un modèle :

$$\begin{cases} y_t = \alpha + \beta x_t + \varepsilon_t \\ x_t = \delta + \gamma y_{t-1} + \nu_t \end{cases} \quad t = 1,2, \dots, T.$$

avec ε et ν indépendant. Est-ce que x_t est strictement exogène dans la première équation ? Quelles sont les conséquences sur les propriétés de l'estimateur des MCO ?

35. Pourquoi utilise-t-on $N - K$ pour le dénominateur de l'estimateur des MCO de la variance : $\widehat{\sigma^2} = SCR/(N - K)$? Montrez que cet estimateur est sans biais ? Quelle est la distribution de cet estimateur sous l'hypothèse de normalité des erreurs ?

36. Démontrez que la trace de la matrice idempotente M est égale à $N - K$.

37. Soit un estimateur $\widetilde{\boldsymbol{\beta}}$ d'un vecteur de paramètre $\boldsymbol{\beta}$,

   a) Prouvez que : $Var(\widetilde{\boldsymbol{\beta}}) = E[Var(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})] + Var[E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})]$.

   b) Prouvez que $Var(\widetilde{\boldsymbol{\beta}}) \geq Var(\widehat{\boldsymbol{\beta}})$.

Indices : Par définition,

$$Var(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) = E\left[\left(\widetilde{\boldsymbol{\beta}} - E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})\right)\left(\widetilde{\boldsymbol{\beta}} - E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})\right)'\middle|\boldsymbol{X}\right]$$

$$Var[E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})] = E\left\{\left(E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) - E(\widetilde{\boldsymbol{\beta}})\right)\left(E(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) - E(\widetilde{\boldsymbol{\beta}})\right)'\right\}$$

Si $Var(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) \geq Var(\widehat{\boldsymbol{\beta}}|\boldsymbol{X})$, alors $E[Var(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X})] \geq E[Var(\widehat{\boldsymbol{\beta}}|\boldsymbol{X})]$

[Hayashi (2000), exercice 4, Section I.3]

38. Prouvez la propriété II.1.i : $Cov(\boldsymbol{e}, \widehat{\boldsymbol{\beta}}|\boldsymbol{X}) = \boldsymbol{0}$
[Hayashi (2000), exercice 6, Section I.3]

39. Démontrez que la variance d'un paramètre estimé par MCO peut s'écrire :

$$V(\widehat{\beta_k}|X) = \frac{\sigma^2}{SCT_k(1 - R^2_{(k)})}$$

avec $SCT_k = \sum_{i=1}^{N}(x_{k,i} - \overline{x_k})^2$ la somme des carrés totaux de la $k^{ème}$ variable explicative, et $R^2_{(k)}$: le R^2 de la régression de x_k sur toutes les autres variables explicatives.

40. Supposons que le vrai modèle de régression multiple soit le suivant :

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

Les hypothèses H1 à H5 sont satisfaites. Cependant on estime le modèle qui omet la variable x_4. Soient $\widetilde{\beta_1}, \widetilde{\beta_2}$, et $\widetilde{\beta_3}$ les estimateurs MCO de cette régression de y sur x_2 et x_3. Montrez que l'espérance de $\widetilde{\beta_2}$ (étant donné les valeurs des variables explicatives dans l'échantillon est :

$$E(\widetilde{\beta}_2) = \beta_2 + \beta_4 \frac{\sum_{i=1}^{N} u_{i2} x_{i4}}{\sum_{i=1}^{N} u_{i2}^2}$$

avec $u_{2,i}$ les résidus de la régression par MCO de x_2 sur x_3 ?
[Wooldridge (2009), exercice 3.10]

41. Considérez le modèle de régression simple : $y = \beta_1 + \beta_2 x + \varepsilon$. Sous les hypothèses de Gauss-Markov, les estimateurs traditionnels des MCO $\widehat{\beta_1}$ et $\widehat{\beta_2}$ sont des estimateurs sans biais des paramètres de la population. Soit $\widetilde{\beta_2}$ l'estimateur de lorsqu'il n'y pas de constante dans la régression.

a) Exprimer $E\left(\widetilde{\beta_2}\right)$ en fonction des x_i, β_1 et β_2. Vérifiez que $\widetilde{\beta_2}$ est un estimateur sans biais de β_2 lorsque la valeur de la constante au sein de la population (β_1) est nulle. Existe-t-il d'autres cas pour lesquels $\widetilde{\beta_2}$ est sans biais ?

b) Calculez la variance de $\widetilde{\beta_2}$.

c) Montrez que $V\left(\widetilde{\beta_2}\right) \leq V\left(\widehat{\beta_2}\right)$.

d) Expliquez le compromis qui existe entre le biais et la variance lorsqu'il s'agit de faire un choix entre $\widehat{\beta_2}$ et $\widetilde{\beta_2}$.

42. Supposons que les variables $y_i, x_{1,i},$ et $x_{2,i}$ satisfont les hypothèses des moindres carrés et que, pour toutes les observations $V\left(\varepsilon_i | x_{1,i}, x_{2,i}\right) = 4$ et $V\left(x_{1,i}\right) = 6$. L'échantillon aléatoire comprend 400 observations.

a) Supposons que x_1 et x_2 soient non corrélés. Calculez la variance de $\hat{\beta}_1$? Pour cela démontrez que :

$$V\left(\widehat{\beta_1}\right) = \frac{1}{N} \frac{1}{1 - \left(\rho_{x_1,x_2}\right)^2} \frac{{\sigma_\varepsilon}^2}{{\sigma_{x_1}}^2}$$

où ρ_{x_1,x_2} est la corrélation théorique entre x_1 et x_2, ${\sigma_{x_1}}^2$ est la variance théorique de x_1.

b) Supposons maintenant que la corrélation entre x_1 et x_2 soit égale à 0.5 : $\rho_{x_1,x_2} = 0.5$. Calculez la variance de $\widehat{\beta_1}$?

c) Commentez l'affirmation : « *si x_1 et x_2 sont corrélés, la variance de $\widehat{\beta_1}$ est plus grande que dans le cas contraire* » ?

[Stock et Watson (2011), exercice 3.8]

43. Considérons le modèle de régression sans constante :

$$y_i = \beta_1 x_{1,i} + \beta_2 x_{2,i} + \varepsilon_i, \qquad i = 1,2, \dots, N.$$

Répondez aux questions suivantes :

a) Précisez la fonction des moindres carrés, minimisée par les MCO ?
b) Calculez les conditions du premier ordre et donnez les équations normales ?
c) Supposons que $\sum_{i=1}^{N} x_{1,i}x_{2,i} = 0$, montrez qu'on peut alors écrire :

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{N} x_{1,i}y_i}{\sum_{i=1}^{N} {x_{1,i}}^2} \quad et \quad \widehat{\beta_2} = \frac{\sum_{i=1}^{N} x_{2,i}y_i}{\sum_{i=1}^{N} {x_{2,i}}^2}.$$

d) Supposons que $\sum_{i=1}^{N} x_{1,i}x_{2,i} \neq 0$, établissez une expression de $\widehat{\beta_1}$ en fonction des variables $y_i, x_{1,i},$ et $x_{2,i}$?
e) Supposons que le modèle comporte une constante, et s'écrit donc : $y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \varepsilon_i$. Démontrez que les estimateurs des moindres carrés satisfont : $\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\overline{x_1} - \widehat{\beta_2}\overline{x_2}$?
f) De la même manière que dans (e), supposons que le modèle comporte une constante et que $\sum_{i=1}^{N}\left(x_{1,i} - \overline{x_1}\right)\left(x_{2,i} - \overline{x_2}\right) = 0$. Montrez que

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{N}\left(x_{1,i} - \overline{x_1}\right)(y_i - \bar{y})}{\sum_{i=1}^{N}\left(x_{1,i} - \overline{x_1}\right)^2}$$

g) Que se passe-t-il si la variable x_2 est omise dans la régression ? Comparez les estimateurs de β_1 des MCO correspondants aux régressions avec et sans la variable x_2 ?

[Stock et Watson (2011), exercice 3.9]

44. Dans une régression simple, $\widehat{\beta_1}$ et $\widehat{\beta_2}$ sont les estimateurs de l'ordonnée à l'origine et de la pente de la droite de régression. Soit $\bar{\varepsilon}$ la moyenne des erreurs (et non des résidus !).

a) Montrez que $\widehat{\beta_2}$ peut s'écrire sous la forme $\widehat{\beta_2} = \beta_2 + \sum_{i=1}^{N} w_i\varepsilon_i$ où $w_i = d_i/SCT$ et $d_i = x_i - \bar{x}$.
b) En partant du point (i), montrez que $\sum_i w_i = 0$ et que $\widehat{\beta_2}$ et $\bar{\varepsilon}$ ne sont pas corrélés.
c) Montrez que $\widehat{\beta_1}$ peut s'écrire sous la forme : $\widehat{\beta_1} = \beta_1 + \bar{\varepsilon} - \left(\widehat{\beta_1} - \beta_1\right)\bar{x}$.
d) Utilisez les points (ii) et (iii) pour montrer que $V\left(\widehat{\beta_1}\right) = \frac{\sigma^2}{N} + \frac{\sigma^2\bar{x}^2}{SCT_x}$.
e) Simplifiez l'expression précédente pour aboutir à l'équation :

$$V\left(\widehat{\beta_1}\right) = \frac{\sigma^2}{N}\frac{\sum_i x_i^2}{\sum_i (x_i - \bar{x})^2}$$

45. Supposons que nous avons deux estimateurs $(\widehat{\theta_1} \text{ et } \widehat{\theta_2})$ d'un paramètre θ.Ceux-ci sont sans biais, et indépendants l'un de l'autre. Leur variance est respectivement v_1 et v_2. Quelle est la combinaison linéaire des deux estimateurs : $\hat{\theta} = c_1\widehat{\theta_1} + c_2\widehat{\theta_2}$ qui est l'estimateur sans biais de variance minimale de θ ?
[Greene (2018), exercice 4.1]

46. Supposons un modèle de régression linéaire classique, mais que la vraie valeur de la constante est zéro.
Comparez la variance de l'estimateur MCO des paramètres de pente calculé sans la constante avec celle de l'estimateur MCO des paramètres de pente calculé avec cette constante inutile.
[Greene (2018), exercice 4.3]

47. Considérons la régression multiple de $\boldsymbol{y}$ sur K variables explicatives $\boldsymbol{X}$ et une variable explicative additionnelle $\boldsymbol{z}$.
Sous les hypothèses H1 à H4 du modèle de régression linéaire classique, prouvez que la vraie variance de l'estimateur des MCO des paramètres de X est plus grande quand la variable z est incluse dans la régression que dans le cas où elle n'est incluse dans la régression. (on suppose que le coefficient de z n'est pas zéro)
[Greene (2018), exercice 4.8]

48. Un chercheur dispose de deux échantillons indépendants des variables (Y_i, X_i) qui désignent respectivement les salaires et le nombre d'années d'étude pour les hommes et pour les femmes.
Pour les hommes, (1er échantillon), considérons la régression :

$$Y_{M,i} = \beta_{1,M} + \beta_{2,M}X_{M,i} + \varepsilon_{M,i}$$

Pour les femmes, (2ème échantillon), considérons la régression :

$$Y_{F,i} = \beta_{1,F} + \beta_{2,F}X_{M,i} + \varepsilon_{F,i}$$

$s_{\widehat{\beta_{2,M}}}$ et $s_{\widehat{\beta_{2,F}}}$ représentent respectivement l'écart-type estimé du paramètre de la variable X dans les deux régressions. Montrez que l'écart-type estimé de $\widehat{\beta_{2,M}} - \widehat{\beta_{2,F}}$ s'écrit :

$$s_{(\widehat{\beta_{2,M}}-\widehat{\beta_{2,F}})} = \sqrt{\left(s_{\widehat{\beta_{2,M}}}\right)^2 + \left(s_{\widehat{\beta_{2,F}}}\right)^2}$$

[Stock et Watson (2011), exercice 2.15]

49. Si les contraintes sont correctes, démontrez l'estimateur des moindres carrés contraint est sans biais. Quelles sont les hypothèses que vous devez supposer ? Montrez que l'estimateur des MCO est aussi sans biais.
Dérivez la matrice de variance – covariance de l'estimateur des MCC. Montrez qu'elle est plus petite (au sens matriciel) que celle des MCO. Interprétez ce résultat.

50. Donnez une expression du biais de l'estimateur des MCC si la contrainte n'est pas correcte ?

51. On sait que la somme des carrés des résidus des MCC est supérieure à celle des MCO. Donnez la condition pour que l'estimateur de la variance de l'erreur pour les MCC (σ^{*2}) soit supérieure à l'estimateur de la variance de l'erreur des MCO ($\hat{\sigma}^2$).

52. Considérons le modèle de régression multiple avec 3 variables explicatives sous les hypothèses usuelles du modèle linéaire classique :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

Vous souhaitez tester l'hypothèse nulle $H_0: \beta_1 - 3\beta_2 = 1$.

a) Soit $\widehat{\beta_1}$ et $\widehat{\beta_2}$ les estimateurs des MCO des paramètres β_1 et β_2 respectivement. Ecrivez $Var\left(\widehat{\beta_1} - 3\widehat{\beta_2}\right)$ en fonction des variances et covariances des paramètres estimés $\widehat{\beta_1}$ et $\widehat{\beta_2}$. Quelle est l'expression de l'écart-type de $\widehat{\beta_1} - 3\widehat{\beta_2}$?
b) Ecrivez la statistique t permettant de tester $H_0: \beta_1 - 3\beta_2 = 1$.
c) On définit $\theta_1 = \beta_1 - 3\beta_2$. Réécrivez l'équation de départ de sorte à faire apparaître les paramètres $\beta_0, \theta_1, \beta_2$ et β_3 dans l'équation et ainsi obtenir directement l'estimateur de $\widehat{\theta_1}$.

d) On définit aussi $\theta_2 = \beta_3 + 2\beta_2$. Quels sont les variances de $\widehat{\theta_1}$ et $\widehat{\theta_2}$, et la covariance entre $\widehat{\theta_1}$ et $\widehat{\theta_2}$ en fonction des variances et covariances des $\widehat{\beta_J}$.

53. Pour un nombre d'observation (N) et un nombre de paramètres donnés (K), quel est le niveau du R^2 qui assure le rejet de l'hypothèse nulle que tous les paramètres de pente soient zéros ? Utilisez pour cela le test F de significativité globale d'une régression. Quel est ce niveau de R^2 pour $K = 5$ et $N = 20, 50, 100$ ou 1000 ?

Chapitre 3 : Les propriétés asymptotiques de l'estimateur des MCO

54. Supposons que $\sqrt{N}(\widehat{\theta_N} - \theta) \overset{d}{\rightarrow} N(0\,, \sigma^2)$. Prouvez que $\widehat{\theta_N} \overset{p}{\rightarrow} \theta$.
    [Hayashi, 2000, Section II.1, Exercice 4]

55. Soit $\{z_i\}$ une séquence de variables aléatoires indépendantes et identiquement distribuées avec $E(z_i) = \mu \neq 0$ et $V(z_i) = \sigma^2$. La moyenne d'échantillon est $\overline{z_N} = \frac{1}{N}\sum_{i=1}^{N} z_i$. Démontrez que

$$\sqrt{N}\left(\frac{1}{\overline{z_N}} - \frac{1}{\mu}\right) \overset{d}{\rightarrow} N\left(0, \frac{\sigma^2}{\mu^4}\right)$$

    [Hayashi, 2000, Section II.1, Exercice 5]

56. Soit un vecteur $K \times 1$ de variables aléatoires $\boldsymbol{z}_t$ stationnaires en covariance avec une matrice des autocovariances $\boldsymbol{\Gamma}_s = Cov(\boldsymbol{z}_t, \boldsymbol{z}_{t-s})$. Montrez que $\boldsymbol{\Gamma}_s = \boldsymbol{\Gamma}'_{-s}$
    [Hayashi, 2000, Section II.2, Exercice 1]

57. Supposons que le processus $\{x_t\}$ soit une martingale par rapport à $\{\boldsymbol{z}_t\}$. Montrez que $E(x_{t+s}|\boldsymbol{z}_{t-1}, \boldsymbol{z}_{t-1}, \dots \boldsymbol{z}_1) = x_{t-1}$ et que $E(x_{t+s+1} - x_{t+s}|\boldsymbol{z}_{t-1}, \boldsymbol{z}_{t-1}, \dots \boldsymbol{z}_1) = 0$ pour $s = 0,1, \dots$
    [Hayashi, 2000, Section II.2, Exercice 3]

58. Soit $\{x_i\}$ une séquence de nombres réels et $\{\varepsilon_i\}$ une séquence de variables aléatoires indépendantes et identiquement distribuées d'espérance nulle et de variance finie. Est-ce que la séquence $\{x_i\varepsilon_i\}$ est indépendante et identiquement distribuée ? A-t-elle des autocorrélations nulles ? Est-elle une martingale en différence ? Est-elle stationnaire ?
    [Hayashi, 2000, Section II.2, Exercice 4]

59. Définissez une marche aléatoire. Quelle est son espérance, sa variance et ses autocorrélations. Montrez que cette marche aléatoire n'est pas stationnaire.
[Hayashi, 2000, Section II.2, Exercice 5]

60. Supposons le modèle de régression sans variables explicatives : $y_i = \mu + \varepsilon_i$ avec $\varepsilon_i \approx i.i.d.N(0, \sigma^2)$. Prouvez que la moyenne d'échantillon $\hat{\mu} = \bar{y} = \frac{1}{N}\sum_i y_i$ est un estimateur convergent et asymptotiquement normal. Maintenant considérez l'estimateur alternatif : $\tilde{\mu} = \sum_i w_i y_i$ avec $w_i = i/\sum_i i$. Montrez que cet estimateur est convergent et calculez se variance asymptotique. Pour rappel : $\sum_i i = n(n+1)/2$ et $\sum_i i^2 = n(n+1)(2n+1)/6$.
[Greene (2018), exercice 4.8]

61. Soit $s_{\widehat{\beta_k}}$ l'écart-type estimé d'un paramètre estimé par MCO. Montrez que $s_{\widehat{\beta_k}} \xrightarrow{p} 0$
[Hayashi, 2000, Section II.4, Exercice 1]

62. Supposons par simplicité que $K = 1$ et que $\hat{\beta}$ est l'estimateur MCO de β. L'écart-type de $\hat{\beta}$ est $\sqrt{Avar(\hat{\beta})/N}$. On définit $\lambda = -\log \beta$ qui est estimé par $\hat{\lambda} = -\log \hat{\beta}$. Vérifiez que l'écart-type de$\hat{\lambda}$ est alors $(1/\hat{\beta})\sqrt{Avar(\hat{\beta})/N}$.
[Hayashi, 2000, Section II.4, Exercice 2]

63. Il n'y a pas une manière unique d'écrire les J hypothèses linéaires $\boldsymbol{R\beta} = \boldsymbol{r}$. On peut représenter les mêmes restrictions par $\widetilde{\boldsymbol{R}}\boldsymbol{\beta} = \tilde{\boldsymbol{r}}$ en prémultipliant $\boldsymbol{R}$ et $\boldsymbol{r}$ par une matrice $\boldsymbol{F}$ de dimension $J \times J$ non singulière : $\widetilde{\boldsymbol{R}} = \boldsymbol{FR}$ et $\tilde{\boldsymbol{r}} = \boldsymbol{Fr}$. Est-ce que les choix de $\boldsymbol{R}$ ou de $\widetilde{\boldsymbol{R}}$ modifient la valeur numérique du test de Wald, sa distribution asymptotique, sa distribution en échantillon fini ?
[Hayashi, 2000, Section II.4, Exercice 3]

64. On veut tester la relation non linéaire entre les paramètres d'un modèle : $\beta_1\beta_2 = 1$ avec le test de Wald. Montrez que la statistique de Wald de l'hypothèse nulle $H_0: \beta_1\beta_2 - 1 = 0$ est mathématiquement différente de la statistique de test de l'hypothèse nulle $H_0: \beta_1 - (1/\beta_2) = 0$. Est-ce que ces deux statistiques sont asymptotiquement équivalentes ?

65. On veut tester la relation non linéaire entre les paramètres d'un modèle : $\beta_1\beta_2 = \beta_3$ avec le test de Wald. Montrez que la statistique de Wald de l'hypothèse nulle $H_0: \beta_1\beta_2 - \beta_3 = 0$ est mathématiquement différente de la statistique de test de l'hypothèse nulle $H_0: \beta_1 - (\beta_3/\beta_2) = 0$. Est-ce que ces deux statistiques sont asymptotiquement équivalentes ?

66. Montrez que la vraie variance de l'estimateur des MCO est plus grande que la vraie variance de l'estimateur des MCG.

67. Parmi les propositions suivantes, quelles sont celles susceptibles d'être causées par la présence d'hétéroscédasticité ? Expliquez.
    a) Les estimateurs des MCO, $\widehat{\beta_k}$, ne sont pas convergents.
    b) La statistique *F* habituelle ne suit plus une distribution *F* de Fisher.
    c) Les estimateurs des MCO ne sont plus les meilleurs estimateurs linéaires sans biais.

68. Avec les hypothèses 2.1 à 2.5, la variance asymptotique de l'estimateur des MCO est $Avar(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Sigma}_{\boldsymbol{xx}}^{-1}\boldsymbol{S}\boldsymbol{\Sigma}_{\boldsymbol{xx}}^{-1}$ avec $\boldsymbol{S} = E(\varepsilon_i^2\boldsymbol{x}_i\boldsymbol{x}_i')$. Est-elle estimée de manière convergente par $\widehat{Avar(\widehat{\boldsymbol{\beta}})} = N\widehat{\sigma^2}\mathbf{S}_{\boldsymbol{xx}}^{-1}$ avec $\boldsymbol{S}_{\boldsymbol{xx}} = \left(\frac{1}{N}\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$ sans l'hypothèse (2.7) d'homoscédasticité conditionnelle ? Est-ce que la statistique de test $t_{\widehat{\beta_k}} = \widehat{\beta_k}/s_{\widehat{\beta_k}}$ est asymptotiquement distribué selon une loi normale standard.
    [Hayashi, 2000, Section II.6, Exercice 1]

69. Sans l'hypothèse (2.7) d'homoscédasticité conditionnelle, est-ce que la statistique de test de J contraintes $(SCR_C - SCR)/\widehat{\sigma^2}$ est asymptotiquement distribué selon une loi du Khi-deux avec J degré de liberté.
[Hayashi, 2000, Section II.6, Exercice 4]

70. Dans une régression avec une constante, considérez l'hypothèse nulle que tous les $K-1$ paramètres de pente soient nuls. Montrez que $NR^2 \overset{d}{\rightarrow} \chi^2(K-1)$ sous les hypothèses 2.1 à 2.5 et l'hypothèse (2.7) d'homoscédasticité conditionnelle. Pourquoi ne peut-on pas utiliser cette expression lorsqu'il y a de l'hétéroscédasticité conditionnelle.
[Hayashi, 2000, Section II.6, Exercice 5]

71. Quelle est la matrice de covariance, $Cov\left(\widehat{\boldsymbol{\beta}_{MCG}}, \widehat{\boldsymbol{\beta}_{MCG}} - \widehat{\boldsymbol{\beta}_{MCO}}\right)$, entre l'estimateur des *MCG* : $\widehat{\boldsymbol{\beta}_{MCG}} = (\boldsymbol{X}'\boldsymbol{\Psi}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Psi}^{-1}\boldsymbol{y}$ et sa différence avec l'estimateur des *MCO* : $\widehat{\boldsymbol{\beta}_{MCO}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ **.**
[Greene (2018), exercice 9.1]

72. *[Théorème de Kruskal]* Dans un modèle de régression par *MCG*, si on peut écrire que les *K* colonnes de $\boldsymbol{X}$ sont les vecteurs caractéristiques (*eigenvector*) de $\boldsymbol{\Omega} = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$, alors les estimateurs des *MCG* et des *MCO* sont identiques.
[Greene (2018), exercice 9.5]

73. Dans un modèle de régression par *MCG* avec la matrice $\boldsymbol{\Omega} = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$ connue,
    a) Quelle est la matrice de covariance entre les estimateurs *MCG* ($\widehat{\boldsymbol{\beta}_{MCG}}$) et *MCO* ($\widehat{\boldsymbol{\beta}_{MCO}}$) ?
    b) Quelle est la matrice de variance-covariance des résidus des *MCO* ($\boldsymbol{e}_{MCO} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}_{MCO}}$) ?
    c) Quelle est la matrice de variance-covariance des résidus des *MCG* ($\boldsymbol{e}_{MCG} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}_{MCG}}$)
    d) Quelle est la matrice de covariance entre les résidus des *MCG* ($\boldsymbol{e}_{MCG}$) et les résidus des *MCO* ($\boldsymbol{e}_{MCO}$) ?

[Greene (2018), exercice 9.6]

74. Considérons le modèle linéaire suivant pour expliquer la consommation mensuelle de bière :

$$bière = \beta_1 + \beta_2 revenu + \beta_3 prix + \beta_4 etude + \beta_5 femme + \varepsilon$$

$$\text{avec : } \begin{cases} E(u|revenu, prix, etude, femme) = 0 \\ V(u|revenu, prix, etude, femme) = \sigma^2 revenu^2 \end{cases}$$

Ecrivez le modèle transformé avec des erreurs homoscédastiques. Interprétez ce modèle.

[Wooldridge (2009), exercice 8.2]

75. Considérons un modèle appliqué à un ensemble d'employés :

$$y_{i,e} = \beta_0 + \beta_1 x_{1,i,e} + \beta_2 x_{2,i,e} e + \cdots + \beta_4 x_{k,i,e} + v_{i,e} + f_i$$

où la variable non observée f_i capture un « effet d'entreprise », c'est-à-dire l'effet que les caractéristiques propres de l'entreprise *i* peuvent avoir sur la variable dépendante. Le terme d'erreur $v_{i,e}$ est spécifique à l'employé *e* de l'entreprise *i*. L'erreur composite $u_{i,e} = v_{i,e} + f_i$ est celle de l'équation (8.28).

a) Supposons que $Var(f_i) = \sigma_f^2$, $Var(v_{i,e}) = \sigma_v^2$ et que $v_{i,e}$ et f_i ne sont pas corrélés. Montrez alors que $Var(u_{i,e}) = \sigma_f^2 + \sigma_v^2$. Appelez cette variance σ^2.

b) Supposons maintenant que $v_{i,e}$ et $v_{i,g}$ ne sont pas corrélés pour $e \neq g$. Montrez que $v(u_{i,e}\ ,\ u_{i,g}) = \sigma_f^2$.

c) Soit la moyenne des erreurs composites au sein d'une entreprise : $\overline{u_\iota} = m_i^{-1} \sum_{e=1}^{m_i} u_{i,e}$ avec m_i le nombre d'employés d'une entreprise *i*. Montrez que $Var(\overline{u_\iota}) = \sigma_f^2 + \sigma_v^2/m_i$.

d) Discutez l'intérêt du point (iii) pour l'application de la méthode des MCP à des données agrégées au niveau de l'entreprise, où le poids utilisé pour l'observation *i* est fonction de la taille de l'entreprise.

[voir Wooldridge (2009), exercice 8.7]

76. Démontrez que la statistique de Lung-Box est supérieure à la statistique de Box-Pierce. Montrez que ces deux statistiques de test sont asymptotiquement équivalentes. Qu'en concluez-vous ?

77. Supposons un processus stochastique $\{y_t\}$ qui est généré par le modèle : $y_t = z + \varepsilon_t$ pour $t = 1,2, \ldots$ avec $\{\varepsilon_t\}$ une séquence *i.i.d.* d'espérance nulle et de variance σ_ε^2. La variable aléatoire z est invariante dans le temps avec une espérance nulle et une variance σ_z^2. On suppose que chaque ε_t n'est pas corrélé avec z.

a) Trouvez l'espérance et la variance de y_t. Est-ce qu'elles dépendent du temps ?
b) Donnez la covariance : $Cov(y_t, y_{t+h})$ pour tout t et h. Est-ce que le processus $\{y_t\}$ est stationnaire en covariance ?
c) Montrez que les autocorrélations sont : $Corr(y_t, y_{t+h}) = \sigma_z^2/(\sigma_z^2 + \sigma_\varepsilon^2)$ pour tout t et h.
d) Est-ce que y_t est asymptotiquement non corrélé ? Expliquez.

[voir Wooldridge (2009), exercice 11.3]

78. Considérez le modèle : $y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + u_t$ avec $u_t = \rho u_{t-1} + \varepsilon_t$ où ε_t est un bruit blanc gaussien. Comparez les autocorrélations de u_t dans le modèle original, avec celles de v_t dans le modèle en différence première : $(y_t - y_{t-1}) = (\boldsymbol{x}_t - \boldsymbol{x}_{t-1})'\boldsymbol{\beta} + (u_t - u_{t-1})$ ou $\Delta y_t = \Delta \boldsymbol{x}_t{}'\boldsymbol{\beta} + v_t$ avec $v_t = u_t - u_{t-1}$.

Est-ce que la transformation en différence première réduit l'autocorrélation ?

[voir Greene (2018), exercice 20.1]

Chapitre 4 : L'estimateur de la Méthode des Moments Généralisée (GMM)

79. Soit un modèle d'offre – demande classique de la forme :

$$\begin{cases} q_i = \alpha_0 + \alpha_1 p_i + u_i & \rightarrow \text{ équation de demande} \\ q_i = \beta_0 + \beta_1 p_i + v_i & \rightarrow \text{ équation d'offre} \end{cases}$$

a) Supposons que la covariance entre les deux erreurs n'est pas nulle : $Cov(u_i, v_i) \neq 0$. est-ce que le prix et le choc de demande (u_i) sont positivement corrélés quand $\alpha_1 < 0$ et $\beta_1 > 0$?

b) On a vu que l'estimateur des MCO du paramètres du prix dans une régression de la quantité sur une constante et la prix est biaisé pour α_1. Est-ce que le paramètre estimé de la constante est aussi biaisé pour α_0 ?

c) Démontrez que l'estimateur du paramètre du prix dans une régression de la quantité sur une constante et le prix ($q_i = \gamma_0 + \gamma_1 p_i + \eta_i$) a pour limite en probabilité :

$$\text{plim } \widehat{\gamma_1} = \frac{\alpha_1 V(v_i) + \beta_1 V(u_i)}{V(v_i) + V(u_i)}$$

[Hayashi (2000), Exercices 3.1.(1), (2) et (3)]

80. Sous quelles hypothèses l'estimateur par variables instrumentales est convergent ? Démontrez la convergence de cet estimateur. Donnez la distribution asymptotique de l'estimateur VI.

81. Montrez que la régression par MCO est un cas particulier de l'estimation par variables instrumentales.

82. Soit $\boldsymbol{\Sigma}_{\boldsymbol{xz}} = E(\boldsymbol{x}_i \boldsymbol{z}_i')$ de dimension $K \times L$ et $\boldsymbol{\Sigma}_{\boldsymbol{xy}} = E(\boldsymbol{x}_i y_i)$ de dimension $K \times 1$. Montrez que :

$$\text{rang}(\boldsymbol{\Sigma}_{\boldsymbol{xz}}) = \text{rang}\big(\boldsymbol{\Sigma}_{\boldsymbol{xz}} \ \vdots \ \boldsymbol{\Sigma}_{\boldsymbol{xy}}\big)$$

[Hayashi (2000), Exercices 3.3(5)]

83. Si on ajoute une autre variable instrumentale (ξ_i) aux instruments ($\boldsymbol{x}_i$). Bien que cette variable supplémentaire soit prédéterminée, elle n'est pas reliée aux régresseurs : $E\left(\xi_i z_{i,l}\right) = 0$ pour tout $l = 1,2, \dots, L$. Est-ce que la condition de rang est toujours satisfaite ?
[Hayashi (2000), Exercices 3.3(6)]

84. Soit la matrice $\boldsymbol{A}$ de dimension $q \times K$ de rang plein ligne avec $q \leq K$ telle que $\boldsymbol{A\Sigma_{xz}}$ est de rang plein colonne. Notons $\widehat{\boldsymbol{x}}_i = \boldsymbol{A}\boldsymbol{x}_i$. Vérifiez que les hypothèses 3.3, 3.4 et 3.5 tiennent pour les instruments transformés $\widehat{\boldsymbol{x}}_i$ si elles sont satisfaites pour les variables instrumentales $\boldsymbol{x}_i$.
[Hayashi (2000), Exercices 3.3(8)]

85. Si l'équation est juste identifiée, quelle est la valeur minimisée de $J\left(\widetilde{\boldsymbol{\delta}}, \widehat{\boldsymbol{W}}\right)$?
[Hayashi (2000), Exercices 3.4(2)]

86. Soit un modèle de régression simple $y_t = \alpha + \beta z_t + \varepsilon_t$ avec $\varepsilon_t \approx i.i.d. (0, \sigma_\varepsilon^2)$. On suppose que $E(z_t \varepsilon_t) \neq 0$, mais qu'on a une variable instrumentale x_t corrélée avec z_t telle que $E(x_t \varepsilon_t) = 0$.
Montrez que l'estimateur par variable instrumentale de β est équivalent à l'estimateur par *MCO* de β dans modèle : $y_t = \alpha + \beta z_t + \gamma v_t + \varepsilon_t$ où v_t est le résidus de la régression de première étape : $z_t = \delta + \theta x_t + \nu_t$.
[Proposition de Durbin (1954) et Wu (1973)]

87. Supposons que l'on recueille par une enquête des données sur la consommation de vin des ménages, ainsi que leurs revenus. On veut estimer l'élasticité de la consommation de vin au revenu. Mais ces deux variables sont mal renseignées, ou avec des erreurs.

    a) Quelle est la conséquence des erreurs de mesure sur la variable dépendante (consommation de vin) et sur la variable explicative (le revenu) ?
    b) Un économètre propose d'utiliser la variable instrumentale : montant total des chèques émis par le ménage. Pensez-vous que cette variable est un bon instrument ? Pourquoi ?

88. Dans un modèle de détermination des salaires, on explique souvent le log ($salaire$) par l'éducation mesurée par le nombre d'années d'étude ($etude$) et par d'autres variables explicatives $\boldsymbol{X}$. Soit le modèle :

$$\log(salaire) = \beta_1 + \beta_2 etude + \boldsymbol{X\gamma} + \varepsilon$$

Or le nombre d'année d'étude ne mesure qu'imparfaitement la qualité de la formation suivie par la personne. Quel est alors la conséquence de prendre la variable $etude$ dans cette régression au lieu de la qualité de la formation ?

89. Dans un modèle où la variable explicative est endogène, l'estimateur des MCO est biaisé et non convergent : plim $\widehat{\boldsymbol{\beta}_{MCO}} = \boldsymbol{\beta} + \boldsymbol{Q}^{-1}\boldsymbol{\gamma} = \boldsymbol{\theta}$ avec $\boldsymbol{Q} = \text{plim}(\boldsymbol{X}'\boldsymbol{X}/N)$ et $\boldsymbol{\gamma} = \text{plim}(\boldsymbol{X}'\boldsymbol{\varepsilon}/N)$.

   a) Dérivez la matrice de variance-covariance asymptotique de l'estimateur $\widehat{\boldsymbol{\beta}_{MCO}}$.

   b) Montrez que l'estimateur $\widehat{\boldsymbol{\beta}_{MCO}}$ est asymptotiquement normalement distribué.

[Greene (2018) : Exercice 8.1]

90. Supposons deux choix de matrice de pondération $\widehat{\boldsymbol{W}_1}$ et $\widehat{\boldsymbol{W}_2}$ tel que $\widehat{\boldsymbol{W}_1} - \widehat{\boldsymbol{W}_2} \overset{p}{\to} \mathbf{0}$. Démontrez que :

$$\sqrt{N}\,\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}_1}) - \sqrt{N}\,\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{W}_2}) \overset{p}{\to} \mathbf{0}$$

[Hayashi (2000), Exercices 3.5(3)]

91. Démontrez que :

$$J\left(\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{S}}^{-1}), \widehat{\boldsymbol{S}}^{-1}\right) = N\boldsymbol{s}_{xy}'\widehat{\boldsymbol{S}}^{-1}\left(\boldsymbol{s}_{xy} - \boldsymbol{S}_{xz}\widehat{\boldsymbol{\delta}}(\widehat{\boldsymbol{S}}^{-1})\right)$$

[Hayashi (2000), Exercices 3.6(2)]

92. En cas d'homoscédasticité conditionnelle, la statistique de test de Sargan est :

$$Sargan = N\frac{\left(\boldsymbol{s}_{xy} - \boldsymbol{S}_{xz}\widehat{\boldsymbol{\delta}_{DMC}}\right)'\boldsymbol{S}_{xx}^{-1}\left(\boldsymbol{s}_{xy} - \boldsymbol{S}_{xz}\widehat{\boldsymbol{\delta}_{DMC}}\right)}{\widehat{\sigma^2}}$$

Démontrez que cette statistique est égale à $N \times R_{nc}^2$, le coefficient de détermination non centré d'une régression de $\boldsymbol{e}_{DMC} = \boldsymbol{y} - \boldsymbol{Z}\widehat{\boldsymbol{\delta}_{DMC}}$ sur $\boldsymbol{X}$.
[Hayashi (2000), Exercices 3.8(6)]

93. Considérez le modèle de régression : $y_i = \alpha + \beta x_i^* + \varepsilon_i$. Mais la variable x_i^* n'est pas observée, on observe seulement une variable avec erreurs : $x_i = x_i^* + \nu_i$. Cependant on dispose d'une indicatrice permettant de classer les observations : $d_i = 1$ ou 0 .

   a) Montrez que l'estimateur par variables instrumentales utilisant cette indicatrice comme variable instrumentale est :

$$\widehat{\beta_{VI}} = \frac{\overline{y_{d=1}} - \overline{y_{d=0}}}{\overline{x_{d=1}} - \overline{x_{d=0}}}$$

   où $\overline{y_{d=1}}$ est la moyenne de y pour les observations avec $d_i = 1$, et ainsi de suite. Interprétez cette expression.

   b) Quelle est la condition pour laquelle cet estimateur est calculable, interprétez cette condition ?

   c) Quel est l'expression pour l'estimateur $\widehat{\alpha_{VI}}$?

[voir Greene (2018) : Exemple 8.2 et Exercice 8.5, ou Wooldridge(2009) : Exercice 15.3]

94. Considérez le modèle de régression simple sur séries temporelles où la variable explicative est sujette à des erreurs de mesure :

$$y_t = \alpha + \beta x_t^* + \varepsilon_t$$
$$x_t = x_t^* + \nu_t$$

   a) Introduisez $x_t^* = x_t - \nu_t$ dans le modèle et montrez que le terme d'erreur de la nouvelle équation (u_t) est négativement corrélé avec x_t si $\beta > 0$. Quelle est l'implication pour l'estimateur des MCO de β dans la régression de y sur x ?

   b) En plus des précédentes hypothèses, on suppose que ε_t et ν_t ne sont pas corrélés avec toutes les valeurs passées de x_t^* et de ν_t, et en particulier avec x_{t-1}^* et ν_{t-1}. Montrez alors que $E(x_{t-1}u_t) = 0$.

   c) Est-ce que x_t et x_{t-1} sont probablement corrélés ? Expliquez.

   d) Donnez une stratégie pour estimer de manière convergente α et β dans ce modèle.

[Wooldridge(2009) : Exercice 15.11]

Chapitre 5 : La méthode du maximum de vraisemblance (MV)

95. Considérons un échantillon aléatoire de 10 observations :

5 ; 0 ; 1 ; 1 ; 0 ; 3 ; 2 ; 3 ; 4 ; 1

provenant d'une distribution de Poisson. La fonction de densité de chaque observation est :

$$f(y_i|\theta) = \frac{e^{-\theta}\theta^{y_i}}{y_i!} \quad \text{avec } y_i = 0,1,2,\ldots \text{ et } \theta > 0$$

Pour une distribution de Poisson : $E(y_i) = V(y_i) = \theta > 0$.

a) Quelle est la vraisemblance de cet échantillon en fonction du paramètre θ.

b) Donnez en général pour des observations suivant une loi de Poisson, la log-vraisemblance et l'équation de vraisemblance pour l'estimateur du maximum de vraisemblance.

c) Calculez l'estimateur du maximum de vraisemblance de θ. Vérifiez que vous avez bien un maximum de la log-vraisemblance.

d) Donnez la probabilité estimée des valeurs de 0 à 5.

[voir Greene (2018), Chapitre XIV.3]

96. Dans un modèle Probit, quel est l'effet marginal d'une variable explicative sur la probabilité de choix ?

a) Si la variable explicative est continue

b) Si la variable explicative est une indicatrice.

97. On a un échantillon de disques durs d'ordinateurs dont on a mesuré la durée de vie y_i avant de tomber en panne. On va supposer que cette variable dépendante y_i est distribuée selon une loi exponentielle. Celle-ci permet de mesurer la durée de vie d'un phénomène dont le taux de panne à un moment est constant dans le temps. Sa fonction de densité est :

$$f(y_i) = \frac{1}{\mu_i} \exp\left(\frac{-y_i}{\mu_i}\right) \quad \text{pour } y_i > 0$$

où $\mu_i = E(y_i) > 0$ est l'espérance de y_i. Par définition, sa variance est $V(y_i) = \mu_i^2$. On suppose que $\mu_i = \exp(x_i'\beta) > 0$ pour assurer la positivité.

a) Donnez la fonction de log-vraisemblance pour un échantillon aléatoire de N observations indépendantes.

b) Donnez les équations de vraisemblance.

c) Peut-on donner une expression analytique pour l'estimateur du maximum de vraisemblance de β ?

d) Donnez une expression pour la matrice de variance-covariance asymptotique de l'estimateur du maximum de vraisemblance.

98. Supposons que la variable x a une distribution de Weibull telle que la fonction de densité est : $f(x) = \alpha\beta x^{\beta-1} \exp(-\alpha x^{\beta})$ pour $x \geq 0$ et $\alpha, \beta > 0$.

a) Donnez la fonction de log-vraisemblance pour un échantillon aléatoire de N observations.

b) Donnez les équations de vraisemblance pour l'estimation par maximum de vraisemblance de α et β.

c) Obtenez la matrice des dérivées secondes de la log-vraisemblance par rapport à α et β. Les espérances exactes des éléments impliquant β impliquent des dérivées de la fonction Gamma. Ce qui est assez difficile à trouver analytiquement. Votre résultat exact donne un estimateur empirique. Comment estimeriez-vous plus facilement la matrice de covariance asymptotique de l'estimateur MV ?

d) Prouvez que $\alpha\beta Cov(\ln x\ , x^{\beta}) = 1$.
[Utilisez le fait que l'espérance des dérivées premières de la log-vraisemblance est nulle]

[voir Greene (2018), exercice 14.4]

99. On veut tester les J hypothèses linéaires $\boldsymbol{R\beta} = \boldsymbol{r}$. On peut utiliser un test de Wald de l'estimateur des multiplicateurs de Lagrange des MCC, soit les hypothèses : $H_0 : \boldsymbol{\lambda} = \boldsymbol{0}$. Le test de Wald est : $W = \boldsymbol{\lambda}^{*\prime}\hat{V}(\boldsymbol{\lambda}^*)^{-1}\boldsymbol{\lambda}^*$.

a) Prouvez que ce test dépend du ratio des estimateurs de la variance de l'erreur :

$$W = (N-K)\left[\frac{\boldsymbol{e}^{*\prime}\boldsymbol{e}^*}{\boldsymbol{e}'\boldsymbol{e}} - 1\right]$$

b) Pourquoi le facteur entre crochets est positif ?

c) Prouvez que cette statistique de test est égale à $J \times F$, où F est la statistique de test des restrictions $\boldsymbol{R\beta} = \boldsymbol{r}$ dans le modèle non contraint.

d) Montrez que la statistique de test du multiplicateur de Lagrange peut se réécrire comme

$$LM = \frac{N \times J}{((N-K)/F) + J}$$

[voir Greene (2018), exercice 5.7]

100. Dans un modèle *Logit* ou *Probit* où il n'y a que la constante : $y_i = \beta_1 + \varepsilon_i$. Notez P : le nombre d'observations avec une valeur $y = 1$ parmi les N observations.

a) Donnez l'estimateur de *MV* de la constante dans le modèle *Logit* et *Probit*.

b) Quelle est la valeur maximale de la log-vraisemblance dans ces deux modèles ?

c) Proposez un test de significativité conjointe des paramètres de pente s'il y a d'autres régresseurs dans le modèle : $y_i = \beta_1 + \beta_2 x_2 + \cdots + \beta_K x_K + \varepsilon_i$.