MODÈLES DE PRÉVISION, ÉTÉ 2014, ACT6420

ARTHUR CHARPENTIER

EXAMEN FINAL (3 heures)

L'épreuve dure 3 heures, les documents ne sont pas autorisés. Les calculatrices standards (cf. plan de cours) sont autorisées.

Il y a 50 exercices, souvent indépendants les uns des autres, sauf indications contraires, 20 questions sur les données individuelles, et 30 sur les données temporelles. 2 points par bonne réponse, 0 par mauvaise réponse, réponse double, et pas de réponse. La note finale sera le total des points.

Pour rappels, si $Z \sim \mathcal{N}(0,1)$, on a les probabilités suivantes

z	-2.575829	-2.326348	-1.959964	-1.644854	-1.281552	1.644854	1.959964
$\mathbb{P}(Z \leq z)$	0.005	0.01	0.025	0.05	0.10	0.95	0.975

Pour rappels enfin, un processus $ARMA(p,q)$ vérifie une relation de la forme

$$X_t = m + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

où (ε_t) est un bruit blanc.Sauf précision contraire, on supposera que $m = 0$. Pour un processus stationnaire, on note $\gamma(\cdot)$ la fonction d'autocovariance, $\rho(\cdot)$ la fonction d'autocorrélation et $\psi(\cdot)$ la fonction d'autocorrélation partielle.

Le problème était long, mais comme annoncé, il s'agissait de questions extraites d'un VEE passé cet hiver, qui contenait 50+50 questions, à faire en 2+2 heures (séparation entre la régression et les séries temporelles). Ici, il y avait 50 questions à faire en 3 heures. Ensuite les conditions étaient les mêmes : aucun document, seulement la calculatrice.

$\boxed{1}$ Le modèle classique de régression est basé sur une écriture

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i} + \varepsilon_i.$$

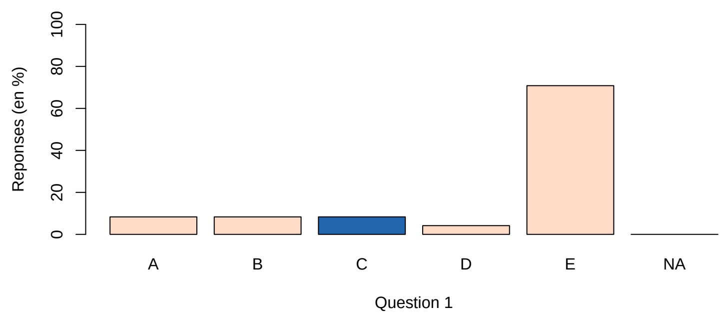
Les affirmations suivantes font partie des hypothèses standards, pour valider la méthode des moindres carrés, sauf une.

A. linéarité ou résidus centré : $\mathbb{E}(\varepsilon_i) = 0$
B. variance constante : $\text{Var}(\varepsilon_i)$ est identique, pour tout i,
C. normalité : ε_i suit une loi normale,
D. indépendance : ε_i et ε_j sont indépendants,
E. valeurs fixes : les valeurs de $\boldsymbol{X} = (X_1, \cdots, X_k)$ sont supposés fixes, ou si ce sont des mesures aléatiores, elles sont effectuées sans erreur, et on suppose alors $\text{cov}(X_j, \varepsilon) = 0$.

Pour faire des moindres carrés, i.e. résoudre

$$\text{argmin}\left\{\sum_{i=1}^{n}[Y_i - (\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i})]^2\right\}$$

on a besoin des conditions A, B, D et E (cf cours). L'hypothèse de normalisté est ici superflue. Elle est utile pour faire des tests, en revanche, surtout si on a peu d'observations. La bonne réponse est la réponse C. L'hypothèse E est aussi centrale ! Sans elle, on pourrait améliorer la prévision... C'est la seconde équation normale (slide #1-49). C'est aussi toute la discussion (que l'on avait eu à l'oral) autour des slides #1-60-62, sur la notion de projection orthogonale : comme on projette sur l'espace linéaire engendré par les X_i, $X_j \perp \varepsilon$ pour tout j. Si les X sont considérés comme des valeurs constantes, pas de soucis, mais on peut plus généralement les voir comme des variables aléatoires, ce qui rend légitime l'écriture en terme d'espérance conditionnelle.



$\boxed{2}$ On a estimé un modèle $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, sur un échantillon. On a obtenu

$$\sum_{i=1}^{n} \widehat{\varepsilon}_i^2 = 50, \sum_{i=1}^{n} X_i = 10, \sum_{i=1}^{n} X_i^2 = 100, \widehat{\beta}_0 = 2 \text{ et } \widehat{\beta}_1 = 1$$

Sur un autre échantillon, on a obtenu

$$\sum_{i=1}^{n} \tilde{\varepsilon}_i^2 = 80, \sum_{i=1}^{n} X_i = 10, \sum_{i=1}^{n} X_i^2 = 100, \tilde{\beta}_0 = 2 \text{ et } \tilde{\beta}_1 = 1$$

Que peut-on dire sur les statistiques de test t pour nos différents estimateurs

A. $t_{\widehat{\beta}_0} \leq t_{\tilde{\beta}_0}$ et $t_{\widehat{\beta}_1} \leq t_{\tilde{\beta}_1}$
B. $t_{\widehat{\beta}_0} \leq t_{\tilde{\beta}_0}$ et $t_{\widehat{\beta}_1} \geq t_{\tilde{\beta}_1}$
C. $t_{\widehat{\beta}_0} \geq t_{\tilde{\beta}_0}$ et $t_{\widehat{\beta}_1} \leq t_{\tilde{\beta}_1}$
D. $t_{\widehat{\beta}_0} \geq t_{\tilde{\beta}_0}$ et $t_{\widehat{\beta}_1} \geq t_{\tilde{\beta}_1}$
E. $t_{\widehat{\beta}_0} = t_{\tilde{\beta}_0}$ et $t_{\widehat{\beta}_1} = t_{\tilde{\beta}_1}$

Comme les estimateurs sont identiques ($\widehat{\beta}_1 = \tilde{\beta}_1$ et $\widehat{\beta}_0 = \tilde{\beta}_0$), comparer les statistiques de test revient à comparer les variances : comme les statistiques sont positives (elles sont toujours du même signe que les estimateurs, qui sont ici positifs), une statistique plus grande correspond à une variance plus petite.

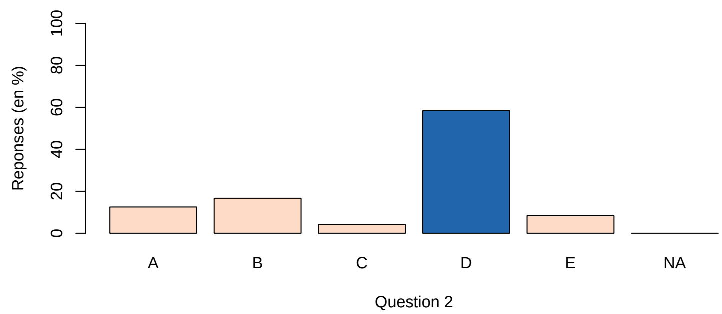
Si on utilise une écriture matricielle, on sait que la variance de nos estimateurs s'écrit $\text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$. Or comme la somme des X_i et la somme des X_i^2 est la même, le terme de droite $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$ ne change pas. Aussi, ce qu'il faut comparer, ce sont juste les estimations de variance des résidus, σ^2. Comme les résidus sont centrés, et que n est la même, on compare la somme des carrés des résidus. Qui est plus grande pour le second modèle que pour le premier. Aussim

$$\text{Var}(\tilde{\beta}_0) > \text{Var}(\hat{\beta}_0) \text{ et } \text{Var}(\tilde{\beta}_1) > \text{Var}(\hat{\beta}_1)$$

de telle sorte que

$$t_{\widehat{\beta}_0} \geq t_{\tilde{\beta}_0} \text{ et } t_{\widehat{\beta}_1} \geq t_{\tilde{\beta}_1}$$

qui est la réponse D.



$\boxed{3}$ À partir de n observations, on construit les estimateurs par moindres carrés de $\boldsymbol{\beta}$, avec $Y = \boldsymbol{X}^\mathsf{T}\boldsymbol{\beta} + \varepsilon$, notés $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_0, \widehat{\beta}_1, \cdots, \widehat{\beta}_k)$. On suppose que $\varepsilon \sim \mathcal{N}(0, 1)$. Sous ces hypothèses, et les hypothèses standard du modèle linéaire, une affirmation parmi les suivantes est fausse : les estimateurs par moindres carrés $\widehat{\beta}_j$ sont

A. indépendants les uns des autres

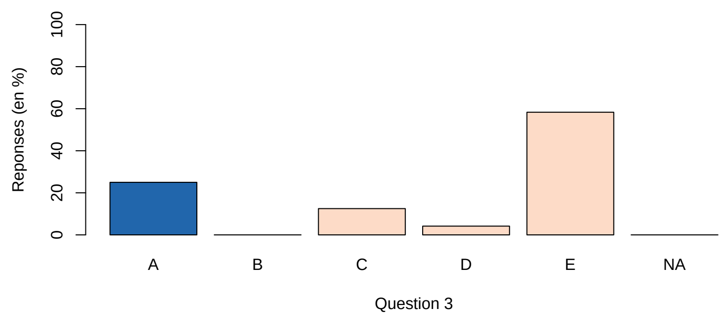
B. des estimateurs sans biais des β_j

C. les estimateurs sans biais de variance minimale

D. les estimateurs du maximum de vraisemblance

E. distribués suivant une loi normale

Les estimateurs ne sont pas indépendants, loin de là ! La bonne réponse est la réponse A. On avait vu dans le modèle de régression simple qu'il nétait pas rare - si les variables X et Y sont toutes deux positives - que $\text{cor}(\widehat{\beta}_0, \widehat{\beta}_1)$ soit négative. Pour rappels, on aurait effectivement l'indépendance si $\text{Var}(\widehat{\boldsymbol{\beta}})$ était une matrice diagonale. Ce qui n'est - a priori - pas le cas. Quant à E, on passe notre temps dans le cours a exploiter cette hypothèse pour faire des tests, par exemple. La propriété de normalité des estimateurs a été montrée en classe et découle du fait qu'une combiniaison linéaire de variables gaussiennes indépendantes est une variable gaussienne.



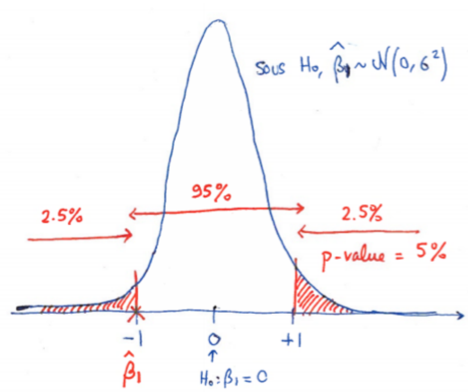
$\boxed{4}$ À partir de n observations, on construit estime le modèle $Y = \beta_0 + \beta_1 X + \varepsilon$. On utilise un test de Student pour tester l'hypothèse $H_0 : \beta_1 = 0$, et l'estimation par moindres carrés est $\widehat{\beta}_1 = -1$. On sait aussi que la p-value (pour ce test bilatéral) est 5%. Quelle affirmation parmi les suivantes est vraie,

A. la probabilité que $\beta_1 = -1$ est 95%
B. la probabilité que $\beta_1 = -1$ est 5%
C. la probabilité que $\beta_1 > -1$ est 5%
D. la probabilité que $\beta_1 < -1$ est 5%
E. la probabilité que $\beta_1 < -1$ est 2.5%

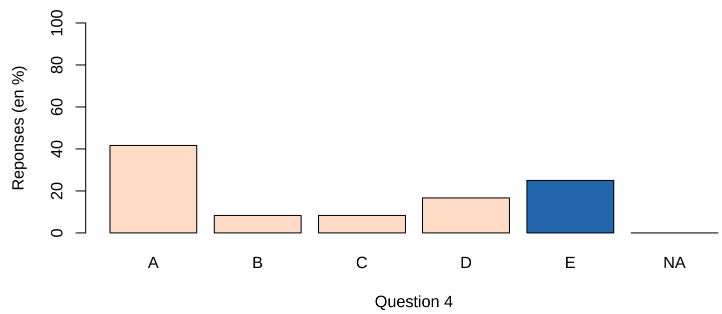
C'est un problème classique de probas, autant faire un dessin : sous H_0, $\beta_1 = 0$ et donc $\widehat{\beta}_1 \sim \mathcal{N}(0, \sigma^2) = Z$. Or

$$p - \text{value} = \mathbb{P}(|Z| > |\widehat{\beta}_1|) = 5\%$$

Donc ici, on a que $\mathbb{P}(\beta < -1) = 2.5\%$, qui est la réponse E.



Je rappelle que cet exercice n'est pas de moi, et je trouve douteux les énoncés, qui mélange des affirmations Bayésiennes (où β_1 est considéré comme une variable aléatoire) avec des tests fréquentistes. Disons que la réponse E est la plus vraisemblable, A et B n'ayant pas de sens (on parle d'un paramètre dans $\mathbb{R}$) par exemple, et parmi les trois autres, E est le seul qu'on peut justifier.



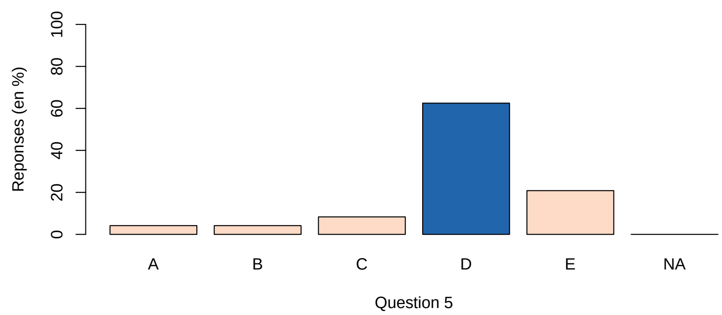
$\boxed{5}$ On a estimé une régression simple, $Y = \beta_0 + \beta_1 X + \varepsilon$, avec n observations. On note $\sigma^2 = \text{Var}(\varepsilon)$ et S^2 la variance de la variable X. Parmi les 5 cas suivant, quel cas correspond au cas où la variance de $\widehat{\beta}_1$ est la plus faible

A. $\sigma^2 = 4$, $n = 12$ et $S^2 = 5$
B. $\sigma^2 = 3$, $n = 11$ et $S^2 = 6$
C. $\sigma^2 = 4$, $n = 11$ et $S^2 = 5$
D. $\sigma^2 = 2$, $n = 13$ et $S^2 = 6$
E. $\sigma^2 = 3$, $n = 13$ et $S^2 = 5$

Pour rappels la variance de $\widehat{\beta}_1$ est

$$\text{Var}(\widehat{\beta}_1) = \frac{\text{Var}(\varepsilon)}{n\text{Var}(X)}$$

(on remontrera cette formule dans quelques questions). On peut calculer cette valeur dans les 5 cas, mais on peut aussi noter que cette variance sera minimal si σ^2 est faible, et si n et S^2 est grand. Ce qui correspond au cas D (pas besoin de faire ici tous les calculs).



Dans les deux questions qui suivent, on a estimé une régression simple, $Y = \beta_0 + \beta_1 X + \varepsilon$, avec 24 observations. On suppose que les moyennes empiriques de X et Y sont respectivement -3.7 et -2.4. On suppose que la variance (empirique) de X est 37, et que la covariance empirique entre X et Y est 74. Enfin, on suppose que ε est centré, de variance constante.

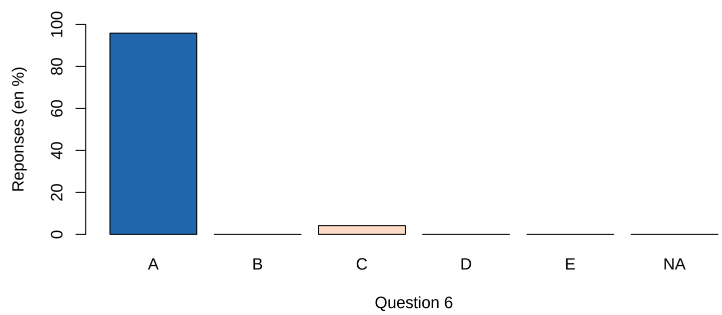
Les formules qu'on va utiliser ici sont les suivantes :

$$\begin{cases} \widehat{\beta}_1 = \dfrac{\text{Cov}(X,Y)}{\text{var}(X)} \\ \widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1 \overline{X} \end{cases}$$

$\boxed{6}$ Que vaut l'estimateur par moindres carrés de β_1 ?

A. $\widehat{\beta}_1 = 2$
B. $\widehat{\beta}_1 = 3$
C. $\widehat{\beta}_1 = 4$
D. $\widehat{\beta}_1 = 5$
E. $\widehat{\beta}_1 = 6$

Pour la pente, on fait le rapport entre la covariance et la variance de X, soit $74/37 = 2$, qui est la réponse A.



$\boxed{7}$ Que vaut l'estimateur par moindres carrés de β_0 ?

A. $\widehat{\beta}_0 = 2$

B. $\widehat{\beta}_0 = 3$

C. $\widehat{\beta}_0 = 4$

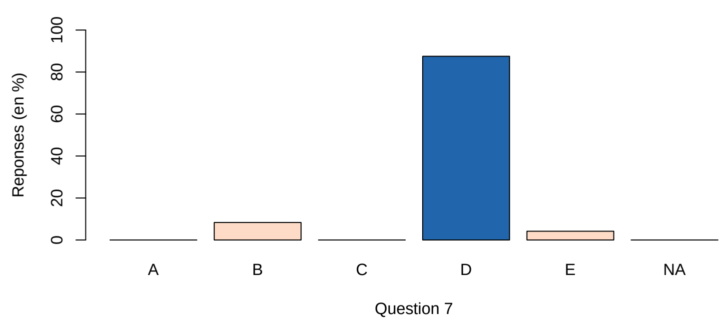
D. $\widehat{\beta}_0 = 5$

E. $\widehat{\beta}_0 = 6$

Pour la constante, on utilise le fait que la droite de régression passe forcément par le centre de gravité du nuage de points (X_i, Y_i), et donc

$$\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{X} = -2.4 - 2 \cdot (-3.7) = 5$$

qui est la réponse D.



Dans les deux questions qui suivent, on a estimé une régression simple, $Y = \beta_0 + \beta_1 X + \varepsilon$, avec 18 observations. On suppose que ε est centré, de variance constante valant $\sigma^2 = 72$, que

$$\sum_{i=1}^n X_i = 6 \text{ et } \sum_{i=1}^n X_i^2 = 10$$

et que les estimateurs par moindres carrés sont $\widehat{\beta}_0 = 9$ et $\widehat{\beta}_1 = 6$.

Un peu de calcul avant de répondre aux questions. Pour rappels, comme on l'a mentionné régulièrement durant le cours, la variance de $\widehat{\boldsymbol{\beta}}$ est

$$\text{Var}(\widehat{\boldsymbol{\beta}}) = \sigma^2 \cdot (\boldsymbol{X}^\top \boldsymbol{X})^{-1}$$

or, pour un modèle simple, cette matrice s'écrit plus simplement

$$\boldsymbol{X}^\top \boldsymbol{X} = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ X_1 & X_2 & \cdots & X_{n-1} & X_n \end{pmatrix} \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_{n-1} \\ 1 & X_n \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}$$

de telle sorte que

$$(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{1}{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & n \end{pmatrix}$$

i.e.

$$(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{1}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 & -\overline{X} \\ -\overline{X} & 1 \end{pmatrix}$$

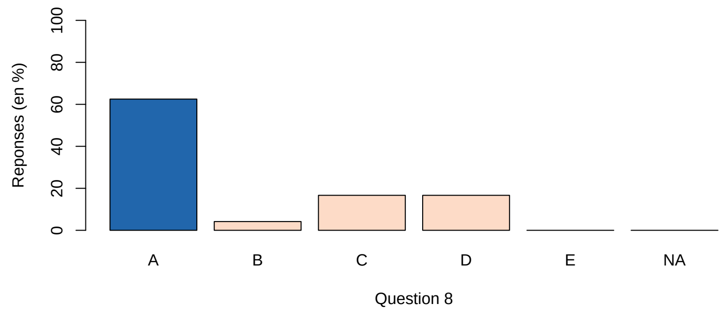
$\boxed{8}$ Que vaut la variance (estimée) de l'estimateur par moindres carrés de β_1 ?

A. $\text{Var}(\widehat{\beta}_1) = 9$
B. $\text{Var}(\widehat{\beta}_1) = 10$
C. $\text{Var}(\widehat{\beta}_1) = 12$
D. $\text{Var}(\widehat{\beta}_1) = 14$
E. $\text{Var}(\widehat{\beta}_1) = 16$

La variance de $\widehat{\beta}_1$ est donnée par le terme diagonal de la matrice ci-dessus, i.e.

$$\text{Var}(\widehat{\beta}_1) = \frac{\sigma^2}{n\text{Var}(X)} = \frac{\sigma^2}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} = \frac{72}{10 - 18 \cdot (6/18)^2} = 9$$

qui est la réponse A.



9 Que vaut la variance (estimée) de l'estimateur par moindres carrés de β_1 ? β_0 ?

A. $\text{Var}(\widehat{\beta}_0) = 5$

B. $\text{Var}(\widehat{\beta}_0) = 9$

C. $\text{Var}(\widehat{\beta}_0) = 13$

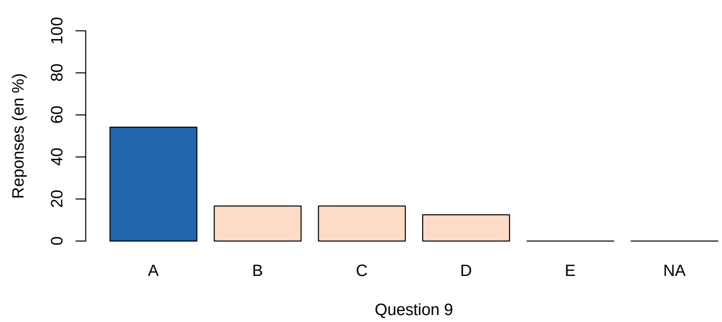
D. $\text{Var}(\widehat{\beta}_0) = 17$

E. $\text{Var}(\widehat{\beta}_0) = 20$

La variance de $\widehat{\beta}_0$ est donnée par le terme diagonal de la matrice ci-dessus, i.e.

$$\text{Var}(\widehat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} = \frac{72 \cdot 10}{18 \cdot 10 - 6^2} = 5$$

qui est - encore - la réponse A.



10 (A) estime un modèle $Y = \beta_0 + \beta_1 X + \varepsilon$ pour expliquer la pression sanguine (Y) en fonction de l'âge (X). Pour cela, (A) dispose de 20 observations, avec 20 valeurs pour $X_i \in \{22, 24, 26, \cdots, 58, 60\}$. La p-value du test de significativité de β_1 vaut 10%. Pour rappel, on dira que β_1 est significatif au seuil α si la p-value est inférieure

à α. (B) estime un modèle semblable, toujours avec 20 observations, mais cette fois $X_i \in \{21, 22, \cdots, 39, 40\}$. (A) et (B) ont exactement les mêmes estimateurs $\widehat{\beta}_0$, $\widehat{\beta}_1$ et $\widehat{\sigma}^2$, par moindres carrés.

Laquelle, parmi les affirmations suivantes, est vraie pour (B) :

A. on ne peut pas dire si β_1 est ou n'est pas significatif au seuil de 10%, mais l'est au seuil de 5%
B. β_1 est significatif au seuil de 10%, mais pas au seuil de 5%
C. β_1 n'est pas significatif au seuil de 10%, mais l'est au seuil de 5%
D. β_1 est significatif au seuil de 10%, et au seuil de 5%
E. β_1 n'est ni significatif au seuil de 10%, ni au seuil de 5%

L'ńoncé est un peu long... Mais on a ici des estimateurs identiques, pour les paramètres, avec une variance plus faible pour les X_i dans le cas (B) que dans le cas (A), avec le même nombre d'observations. Or on a vu que

$$\text{Var}(\widehat{\beta}_1) = \frac{\sigma^2}{n\text{Var}(X)}.$$

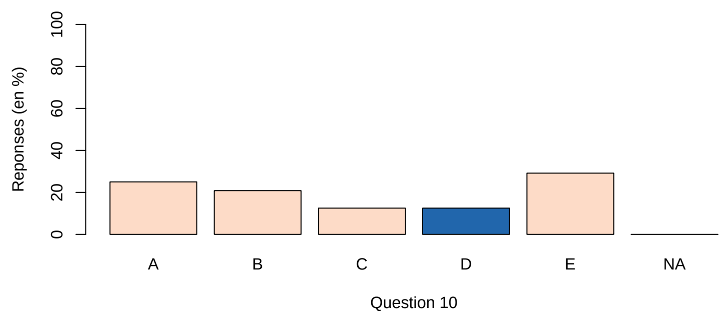
Aussi, dans le cas B, cette valeur est plus petite que dans le cas A,

$$\text{Var}(\widehat{\beta}_1^{(B)}) < \text{Var}(\widehat{\beta}_1^{(A)}).$$

En fait, la variance est 4 fois plus grande. En effet, $X^{(A)} = 20 + 2 \cdot (X^{(B)} - 20)$. Aussi, la statistique de test est

$$t^{(A)} = \frac{\widehat{\beta}_1^{(A)}}{\sqrt{\text{Var}(\widehat{\beta}_1^{(A)})}} = \frac{\widehat{\beta}_1^{(B)}}{\sqrt{4 \cdot \text{Var}(\widehat{\beta}_1^{(B)})}} = \frac{t^{(B)}}{2}$$

ou encore $t^{(B)} = 2 \cdot t^{(A)}$. On sait que la p-value pour A est de 10% (pour le test de significativité sur β_1). Donc $t^{(A)}$ est de l'ordre de ± 1.64, et donc $t^{(B)}$ est de l'ordre de ± 3.2, donc β_1 est significatif au seuil de 10%, et au seuil de 5%, même 1%, voire presque 0.1%. On va donc retenir la réponse D.



$\boxed{11}$ Toujours dans un modèle de régression simple, $Y = \beta_0 + \beta_1 X + \varepsilon$, on utilise 6 observations. L'intervalle de confiance à 99% pour β_1 est $[0, 2]$. Quel serait l'intervalle de confiance à 95% pour β_1 ?

A. [0.232, 1.768]
B. [0.397, 1.603]
C. [0.537, 1.463]
D. [0.216, 1.784]
E. [0.362, 1.638]

L'intervalle de confiance à 99% pour β_1 est

$$\left[\widehat{\beta}_1 \pm 2.575829\sqrt{\text{Var}(\widehat{\beta}_1)}\right]$$

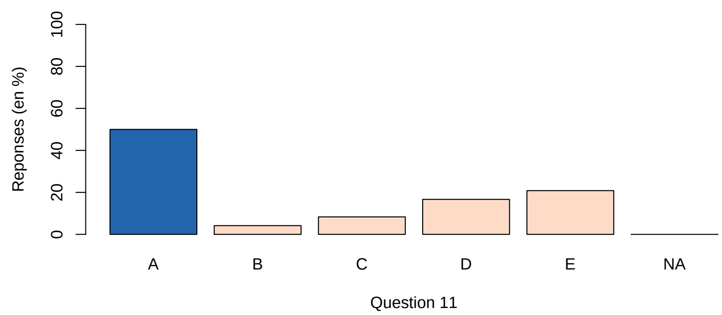
car $\mathbb{P}(Z < -2.576) = 0.5\%$, si $Z \sim \mathcal{N}(0, 1)$, i.e. $\mathbb{P}(Z \in [-2.576; 2.576]) = 99\%$, compte tenu des valeurs indiquées sur la page de garde. Aussi, comme l'intervalle de confiance est ici centré sur 1 on en dd́uit que $\widehat{\beta}_1 = 1$, et que $\sqrt{\text{Var}(\widehat{\beta}_1)} = 1/2.575829$. Maintenant, on sait aussi que l'intervalle de confiance à 95% pour β_1 est

$$\left[\widehat{\beta}_1 \pm 1.959964\sqrt{\text{Var}(\widehat{\beta}_1)}\right]$$

donc, par substitution,

$$\left[1 \pm \frac{1.959964}{2.575829}\right] = [1 \pm 0.760906] = [0.239094; 1.760906]$$

qui ressemble à la réponse A.



$\boxed{12}$ Parmi les transformations suivantes, laquelle n'est pas dans la famille de Box-Cox,

A. $g(x) = x - 1$

B. $g(x) = x^2 - 2$

C. $g(x) = \log(x)$

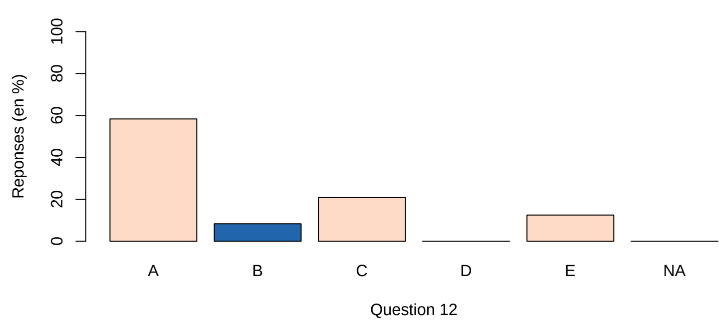
D. $g(x) = x^2/2 - 1/2$

E. $g(x) = 2\sqrt{x} - 2$

On avait vu dans le cours (slide #2-92) que la transformation de Box Cox est de la forme

$$g_\lambda(x) = \frac{x^\lambda - 1}{\lambda} \text{ pour } \lambda \neq 0$$

avec la limite $g_0(x) = \log(x)$. Bref, C est dans cette famille, ainsi que A, D et E. C'est B qui n'est pas dans cette famille.



$\boxed{13}$ On ajuste un modèle linéaire sur $n = 100$ observations, $Y = \beta_0 + \beta_1 X + \varepsilon$, et on obtient

$$\widehat{\beta}_1 = -35.69, \sum_{i=1}^{n}(X_i - \overline{X})^2 = 1.62 \text{ et } \sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = 2934$$

Quel est l'intervalle de confiance (symmétrique) à 90% pour β_1 ?

A. $[-43.23; -28.14]$
B. $[-42.01; -29.36]$
C. $[-111.12; 39.74]$
D. $[-99.00; 27.62]$
E. $[-54.68; -16.69]$

L'estimateur de σ^2 est

$$\widehat{\sigma}^2 = \frac{1}{100}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = \frac{2934}{100} = 23.94$$

(en principe, on divise par $100 - 2$ si on veut un estimateur sans biais, mais ça ne changera pas grand chose). Aussi, l'écart type de notre estimateur $\widehat{\beta}_1$ est ici (on a déjà fait plusieurs fois le calcul depuis le début)

$$\sqrt{\text{Var}(\widehat{\beta}_1)} = \sqrt{\frac{\widehat{\sigma}^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}} = \sqrt{\frac{23.94}{1.62}} = 3.847.$$

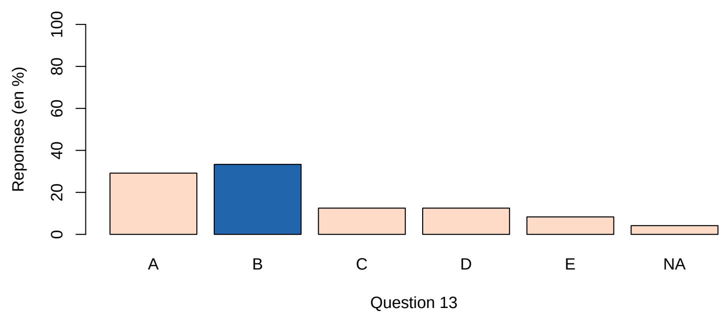
Aussi, l'intervalle de confiance à 90% est

$$[-35.69 \pm u \cdot 3.847]$$

où u est le quantile à 95% de la loi normale centré réduite (90% de la masse est comprise entre le quantile à 5% et à 95%), soit 1.6448, aussi l'intervalle est

$$[-35.69 \pm u \cdot 1.6448] = [-42.01841; -29.36159]$$

qui correspond au cas B.



$\boxed{14}$ On ajuste un modèle $Y = \beta_0 + \beta_1 X + \varepsilon$ sur $n = 100$ observations, où X est une variable prenant les valeurs 0 ou 1. Dans 40% des cas, X_i a pris la valeur 1. On nous dit que

$$\widehat{\beta}_1 = 1.4 \text{ et } \sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = 920.$$

Que vaut la statistique du test de Student associé au test de significativité $H_0 : \beta_1 = 0$?

A. 1.15

B. 1.78

C. 2.26

D. 2.46

E. 3.51

L'estimateur de σ^2 est

$$\widehat{\sigma}^2 = \frac{1}{100}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = \frac{920}{100} = 9.2$$

(là encore, en principe, on divise par $100 - 2$ si on veut un estimateur sans biais, mais ça ne changera pas grand chose). Pour avoir l'écart type de notre estimateur $\widehat{\beta}_1$, il nous manque le terme

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$$

Comme ici X_i prend les valeurs 0 ou 1, $X_i^2 = X_i$. Donc la relation précédante se simplifie,

$$\sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n X_i - \overline{X}^2 = n\overline{X} - n\overline{X}^2 = 40 - 16 = 24$$

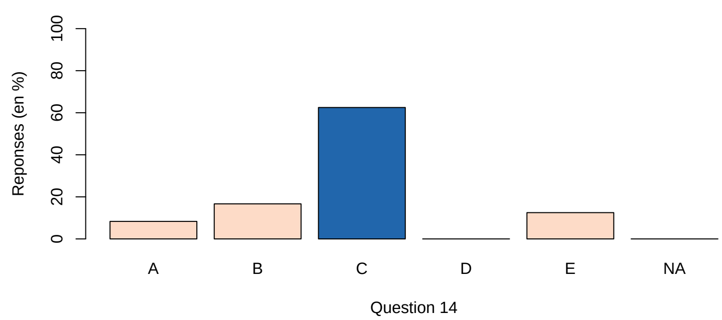
(qui est juste la variance d'une loi binomiale, $n\overline{X}(1-\overline{X})$). Aussi, l'écart type de notre estimateur $\widehat{\beta}_1$ s'écrit

$$\sqrt{\text{Var}(\widehat{\beta}_1)} = \sqrt{\frac{\widehat{\sigma}^2}{\sum_{i=1}^n (X_i - \overline{X})^2}} = \sqrt{\frac{9.2}{24}} = 0.6191.$$

Aussi, la statistique de test est

$$t = \frac{\widehat{\beta}_1}{\sqrt{\text{Var}(\widehat{\beta}_1)}} = \frac{1.4}{0.6191} = 2.2612$$

qui correspond au cas C.



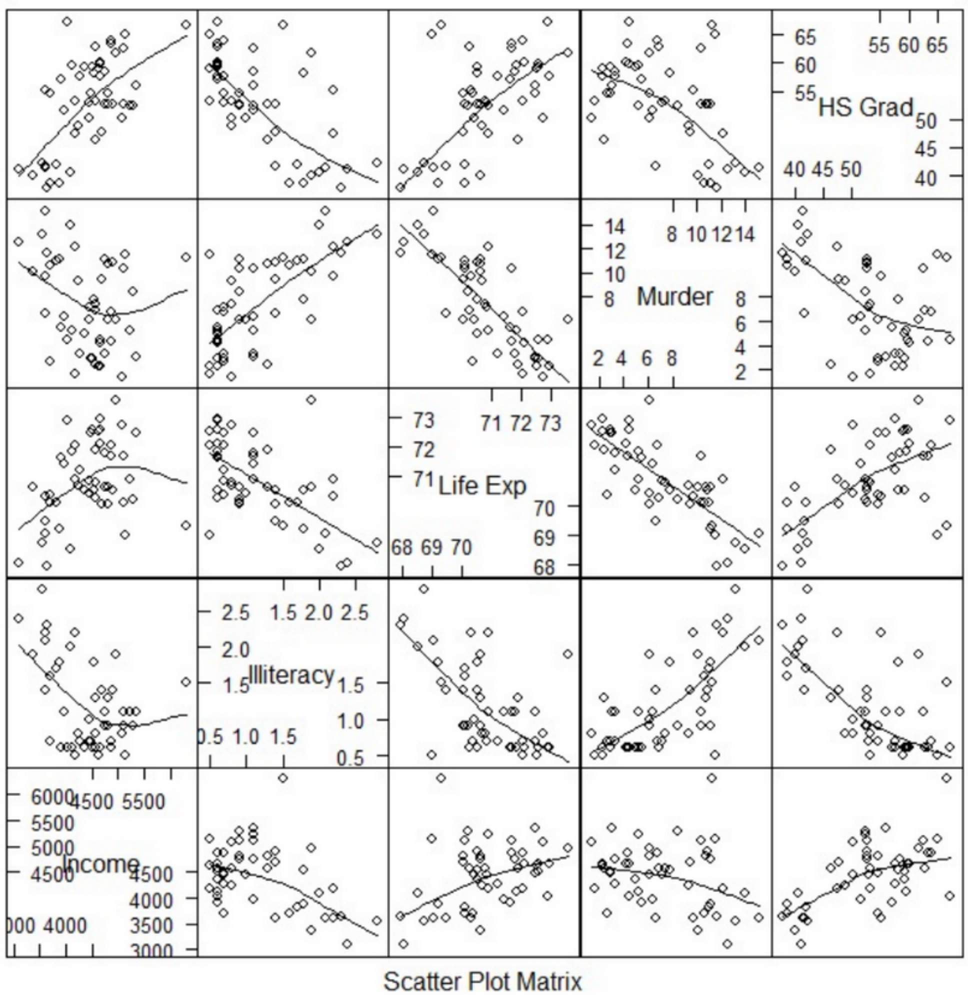
$\boxed{15}$ On dispose d'une base de données contenant 5 variables quantitatives,

Income = income per capita (in 1974)
Illiteracy = illiteracy as percent of population (in 1970)
Life exp = life expectancy in years (in 1969–71)
Murder = murder and non-negligent manslaughter rate per 100,000 population (in 1976)
HS Grad = high-school graduation percentage (in 1970)

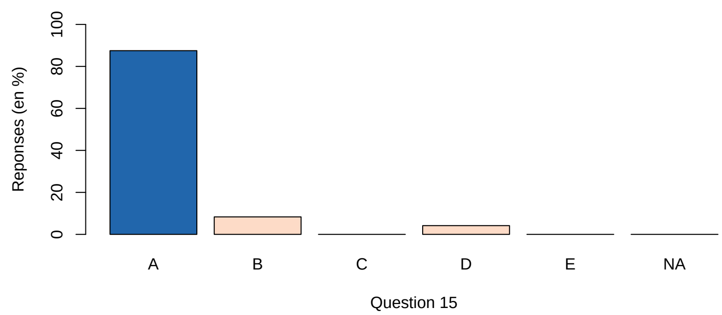
On a visualisé les pairs de variables sur le graphique ci-dessous. Quelles sont les paires qui sont positivement corrélées

A. High school graduation rate et income
B. Murder rate et life expectancy
C. Income et illiteracy rate
D. High school graduation rate et illiteracy rate

E. Illiteracy rate et life expectancy



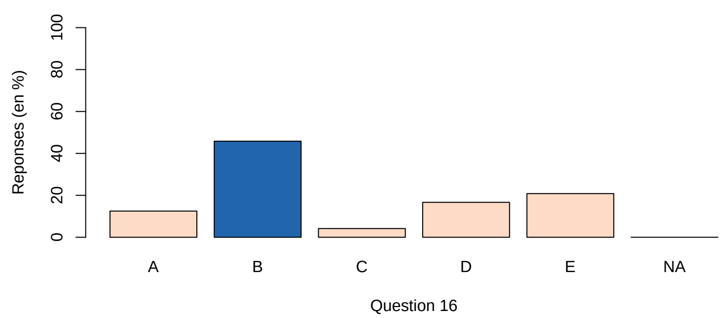
On ne va pas y passer des heures, c'est la réponse A (en bas à droite, et en haut à gauche).



$\boxed{16}$ On estime ici un modèle linéaire. Quelle relation est la bonne, parmi les cinq suivantes

A. $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \widehat{\varepsilon}_i$
B. $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i$
C. $Y_i = \beta_0 + \beta_1 X_i$
D. $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \varepsilon_i$
E. $Y_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i + \varepsilon_i$

Réponse B.



Les deux questions suivantes portent sur le modèle suivant: on suppose que $Y = \beta_0 + \beta_1 X + \varepsilon$, avec les hypothèses standards, et on dispose de 5 observations, avec $X_i = \{1, 2, 3, 4, 5\}$. On suppose aussi que $\sigma^2 = \mathbb{E}(\varepsilon^2) = 50$.

Pour rappels, on avait montré tantôt que

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} = \frac{1}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 & -\overline{X} \\ -\overline{X} & 1 \end{pmatrix}$$

Ce qui nous intéresse, ici, ce sont juste les termes diagonaux de la matrice de variance-covariance,

$$\text{Var}(\widehat{\beta}_1) = \frac{\sigma^2}{n\text{Var}(X)} = \frac{\sigma^2}{\sum_{i=1}^n X_i^2 - n\overline{X}^2}$$

et

$$\text{Var}(\widehat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}$$

Maintenant, notons qu'ici

$$n = 5, \sum_{i=1}^n X_i = 15 \text{ et } \sum_{i=1}^n X_i^2 = 55.$$

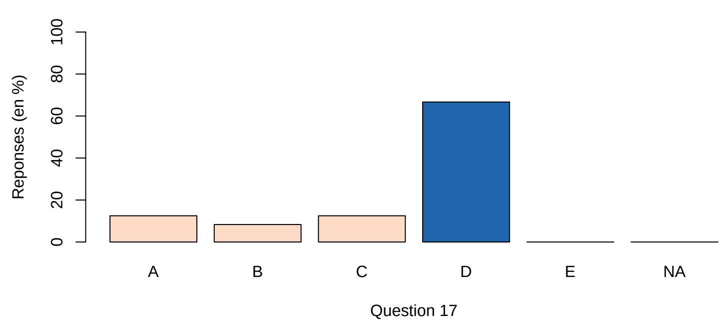
$\boxed{17}$ Quelle est la variance de $\widehat{\beta}_1$?

A. 2
B. 3
C. 4
D. 5
E. 6

Lançons nous dans les calculs (qui sont toujours les mêmes), et on obtient

$$\text{Var}(\widehat{\beta}_1) = \frac{50}{55 - 5 \cdot [15/5]^2} = \frac{50}{10} = 5$$

qui est la réponse D.



$\boxed{18}$ Quelle est la variance de $\widehat{\beta}_0$?

A. 55

B. 77

C. 91

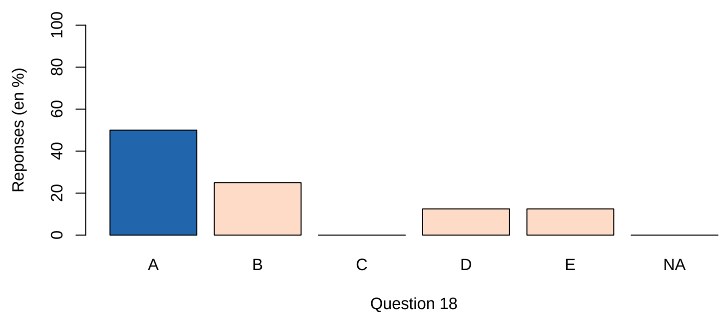
D. 95

E. 145

Là encore, on remplace,

$$\mathrm{Var}(\widehat{\beta}_0) = \frac{50 \cdot 55}{5 \cdot 55 - 15^2} = \frac{50 \cdot 55}{50} = 55$$

qui est la réponse A.



$\boxed{19}$ Suite à l'estimation d'un modèle linéaire, sur $n = 120$ observations, on nous dit que

$$\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = 435 \text{ et } R^2 = 0.65$$

Que vaut $\widehat{\sigma}^2$?

A. 0.657

B. 1.110

C. 1.233

D. 1.522

E. 2.318

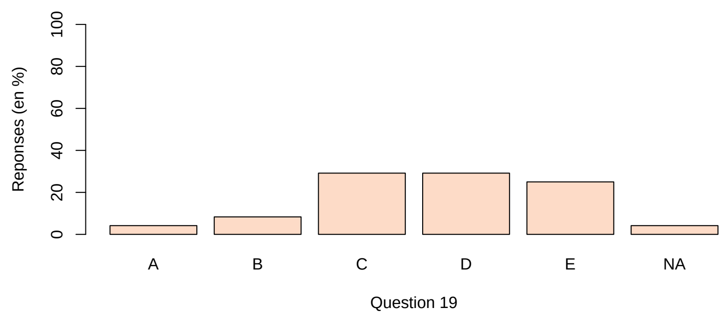
On se souvient que

$$R^2 = 1 - \frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(Y)}$$

aussi,

$$\mathrm{Var}(\varepsilon) = (1 - R^2) \cdot \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

Maintenant, avec les chiffres donnés, ça ne collait pas. En fait, j'attendais la réponse D, mais pour ça il fallait $n = 100$. Désolé, j'enlève la question.



$\boxed{20}$ Suite à l'estimation d'un modèle linéaire $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$, sur $n = 120$ observations, on nous dit que

$$(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \begin{pmatrix} 188.983 & 0.858 & -28.027 \\ 0.858 & 0.250 & -0.600 \\ -28.027 & -0.600 & 5.062 \end{pmatrix}, \widehat{\boldsymbol{\beta}} = \begin{pmatrix} 87 \\ 0.14 \\ -7.2 \end{pmatrix}, \text{ et } \widehat{\sigma}^2 = 0.0361$$

Quel est le plus petit intervalle de confiance à 95% de β_1

A. $[-0.046; 0.326]$
B. $[-0.232; 0.512]$
C. $[0.123; 0.156]$
D. $[0.138; 0.141]$
E. $[0.046; 0.233]$

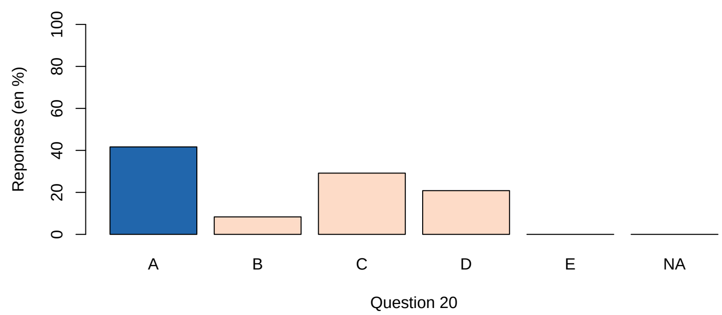
Le plus petit intervalle de confiance est de la forme

$$\left[\widehat{\beta}_1 \pm 1.96\sqrt{\sigma^2((\boldsymbol{X}^\top \boldsymbol{X})^{-1})_{1,1}}\right]$$

soit ici

$$0.14 \pm 1.96\sqrt{0.0361 \cdot 0.250} = [-0.04619658; 0.32619658]$$

qui est la réponse A.



$\boxed{21}$ Considérons la série temporelle $Y_t = Y_{t-1} + \varepsilon_t$, où (ε_t) est un bruit blanc, avec $\sigma_\varepsilon^2 = 0.18$, et $Y_t = 0$ pour $t \leq 0$. Quelle est l'écart-type de Y_2 ?

A. 0.6

B. 1.2

C. 3.0

D. 3.6

E. 7.2

On a ici une marche aléatoire,

$$Y_t = Y_{t-1} + \varepsilon_t = Y_{t-2} + \varepsilon_{t-1} + \varepsilon_t = \cdots$$

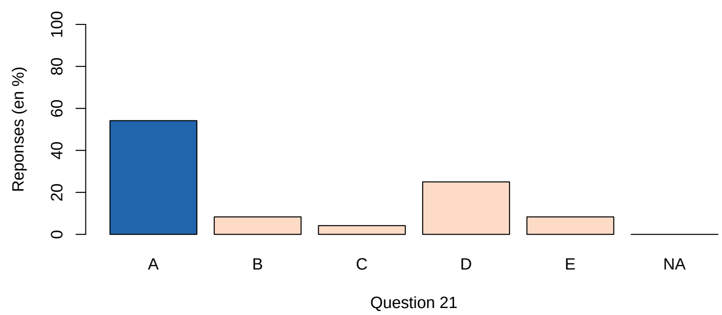
et donc, en particulier,

$$Y_2 = Y_0 + \varepsilon_{t-1} + \varepsilon_t.$$

Aussi, comme $\text{Cov}(\varepsilon_{t-1}, \varepsilon_t) = 0$ on en déduit

$$\text{Var}(Y_2) = \text{Var}(\varepsilon_{t-1}) + \text{Var}(\varepsilon_t) = 2\sigma_\varepsilon^2 = 2 \cdot 0.18 = 0.36$$

donc l'écart type est $0.36 = 0.6$, qui est la réponse A.



$\boxed{22}$ Considérons la série temporelle

$$Y_t = \frac{\varepsilon_t + \varepsilon_{t-1}}{2}$$

où (ε_t) est un bruit blanc, avec $\sigma_\varepsilon^2 = 8$. Quelle est la covariance entre Y_t et Y_{t-1} ?

A. 0

B. 1

C. 2

D. 3

E. 4

On se lance :

$$\text{Cov}(Y_t, Y_{t-1}) = \text{Cov}\left(\frac{\varepsilon_t + \varepsilon_{t-1}}{2}, \frac{\varepsilon_{t-1} + \varepsilon_{t-2}}{2}\right)$$

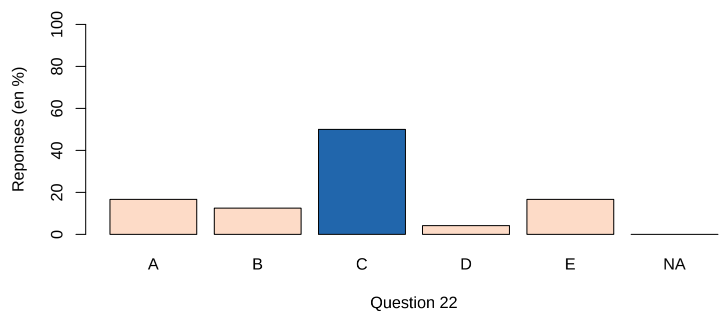
qui donne, après développement

$$\text{Cov}(Y_t, Y_{t-1}) = \text{Cov}\left(\frac{\varepsilon_{t-1}}{2}, \frac{\varepsilon_{t-1}}{2}\right)$$

les autres covariances étant nulle, puisque (ε_t) est un bruit blanc. Bref,

$$\text{Cov}(Y_t, Y_{t-1}) = \text{Var}\left(\frac{\varepsilon_{t-1}}{2}\right) = \frac{1}{4}\sigma_\varepsilon^2 = \frac{8}{4} = 2,$$

qui est la réponse C.



Pour les deux questions qui suivent, considérons la série temporelle

$$Y_t = \frac{\varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_{t-7}}{8}$$

où (ε_t) est un bruit blanc, avec $\sigma_\varepsilon^2 = 0.64$.

$\boxed{23}$ Quelle est la covariance entre Y_t et Y_{t-1} ?

A. 0.03
B. 0.04
C. 0.05
D. 0.06
E. 0.07

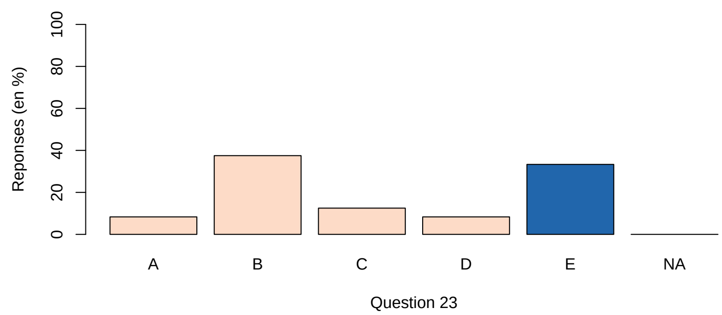
Pour la covariance, la encore, il faut avoir des termes identiques pour la covariance ne s'annule pas (cf question précédante)

$$\text{Cov}(Y_t, Y_{t-1}) = \frac{7}{8^2} \cdot \sigma_\varepsilon^2$$

car on a juste 7 termes croisés (ε_{t-1}, ..., ε_{t-7}). Aussi,

$$\gamma(1) = \text{Cov}(Y_t, Y_{t-1}) = \frac{7}{8^2} \cdot \sigma_\varepsilon^2 = \frac{7 \cdot 0.64}{8^2} = 0.07$$

qui est la réponse E.



$\boxed{24}$ Quelle est la corrélation entre Y_t et Y_{t-1} ?

A. 0.667

B. 0.750

C. 0.800

D. 0.833

E. 0.875

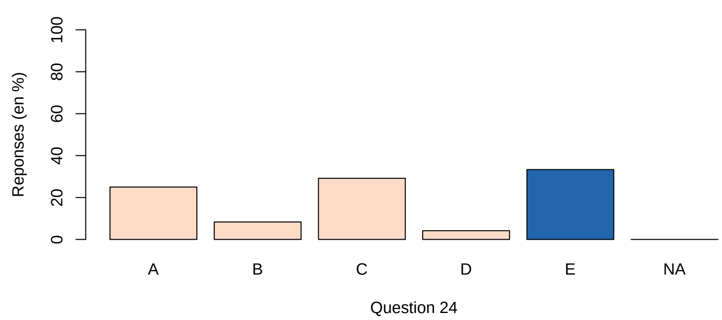
Il nous manque ici la variance. Or

$$\gamma(0) = \text{Var}(Y_t) = \left(\frac{1}{8^2} + \cdots \frac{1}{8^2}\right)\sigma_\varepsilon^2 = \frac{8}{8^2} \cdot \sigma_\varepsilon^2 = \frac{0.64}{8} = 0.08.$$

donc

$$\text{Cor}(Y_t, Y_{t-1}) = \frac{\gamma(1)}{\gamma(0)} = \frac{0.07}{0.08} = \frac{7}{8} = 0.875$$

qui est la réponse E.



$\boxed{25}$ Considérons un processus $AR(1)$ stationnaire $AR(1)$ avec 200 observations. On observe $\widehat{\gamma}(0) = 0.348$ et $\text{Var}(\overline{Y}) = 0.006$. Quelle est la valeur de ϕ ?

A. 0.50

B. 0.55

C. 0.60

D. 0.65

E. 0.70

Je dois avouer que la question était difficile, et qu'en étant raisonnable, mieux valait sauter la question (à moins de l'avoir travaillé en préparant le cours, car la résolution prend un peu de temps) ! Mais encore une fois, il s'agit de questions réllement posée dans un VEE, et qui pouvait être faite (mais il fallait un peu de temps).

Question 1.5: Variance of $\overline{y}$

A stationary autoregressive AR(1) time series of 200 observations has γ_0 = 0.348 and Var($\overline{y}$) = 0.006. What is the autoregressive parameter ϕ of this time series?

A. 0.500
B. 0.550
C. 0.600
D. 0.650
E. 0.700

$$\text{Var}\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) = \frac{1}{n^2}\left(\sum_{t=1}^{n}\text{Var}(Y_t) + \sum_{s\neq t}\text{Cov}(Y_s, Y_t)\right)$$

de telle sorte que

$$\text{Var}\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) = \frac{1}{n^2}\left(n\gamma(0) + \sum_{h=1}^{n-1} 2(n-h)\gamma(h)\right)$$

Or on sait que $\gamma(h) = \phi^h\gamma(0)$ donc finalement

$$\text{Var}\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) = \frac{\gamma(0)}{n^2}\left(n\gamma(0) + \sum_{h=1}^{n-1} 2(n-h)\phi^h\right)$$

Le terme de droite peut toutefois se simplifier un peu,

$$\sum_{h=1}^{n-1} 2(n-h)\phi^h = 2\phi^{n-1}\sum_{h=1}^{n-1}(n-h)\left(\frac{1}{\phi}\right)^{n-h-1} = 2\phi^{n-1}\frac{\partial}{\partial x}\sum_{i=1}^{n} x^{n-i}\bigg|_{x=1/\phi}$$

de telle sorte que

$$\sum_{h=1}^{n-1} 2(n-h)\phi^h = 2\phi^{n-1}\frac{(n-1)\phi^{-(n+1)} - n\phi^{-n} + \phi^{-1}}{\phi^{-1}(\phi^{-1}-1)^2} = 2\frac{(n-1)\phi^{-1} - n + \phi^{n-1}}{(\phi^{-1}-1)^2}$$

Aussi

$$V = \text{Var}\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) = \frac{\gamma(0)}{n^2}\left[n + 2\frac{(n-1)\phi^{-1} - n + \phi^{n-1}}{(\phi^{-1}-1)^2}\right]$$

Maintenant, je ne me sens pas très inspiré pour résoudre cette équation (non linéaire), mais numériquement on peut voir que pour $\phi = 0.55$, on a une variance de l'ordre de 0.00594, soit environ 0.006.

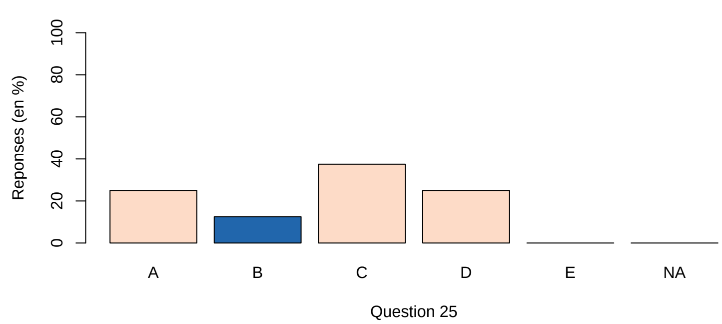
Maintenant, on pouvait aussi faire un peu de calculs, en notant qu'avec $n = 200$, $\phi^{n-1} \sim 0$. Et dans ce cas, on peut simplifier, en remplaçant aussi $n-1$ par n, de telle sorte que

$$\frac{\gamma(0)}{n^2}\left[n + 2\frac{(n-1)\phi^{-1} - n + \phi^{n-1}}{(\phi^{-1}-1)^2}\right] \sim \frac{\gamma(0)}{n}\frac{1+\phi}{1-\phi}$$

i.e.

$$\phi \sim \sqrt{\frac{\gamma(0)}{\gamma(0) + nV}}$$

soit numériquement une valeur proche de 0.54. La bonne réponse était la réponse B.



Dans les trois questions qui suivent, on considère un processus (Y_t), $MA(2)$, avec $\theta_1 = +0.2$, $\theta_2 = -0.7$ et $\sigma_\varepsilon^2 = 1$.

$\boxed{26}$ Quelle est la variance de Y_t ?

A. 1.34

B. 1.40

C. 1.45

D. 1.49

E. 1.53

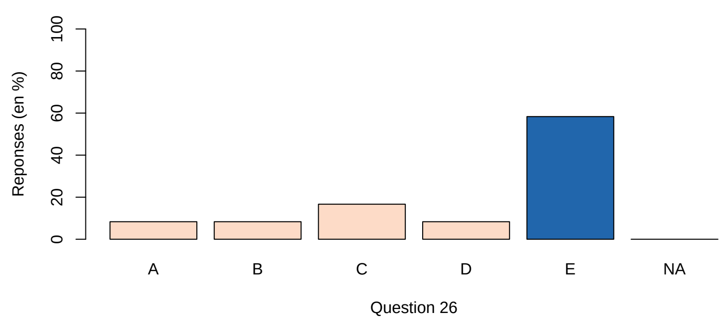
On sait que $Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$ oú (ε_t) est un bruit blanc, et donc $\text{cov}(\varepsilon_s, \varepsilon_t) = 0$ si $s \neq t$. Aussi

$$\text{Var}(Y_t) = [1 + \theta_1^2 + \theta_2^2] \cdot \text{Var}(\varepsilon_t)$$

soit ici

$$\text{Var}(Y_t) = [1 + 0.2^2 + 0.7^2] = 1.53$$

qui est la réponse E.



$\boxed{27}$ Quelle est la corrélation entre Y_t et Y_{t-1} ?

A. -0.222

B. -0.123

C. -0.039

D. $+0.039$

E. $+0.222$

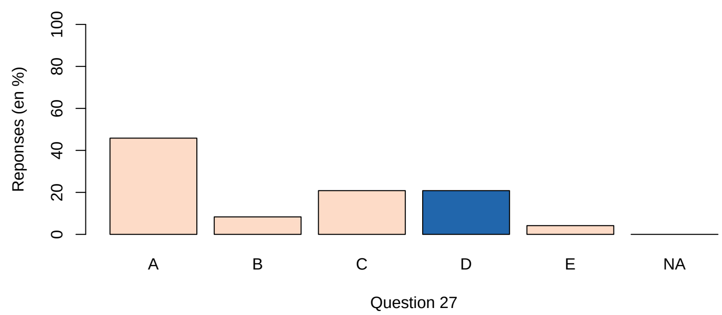
On sait que

$$\text{Cov}(Y_t, Y_{t-1}) = \text{Cov}(\theta_1\varepsilon_{t-1}, \varepsilon_{t-1}) + \text{Cov}(\theta_2\varepsilon_{t-2}, \theta_1\varepsilon_{t-2}) = \theta_1[1 + \theta_2] \cdot \text{Var}(\varepsilon_t)$$

tous les autres termes seront nuls. Aussi

$$\text{Cor}(Y_t, Y_{t-1}) = \frac{\theta_1[1 + \theta_2]}{1 + \theta_1^2 + \theta_2^2} = \frac{0.2 \cdot (1 - 0.7)}{1 + 0.2^2 + 0.7^2} = 0.039$$

qui est la réponse D.



$\boxed{28}$ Quelle est la corrélation entre Y_t et Y_{t-2} ?

A. -0.458

B. -0.429

C. $+0.429$

D. $+0.441$

E. $+0.458$

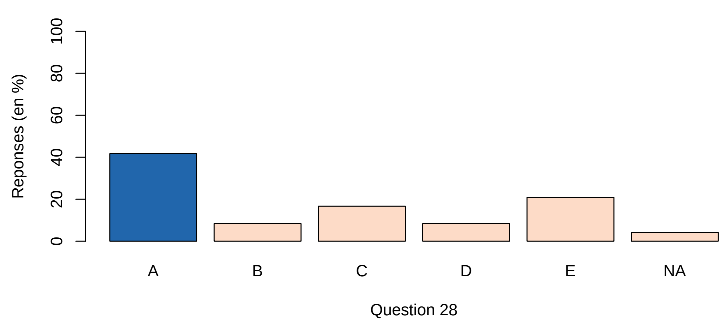
On sait que

$$\text{Cov}(Y_t, Y_{t-2}) = \text{Cov}(\theta_2 \varepsilon_{t-2}, \varepsilon_{t-2}) = \theta_2 \cdot \text{Var}(\varepsilon_t)$$

tous les autres termes seront nuls. Aussi

$$\text{Cor}(Y_t, Y_{t-2}) = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} = \frac{-0.7}{1 + 0.2^2 + 0.7^2} = -0.457516$$

qui est la réponse A.



Dans les deux questions qui suivent, on considère un processus $AR(1)$, avec $\phi = -0.5$ et $\sigma_\varepsilon = 8.66$.

$\boxed{29}$ Quelle est la covariance entre Y_t et Y_{t-1} ?

A. -30

B. -40

C. -50

D. -70

E. -90

On sait que

$$\text{Cov}(Y_t, Y_{t-1}) = \text{Cov}(\phi Y_{t-1} + \varepsilon_t, Y_{t-1})$$

où (ε_t) est le processus d'innovation, et donc $\text{Cov}(\varepsilon_t, Y_{t-1}) = 0$. Aussi,

$$\text{Cov}(Y_t, Y_{t-1}) = \phi \text{Var}(Y_{t-1})$$

Bon, on y est presque, il manque juste la variance... Mais

$$\text{Var}(Y_t) = \text{Var}(\phi Y_{t-1} + \varepsilon_t) = \phi^2 \underbrace{\text{Var}(Y_{t-1})}_{\text{Var}(Y_t)} + 2\phi \underbrace{\text{Cov}(Y_{t-1}, \varepsilon_t)}_{=0} + \text{Var}(\varepsilon_t)$$

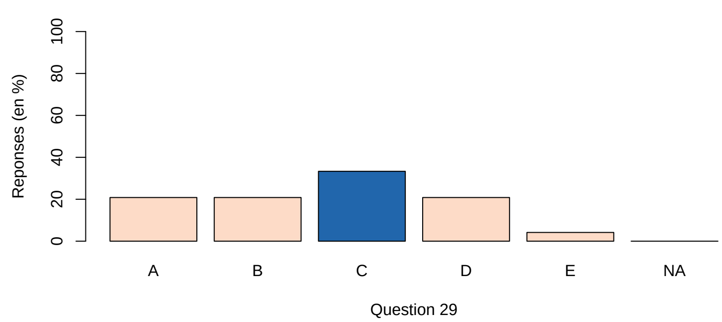
de telle sorte que $\text{Var}(Y_t)$ est donné par

$$[1 - \phi^2]\text{Var}(Y_t) = \sigma^2$$

Bref,

$$\text{Cov}(Y_t, Y_{t-1}) = \phi \cdot \frac{\sigma^2}{1-\phi^2} = -0.5 \cdot \frac{8.66^2}{1-0.5^2} = -50,$$

qui est la réponse C.



$\boxed{30}$ Quelle est la covariance entre Y_t et Y_{t-2} ?

A. 9

B. 16

C. 25

D. 49

E. 81

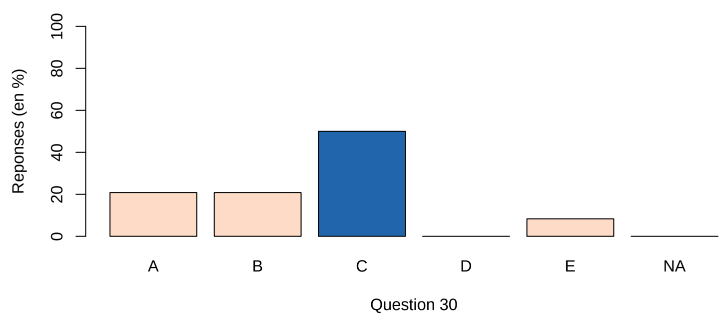
On sait que

$$\mathrm{Cov}(Y_t, Y_{t-2}) = \mathrm{Cov}(\phi Y_{t-1} + \varepsilon_t, Y_{t-2}) = \phi \cdot \underbrace{\mathrm{Cov}(Y_{t-1}, Y_{t-2})}_{\mathrm{Cov}(Y_t, Y_{t-1})}$$

le dernier terme ayant été calculé dans la question précédante. Bref,

$$\mathrm{Cov}(Y_t, Y_{t-2}) = \phi \cdot \frac{\phi\sigma^2}{1-\phi^2} = 0.5^2 \cdot \frac{8.66^2}{1-0.5^2} = 25,$$

qui est la réponse C.



Dans les deux questions qui suivent, on considère un processus $AR(1)$, avec $\phi > 0$, $\sigma_\varepsilon^2 = 5.333$, et $\gamma(2) = 3$.

$\boxed{31}$ Quelle est la valeur de ϕ, coefficient d'autororrélation ?

A. 0.5

B. 0.6

C. 0.7

D. 0.8

E. 0.9

On vient de montrer que

$$\mathrm{Cov}(Y_t, Y_{t-2}) = \gamma(2) = \sigma^2 \cdot \frac{\phi^2}{1-\phi^2}$$

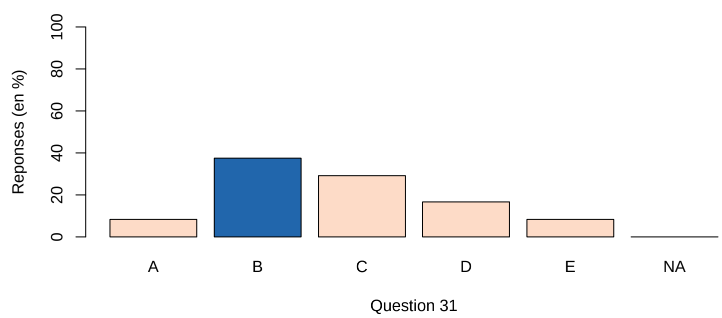
donc ici, on cherche ϕ^2 tel que

$$\frac{\phi^2}{1-\phi^2} = \frac{\gamma(2)}{\sigma^2} = \frac{3}{16/3} = \frac{9}{16}$$

soit

$$16\phi^2 = 9 - 9\phi^2 \text{ ou encore} \phi^2 = \frac{9}{25} = \frac{3}{5} = 0.6$$

qui est la réponse B.



$\boxed{32}$ Quelle est la valeur de l'autocovariance $\gamma(1)$?

A. 4

B. 5

C. 6

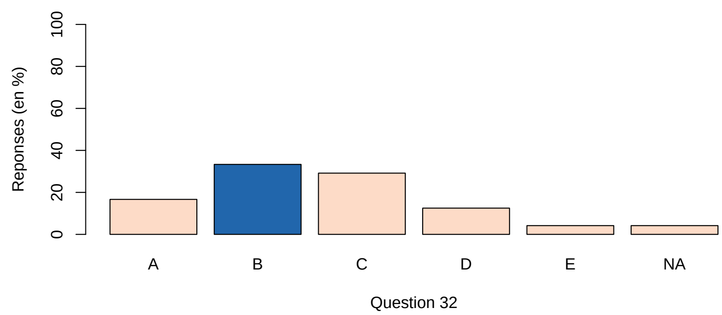
D. 7

E. 8

Je pense que le plus simple est d'utiliser la relation $\gamma(2) = \phi\gamma(1)$, i.e.

$$\gamma(1) = \frac{\gamma(2)}{\phi} = \frac{3}{3/5} = 5$$

qui est la réponse B, encore une fois.



$\boxed{33}$ On observe la série suivante

t	1	2	3	4	5
Y_t	1.2	1.1	0.9	1.3	1.5

Que vaut $\widehat{\rho}(2)$?

A. -0.50

B. -0.02

C. 0.15

D. 0.22

E. 0.52

La moyenne empirique est ici

$$\overline{Y} = \frac{1.2 + 1.1 + .9 + 1.3 + 1.5}{5} = 1.2$$

La série centrée est alors

t	1	2	3	4	5
$Y_t - \overline{Y}$	0	-0.1	-0.3	0.1	0.3

Rappelons que la corrélation est estimée par

$$\widehat{\rho}(2) = \frac{\sum_{t=1}^{3}(Y_t - \overline{Y})(Y_{t+2} - \overline{Y})}{\sum_{t=1}^{5}(Y_t - \overline{Y})^2}$$

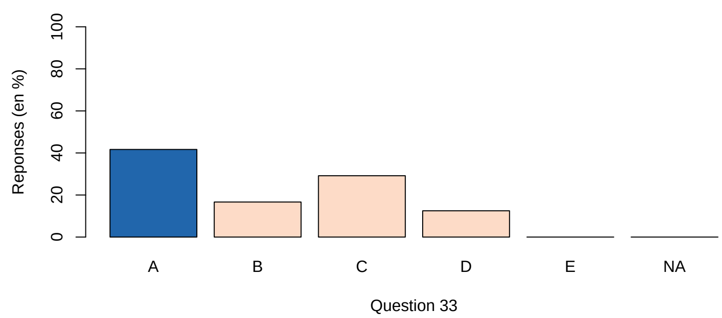
soit ici

$$\widehat{\rho}(2) = \frac{0 \cdot (-0.3) + (-0.1) \cdot 0.1 + (-0.3) \cdot 0.3}{0^2 + (-0.1)^2 + (-0.3)^2 + (0.1)^2 + (0.3)^2} = \frac{-1}{2}$$

qui est la réponse A. Maintenant, on pourrait aussi prétendre que la corrélation s'estime par

$$\widehat{\rho}(2) = \frac{\frac{1}{3}\sum_{t=1}^{3}(Y_t - \overline{Y})(Y_{t+2} - \overline{Y})}{\frac{1}{5}\sum_{t=1}^{5}(Y_t - \overline{Y})^2}$$

pour tenir compte du nombre de terme dans chacune des sommes. Il faudrait donc multiplier la valeur précédante par 5/3, ce qui donnerait -0.833 qui n'était pas proposé. Le soucis de cette correction par le nombre de terme est que l'on peut alors dépasser 1 (en valeur absolue). On va donc rester sur notre calcul précédant.



Dans les quatre questions qui suivent, on considère un processus $ARMA(1,1)$, avec $\phi = -0.5$, $\theta = +0.6$ et $\sigma_\varepsilon^2 = 4$.

$\boxed{34}$ Quelle est la valeur de $\gamma(0)$?

A. +4.053
B. +4.063
C. +4.429
D. +4.706
E. +5.099

C'est un peu toujours pareil, mais faisons le tranquillement pour éviter les erreurs de calculs. Pour rappel, on a

$$Y_t = \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$$

Si on multiplie par Y_{t-k} et que l'on prend les espérances, des deux côtés, on obtient

$$\gamma(k) = \phi\gamma(k-1) + \mathbb{E}(\varepsilon_t Y_{t-k}) + \theta\mathbb{E}(\varepsilon_{t-1} Y_{t-k})$$

Si $k \geq 2$, on retrouve la relation récursive des $AR(1)$,

$$\gamma(k) = \phi_1 \gamma(k-1)$$

Pour les calculs avec $k = 0$, notons que

$$\mathbb{E}(\varepsilon_t Y_t) = \mathbb{E}(\varepsilon_t \cdot [\phi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}])$$

or seul un terme est non nul,

$$\mathbb{E}(\varepsilon_t Y_t) = \sigma_\varepsilon^2$$

et pour l'autre terme,

$$\mathbb{E}(\varepsilon_{t-1} Y_t) = \mathbb{E}(\varepsilon_{t-1} \cdot [\phi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}])$$

or cette fois deux termes sont non nuls,

$$\mathbb{E}(\varepsilon_{t-1} Y_t) = [\phi + \theta]\sigma_\varepsilon^2$$

On obtient ainsi deux équations,

$$\gamma(0) = \phi\gamma(1) + \sigma_\varepsilon^2 + \theta[\phi + \theta]\sigma_\varepsilon^2$$

et

$$\gamma(1) = \phi\gamma(0) + 0 + \theta\sigma_\varepsilon^2$$

On a ainsi un système (linéaire) de deux équations à deux inconnues. On en déduit

$$\gamma(0) = \frac{1 + \theta^2 + 2\phi\theta}{1 - \phi^2}\sigma_\varepsilon^2$$

$$\gamma(1) = \frac{(1 + \phi\theta)(\theta + \phi)}{1 - \phi^2}\sigma_\varepsilon^2$$

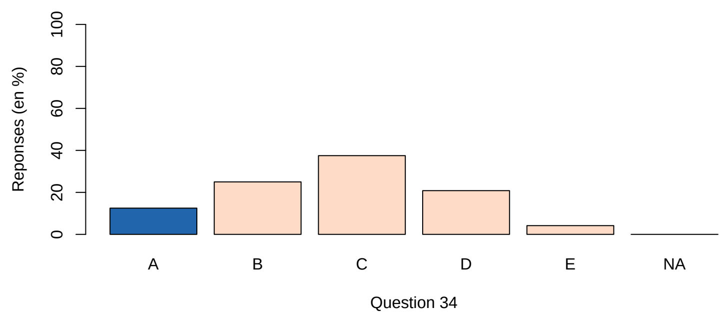
et

$$\gamma(k) = \phi\gamma(k-1)$$

pour $k \geq 2$. Numériquement, la première équation donne

$$\gamma(0) = \frac{1 + \theta^2 + 2\phi\theta}{1 - \phi_1^2}\sigma_\varepsilon^2 = \frac{1 + (0.6)^2 + 2(-0.5)(0.6)}{1 - (-0.5)^2} \cdot 4 = 4.05333$$

qui est la réponse A.



$\boxed{35}$ Quelle est la valeur de $\gamma(1)$?

A. -1.694

B. -0.438

C. $+0.373$

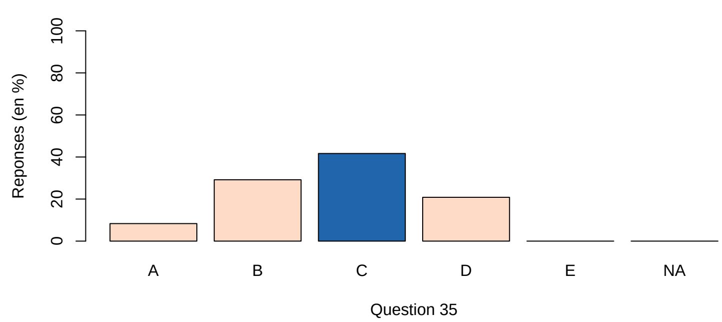
D. $+1.029$

E. $+1.670$

Si on utilise maintenant la seconde équation, on obtient, numériquement,

$$\gamma(1) = \phi\gamma(0) + 0 + \theta\sigma_\varepsilon^2 = \frac{(1+\phi\theta)(\theta+\phi)}{1-\phi_1^2}\sigma_\varepsilon^2 = 0.37333$$

qui est la réponse C.



$\boxed{36}$ Quelle est la valeur de $\rho(1)$?

A. -0.360

B. -0.108

C. $+0.092$

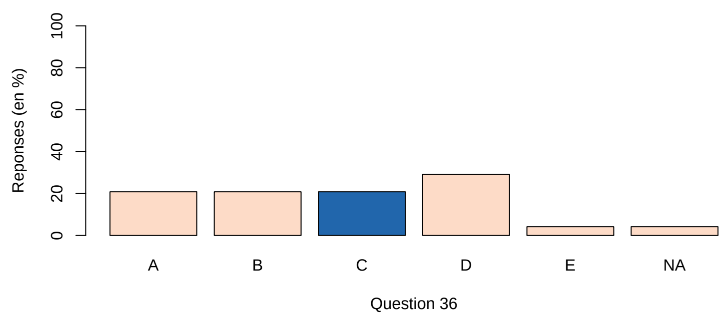
D. $+0.232$

E. $+0.328$

Compte tenu des calculs précédents,

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{1+\theta^2+2\phi\theta}{(1+\phi\theta)(\theta+\phi)} = 0.0921052$$

(je passe les calculs), et on obtient la réponse C.



$\boxed{37}$ Quelle est la valeur de $\rho(2)$?

A. -0.098

B. -0.093

C. -0.046

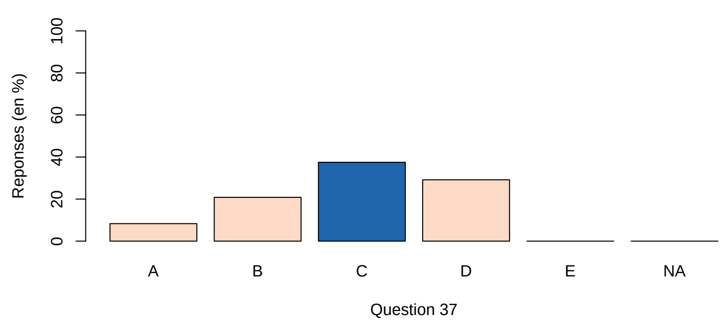
D. $+0.065$

E. $+0.252$

De $\gamma(k) = \phi\gamma(k-1)$, on en déduit que $\rho(k) = \phi\rho(k-1)$, aussi

$$\rho(2) = \phi \cdot \rho(1) = \phi\frac{1+\theta^2+2\phi\theta}{(1+\phi\theta)(\theta+\phi)}$$

soit, numériquement, -0.046052, qui est la réponse C.



Dans les deux questions qui suivent, on considère un processus $ARIMA(2,1,2)$,

$$Y_t = 1 + 1.8Y_{t-1} - 1.4Y_{t-2} + 0.6Y_{t-3} + \varepsilon_t - 0.4\varepsilon_{t1} + 0.1\varepsilon_{t-2}$$

Faisons ici les calculs. On nous dit que c'est un processus $ARIMA(2,1,2)$ autrement dit, la comopsante autoréssive est d'ordre 3, et un 1 est racine du polynôme $\Phi(\cdot)$. Aussi, on sait que

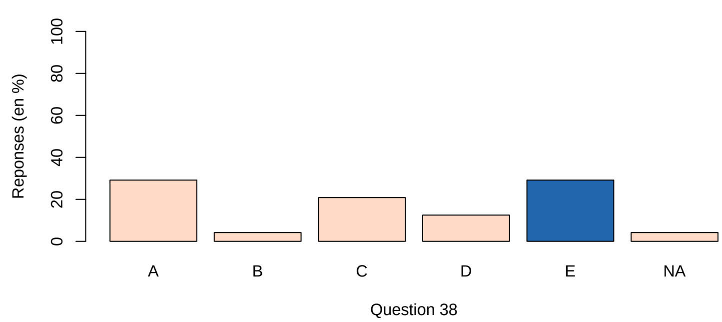
$$Y_t-(1.8Y_{t-1}-1.4Y_{t-2}+0.6Y_{t-3}) = (1-1.8L+1.4L^2+0.6L^3)Y_t = (1-L)(1-\phi_1L-\phi_2L^2)Y_t$$

Soit on sait résoudre ce genre d'équation, soit on essaye juste de trouver les seules réponses possibles. La composante en L est $-\phi_1 - 1$ qui doit valoir -1.8 et la composante en L^3 est ϕ_2 qui doit valoir -0.6.

$\boxed{38}$ Quelle est la valeur de ϕ_1 ?

A. +0.4
B. +0.5
C. +0.6
D. +0.7
E. +0.8

D'après les calculs précédants, $-\phi_1 - 1 = -1.8$ donc $\phi_1 = +0.8$, qui est la réponse E.



$\boxed{39}$ Quelle est la valeur de ϕ_2 ?

A. −0.6

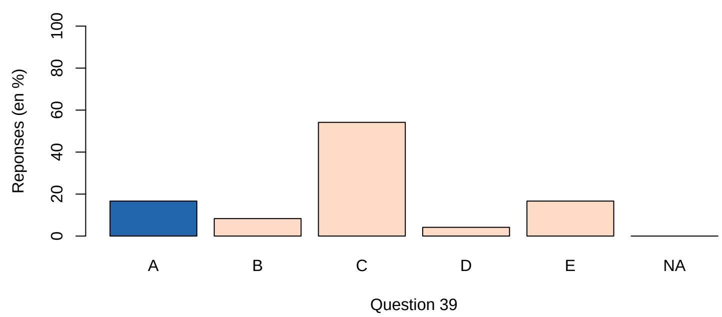
B. -0.5

C. -0.4

D. -0.3

E. -0.2

D'après les calculs précédants, $\phi_2 = -0.6$, qui est la réponse A.



Dans les deux questions qui suivent, on suppose que $Y_t = X_t + \varepsilon_t$ avec $X_t = X_{t-1} + u_t$, où les bruits blancs (ε_t) et (u_t) sont indépendants, avec $\sigma_\varepsilon^2 = 2$ et $\sigma_u^2 = 1$.

$\boxed{40}$ Quelle est la variance de ΔY_t ?

A. 2

B. 3

C. 4

D. 5

E. 6

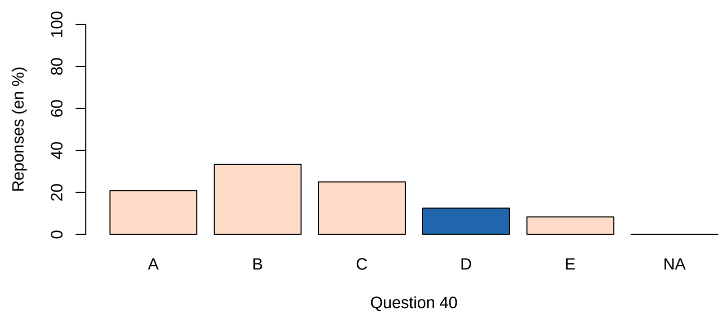
On sait que $Y_t = X_t + \varepsilon_t$ et $Y_{t-1} = X_{t-1} + \varepsilon_{t-1}$, et telle sorte que $X_{t-1} = Y_{t-1} - \varepsilon_{t-1}$. Or $X_t = X_{t-1} + u_t$. Si on substitue,

$$Y_t = X_t + \varepsilon_t = X_{t-1} + u_t + \varepsilon_t = Y_{t-1} - \varepsilon_{t-1} + u_t + \varepsilon_t$$

aussi, $\Delta Y_t = u_t - \varepsilon_{t-1} + \varepsilon_t$. Aussi,

$$\text{Var}(\Delta Y_t) = \text{Var}(u_t + \varepsilon_t - \varepsilon_{t-1}) = \text{Var}(u_t) + \text{Var}(\varepsilon_t) + \text{Var}(\varepsilon_{t-1}) = 1 + 2 + 2 = 5$$

qui est la réponse D. Notons au passage que $\mathbb{E}(\Delta Y_t) = 0$.



$\boxed{41}$ Quelle est la corrélation entre de ΔY_t et ΔY_{t-1}?

A. -0.455

B. -0.444

C. -0.429

D. -0.400

E. -0.364

Reprenons les calculs précédants,

$$\text{Cov}(\Delta Y_t, \Delta Y_{t-1}) = \mathbb{E}(\Delta Y_t \Delta Y_{t-1}) = \mathbb{E}([u_t + \varepsilon_t - \varepsilon_{t-1}] \cdot [u_{t-1} + \varepsilon_{t-1} - \varepsilon_{t-2}])$$

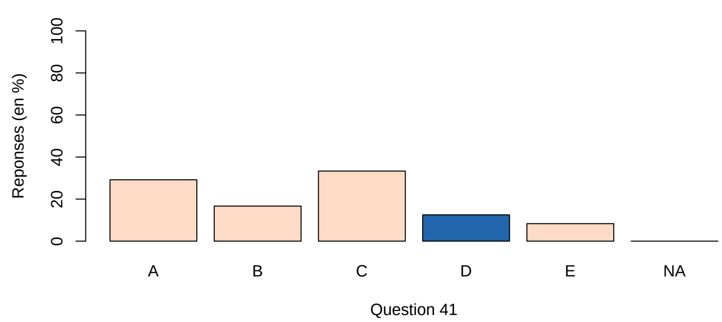
Beaucoup de termes croisés sont nuls, sauf un,

$$\text{Cov}(\Delta Y_t, \Delta Y_{t-1}) = \mathbb{E}(-\varepsilon_{t-1}^2) = -2.$$

Aussi,

$$\text{Corr}(\Delta Y_t, \Delta Y_{t-1}) = \frac{\text{Cov}(\Delta Y_t, \Delta Y_{t-1})}{\text{Var}(\Delta Y_t)} = \frac{-2}{5} = -0.4$$

C'est la réponse D. Encore.



Dans les trois questions qui suivent, on considère un processus $AR(2)$, centré, avec $\phi_1 = -0.5$ et $\phi_2 = -0.3$.

$\boxed{42}$ Quelle est la valeur de $\rho(1)$?

A. -0.714

B. -0.385

C. $+0.135$

D. $+0.385$

E. $+0.714$

Cette valeur peut s'obtenir facilement ! Pour rappel notre $AR(2)$ s'écrit

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

Si on multiplie par X_{t-1}, des deux côtés,

$$X_{t-1}X_t = \phi_1 X_{t-1}X_{t-1} + \phi_2 X_{t-1}X_{t-2} + X_{t-1}\varepsilon_t$$

et si on prend ensuite les espérances,

$$\underbrace{\mathbb{E}[X_{t-1}X_t]}_{\gamma(1)} = \phi_1 \underbrace{\mathbb{E}[X_{t-1}X_{t-1}]}_{\gamma(0)} + \phi_2 \underbrace{\mathbb{E}[X_{t-1}X_{t-2}]}_{\gamma(1)} + \underbrace{\mathbb{E}[X_{t-1}\varepsilon_t]}_{=0}$$

soit

$$\gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1)$$

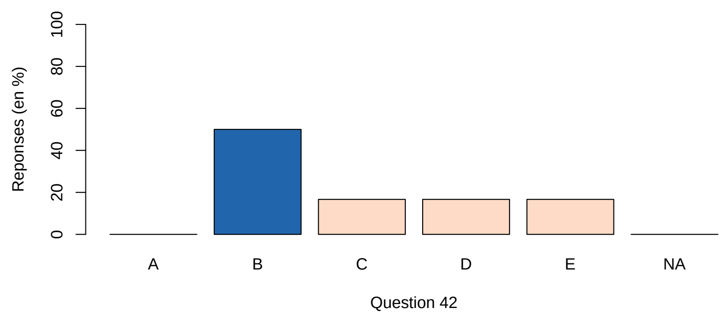
Maintenant, on ne nous demande pas les autocovariances, mais des autocorrélations, donc si on divise par $\gamma(0)$,

$$\frac{\gamma(1)}{\gamma(0)} = \phi_1 + \phi_2\frac{\gamma(1)}{\gamma(0)} \text{ i.e. } \rho(1) = \phi_1 + \phi_2\rho(1)$$

aussi,

$$\rho(1) = \frac{\phi_1}{1-\phi_2} = \frac{-0.5}{1+0.3} = -0.3846154$$

qui correspond à la réponse B.



$\boxed{43}$ Quelle est la valeur de $\rho(2)$?

A. -0.385

B. -0.300

C. -0.108

D. $+0.300$

E. $+0.657$

On va croiser les doigts, et espérer que notre réponse à la question précédante était bonne. Parce qu'on va en avoir besoin ! Pour rappel notre $AR(2)$ s'écrit

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

Cette fois, on multiplie par X_{t-2}, des deux côtés,

$$X_{t-2}X_t = \phi_1 X_{t-2}X_{t-1} + \phi_2 X_{t-2}X_{t-2} + X_{t-2}\varepsilon_t$$

et si on prend ensuite les espérances,

$$\underbrace{\mathbb{E}[X_{t-2}X_t]}_{\gamma(2)} = \phi_1 \underbrace{\mathbb{E}[X_{t-2}X_{t-1}]}_{\gamma(1)} + \phi_2 \underbrace{\mathbb{E}[X_{t-2}X_{t-2}]}_{\gamma(0)} + \underbrace{\mathbb{E}[X_{t-2}\varepsilon_t]}_{=0}$$

soit

$$\gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0)$$

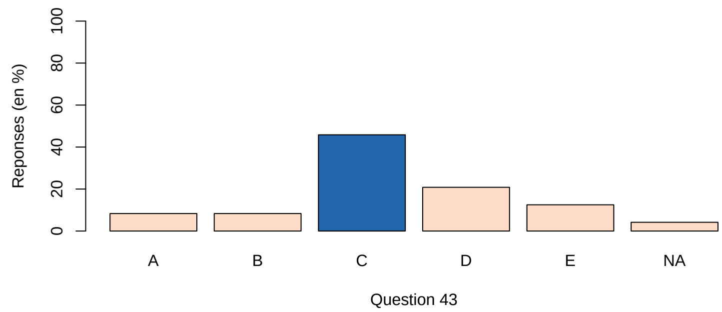
Maintenant, si on divise par $\gamma(0)$,

$$\frac{\gamma(2)}{\gamma(0)} = \phi_1\frac{\gamma(1)}{\gamma(0)} + \phi_2 \text{ i.e. } \rho(2) = \phi_1\rho(1) + \phi_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2$$

aussi,

$$\rho(2) = \frac{\phi_1^2 + \phi^2(1-\phi_2)}{1-\phi_2} = \frac{0.5^2 - 0.3\cdot 1.3}{1+0.3} = -0.1076923$$

qui correspond à la réponse C.



44 Quelle est la valeur de $\psi(2)$?

A. -0.385

B. -0.300

C. -0.108

D. $+0.300$

E. $+0.657$

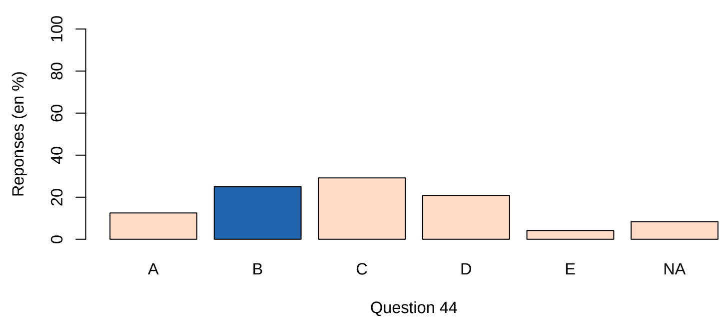
De manière générale, la formule de Bartlett nous dit que

$$\psi(2) = \frac{[\rho(2) - \rho(1)^2]}{[1 - \rho(1)^2]}$$

que l'on retrouve très facilement à partir de l'écriture matricielle. Mais si on creuse un peu plus loin (cf aussi le cours), on a

$$\psi(2) = \phi_2$$

C'est donc la réponse B.



$\boxed{45}$ Considérons 100 observation d'un processus $AR(2)$ (Y_t). On suppose que $\widehat{\gamma}(0) = 4$, $\widehat{\rho}(1) = 0.5$, $\widehat{\rho}(2) = 0.5$ et $\overline{Y} = 1$. Quelle est la valeur de σ_ε^2

A. 2.667

B. 3.267

C. 3.345

D. 3.669

E. 3.733

On sait - ce sont les équations de Yule-Walker - que

$$\begin{cases} \rho(1) = \phi_1 + \rho(1)\phi_2 \\ \rho(2) = \rho(1)\phi_1 + \phi_2 \end{cases}$$

soit numériquement, comme $\rho(1) = \rho(2) = 1/2$,

$$\begin{cases} 1 = 2\phi_1 + \phi_2 \\ 2 = 2\phi_1 + 4\phi_2 \end{cases} \text{ et } \begin{cases} 2 = 4\phi_1 + 2\phi_2 \\ 1 = \phi_1 + 2\phi_2 \end{cases}$$

aussi, $3\phi_2 = 1$ et $3\phi_1 = 1$, de telle sorte que $\phi_1 = \phi_2 = 1/3$. Pour rappel, notre $AR(2)$ s'écrit

$$Y_t = m + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

En prenant les espérances, notons que

$$\mathbb{E}(Y_t) = m + \phi_1 \mathbb{E}(Y_{t-1}) + \phi_2 \mathbb{E}(Y_{t-2})$$

donc si le processus est stationnaire (ce que l'on va toujours supposer), sa moyenne vérifie

$$\mathbb{E}(Y) = m + \frac{1}{3}\mathbb{E}(Y) + \frac{1}{3}\mathbb{E}(Y_t), \text{ i.e. } \mathbb{E}(Y) = 3m = \overline{Y} = 1,$$

donc $m = 1/3$. Si on multiplie par Y_t, des deux côtés,

$$Y_t Y_t = m Y_t + \phi_1 Y_t Y_{t-1} + \phi_2 Y_t Y_{t-2} + Y_t \varepsilon_t$$

Avant de prendre les espérances, rappelons que le dernier terme peut aussi s'écrire

$$Y_t \varepsilon_t = [\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t]\varepsilon_t$$

Si on prend maintenant les espérances

$$\mathbb{E}[Y_t Y_t] = m\mathbb{E}[Y_t] + \phi_1 \mathbb{E}[Y_t Y_{t-1}] + \phi_2 \mathbb{E}[Y_t Y_{t-2}] + 0 + \mathbb{E}[\varepsilon_t^2]$$

Soit

$$\underbrace{\mathbb{E}[Y_tY_t]-\overline{Y}^2}_{\gamma(0)}=\underbrace{m\mathbb{E}[Y_t]-\frac{1}{3}\overline{Y}^2}_{=0}+\phi_1\underbrace{(\mathbb{E}[Y_tY_{t-1}]-\overline{Y}^2)}_{\gamma(1)}+\phi_2\underbrace{(\mathbb{E}[Y_tY_{t-2}]-\overline{Y}^2)}_{\gamma(2)}+\underbrace{\mathbb{E}[\varepsilon_t^2]}_{=\sigma_\varepsilon^2}$$

Bref,

$$\sigma_\varepsilon^2=\gamma(0)-\phi_1\gamma(1)-\phi_2\gamma(2)$$

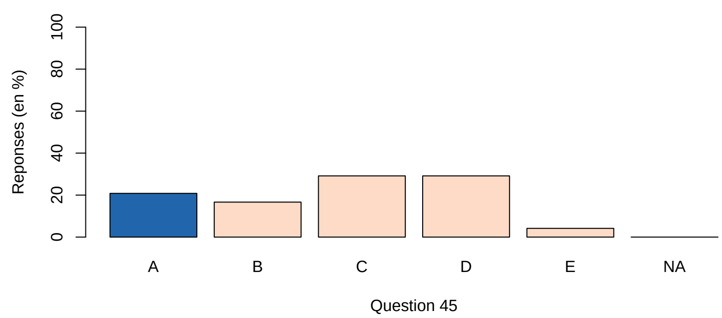
soit encore

$$\sigma_\varepsilon^2=\gamma(0)-\phi_1\rho(1)\gamma(0)-\phi_2\rho(2)\gamma(0)=\gamma(0)\left(1-\phi_1\rho(1)-\phi_2\rho(2)\right)$$

On a ici toutes les quantités mentionnées, ou calculées auparavant,

$$\sigma_\varepsilon^2=4\left(1-\frac{.5}{3}-\frac{.5}{3}\right)=\frac{4\cdot 2}{3}$$

ce qui me donne 2.66667. On obtient ici la réponse A.



Pour les deux questions suivantes, on a modélisé (Y_t) par un processus $AR(2)$. On suppose que $\overline{Y}=80$, et on suppose qu'on a les prévisions à horizon de 1 suivantes

h	1	2	3	4
${}_{t+h-1}Y_{t+h}$	80	80	80.8	80.1
Y_{t+h}	80.0	81.0	81.0	80.0

Centrons le processus, histoire de simplifier les notations

h	1	2	3	4
${}_{t+h-1}Y_{t+h}$	0	0	0.8	0.1
Y_{t+h}	0.0	1.0	1.0	0.0

Comme le processus est centré, rappelons que

$$\begin{cases} {}_tY_{t+1} = \phi_1 Y_t + \phi_2 Y_{t-1} \\ {}_{t+1}Y_{t+2} = \phi_1 Y_{t+1} + \phi_2 Y_t \\ {}_{t+2}Y_{t+3} = \phi_1 Y_{t+2} + \phi_2 Y_{t+1} \\ {}_{t+3}Y_{t+4} = \phi_1 Y_{t+3} + \phi_2 Y_{t+2} \end{cases}$$

$\boxed{46}$ Quelle est la valeur de ϕ_1?

A. 0.2

B. 0.4

C. 0.5

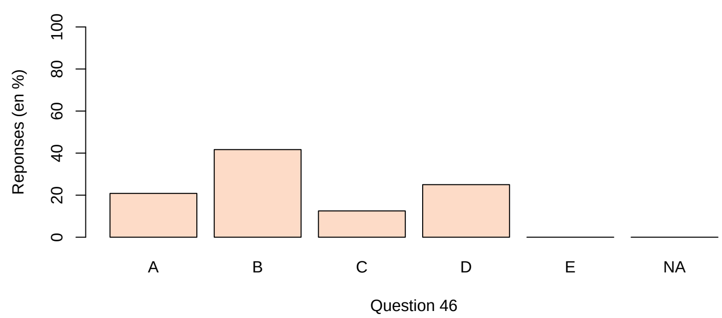
D. 0.6

E. 0.8

La première et la seconde relations ne servent à rien car Y_{t-1} et Y_t ne sont pas données. On va donc utiliser nos deux équations, pour calculer les deux inconnues.

$$\begin{cases} 0.8 = \phi_1 1 + \phi_2 0 \\ 0.1 = \phi_1 1 + \phi_2 1 \end{cases}$$

De la première équation, on en déduit $\phi_1 = 0.8$, qui est la réponse E.



$\boxed{47}$ Quelle est la valeur de ϕ_2?

A. -0.8

B. -0.7

C. -0.5

D. -0.4

E. -0.3

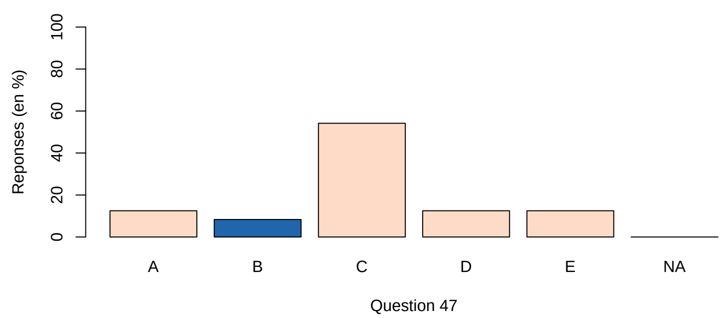
On a donc

$$\begin{cases} 0.8 = \phi_1 1 + \phi_2 0 \\ 0.1 = \phi_1 1 + \phi_2 1 \end{cases}$$

En prenant la seconde équation moins la première, on obtient

$$0.1 - 0.8 = -0.7 = \phi_2$$

qui est la réponse B.



On considère dans les trois prochains questions un modèle MA saisonnier, avec

$$\Theta(L) = (1 + \theta_1 L)(1 + \theta_{12} L^{12})$$

et $Y_t = \Theta(L)\varepsilon_t$. On suppose que $\theta_1 = 0.3$, $\theta_{12} = 0.6$ et $\sigma^2 = 4$.

Si on développe, on note que

$$\Theta(L) = 1 + \theta_1 L + \theta_{12} L^{12} + \theta_1 \theta_{12} L^{13}$$

$\boxed{48}$ Quelle est la valeur de $\gamma(0)$?

A. 5.93

B. 6.26

C. 6.42

D. 6.96

E. 7.50

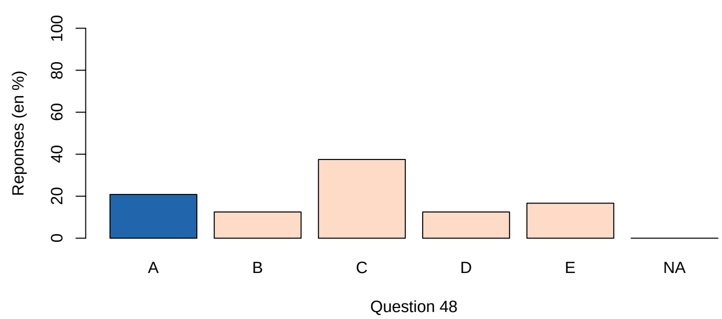
Pour calculer la variance, on se souvient que $\text{Cov}(\varepsilon_s, \varepsilon_t) = 0$ dès lors que $s \neq t$. Bref,

$$\text{Var}(Y_t) = \text{Var}(\Theta(L)\varepsilon_t) = (1 + \theta_1^2 + \theta_{12}^2 + [\theta_1\theta_{12}]^2)\sigma^2$$

soit, numériquement

$$\gamma(0) = \text{Var}(Y_t) = (1 + 0.3^2 + 0.6^2 + [0.3 \cdot 0.6]^2) \cdot 4 = 5.9296$$

qui est la réponse A.



$\boxed{49}$ Quelle est la valeur de $\rho(1)$?

A. 1.236

B. 1.326

C. 1.632

D. 2.136

E. 2.631

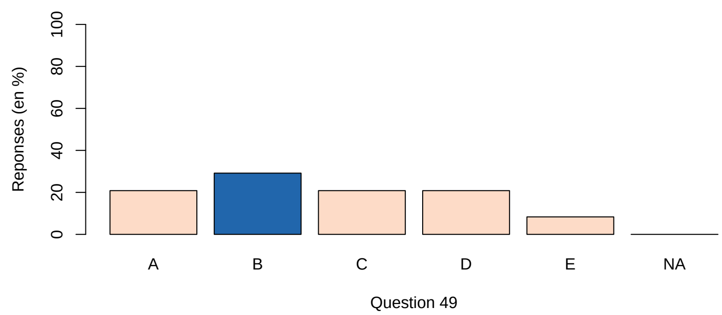
Rappelons que $\gamma(1) = \text{Cov}(Y_t, Y_{t-1})$ i.e.

$$\gamma(1) = \text{Cov}(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_{12}\varepsilon_{t-12} + \theta_1\theta_{12}\varepsilon_{t-13}, \varepsilon_{t-1} + \theta_1\varepsilon_{t-2} + \theta_{12}\varepsilon_{t-13} + \theta_1\theta_{12}\varepsilon_{t-14})$$

quand on développe, seuls les termes $\text{Cov}(\varepsilon_s, \varepsilon_s)$ vont rester, soit ici

$$\gamma(1) = \theta_1\text{Var}(\varepsilon_{t-1}) + \theta_1\theta_{12}^2\text{Var}(\varepsilon_{t-13})$$

i.e. $\gamma(1) = \theta_1[1 + \theta_{12}^2]\sigma^2$, soit, numériquement 1.632, qui est la réponse B.



$\boxed{50}$ On considère (Y_t) un processus $AR(2)$ avec $\phi_1 = 0.5$, $\phi_2 = 0.3$ de moyenne $\mu = 5$. Si $Y_1 = 5$, $Y_2 = 7$ et $Y_3 = 5$, quelle serait la valeur de ε_3 ?

A. -1.0

B. -0.8

C. -0.6

D. -0.4

E. -0.2

La relation de récurence s'écrit

$$(Y_3 - \mu) = \phi_1(Y_2 - \mu) + \phi_2(Y_1 - \mu) + \varepsilon_3$$

en centrant le processus. Numériquement, si on substitut,

$$(5 - 5) = 0.5 \cdot (7 - 5) + 0.3 \cdot (5 - 5) + \varepsilon_3$$

de telle sorte que $1 + \varepsilon_3 = 0$. C'est donc la réponse A.