Exercice 1 : Soit $(X_t)_t$ la série temporelle définie pour tout $t \in \mathbb{Z}$ par

$$X_t = m_t + s_t + \varepsilon_t$$

où $(m_t)_t$ est une tendance affine, $(\varepsilon_t)_t$ est une suite i.i.d suivant $\mathcal{N}\left(0, \sigma^2\right)$ et $s_t = u_t + v_t$, avec

$$u_t = \cos\left(\frac{2\pi}{3}t\right) \qquad \text{et} \qquad v_t = \cos\left(\frac{\pi}{3} + \frac{\pi}{2}t\right).$$

1) Déterminer les périodes de $(u_t)_t$, $(v_t)_t$ et $(s_t)_t$.
2) Montrer que $(u_t)_t$, $(v_t)_t$ et $(s_t)_t$ sont chacune de moyenne nulle sur une période.
3) Déterminer une moyenne mobile centrée et symétrique qui laisse invariante la suite $(m_t)_t$ et absorbe $(u_t)_t$ (resp. $(v_t)_t$, $(s_t)_t$).

Exercice 2 : Considérons les moyennes mobiles suivantes

$$M = \frac{1}{3}\left(I + B + B^2\right), \quad N = 2M - MM.$$

1) Montrer que N laisse invariantes les tendances linéaires.
2) Montrer que M et N absorbent la série $\left(\sin\left(\frac{2\pi}{3}t\right)\right)_t$.

Exercice 3 : On étudie la série temporelle `nottem` des températures mensuelles enregistrées à Nottingham Castle entre janvier 1920 et décembre 1939. Cette série est disponible sous R.

1) Regarder la structure de ces données, puis tracer la série à l'aide de la fonction `plot.ts`. Observe-t-on une tendance ? une saisonnalité ?
2) On tente de la désaisonnaliser par régression sur les fonctions trigonométriques, $t \mapsto \cos\left(\frac{2\pi j}{12}t\right)$ pour $j = 1, \ldots, 6$ et $t \mapsto \sin\left(\frac{2\pi j}{12}t\right)$ pour $j = 1, \ldots, 5$. Créer la base trigonométrique voulue à l'aide du code suivant :

```
f=t(as.matrix(1:6))/12
temps=as.matrix(1:length(nottem))
X=cbind(cos(2*pi*temps%*%f),sin(2*pi*temps%*%f))[,-12]
X=as.data.frame(X)
colnames(X)=c('cos1','cos2','cos3','cos4','cos5','cos6',
'sin1','sin2','sin3','sin4','sin5')
```

3) Effectuer la régression sur la base créée en utilisant la fonction `lm` puis faire un résumé (`summary`) des résultats obtenus.
4) Recommencer en ne gardant que les éléments significatifs de la base.

5) Créer une série temporelle contenant la partie saisonnière obtenue sur cette base puis une contenant les résidus de la régression. On fera appel aux commandes `$fitted` et `$residuals`.

6) Tracer les trois séries dans une même fenêtre.

7) Filtrer la série de départ pour éliminer la saisonnalité en utilisant le filtre de différenciation saisonnière via la fonction `diff`. Superposer la série ainsi obtenue avec les résidus précédents. Commenter.

8) Même question en utilisant cette fois la moyenne mobile arithmétique d'ordre 13 modifiée $M_6^*(B)$. On utilisera la fonction `filter` qui prend en argument la série à filter et un vecteur contenant les coefficients θ_i de la moyenne mobile à appliquer.

9) Pour terminer, utiliser la fonction `decompose`. Que fait-elle ? Faire un graphe de sa sortie. Commenter.

Exercice 4 : Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ un bruit blanc de variance $\sigma^2 > 0$. Etudier la stationnarité faible des processus suivants :

1) $Y_t = \varepsilon_t \cos(\omega t) + \varepsilon_{t-1} \sin(\omega t)$ où $0 \leq \omega < 2\pi$;

2) $Z_t = at + b + \varepsilon_t$ où $a, b \in \mathbb{R}$;

3) $W_t = Z_t - Z_{t-1}$ (W est le processus *différences premières* du processus Z).

Exercice 5 : Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ un bruit blanc de variance $\sigma^2 > 0$ et soit $(X_t)_{t\in\mathbb{Z}}$ le processus défini par

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}, \qquad t \in \mathbb{Z}, \qquad \theta \in \mathbb{R}^*$$

1) Calculer la fonction d'autocorrélation de $(X_t)_{t\in\mathbb{Z}}$ puis tracer son autocorrélogramme.

2) Soit $\varphi \in \mathbb{R}^*$ tel que $|\varphi| \leq 1/2$ et soit $\rho : \mathbb{Z} \to \mathbb{R}$ la fonction donnée par

$$\rho(h) = \begin{cases} 1 & \text{si} \quad h = 0, \\ \varphi & \text{si} \quad |h| = 1, \\ 0 & \text{sinon.} \end{cases}$$

De quel processus stationnaire ρ est-elle la fonction d'autocorrélation ?

Exercice 6 : Soit $(X_t)_{t\in\mathbb{Z}}$ un processus ayant la représentation MA(1)

$$X_t = \varepsilon_t - 2\,\varepsilon_{t-1}$$

où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un bruit blanc fort de variance 1 avec $\mathbb{E}(|\varepsilon_t|^3) < \infty$ et $\mathbb{E}(\varepsilon_0^3) = \mu$ non nul.
Soit $(Y_t)_{t\in\mathbb{Z}}$ le processus défini par

$$X_t = Y_t - \frac{1}{2}\,Y_{t-1}.$$

1) Montrer que le processus Y vérifie

$$Y_t = \sum_{i=0}^{+\infty} \frac{1}{2^i}\,X_{t-i}.$$

2) Vérifier que $(Y_t)_{t\in\mathbb{Z}}$ est un bruit blanc et calculer sa variance.

3) Déterminer la projection orthogonale de X_t sur $\overline{\text{Vect}}\,(X_{t-1}, X_{t-2}, \cdots)$.

4) Montrer que le processus Y vérifie

$$Y_t = \varepsilon_t - 3\sum_{i=1}^{+\infty} \frac{1}{2^i}\varepsilon_{t-i}.$$

5) Calculer $\mathbb{E}(Y_1^2 Y_2)$. Les variables Y_1 et Y_2 sont-elles indépendantes ? Interpréter ces résultats.

Exercice 7 : L'objectif de cet exercice est de calculer les autocorrélations partielles d'un processus MA(1). Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ un bruit blanc de variance σ^2

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1} \qquad t\in\mathbb{Z}, \qquad 0<|\theta|<1.$$

On souhaite calculer les coefficients $\phi_1, \ldots, \phi_n$ tels que $\sum_{j=1}^n \phi_j X_{n+1-j}$ soit la projection de X_{n+1} sur le sous-espace vectoriel de $\mathbb{L}^2$ engendré par $X_1, \ldots, X_n$.

1) Montrer que les coefficients cherchés vérifient les équations

$$\begin{cases} -\theta\phi_{j-1} + \left(1+\theta^2\right)\phi_j - \theta\phi_{j+1} = 0, \quad 2\le j\le n-1, \\ \left(1+\theta^2\right)\phi_n - \theta\phi_{n-1} = 0, \\ \left(1+\theta^2\right)\phi_1 - \theta\phi_2 = -\theta \end{cases}$$

2) Résoudre ce système et montrer que l'autocorrélation partielle à l'ordre $n\ge 1$ est donnée par

$$r(n) = -\frac{\theta^n\left(1-\theta^2\right)}{1-\theta^{2n+2}}.$$

Exercice 8 : [*Partiel 2017*] On considère la moyenne mobile $M = (I-B)\left(I-B^2\right)$ et le processus $(X_t)_t$ défini pour tout $t\in\mathbb{Z}$ par

$$X_t = a + bt + ct^2 + \cos(\pi t) + \varepsilon_t$$

où a, b et c sont des réels non nuls et $(\varepsilon_t)_t$ est un bruit blanc de variance σ^2.

1) Montrer que $M(a+bt+ct^2) = 4c$, *i.e.* M transforme les polynômes de degré 2 en une constante.

2) Montrer que M absorbe la saisonnalité $s(t) = \cos(\pi t)$.

3) Le processus $(X_t)_t$ est-il faiblement stationnaire ? Justifier.

4) On note $Z_t = M(X_t)$, le processus obtenu en appliquant la moyenne mobile M au processus $(X_t)_t$. Montrer que (Z_t) s'écrit comme un processus MA(3) dont on précisera les coefficients.

5) *Bonus :* En déduire la moyenne de $(Z_t)_t$ et sa fonction d'autocovariance γ_Z.

Exercice 9 : Soient (X_t) et (Y_t) deux processus MA définis pour tout $t \in \mathbb{Z}$ par

$$X_t = 3 + \varepsilon_t + 2\varepsilon_{t-1} \quad \text{et} \quad Y_t = \eta_t - \frac{1}{2}\eta_{t-1} + \frac{1}{3}\eta_{t-2}$$

où (ε_t) et (η_t) sont deux bruits blancs indépendants de variance respective $\sigma_\epsilon^2 = 1$ et $\sigma_\eta^2 = 2$. On pose $Z_t = X_t + Y_t$.

1) Simuler une trajectoire de X, de Y et de Z, de taille $n = 1000$.

2) A l'aide des fonctions `acf` et `pacf`, tracer les fonctions d'autocorrélation et d'autocorrélation partielle empirique de chacun des trois processus.

Exercice 10 : Soit le processus AR(1) défini, pour tout $t \in \mathbb{Z}$, par

$$X_t = \mu + \phi X_{t-1} + \varepsilon_t$$

où (ε_t) est un bruit blanc de variance σ^2.

1) Simuler une trajectoire de ce processus de taille $n = 1000$ pour $\mu = 1$, $\sigma^2 = 0.5$ et $\phi = 0.45$ puis pour $\phi = 0.95$ et enfin pour $\phi = 1$.

2) Pour chaque valeur de ϕ représenter graphiquement la trajectoire simulée.

3) A l'aide de la fonction `acf`, tracer la fonction d'autocorrelation empirique du processus. Dans quels cas est-il stationnaire ?

4) Soit $10 \leq j \leq n$. Dans les cas stationnaires, estimer μ, ϕ et σ^2 à partir de la trajectoire $X_1, \ldots, X_j$ (on utilisera la moyenne empirique et les équations de Yule-walker). Comparer graphiquement les estimations avec les vraies valeurs des paramètres en fonction de la valeur de j.

5) *Bonus* : Pour chaque $10 \leq j \leq n$, proposer un estimateur de la prévision de X_{j+1} à partir de l'observation de la trajectoire $(X_1, \ldots, X_j)$. Superposer la vraie trajectoire et sa prévision en fonction de j. Que constatez-vous selon la valeur de ϕ ?

Exercice 11 : Soit (X_t) un processus MA(2) de paramètres $m = 3$, $\theta_1 = \frac{1}{2}$ et $\theta_2 = -\frac{1}{3}$.

1) Simuler une trajectoire pour $t = 1, \ldots, 300$.

2) Utiliser la fonction `arima` pour estimer les paramètres m, θ_1 et θ_2 par maximum de vraisemblance à partir des 200 premières observations.

3) *Bonus* : Ecrire un algorithme pour prédire les valeurs de X_t pour $t = 201, \ldots, 300$ et les comparer graphiquement avec les vraies valeurs observées.

Exercice 12 : Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ un bruit blanc de variance $\sigma^2 > 0$. On définit le processus $(X_t)_{t\in\mathbb{Z}}$ par

$$X_t = \frac{1}{2}X_{t-1} - \frac{1}{4}X_{t-2} + \varepsilon_t, \qquad t \in \mathbb{Z}.$$

1) Montrer que $(\varepsilon_t)_{t\in\mathbb{Z}}$ est le processus innovation de $(X_t)_t$.

2) Montrer que la fonction d'autocovariance $\gamma(\cdot)$ vérifie la récurrence

$$\gamma(h) = \frac{1}{2}\gamma(h-1) - \frac{1}{4}\gamma(h-2), \qquad h > 0.$$

3) Exprimer $\gamma(1)$ et $\gamma(2)$ en fonction de $\gamma(0)$.

4) Résoudre l'équation de récurrence et exprimer la solution en fonction de $\gamma(0)$.

5) Calculer $\gamma(0)$ en fonction de σ^2.

6) Donner également l'expression de $(X_t)_{t\in\mathbb{Z}}$ sous la forme d'une moyenne mobile infinie.

Exercice 13 : On observe une trajectoire $(x_1, \dots, x_T)$ d'un processus $(X_t)_t$. On souhaite le modéliser par un AR(2). On cherche donc à estimer les paramètres ϕ_1 et ϕ_2.

1) Ecrire les équations de Yule-Walker correspondantes et les estimateurs qui en découlent.

2) A partir des observations, on calcule les autocovariances empiriques suivantes : $\hat{\gamma}(0) = 1,774$, $\hat{\gamma}(1) = 1,186$ et $\hat{\gamma}(2) = 0,692$. En déduire une estimation de ϕ_1 et ϕ_2.

Exercice 14 : [*Session 2 2017*] On a relevé la température à Bordeaux et à Nancy chacun des 31 jours du mois de mai 2017. On note U_t (respectivement V_t) la température à Bordeaux (resp. Nancy) le jour t avec $t = 1, 2, \dots, 31$, et $\overline{U}$ et $\overline{V}$ leur moyenne empirique respective. De plus, on a calculé :

$$\sum_{t=1}^{31} U_t \approx 714.2, \quad \sum_{t=1}^{31} V_t \approx 441.6, \quad \sum_{t=1}^{31}(U_t - \overline{U})^2 \approx 70.81, \quad \sum_{t=1}^{31}(V_t - \overline{V})^2 \approx 50.13,$$

$$\sum_{t=2}^{31}(U_{t-1} - \overline{U})(U_t - \overline{U}) \approx 32.31, \quad \sum_{t=2}^{31}(V_{t-1} - \overline{V})(V_t - \overline{V}) \approx 11.68,$$

$$\sum_{t=3}^{31}(U_{t-2} - \overline{U})(U_t - \overline{U}) \approx -15, \quad \text{et} \quad \sum_{t=3}^{31}(V_{t-2} - \overline{V})(V_t - \overline{V}) \approx -9.656.$$

1) On considère que la température du jour dépend fortement des températures de la veille et de l'avant-veille. Proposer une modélisation adéquate pour $(U_t)_t$ et $(V_t)_t$.

2) Estimer les paramètres de ces modèles par la méthode de Yule-Walker.

3) Sachant que $U_{30} = 23.51$, $U_{31} = 22.82$, $V_{30} = 15.08$ et $V_{31} = 14.28$, donner une prévision de la température le 1er juin à Bordeaux et à Nancy.

Exercice 15 : Soit (ε_t) un bruit blanc. Pour chacun des processus suivants, dire s'il s'agit d'un processus ARMA stationnaire. Si oui, déterminer les ordres p et q et préciser si le processus admet une représentation AR(∞) et/ou MA(∞) et si la représentation ARMA est minimale.

1) $X_t = \dfrac{5}{6}\, X_{t-1} - \dfrac{1}{6}\, X_{t-2} + \varepsilon_t + \varepsilon_{t-1}$

2) $\displaystyle\sum_{k=0}^{+\infty} \frac{1}{2^k} X_{t-k} = \varepsilon_t$

3) $X_t = \dfrac{5}{4}\, X_{t-1} - \dfrac{1}{4}\, X_{t-2} + \varepsilon_t + 2\,\varepsilon_{t-1}$

4) $X_t = 2X_{t-1} - X_{t-2} + \varepsilon_t + \dfrac{1}{4}\,\varepsilon_{t-1}.$

Exercice 16 : [*Examen 2015*]

On considère le processus ARMA$(1,1)$, $X_t = \frac{1}{2}X_{t-1} + \varepsilon_t - \frac{1}{4}\varepsilon_{t-1}$, où $(\varepsilon_t)_{t\in\mathbb{Z}}$ est une bruit blanc de variance σ^2.

1) La représentation ARMA est-elle minimale ?

2) Montrer que $X_t = \varepsilon_t + \displaystyle\sum_{j\geq 1} \frac{1}{2^{j+1}}\varepsilon_{t-j}$.

3) Calculer la fonction d'autocorrélation γ du processus.

4) Lorsque $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un bruit blanc gaussien, déterminer la loi du couple (X_t, X_{t+1}).

5) Donner une expression de la projection orthogonale de X_t sur le sous-espace vectoriel fermé de $\mathbb{L}^2$ engendré par $X_{t-1}, X_{t-2}, \ldots$.

Exercice 17 : On considère $(\nu_t)_t$ un bruit blanc de variance $\sigma_\nu^2 = 5/18$ et le processus $(Y_t)_t$ défini par

$$(I - 2B)Y_t = \nu_t.$$

On suppose que l'observation de Y_t est entachée d'une erreur et que l'on observe en réalité le processus $X_t = Y_t + \eta_t$ où $(\eta_t)_t$ est un autre bruit blanc décorrélé du précédent, de variance $\sigma_\eta^2 = 1/6$.

1) Montrer, en raisonnant sur la fonction d'autocovariance, que le processus $(\omega_t)_t$ défini par $\omega_t = \nu_t + \eta_t - 2\eta_{t-1}$ admet une représentation MA inversible.

2) En déduire que (X_t) admet une représentation ARMA dont on précisera les ordres. Donner sa forme minimale et son processus innovation dont on calculera la variance σ^2.

3) Déterminer l'écriture AR(∞) du processus en fonction de son innovation.

Exercice 18 : Pour chacun des processus suivants, donner l'écriture explicite :

1) $(X_t)_t \sim$ ARIMA$(1,1,1)$

2) $(X_t)_t \sim$ ARI$(2,1)$

3) $(X_t)_t \sim$ I(2)

4) $(X_t)_t \sim$ SARIMA$(2,0,1) \times (1,1,0)_7$

5) $(X_t)_t \sim$ SARMA$(2,0) \times (1,2)_{12}$

Exercice 19 : L'objectif est d'utiliser la méthode de Box-Jenkins pour modéliser les deux séries disponibles sur moodle. On importera les données de la façon suivante :

```
s1=scan("serie1.dat")
s1=ts(s1,frequency=1)
```

On commencera par regarder l'allure de la série. Si besoin, on utilisera la fonction `diff` pour la différencier. On déterminera les différents modèles possibles en utilisant l'acf et la pacf. On fera appel à la fonction `arima` pour chacun des modèles retenus. On utilisera `Box.test` pour vérifier la blancheur des résidus (`$res`) et on tracera leur acf et pacf. On comparera les différents modèles via leur AIC (`$aic`), pour choisir celui que l'on retiendra. Pour finir, on utilisera la fonction `auto.arima` de la librairie `forecast` pour comparer.

Exercice 20 : On s'intéresse aux précipitations mensuelles à San Fransisco entre 1932 et 1966, contenues dans le fichier `san_fran.txt`.

1) Utiliser la méthode de Box-Jenkins pour proposer une modélisation de type SARIMA pour cette série.

2) Créer une série d'apprentissage contenant les données jusqu'en 1963, puis une série test avec le reste.

3) Utiliser le modèle choisi pour prédire les données des 3 dernières années.

4) Même question avec un lissage exponentiel de Holt-Winters puis en faisant appel à la fonction `auto.arima` de la librairie `forecast`.

5) Comparer les trois méthodes graphiquement puis à l'aide de l'erreur quadratique moyenne.

6) Faire de la vraie prédiction avec le meilleur modèle.