

Séries temporelles univariées

M1 ECAP – TD5 – Année 2024/2025

TD5 : Evaluation de Modèles de prévision

Responsable d'enseignement : Benoît SÉVI

benoit.sevi@univ-nantes.fr

DAËRON Djayan, PAPEIX Thomas

April 3, 2025

Table of contents

1	Chargement des packages nécessaires	2
2	Chargement des données	2
3	Exercice : Réalisation de prévisions et évaluation de leurs performances	3
3.1	Visualisation de la série	3
3.2	Vérifier la stationnarité	3
3.3	Déterminer le modèle ARMA	4
3.3.1	Visualiser ACF et PACF	4
3.3.2	Utilisation de autoarima	5
3.4	Estimations en $t+1$ et $t+5$	6
3.4.1	Estimations	6
3.4.2	Représentation graphique	7
3.5	Régression Mincer-Zarnowitz pour chaque modèle	9
3.5.1	Code	9
3.5.2	Interprétation	12
3.6	Statistique de Diebold et Mariano	13
3.6.1	Code	13
3.6.2	Interprétation	16
3.7	Conclusion finale :	17

1 Chargement des packages nécessaires

```
library(tseries)
library(forecast) # utilisé
library(FinTS)
library(Metrics)
library(TSA)

library(readxl) # utilisé
library(ggplot2) # utilisé
library(dplyr) # utilisé
```

2 Chargement des données

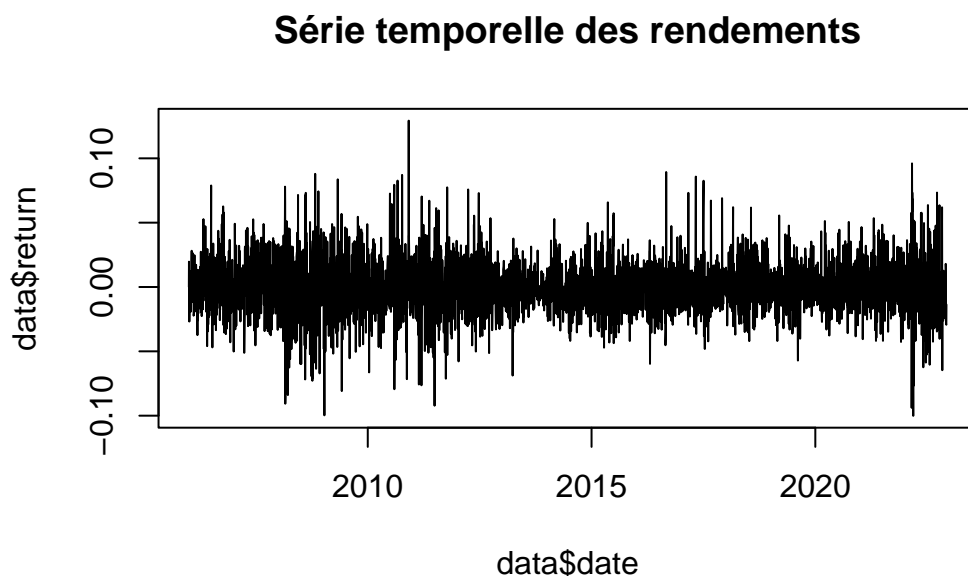
```
data <- read_excel("data/wheat_support5_STU.xlsx")

data$date <- as.Date(data$date)
```

3 Exercice : Réalisation de prévisions et évaluation de leurs performances

3.1 Visualisation de la série

```
plot(data$date, data$return, type = "l", main = "Série temporelle des rendements")
```



3.2 Vérifier la stationnarité

```
adf.test(data$return)
```

Warning in adf.test(data\$return): p-value smaller than printed p-value

Augmented Dickey-Fuller Test

data: data\$return

Dickey-Fuller = -16.304, Lag order = 16, p-value = 0.01

alternative hypothesis: stationary

Le test de Dickey-Fuller augmenté (ADF) vérifie si la série est stationnaire.

Hypothèses :

- H_0 : La série a une racine unitaire (non stationnaire).

- H1 : La série est stationnaire.

Résultats :

- Statistique de test : -16.304
- p-value : < 0.01 (très faible)

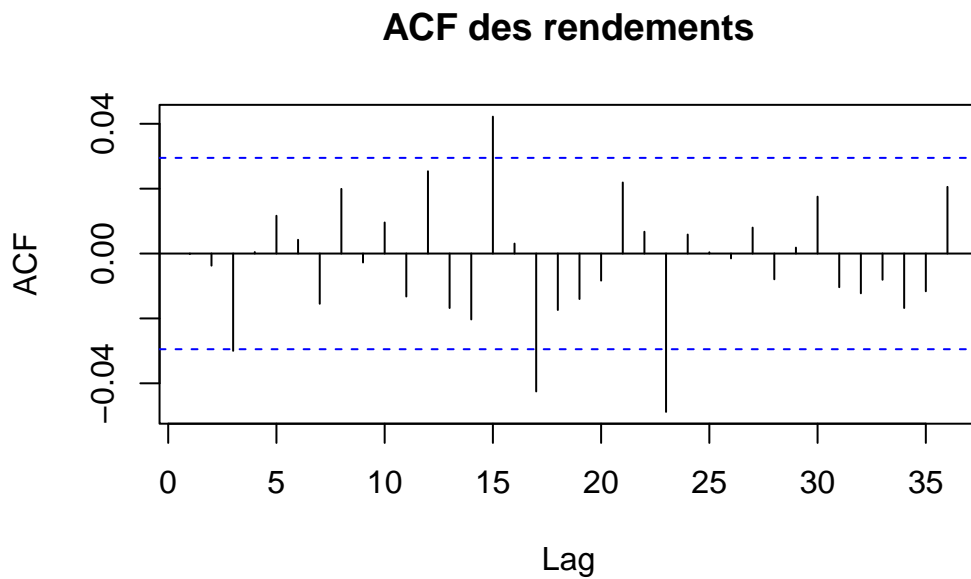
Interprétation :

- On rejette H0 à un niveau de 1%.
- La série des rendements est donc stationnaire.

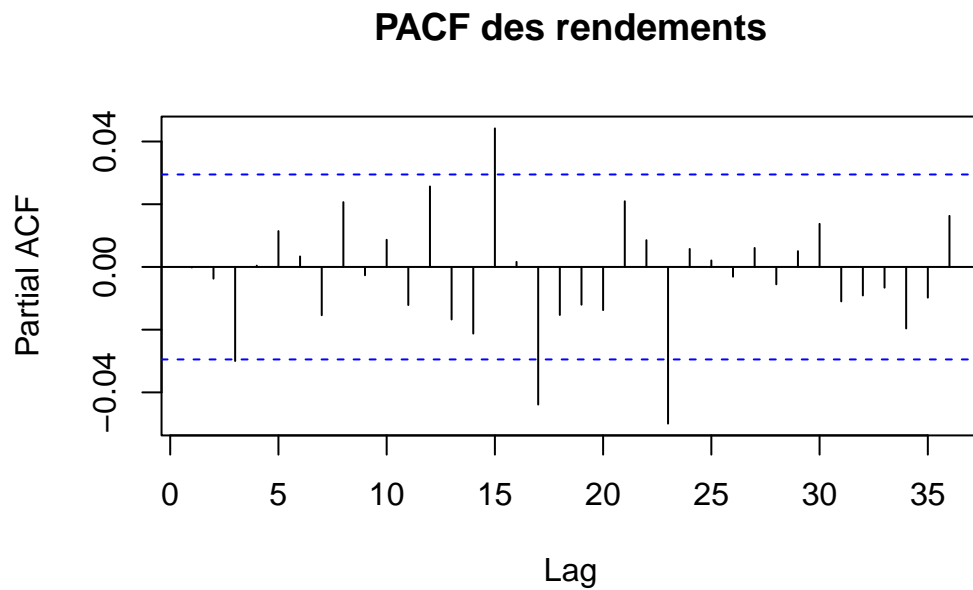
3.3 Déterminer le modèle ARMA

3.3.1 Visualiser ACF et PACF

```
acf(data$return, main = "ACF des rendements")
```



```
pacf(data$return, main = "PACF des rendements")
```



Analyse des autocorrélations :

1. ACF :

- Les coefficients sont faibles et proches de zéro à tous les retards.
- Très peu de valeurs dépassent les bandes de confiance.
- Cela suggère une absence d'autocorrélation significative.

2. PACF :

- Quelques pics isolés (notamment autour des lags 15 et 23).
- Pas de structure claire ni de décroissance caractéristique.
- Aucune justification pour inclure un terme AR important.

Conclusion :

- D'après les graphiques, la série est compatible avec un bruit blanc.
- Cela correspond à un modèle $ARMA(0,0)$, soit un processus purement aléatoire.
- En théorie financière, cela appuie l'idée que les rendements suivent un processus imprévisible.

3.3.2 Utilisation de autoarima

```
meilleur_modele <- auto.arima(data$return, seasonal = FALSE)
summary(meilleur_modele)
```

Series: data\$return

ARIMA(0,0,0) with zero mean

sigma^2 = 0.0004295: log likelihood = 10854.81

AIC=-21707.61 AICc=-21707.61 BIC=-21701.22

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.0001731801	0.02072437	0.01521257	100	100	0.6919066

ACF1

Training set -0.0001657512

Le modèle sélectionné est un ARMA(0,0), soit un bruit blanc. Cela confirme l'absence de structure significative dans les autocorrélogrammes. Toutefois, selon la consigne du professeur, nous allons imposer un AR(1).

3.4 Estimations en $t+1$ et $t+5$

3.4.1 Estimations

```
set.seed(123)

# Boucle faite main et pas par IA

resultats <- matrix(nrow = (length(data$date) - 2520), ncol = 9)
colnames(resultats) <- c("col_rang",
                        "col_realisation",
                        "col_RW",
                        "col_phi_10",
                        "col_phi_3",
                        "col_A10_1",
                        "col_A3_1",
                        "col_A10_5",
                        "col_A3_5" )

for (i in 1:(length(data$date) - 2520)) {

  rang <- 2520 + i
  RW <- data$return[rang - 1]
  realisation <- data$return[rang]

  phi_10 <- coef(arima(data$return[(i):(rang - 1)], order = c(1, 0, 0)))[1]
```

```

phi_3 <- coef(arima(data$return[(rang - 756):(rang - 1)], order = c(1, 0, 0)))[1]

A10_1 <- phi_10 * RW
A3_1 <- phi_3 * RW

A10_5 <- phi_10^5 * RW
A3_5 <- phi_3^5 * RW

resultats[i, "col_rang"] <- rang
resultats[i, "col_realisation"] <- realisation
resultats[i, "col_RW"] <- RW
resultats[i, "col_phi_10"] <- phi_10
resultats[i, "col_phi_3"] <- phi_3
resultats[i, "col_A10_1"] <- A10_1
resultats[i, "col_A3_1"] <- A3_1
resultats[i, "col_A10_5"] <- A10_5
resultats[i, "col_A3_5"] <- A3_5

}

head(resultats)

```

	col_rang	col_realisation	col_RW	col_phi_10	col_phi_3
[1,]	2521	0.011464430	-0.015084811	-0.007421575	0.010876574
[2,]	2522	0.007741974	0.011464430	-0.007562300	0.010263012
[3,]	2523	-0.015022298	0.007741974	-0.007500039	0.010693413
[4,]	2524	-0.029125702	-0.015022298	-0.007587613	0.009874042
[5,]	2525	0.005359070	-0.029125702	-0.007649498	0.011295722
[6,]	2526	0.000000000	0.005359070	-0.008061768	0.013759616

	col_A10_1	col_A3_1	col_A10_5	col_A3_5
[1,]	1.119531e-04	-1.640711e-04	3.396412e-13	-2.296152e-12
[2,]	-8.669746e-05	1.176596e-04	-2.835450e-13	1.305349e-12
[3,]	-5.806511e-05	8.278812e-05	-1.837254e-13	1.082514e-12
[4,]	1.139834e-04	-1.483308e-04	3.778003e-13	-1.409975e-12
[5,]	2.227970e-04	-3.289958e-04	7.628536e-13	-5.356071e-12
[6,]	-4.320358e-05	7.373874e-05	-1.824907e-13	2.643145e-12

3.4.2 Représentation graphique

3.4.2.1 Code

```

resultats_df <- as.data.frame(resultats)

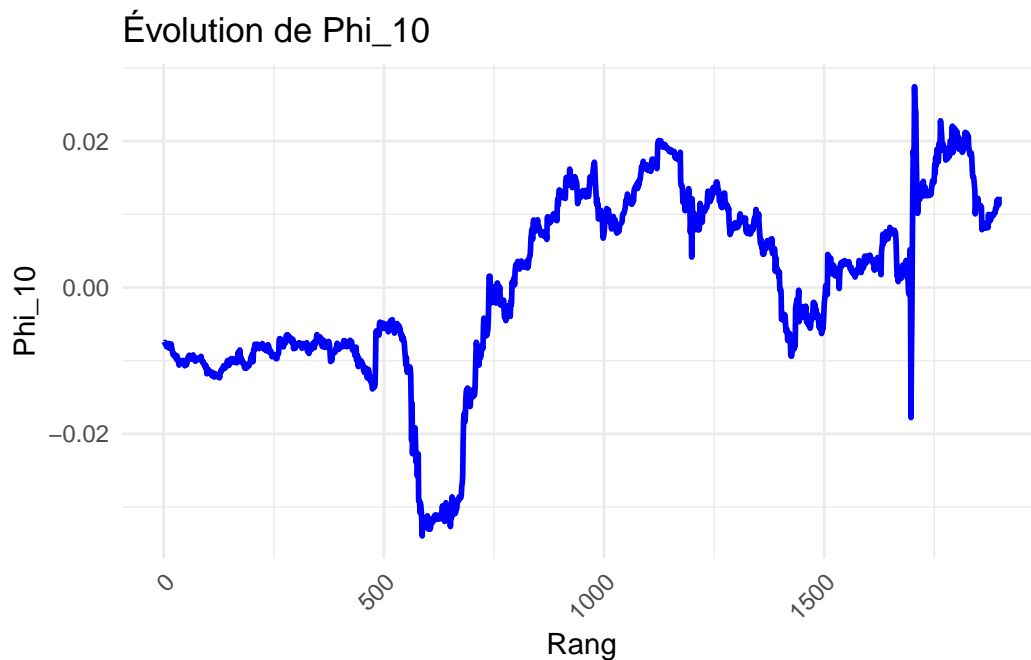
# Plot pour col_phi_10

ggplot(resultats_df, aes(x = 1:nrow(resultats_df), y = col_phi_10)) +
  geom_line(color = "blue", linewidth = 1) +

```

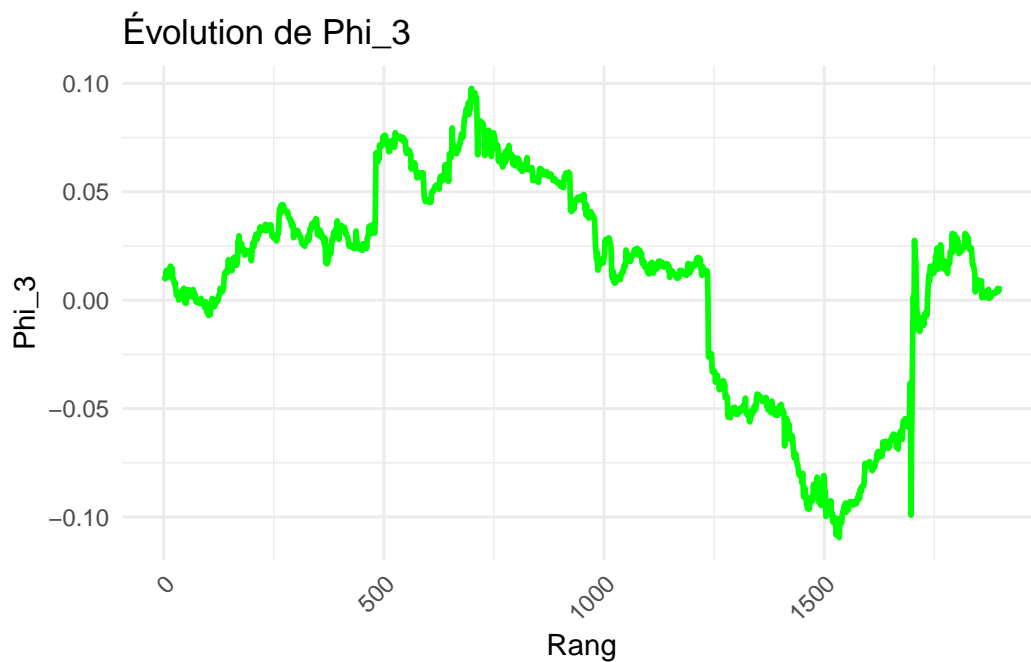


```
ggtitle("Évolution de Phi_10") +
xlab("Rang") +
ylab("Phi_10") +
theme_minimal() +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```
# Plot pour col_phi_3

ggplot(resultats_df, aes(x = 1:nrow(resultats_df), y = col_phi_3)) +
  geom_line(color = "green", linewidth = 1) +
  ggtitle("Évolution de Phi_3") +
  xlab("Rang") +
  ylab("Phi_3") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



3.4.2.2 Interprétation

1. Plus de données, moins de volatilité

- Le modèle basé sur 10 ans (graphique bleu) est plus stable, avec une amplitude réduite (-0.03 à 0.03).
- Celui basé sur 3 ans (graphique vert) présente des variations plus marquées (-0.1 à 0.1), signe d'une plus grande sensibilité.

2. Effet de lissage

- Avec plus d'années, les variations extrêmes s'atténuent, car les tendances long terme amortissent les chocs courts-termes.
- Avec seulement 3 ans, le modèle capte davantage les fluctuations locales et réagit plus fortement.

3. Impact sur la prévision

- Approche sur 10 ans : plus fiable pour capter des tendances globales mais moins réactive aux changements récents.
- Approche sur 3 ans : plus réactive, mais plus volatile, utile pour détecter des tendances rapides.

3.5 Régression Mincer-Zarnowitz pour chaque modèle

3.5.1 Code

```

mz_test_10_1 <- lm(resultats_df$col_realisation ~ resultats_df$col_A10_1)
mz_test_3_1 <- lm(resultats_df$col_realisation ~ resultats_df$col_A3_1)
mz_test_10_5 <- lm(resultats_df$col_realisation ~ resultats_df$col_A10_5)
mz_test_3_5 <- lm(resultats_df$col_realisation ~ resultats_df$col_A3_5)
mz_test_rw <- lm(col_realisation ~ col_RW, data = resultats_df)

# Affichage des résultats
summary(mz_test_10_1)

```

Call:

```
lm(formula = resultats_df$col_realisation ~ resultats_df$col_A10_1)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.106476	-0.011259	-0.000481	0.010160	0.097335

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0001998	0.0004351	0.459	0.64611
resultats_df\$col_A10_1	-4.8546783	1.7591139	-2.760	0.00584 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01895 on 1895 degrees of freedom

Multiple R-squared: 0.004003, Adjusted R-squared: 0.003477

F-statistic: 7.616 on 1 and 1895 DF, p-value: 0.005841

```
summary(mz_test_3_1)
```

Call:

```
lm(formula = resultats_df$col_realisation ~ resultats_df$col_A3_1)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.100373	-0.011284	-0.000357	0.010445	0.097584

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0002148	0.0004360	0.493	0.622
resultats_df\$col_A3_1	-0.2997125	0.4517002	-0.664	0.507

Residual standard error: 0.01898 on 1895 degrees of freedom

Multiple R-squared: 0.0002323, Adjusted R-squared: -0.0002953

F-statistic: 0.4403 on 1 and 1895 DF, p-value: 0.5071

```
summary(mz_test_10_5)
```

Call:

```
lm(formula = resultats_df$col_realisation ~ resultats_df$col_A10_5)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.100972	-0.011268	-0.000225	0.010450	0.095850

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.148e-04	4.358e-04	0.493	0.622
resultats_df\$col_A10_5	-4.144e+06	3.833e+06	-1.081	0.280

Residual standard error: 0.01898 on 1895 degrees of freedom

Multiple R-squared: 0.0006164, Adjusted R-squared: 8.907e-05

F-statistic: 1.169 on 1 and 1895 DF, p-value: 0.2798

```
summary(mz_test_3_5)
```

Call:

```
lm(formula = resultats_df$col_realisation ~ resultats_df$col_A3_5)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.100360	-0.011333	-0.000224	0.010459	0.096807

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.108e-04	4.358e-04	0.484	0.629
resultats_df\$col_A3_5	-1.143e+04	8.433e+03	-1.356	0.175

Residual standard error: 0.01898 on 1895 degrees of freedom

Multiple R-squared: 0.0009692, Adjusted R-squared: 0.000442

F-statistic: 1.838 on 1 and 1895 DF, p-value: 0.1753

```
summary(mz_test_rw)
```

Call:

```
lm(formula = col_realisation ~ col_RW, data = resultats_df)
```

Residuals:

Min	1Q	Median	3Q	Max
-----	----	--------	----	-----

-0.099500 -0.011320 -0.000403 0.010412 0.097039

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0002208	0.0004359	0.506	0.613
col_RW	0.0127203	0.0229692	0.554	0.580

Residual standard error: 0.01898 on 1895 degrees of freedom

Multiple R-squared: 0.0001618, Adjusted R-squared: -0.0003658

F-statistic: 0.3067 on 1 and 1895 DF, p-value: 0.5798

3.5.2 Interprétation

1. Régression MZ pour le modèle A10 à horizon 1 :

- Coefficient de l'Intercept (α) = 0.0001998 : L'intercept est proche de zéro, ce qui est un bon signe.
- Coefficient pour col_A10_1 (β) = -4.8546783 : Ce coefficient est significativement différent de 1 (p-value = 0.00584). Cela suggère que le modèle A10 à horizon 1 est biaisé et ne fournit pas de bonne prévision.
- R^2 ajusté = 0.003477 : La valeur très faible de R^2 montre que ce modèle A10 n'explique pas bien la variation des réalisations.

2. Régression MZ pour le modèle A3 à horizon 1 :

- Coefficient de l'Intercept (α) = 0.0002148 : L'intercept est proche de zéro.
- Coefficient pour col_A3_1 (β) = -0.2997125 : Ce coefficient est non significativement différent de 1 (p-value = 0.507). Cela signifie que le modèle A3 à horizon 1 n'est pas meilleur qu'une prédiction aléatoire, et donc il n'est pas performant dans ce cas.
- R^2 ajusté = -0.0002953 : Très faible, ce qui indique que ce modèle n'explique pratiquement aucune variation.

3. Régression MZ pour le modèle A10 à horizon 5 :

- Coefficient de l'Intercept (α) = 2.148e-04 : Très proche de zéro.
- Coefficient pour col_A10_5 (β) = -4.144e+06 : Ce coefficient est non significativement différent de 1 (p-value = 0.280). Cela indique une prédiction biaisée et très éloignée de la réalité, ce qui montre que le modèle A10 à horizon 5 est très mauvais.
- R^2 ajusté = 8.907e-05 : Très faible, ce modèle ne parvient pas à expliquer la variation des données.

4. Régression MZ pour le modèle A3 à horizon 5 :

- Coefficient de l'Intercept (α) = 2.108e-04 : L'intercept est proche de zéro.

- Coefficient pour `col_A3_5` (β) = -1.143e+04 : Ce coefficient est non significativement différent de 1 (p-value = 0.175). Ce modèle A3 à horizon 5 n'est pas plus performant que la marche aléatoire.
- R^2 ajusté = 0.000442 : Très faible, ce modèle est encore moins performant.

5. Régression MZ pour la marche aléatoire :

- Coefficient de l'Intercept (α) = 0.0002208 : L'intercept est proche de 0.
- Coefficient pour `col_RW` (β) = 0.0127203 : Ce coefficient est très loin de 1, avec une p-value de 0.580, ce qui montre que la marche aléatoire est aussi un mauvais modèle et ne prédit pas mieux que les autres modèles AR dans ce cas.
- R^2 ajusté = -0.0003658 : Très faible, ce modèle n'explique pas bien la variation des données.

Conclusion :

- Aucune des régressions AR(1) (10 et 3) ni la marche aléatoire ne montre de modèle performant.
- Les modèles AR avec 10 et 3 ans de données à horizon 1 et horizon 5 présentent des coefficients biaisés et non significatifs, ce qui signifie qu'ils ne sont pas adaptés aux données.
- La marche aléatoire (RW) a aussi un coefficient loin de 1, ce qui suggère que les prévisions basées sur la marche aléatoire ne sont pas plus performantes que les modèles AR.
- **Le “moins pire” semble être le modèle avec 10 ans de données à horizon 1**, bien qu'il ne soit pas optimal, car il a un coefficient significatif et une p-value relativement faible. Cependant, son faible R^2 montre qu'il n'est pas performant dans l'ensemble.

3.6 Statistique de Diebold et Mariano

3.6.1 Code

```
# Calcul des erreurs pour chaque modèle

erreur_A10_1 <- resultats_df$col_realisation - resultats_df$col_A10_1
erreur_A3_1 <- resultats_df$col_realisation - resultats_df$col_A3_1
erreur_A10_5 <- resultats_df$col_realisation - resultats_df$col_A10_5
erreur_A3_5 <- resultats_df$col_realisation - resultats_df$col_A3_5
erreur_RW <- resultats_df$col_realisation - resultats_df$col_RW

# Calcul des pertes (MSE et MAE) pour chaque observation
```

```

mse_A10_1 <- erreur_A10_1^2
mse_A3_1 <- erreur_A3_1^2
mse_A10_5 <- erreur_A10_5^2
mse_A3_5 <- erreur_A3_5^2
mse_RW <- erreur_RW^2

mae_A10_1 <- abs(erreur_A10_1)
mae_A3_1 <- abs(erreur_A3_1)
mae_A10_5 <- abs(erreur_A10_5)
mae_A3_5 <- abs(erreur_A3_5)
mae_RW <- abs(erreur_RW)

# Différences des pertes pour chaque observation

diff_mse_A10_1 <- mse_A10_1 - mse_RW
diff_mse_A3_1 <- mse_A3_1 - mse_RW
diff_mse_A10_5 <- mse_A10_5 - mse_RW
diff_mse_A3_5 <- mse_A3_5 - mse_RW

diff_mae_A10_1 <- mae_A10_1 - mae_RW
diff_mae_A3_1 <- mae_A3_1 - mae_RW
diff_mae_A10_5 <- mae_A10_5 - mae_RW
diff_mae_A3_5 <- mae_A3_5 - mae_RW

# Moyenne et écart-type des différences pour chaque modèle

mean_diff_mse_A10_1 <- mean(diff_mse_A10_1)
sd_diff_mse_A10_1 <- sd(diff_mse_A10_1)

mean_diff_mse_A3_1 <- mean(diff_mse_A3_1)
sd_diff_mse_A3_1 <- sd(diff_mse_A3_1)

mean_diff_mse_A10_5 <- mean(diff_mse_A10_5)
sd_diff_mse_A10_5 <- sd(diff_mse_A10_5)

mean_diff_mse_A3_5 <- mean(diff_mse_A3_5)
sd_diff_mse_A3_5 <- sd(diff_mse_A3_5)

mean_diff_mae_A10_1 <- mean(diff_mae_A10_1)
sd_diff_mae_A10_1 <- sd(diff_mae_A10_1)

mean_diff_mae_A3_1 <- mean(diff_mae_A3_1)
sd_diff_mae_A3_1 <- sd(diff_mae_A3_1)

mean_diff_mae_A10_5 <- mean(diff_mae_A10_5)

```

```

sd_diff_mae_A10_5 <- sd(diff_mae_A10_5)

mean_diff_mae_A3_5 <- mean(diff_mae_A3_5)
sd_diff_mae_A3_5 <- sd(diff_mae_A3_5)

# Calcul de la statistique DM (pour MSE)

d_mse_A10_1 <- mean_diff_mse_A10_1 / sd_diff_mse_A10_1
d_mse_A3_1 <- mean_diff_mse_A3_1 / sd_diff_mse_A3_1
d_mse_A10_5 <- mean_diff_mse_A10_5 / sd_diff_mse_A10_5
d_mse_A3_5 <- mean_diff_mse_A3_5 / sd_diff_mse_A3_5

# Calcul de la statistique DM (pour MAE)

d_mae_A10_1 <- mean_diff_mae_A10_1 / sd_diff_mae_A10_1
d_mae_A3_1 <- mean_diff_mae_A3_1 / sd_diff_mae_A3_1
d_mae_A10_5 <- mean_diff_mae_A10_5 / sd_diff_mae_A10_5
d_mae_A3_5 <- mean_diff_mae_A3_5 / sd_diff_mae_A3_5

# Affichage des résultats

print(paste("Diebold-Mariano pour MSE (A10_1) : ", d_mse_A10_1))

```

```
[1] "Diebold-Mariano pour MSE (A10_1) : -0.26957436870972"
```

```
print(paste("Diebold-Mariano pour MSE (A3_1) : ", d_mse_A3_1))
```

```
[1] "Diebold-Mariano pour MSE (A3_1) : -0.266176001851699"
```

```
print(paste("Diebold-Mariano pour MSE (A10_5) : ", d_mse_A10_5))
```

```
[1] "Diebold-Mariano pour MSE (A10_5) : -0.269415501772725"
```

```
print(paste("Diebold-Mariano pour MSE (A3_5) : ", d_mse_A3_5))
```

```
[1] "Diebold-Mariano pour MSE (A3_5) : -0.269415386828131"
```

```
print(paste("Diebold-Mariano pour MAE (A10_1) : ", d_mae_A10_1))
```

```
[1] "Diebold-Mariano pour MAE (A10_1) : -0.386592361638786"
```



```
print(paste("Diebold-Mariano pour MAE (A3_1) : ", d_mae_A3_1))
```

```
[1] "Diebold-Mariano pour MAE (A3_1) : -0.385899542093993"
```

```
print(paste("Diebold-Mariano pour MAE (A10_5) : ", d_mae_A10_5))
```

```
[1] "Diebold-Mariano pour MAE (A10_5) : -0.386733962376636"
```

```
print(paste("Diebold-Mariano pour MAE (A3_5) : ", d_mae_A3_5))
```

```
[1] "Diebold-Mariano pour MAE (A3_5) : -0.386733811221832"
```

3.6.2 Interprétation

1. MSE (Mean Squared Error) :

- A10_1 : La statistique DM pour MSE est de -0.2696.
- A3_1 : La statistique DM pour MSE est de -0.2662.
- A10_5 : La statistique DM pour MSE est de -0.2694.
- A3_5 : La statistique DM pour MSE est de -0.2694.

Ces valeurs de la statistique DM pour tous les modèles sont négatives et proches de zéro, ce qui suggère que, par rapport à la marche aléatoire (RW), aucun modèle ne présente une amélioration significative en termes de MSE. En d'autres termes, la performance des modèles pour prédire les rendements, mesurée par l'erreur quadratique moyenne, n'est pas meilleure que celle de la marche aléatoire. Une statistique DM proche de zéro indique qu'il n'y a pas de différence significative entre les modèles et la marche aléatoire, ce qui peut signifier qu'ils ne sont pas réellement utiles pour faire des prévisions précises.

2. MAE (Mean Absolute Error) :

- A10_1 : La statistique DM pour MAE est de -0.3866.
- A3_1 : La statistique DM pour MAE est de -0.3859.
- A10_5 : La statistique DM pour MAE est de -0.3867.
- A3_5 : La statistique DM pour MAE est de -0.3867.

De manière similaire, les valeurs négatives de la statistique DM pour le MAE sont proches de zéro, ce qui signifie que les modèles n'améliorent pas de manière significative la précision des prévisions par rapport à la marche aléatoire en termes d'erreur absolue. Tout comme pour le MSE, cela indique que les modèles n'ont pas réussi à dépasser la marche aléatoire en termes de précision dans la prédiction des rendements.

Conclusion : Les résultats de la statistique de Diebold-Mariano suggèrent que, sur la base des MSE et MAE, aucun des modèles (A10_1, A3_1, A10_5, A3_5) ne surperforme la marche aléatoire. En d'autres termes, les modèles ne sont pas suffisamment performants pour surpasser la marche aléatoire dans la prédiction des rendements financiers.

Cela peut aussi indiquer que les prédictions basées sur ces modèles sont relativement similaires à celles de la marche aléatoire, ou que les modèles sont trop simples ou mal ajustés pour capturer les dynamiques complexes des rendements.

3.7 Conclusion finale :

Après avoir examiné les résultats des tests de Mincer-Zarnowitz et de la statistique de Diebold-Mariano, il est évident qu'aucun modèle ne surpasse de manière significative la marche aléatoire en termes de précision de prédiction.

- Dans les tests de Mincer-Zarnowitz, le modèle A10 à horizon 1 (A10_1) a montré une signification statistique (p-value faible), ce qui pourrait suggérer une relation entre les prédictions du modèle et les rendements réels. Cependant, le R^2 très faible indique que la performance du modèle reste limitée.
- Quant à la statistique de Diebold-Mariano, les résultats pour les MSE et MAE pour tous les modèles (A10_1, A3_1, A10_5, A3_5) indiquent que les modèles n'offrent pas d'amélioration significative par rapport à la marche aléatoire. Les statistiques de Diebold-Mariano sont proches de zéro, ce qui suggère que les modèles ne surclassent pas la marche aléatoire en termes d'erreur quadratique moyenne ou d'erreur absolue.

Le modèle AR(1) avec 10 ans de données à horizon 1 (A10_1) semble être le "meilleur" parmi les modèles testés, bien qu'il ne soit pas optimal. Il montre une relation statistiquement significative avec les rendements réels, mais avec un R^2 faible. Cependant, tous les modèles restent globalement peu performants par rapport à la marche aléatoire.

En conclusion, bien que le(A10_1) puisse offrir un léger avantage par rapport aux autres, aucun modèle ne semble suffisamment performant pour surclasser la marche aléatoire dans la prédiction des rendements financiers.