

Value at Risk : Projet 1

Matthieu DHENAIN

Octobre 2023

Contents

1	Introduction	2
2	Cours et rendement	2
3	Propriétés	3
3.1	Asymétrie Perte et Gain	3
3.2	Queues de distribution épaisses	3
3.3	Autocorrélations des carrés des rendements fortes et faibles pour les rendements	4
3.3.1	Corrélogrammes	4
3.3.2	La statistique de Ljung-Box	4
3.4	Clusters de volatilité	7
3.5	Queues épaisses conditionnelles	7
3.6	Effet de levier	9
3.7	Saisonnalité	9
3.8	Stationnarité	11
4	Conclusion	14
5	Annexe	15
5.1	Résultats des tests d'ARCH pour les clusters de volatilité	15
5.2	Résultats des tests d'ARCH sur les résidus du GARCH(1,1)	17
5.3	Résultats des tests d'ARCH sur les résidus du GARCH(1,15)	20
5.4	Dickey-Fuller	22
5.4.1	test de RU quand il existe une constante mais pas de tendance dans la série	23
6	Annexe 2 : Comparaison avec rtt	24
6.1	Code et Output R avec rtt	24

1 Introduction

AmBev, ou *Companhia de Bebidas das Américas* (en français : compagnie des boissons des Amériques) est le plus gros brasseur d'Amérique du Sud, notamment du marché émergent qu'est le Brésil. Elle est cotée sur le marché de São Paulo. L'étude de cette valeur m'intéresse, car il s'agit d'une des principales valeurs de la bourse de São Paulo et j'envisage de travailler au Brésil.

2 Cours et rendement

Au cours de cette étude, le rendement logarithmique r_t de l'action AmBev sera étudié. Il est défini comme $r_t = \log(p_t) - \log(p_{t-1})$, où p_t représente le cours de l'action le jour t . Par la suite on notera rte la série des rendements s'étendant du 04/01/2010 au 31/12/2018. La figure 1 montre que le cours de l'action AmBev a une tendance croissante tandis que ses variations et son rendement logarithmique fluctuent autour de 0 et restent assez constants dans le temps. Il semble y avoir des paquets de volatilité en 2013 et en 2020, car on observe des zones avec plus de volatilité ces années là.

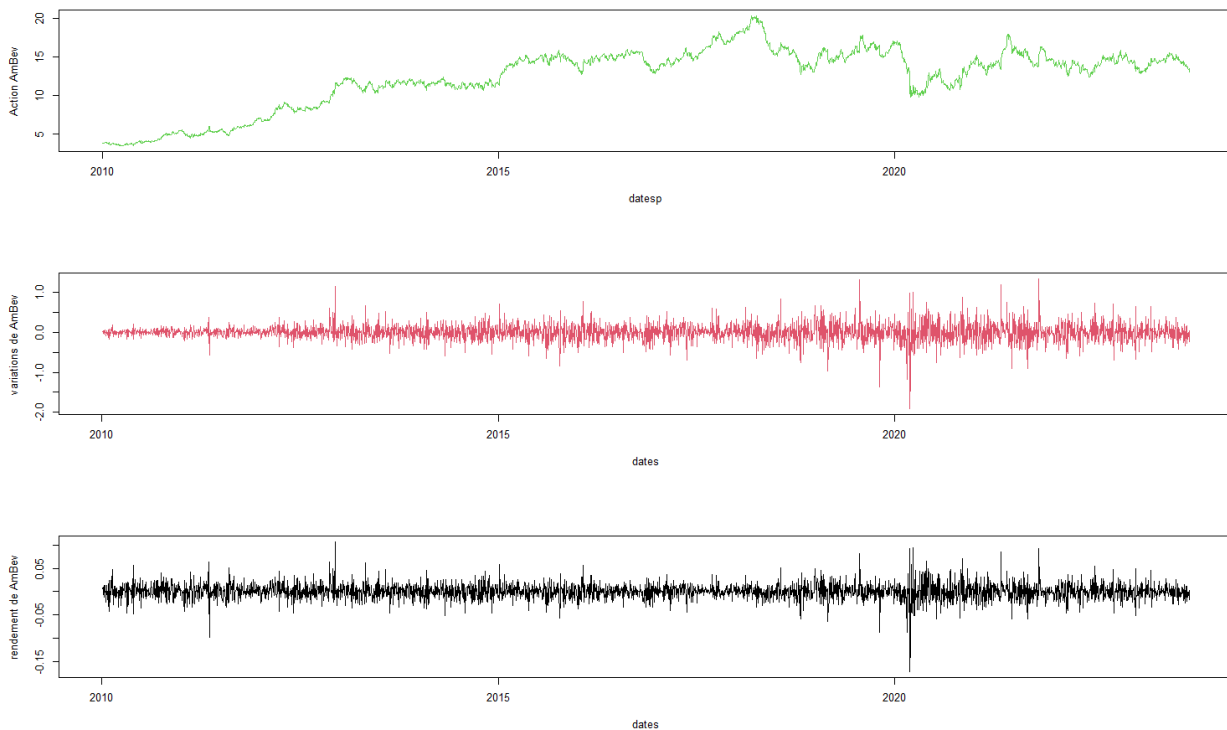


Figure 1: Valeurs, variations et rendement logarithmique de l'action AmBev

On cherche ensuite si la moyenne empirique (très proche de 0) est statistiquement nulle. C'est-à-dire que l'on cherche à déterminer si l'espérance du processus qui a généré rte , notée μ est nulle. Pour cela on teste $H_0 : E(rte_t) = \mu = 0$ versus $E(rte_t) = \mu \neq 0$. Alors on calcule la statistique t , tel que $t = \frac{rte}{\sigma\sqrt{T}}$.

Ici, on considère un seuil de risque $\alpha = 5\%$ et on a $T = 2232 > 200$, donc on peut raisonnablement affirmer que la loi de Student à $T-1$ degrés de liberté suivit par t est confondu avec la loi normale, dont la valeur tabulée est 1.96. La statistique t de l'action AmBev est de 1.895. Comme $|1.89| < 1.96$, on accepte H_0 et on conclut que l'espérance du processus générateur de rte_t est nulle.

3 Propriétés

3.1 Asymétrie Perte et Gain

On cherche à étudier la répartition des rendements entre perte et gain. Pour cela on fait un test de l'hypothèse de symétrie pour vérifier la nullité de la *skewness*, défini comme le moment centré d'ordre 3 normalisé de la distribution. C'est à dire que l'on teste l'hypothèse nulle H_0 telle que :

$$H_0 : E[(\frac{X - E[X]}{\sigma_X})^3] = 0$$

et H_a à l'inverse :

$$H_a : E[(\frac{X - E[X]}{\sigma_X})^3] \neq 0$$

```
> agostino.test(rte)
```

```
D'Agostino skewness test
data:  rte
skew = 0.11743, z = 2.26479, p-value = 0.02353
alternative hypothesis: data have a skewness
```

Pour cela on effectue un test d'Agostino sur les données *rte*. On trouve une p-value de $0.024 < 0.05$, donc on rejette l'hypothèse nulle. Cela signifie que les données ont une *skewness* significative. De plus, le test nous montre qu'elle est positive (*skewness* = 0.11743). On peut en déduire que de nombreux rendements sont négatifs, mais de faibles amplitudes et quelques rendements positifs ont de fortes valeurs. La probabilité de gains est inférieure à la probabilité de pertes mais les pertes sont faibles et les (rares) gains sont élevés.

3.2 Queues de distribution épaisses

Il s'agit ici d'étudier la fréquence d'apparition des événements extrêmes. Pour cela on compare l'épaisseur des queues de distribution des rendements logarithmiques à celles d'une loi normale, dont le *kurtosis* est de 3. En pratique on calcule le *kurtosis*, c'est à dire le moment centré d'ordre 4 d'une variable aléatoire : $\mu_4 = E[(X - \mu)^4]$, à l'aide de la statistique d'Anscombe modifiée par Agostino et Zar. L'hypothèse nulle H_0 de ce test est :

$$H_0 : E[(\frac{X - E[X]}{\sigma_X})^4] = 3$$

A l'inverse,

$$H_a : E[(\frac{X - E[X]}{\sigma_X})^4] \neq 3$$

```
> anscombe.test(rte)
```

```
Anscombe-Glynn kurtosis test
data:  rte
kurt = 6.5182, z = 13.7170, p-value < 2.2e-16
alternative hypothesis: kurtosis is not equal to 3
```

Dans le cas des rendements de AmBev, on a bien une distribution leptokurtique des rendements de l'action AmBev, puisque le kurtosis calculé est de 6.15 (> 3) et significatif, car la p-valeur du test d'Anscombe est inférieure à $2.2e-16$ (< 0.05).

Cela signifie que les rendements extrêmes, à la hausse comme à la baisse, sont plus fréquents que si leur répartition suivait une loi normale.

3.3 Autocorrélations des carrés des rendements fortes et faibles pour les rendements

Afin d'étudier l'autocorrélation de la série rte , on détermine les ratios des coefficients d'autocorrélations totales définis pour un retard k comme :

$$\rho(k) = \frac{\text{cov}(rte_t, rte_{t-k})}{\text{cov}(rte_t, rte_t)}$$

On trace ensuite la suite des ρ qui forment la fonction d'autocorrélation. Si la suite des rte est indépendamment et identiquement distribuée et que l'espérance des rendements au carré est finie alors $\hat{\rho}(k)$ suit asymptotiquement la loi normale centrée réduite.

3.3.1 Corrélogrammes

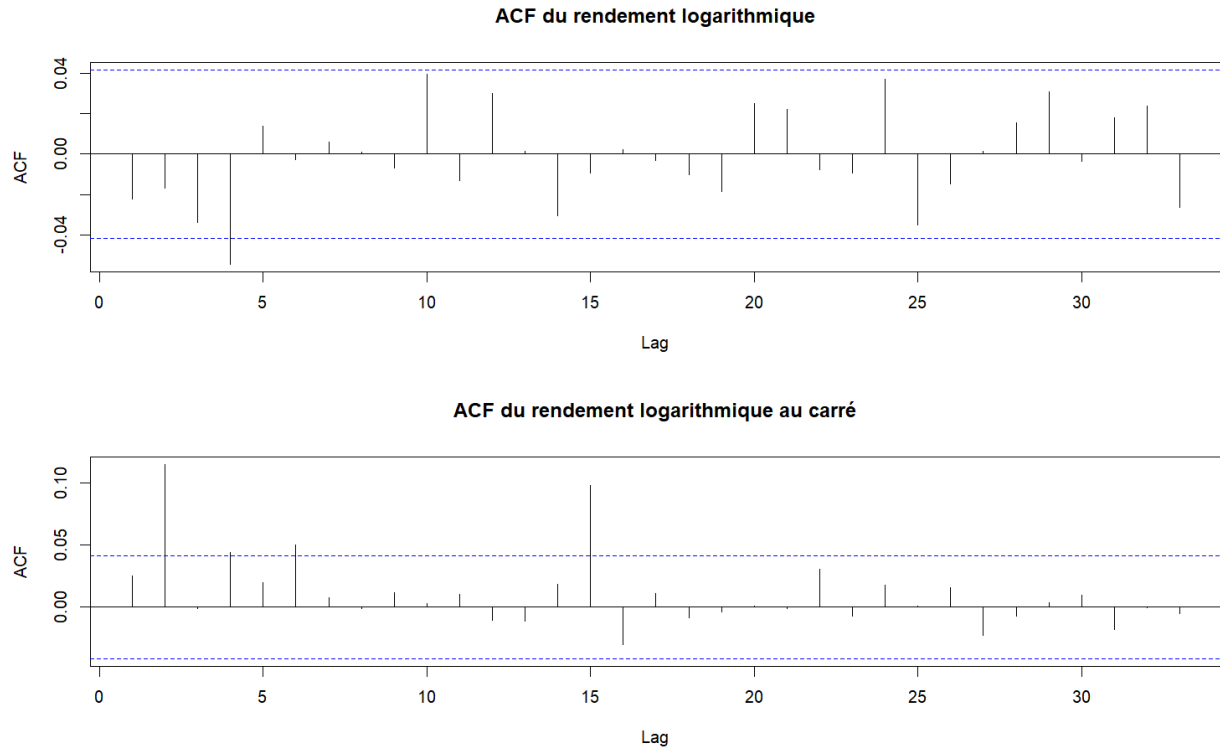


Figure 2: Corrélogrammes des rendements et des rendements au carré d'AmBev

Les corrélogrammes de la figure 2 de la série des rendements et des rendements au carré de l'action AmBev permet de vérifier visuellement que la majorité des coefficients d'auto-corrélations sont statistiquement nuls, car les barres se trouvent entre les traits en pointillé représentant l'intervalle de confiance. On remarque aussi que l'autocorrélation est faible, car la seule barre qui dépasse de l'intervalle de confiance à une très faible valeur ($| - 0.05| < 0.1$). Enfin on observe de l'autocorrélation sur les rendements au carré un peu plus élevée, avec des ρ significatifs (mais qui reste de l'ordre de 0.1).

3.3.2 La statistique de Ljung-Box

Pour aller un peu plus loin, on souhaite quantifier l'autocorrélation, pour cela on utilise la statistique de Ljung-Box.

Ce test est utilisé pour déterminer si les autocorrélations dans une série temporelle sont significativement différentes de zéro jusqu'à un certain retard k . L'hypothèse nulle du test de Ljung-Box stipule qu'il n'y a pas d'autocorrélation significative dans la série temporelle à aucun des retards jusqu'à un certain retard k .

La statistique de test de Ljung-Box est calculée en utilisant les autocorrélations. La statistique de test est basée sur la somme des carrés des autocorrélations jusqu'au retard k , multipliée par la taille de l'échantillon (n).

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k}$$

où K est le nombre de retards pris en compte, $\hat{\rho}_k^2$ est l'auto-corrélation à retard k , et n est la taille de l'échantillon.

On compare la statistique de test Q à la distribution chi-carrée (χ^2) avec K degrés de liberté. Si Q est significativement plus grand que les valeurs critiques de la distribution chi-carrée pour un certain niveau de signification, l'hypothèse nulle d'absence d'autocorrélation est rejetée, indiquant une autocorrélation significative.

Un test statistique de Ljung-Box a été effectué sur les rendements rte et sur les rendements au carré rte^2 .

Les p-valeurs du test de Ljung-Box pour les rendements de AmBev sont les suivantes :

```
[1] 0.29286000 0.42043950 0.23483567 0.02802472 0.04565595 0.07885987 0.12190944 0.17970930
[9] 0.24214799 0.13132183 0.16438772 0.13434954 0.18050251 0.14679363 0.18402567 0.23375735
[17] 0.28843448 0.33506292 0.35131214 0.33331906 0.33064289 0.37992067 0.42732359 0.32070212
[25] 0.24532717
```

La 4ème p-valeur est inférieure à 0.05 (et la 5ème aussi), donc il y a de l'autocorrélation.

Les p-valeurs du test de Ljung-Box pour les rendements au carré sont les suivantes :

```
[1] 2.370526e-01 1.918177e-07 8.766096e-07 4.095387e-07 8.849146e-07 2.072995e-07 5.499993e-07
[8] 1.428928e-06 3.032061e-06 6.870694e-06 1.343847e-05 2.508890e-05 4.457560e-05 6.392865e-05
[15] 2.676651e-08 2.548620e-08 4.858510e-08 9.385920e-08 1.866431e-07 3.644409e-07 6.936276e-07
[22] 6.100067e-07 1.073613e-06 1.502041e-06 2.660269e-06
```

Toutes les p-valeurs pour les rendements au carré sont inférieures à 0.05 (significatifs), donc il y a autocorrélations des rendements au carré.

Pour modéliser l'autocorrélation, un modèle ARMA(p,q) est utilisé. La formule du modèle ARMA(p,q) est :

$$\forall t : X_t = \sum_{i=1}^p \alpha_i X_{t-i} + \epsilon_t + \sum_{i=1}^q \beta_i \epsilon_{t-i}$$

où ϵ_t est l'erreur à l'instant t et les α_i et β_i sont des coefficients à déterminer. On commence par faire appel à l'acaf pour déterminer le p et q du ARMA(p,q).

```
> eacf(rte)
```

```
AR/MA
```

```
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 o o o x o o o o o o o o o
1 x o o x o o o o o o o o o
2 x x o x o o o o o o o o o
3 x x x x o o o o o o o o o
4 x x o o o o o o o o o o o
5 x o o o o o o o o o o o o
6 x x o o o x o o o o o o o
7 x o x o o x o o o o o o o
```

On trouve $p=0$ et $q=4$. Donc un MA(4), dont la formule est :

$$\forall t : X_t = \sum_{i=1}^4 \alpha_i X_{t-i} + \epsilon_t$$

Les coefficients du modèle MA(4) sont ensuite estimés avec la fonction *Arima()* du package forecast.

```
> reg<-Arima(rte, order=c(0,0,4))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ma1	-0.02394836	0.02120039	-1.1296	0.25864
ma2	-0.01785802	0.02118650	-0.8429	0.39929
ma3	-0.03413831	0.02112975	-1.6157	0.10617
ma4	-0.05340456	0.02108645	-2.5326	0.01132 *
intercept	0.00057857	0.00026628	2.1728	0.02980 *

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

On remarque que le coefficient ma4 est significatif (i.e. p-valeur < 0.05). C'est bon signe pour le modèle ma(4). Les autres coefficients ne sont pas tous significatifs. Ils sont un à un fixés à 0. Finalement :

```
> reg<-Arima(rte, order=c(0,0,4),fixed=c(0,0,0,NA,NA))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ma4	-0.05438711	0.02105492	-2.5831	0.009791 **
intercept	0.00057909	0.00028934	2.0014	0.045349 *

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

On détermine finalement un modèle MA(4) où les coefficients ma1, ma2 et ma3 sont nuls et les coefficients d'intercept et ma4 sont non nuls.

Pour que le modèle MA(4) puisse modéliser l'auto-corrélation des données de *rte*, il est nécessaire que les aléas du modèle MA(4) aient une espérance nulle et qu'ils ne soit pas auto-corrélés. Cela est vérifié avec un test d'espérance nulle des aléas du modèle MA sans la constante, dont l'hypothèse nulle H_0 est la nullité de l'espérance des aléas. L'hypothèse alternative H_a est que l'espérance des aléas est différente de zéro.

```
> t.test(residu)
```

One Sample t-test

```
data: residu
t = -0.0043067, df = 2231, p-value = 0.9966
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0005998174 0.0005971886
sample estimates:
mean of x
-1.314404e-06
```

La p-valeur est égale à 0.99, donc est supérieur à 0.05, c'est pourquoi H_0 est accepté, c'est à dire que l'espérance des aléas est considérée nulle et on passe au test d'absence d'auto-corrélation.

```
>residuv=(residu-mean(residu))/sd(residu)
>K<-20
>tmp<-rep(0,K)
>for(i in 1:K){
  tmp[i]<-Box.test(residuv,lag=i,type="Ljung-Box")$p.value
```

```
}
>tmp
```

```
[1] 0.2675477 0.3799019 0.1999706 0.3260184 0.4189761 0.5461192 0.6590141 0.7551910 0.8249853
[10] 0.5960112 0.6450476 0.5440773 0.6254199 0.5481709 0.6018622 0.6673172 0.7307158 0.7664317
[19] 0.7666130 0.7181860
```

Il est ici vérifiable que toutes les p-valeurs sont supérieures à 0.05, donc les résidus du MA(4) ne sont pas auto-corrélés : c'est bon signe concernant la validité de ce modèle.

3.4 Clusters de volatilité

Un test ARCH d'Engle a été effectué sur les données *rte* pour étudier l'hétéroscédasticité conditionnelle.

Le modèle ARCH de l'équation de volatilité est :

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_m \epsilon_{t-m}^2$$

où ϵ_t est le t-ième aléa du modèle ARMA(p,q) représentant l'équation de la moyenne conditionnelle de *rte* et e les résidus associés à son estimation.

Le test ARCH d'Engle consiste à tester l'hypothèse nulle $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ c'est à dire qu'il y a homoscédasticité conditionnelle. Sous H_0 , LM suit la répartition $\chi^2(m)$. A l'inverse H_a : au moins un des coefficients α_i est différent de 0 avec $i \neq 0$ donc il y a hétéroscédasticité conditionnelle.

```
> for (i in seq(1, 20)) {
+   LM[i] <- ArchTest(as.numeric(rte), lag = i)
+   cat("LM d'ordre :", i, "\n")
+   print(LM[i])
+ }
```

Hormis pour le premier test qui présente une p-valeur de 0.2354, toutes les p-valeurs sont inférieures 0.05 (ces résultats sont visibles en annexe 5.1). On peut rejeter l'hypothèse nulle en faveur de l'hypothèse alternative. Cela indique la présence d'effets ARCH dans les résidus, montrant ainsi une hétéroscédasticité conditionnelle.

3.5 Queues épaisses conditionnelles

Après avoir détecté des clusters de volatilité, le test d'Engle est réalisé sur *rte* et le modèle GARCH(1,1) suivant est estimé sur les résidus du ARMA :

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Le test sur un coefficient a pour hypothèse nulle H_0 la nullité de ce coefficient et H_a est que ce coefficient est nul. Les coefficients a_0 ($= \alpha_0$), a_1 ($= \alpha_1$) et b_1 ($= \beta_1$) du GARCH sont tous significatifs (p-valeurs inférieures à 0.05) et sont estimés tels que :

```
Coefficient(s):
Estimate Std. Error t value Pr(>|t|)
a0 0.09231 0.01846 5.001 5.69e-07 ***
a1 0.07640 0.01020 7.487 7.04e-14 ***
b1 0.83385 0.02536 32.887 < 2e-16 ***
```

On vérifie ensuite que toute l'hétéroscédasticité conditionnelle a été prise en compte avec un test d'Engle sur les résidus du GARCH, dont H_0 est une homoscédasticité conditionnelle. L'hypothèse alternative est qu'on a réussi à prendre en compte toute l'hétéroscédasticité conditionnelle.

```
> ArchTest(volat$res, lag=1)
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
data: volat$res
Chi-squared = 1.0383, df = 1, p-value = 0.3082
```

On remarque que la p-valeur(0.3082) est bien supérieure à 0.05 pour le premier ordre. Et une boucle effectue des tests d'ARCH pour tous les degrés entre 2 et 20.

```
> for (i in seq(2, 20)) {  
+   volati <- ArchTest(volat$res, lag = i)  
+   cat("LM d'ordre :", i, "\n")  
+   print(volati)  
+ }
```

Entre les degrés 1 et 14, toutes les p-valeurs sont supérieures à 0.05. Mais à partir de l'ordre 15, les p-valeurs sont inférieures à 0.05. Cela signifie qu'une partie de l'hétéroscédasticité conditionnelle n'a pas été prise en compte.

Les coefficients m et n du modèle GARCH(m,n) ont alors été augmentés, avec des paramètres m et n entre 1 et 15. Malheureusement les coefficients du MA(4) n'étaient pas significatifs dans ces conditions. Des modèles MA(3), MA(5) et MA(15) ont également été testés sur ces données, en vain. Cela m'a alors laissé penser que le modèle GARCH n'est peut être pas adapté sur cette action pendant cette période. C'est pourquoi j'ai réalisé le test d'Anscombe pour déterminer le kurtosis de ce modèle.

```
> anscombe.test(volat$res)  
  
Anscombe-Glynn kurtosis test  
data:  volat$res  
kurt = 5.2139, z = 10.9149, p-value < 2.2e-16  
alternative hypothesis: kurtosis is not equal to 3
```

Le test d'Anscombe sur les résidus des aléas du modèle ARMA-GARCH montre que les queues de distributions conditionnelles sont plus épaisses que celles d'une loi normale, car le kurtosis de 5.2 est supérieur à la valeur de référence égale à 3.

J'ai finalement remarqué que le modèle GARCH(1,15), qui a un coefficient β_1 non significatif, (donc équivalent à GARCH(0,15)), prend bien en compte l'autocorrélation.

```
> ArchTest(volat$res,lag=1)  
  
ARCH LM-test; Null hypothesis: no ARCH effects  
data:  volat$res  
Chi-squared = 0.18173, df = 1, p-value = 0.6699  
  
> ArchTest(volat$res,lag=20)  
  
ARCH LM-test; Null hypothesis: no ARCH effects  
data:  volat$res  
Chi-squared = 17.907, df = 20, p-value = 0.5935
```

En effet, toutes les p-valeurs du test d'ARCH entre 1 et 20 sont supérieures à 0.05. Les sorties R de ces tests sont visibles en annexe 5.3.

```
> anscombe.test(volat$res)  
  
Anscombe-Glynn kurtosis test  
data:  volat$res  
kurt = 4.8734, z = 9.9198, p-value < 2.2e-16  
alternative hypothesis: kurtosis is not equal to 3
```

Le test d'Anscombe sur les résidus des aléas du modèle ARMA-GARCH(1,15) montre que les queues de distributions conditionnelles sont épaisses.

3.6 Effet de levier

L'effet de levier est la différence de volatilité entraînée par des rendements positifs et négatifs, car la volatilité augmente lorsque le cours de l'action baisse. Un graphique des prix de l'action Ambev et de leur volatilité a été tracé figure 3 pour mettre en évidence cet effet. L'estimateur de la volatilité en t utilisé est l'écart-type des 22 derniers jours :

$$\sigma_t^{(22)} = \sqrt{\sum_{i=t-22}^t (y_i - (\sum_{i=t-22}^t y_i / 22))^2 / 22}$$

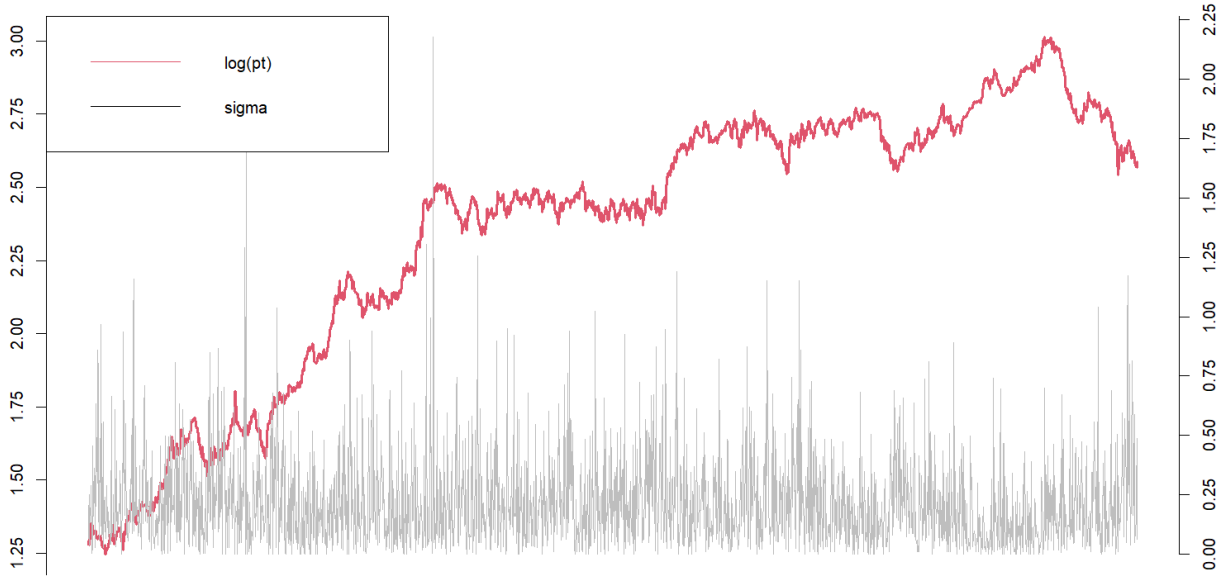


Figure 3: Effet de levier d'AmBev: relation entre le logarithme et l'écart-type des rendements.

J'ai mesuré graphiquement sur un gain, un ratio $\frac{\sigma}{\log(pt)}$ égale à 45. Ce même ratio sur une perte valait 30. Basé sur cette mesure, l'effet de levier semble profiter à la hausse. Seulement ce cas est rare dans la réalité et cette conclusion est sûrement due aux choix des dates sur lesquelles ont été mesurées ces ratios. Pour mieux estimer cet effet de levier, il faudrait sûrement moyenner les ratios calculés sur tous les jours de gain et sur tous les jours de perte.

Sur ce graphique, il n'est pas évident qu'il y avait des périodes de volatilités exceptionnelles simultanément à, ou dans les instants suivants, un cours en baisse. En effet les périodes de plus fortes volatilités sur la période étudiée font parfois suite à des baisses de cours, qui peuvent être faibles, mais aussi en fin de période de hausse. On remarque aussi que certaines phases de baisses des cours assez significatives n'ont eu que peu d'impact sur la volatilité.

3.7 Saisonnalité

Dans cette partie, le phénomène de saisonnalité, selon lequel les cours évoluent différemment sur certaines périodes.

Saisonnalité hebdomadaire :

Moyenne en % :

Les vendredis ont des moyennes positives de 0.0867 plus élevé que les autres jours, indiquant une tendance haussière légère ces jours-là. Les lundis, mardis et jeudis ont des moyennes plus faible que les vendredi, entre 0.05 et 0.065, indiquant des tendances relativement stables ces jours-là. Les mercredis se démarquent des autres jours avec

une moyenne bien plus proche de 0 (0.0249)

Écart-type annuel en % : Les mercredis ont la volatilité annuelle la plus élevée, indiquant des variations plus importantes dans les rendements quotidiens ce jour-là. Les mardis et vendredis ont des volatilités annuelles légèrement plus basses que les mercredis mais plus élevées que les lundis et jeudis. Cependant les variations de volatilité d'un jour à l'autre semblent trop faibles pour être véritablement significatives.

Table 1: Statistiques des rendements de AmBev pour chaque jour de la semaine

	lundi	mardi	mercredi	jeudi	vendredi
Moyenne en %	0.0649	0.0640	0.0249	0.0500	0.0867
Ecart-type annuel en %	22.07	22.65	23.85	23.00	23.07
Skewness	1.08	-0.18	-0.56	0.11	0.33
Kurtosis	8.19	1.30	4.90	2.14	1.23

De façon optionnelle, on peut aussi étudier la *Skewness*

Skewness : Les lundis et vendredis ont une skewness positive, ce qui signifie que les rendements positifs sont légèrement plus fréquents que les rendements négatifs ces jours-là. Les mardis et jeudis ont des skewness proches de zéro, indiquant des distributions relativement symétriques. Tandis que les mercredis présente une skewness légèrement négative, impliquant des rendements négatifs plus fréquents ces jours là.

Saisonnalité mensuelle :

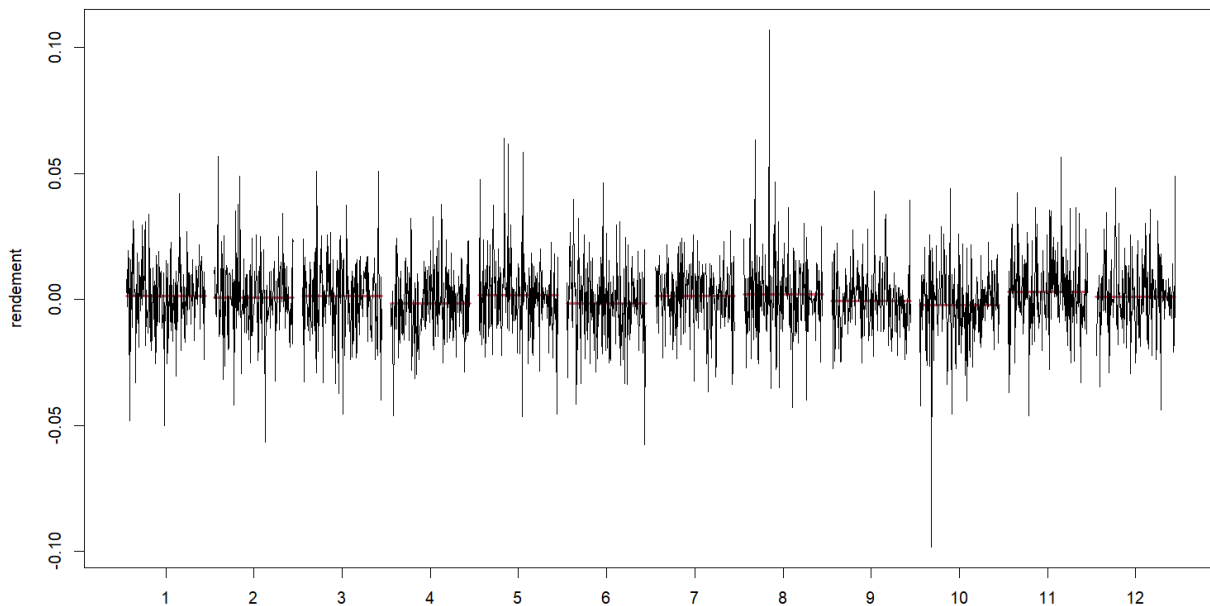


Figure 4: Rendement logarithmique d'AmBev par mois

On voit sur la figure 4 un léger effet de saisonnalité mensuelle en moyenne sur les rendements de l'action AmBev. On remarque notamment un pic de rentabilité positive exceptionnelle à la mi-août, et à l'inverse un pic de rentabilité

négative début octobre. Sur les autres mois, il n'y a pas de rendements statistiquement anormaux, positifs ou négatifs, particulièrement visible. Sinon, on peut remarquer une faible volatilité des rendements en juillet, qui est le mois des vacances d'hiver au Brésil, en même temps qu'une période de grandes vacances dans l'hémisphère nord (où se trouve une partie des investisseurs étrangers).

3.8 Stationnarité

On étudie la stationnarité des rendements de l'action AmBev.

Pour rappel, un processus stochastique $(y_t, t \in Z)$ est dit stationnaire au second ordre si la variance est finie et si la moyenne et la fonction d'autocovariance sont indépendantes du temps.

On cherche à déterminer si la série est stationnaire. Pour cela on effectue un test de racine unitaire de Dickey-Fuller. L'hypothèse nulle est que la série temporelle a une racine unitaire, ce qui signifie qu'elle est non stationnaire. Si le test rejette l'hypothèse nulle (donc accepte H_a), cela suggère que la série est stationnaire après différenciation, c'est-à-dire qu'elle ne possède pas de racine unitaire et peut être analysée à l'aide de modèles statistiques classiques.

On commence par effectuer le test dans les cas où il existe une tendance déterministe dans la série. L'hypothèse nulle est la nullité du coefficient β_1 . On trouve que la valeur absolue de la t-value de β_1 , associée à β_1 n'est pas supérieure à la valeur tabulée, ce qui conforte l'hypothèse nulle affirmant que $\beta_1 = 0$. Les résultats du code sont visibles en annexe 5.4.

On passe au modèle contenant une constante mais pas de tendance. Les résultats fournis par le code sont visibles en annexe (5.4.1). La valeur calculée du t est supérieure à la valeur tabulée, donc on ne rejette pas $H_0 : (\rho - 1) = 0$ et $\beta_0 = 0$.

Dans ce cas de figure, on s'intéresse au test de RU quand il n'existe ni tendance ni constante dans la série. Ce test consiste à tester l'hypothèse nulle $H_0 : (\rho - 1) = 0$ face à l'hypothèse alternative $H_a : \rho < 1$ dans l'équation suivante :

$$\Delta y_t = (\rho - 1)y_{t-1} + \epsilon_t$$

```
> summary(ur.df(rte,type= "none",lags=0))

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1)

Residuals:
    Min       1Q   Median       3Q      Max
-0.098315 -0.007482  0.000328  0.009248  0.106620

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1 -1.02078     0.02123  -48.09  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01445 on 2230 degrees of freedom
Multiple R-squared:  0.5091, Adjusted R-squared:  0.5089
F-statistic: 2313 on 1 and 2230 DF, p-value: < 2.2e-16

Value of test-statistic is: -48.0927

Critical values for test statistics:
    1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

La valeur du test statistique déterminée ici est $t = -48 < -1.95$ (où 1.95 est la valeur tabulée au seuil de 5%), donc on rejette H_0 donc le PGD (processus générateur de données) n'est pas DS. On en conclut que le PGD est stationnaire. Cette conclusion n'est valide que si les aléas de la régression de Dickey et Fuller ne sont pas auto-corrélés. C'est ce que nous vérifions avec le test suivant :

```
> plot(ur.df(rte,lag=0,type="none"))
```

Ce graphique indique que les aléas sont auto-corrélés à l'ordre 4 cela implique que l'on ne peut pas valider notre conclusion d'absence de racine unitaire dans le rendement logarithmique de AmBev. Afin de tenter d'y remédier, un test de racine unitaire dans le cadre de la régression de Dickey-Fuller augmentée (ADF) est effectué.

Le test ADF, de Dickey-Fuller Augmenté, est un test de Dickey-Fuller dans lequel sont ajoutées k variables explicatives (pour augmenter la régression DF), ce sont les différences premières retardées de la variable dépendante.

Puisque notre série n'a pas de tendance et fluctue autour de zéro, on estime l'équation suivante :

$$\Delta y_t = (\rho - 1)y_{t-1} + \sum_{j=1}^k \gamma_j \Delta y_{t-j} + \epsilon_t$$

or il faut déterminer le nombre k de variables explicatives nécessaires. Ce nombre est important pour que H_0 ne soit pas rejetée trop souvent ou trop peu souvent. Pour cela, on utilise la formule de Schwert :

$$k = \lfloor 12 * (\frac{T}{100})^{0.25} \rfloor$$

où T est le nombre de données de la série : c'est à dire 2232 pour *rte*.

```
> Schwert<-as.integer(12*(T/100)^(0.25))
```

On estime ainsi $k = 26$. Et le test ADF est effectué :

```
> summary(ur.df(rte,type="none",lags=Schwert))
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
    Min       1Q   Median       3Q      Max
-0.094928 -0.007414  0.000646  0.008817  0.105526
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.053953   0.118323  -8.907   <2e-16 ***
z.diff.lag1    0.029129   0.116134   0.251   0.802
z.diff.lag2    0.018350   0.113967   0.161   0.872
z.diff.lag3   -0.014390   0.111895  -0.129   0.898
z.diff.lag4   -0.067251   0.109591  -0.614   0.540
z.diff.lag5   -0.052351   0.107284  -0.488   0.626
z.diff.lag6   -0.056664   0.104973  -0.540   0.589
z.diff.lag7   -0.056249   0.102502  -0.549   0.583
z.diff.lag8   -0.053982   0.099827  -0.541   0.589
z.diff.lag9   -0.050274   0.097179  -0.517   0.605
z.diff.lag10  -0.009587   0.094506  -0.101   0.919
z.diff.lag11  -0.021533   0.091761  -0.235   0.814
z.diff.lag12   0.009811   0.088875   0.110   0.912
z.diff.lag13   0.015062   0.085934   0.175   0.861
z.diff.lag14  -0.011068   0.083023  -0.133   0.894
z.diff.lag15  -0.016768   0.079948  -0.210   0.834
```

```

z.diff.lag16 -0.006138  0.076546 -0.080  0.936
z.diff.lag17 -0.008873  0.073076 -0.121  0.903
z.diff.lag18 -0.021455  0.069159 -0.310  0.756
z.diff.lag19 -0.040018  0.065018 -0.615  0.538
z.diff.lag20 -0.012078  0.060584 -0.199  0.842
z.diff.lag21  0.013895  0.055740  0.249  0.803
z.diff.lag22  0.004921  0.050416  0.098  0.922
z.diff.lag23 -0.001438  0.044260 -0.032  0.974
z.diff.lag24  0.042169  0.037816  1.115  0.265
z.diff.lag25  0.009944  0.030604  0.325  0.745
z.diff.lag26 -0.004891  0.021394 -0.229  0.819

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.01441 on 2178 degrees of freedom
Multiple R-squared:  0.5178, Adjusted R-squared:  0.5118
F-statistic: 86.62 on 27 and 2178 DF,  p-value: < 2.2e-16

```

```

Value of test-statistic is: -8.9074
Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62

```

Ensuite on diminue la valeur de "lags" jusqu'à avoir tous les gammas avec une valeur calculée de la statistique t supérieure en valeur absolue à 1.6. Après plusieurs essais on trouve que cette condition est réunie pour lags=3.

```
> summary(ur.df(rte,type= "none",lags=3))
```

```

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
      Min       1Q   Median       3Q      Max
-0.094691 -0.007435  0.000639  0.009115  0.105516
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.13021    0.04388 -25.757 < 2e-16 ***
z.diff.lag1   0.10654    0.03754  2.838  0.00458 **
z.diff.lag2   0.08908    0.03037  2.933  0.00339 **
z.diff.lag3   0.05487    0.02123  2.584  0.00982 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.01444 on 2224 degrees of freedom
Multiple R-squared:  0.5113, Adjusted R-squared:  0.5105
F-statistic: 581.8 on 4 and 2224 DF,  p-value: < 2.2e-16

```

```

Value of test-statistic is: -25.7566
Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62

```

Pour $k=3$, toutes les valeurs t sont supérieures en valeur absolue à 1.6 et la valeur du test statistique est $t = -25.7566 < -1.95$, donc le test ADF permet de conclure à la stationnarité du PGD qui a généré notre série du rendement. La variance est finie et la moyenne et la fonction d'autocovariance sont indépendantes du temps.

4 Conclusion

Au cours de ce projet, 8 propriétés principales des rendements logarithmiques ont pu être étudiées à partir du cours de l'action de la société brésilienne AmBev. Cela a permis l'appropriation d'outils statistiques et d'analyse de données. Les limites de certains modèles ont aussi pu être observées, notamment la difficulté du modèle GARCH(1,1) à prendre en compte toute l'autocorrélation des données.

5 Annexe

5.1 Résultats des tests d'ARCH pour les clusters de volatilité

LM d'ordre : 1

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 1.408, df = 1, p-value = 0.2354

LM d'ordre : 2

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 30.826, df = 2, p-value = 2.024e-07

LM d'ordre : 3

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 30.881, df = 3, p-value = 9.007e-07

LM d'ordre : 4

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 33.007, df = 4, p-value = 1.191e-06

LM d'ordre : 5

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 33.805, df = 5, p-value = 2.603e-06

LM d'ordre : 6

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 37.479, df = 6, p-value = 1.42e-06

LM d'ordre : 7

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)

Chi-squared = 37.463, df = 7, p-value = 3.832e-06

LM d'ordre : 8

ARCH LM-test; Null hypothesis: no ARCH effects

```

data: as.numeric(rte)
Chi-squared = 37.809, df = 8, p-value = 8.165e-06

LM d'ordre : 9

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 38.008, df = 9, p-value = 1.737e-05

LM d'ordre : 10

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 37.963, df = 10, p-value = 3.852e-05

LM d'ordre : 11

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 37.982, df = 11, p-value = 7.879e-05

LM d'ordre : 12

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 38.447, df = 12, p-value = 0.0001298

LM d'ordre : 13

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 38.833, df = 13, p-value = 0.0002126

LM d'ordre : 14

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 39.931, df = 14, p-value = 0.0002615

LM d'ordre : 15

ARCH LM-test; Null hypothesis: no ARCH effects

data: as.numeric(rte)
Chi-squared = 62.212, df = 15, p-value = 1.047e-07

LM d'ordre : 16

ARCH LM-test; Null hypothesis: no ARCH effects

```



```
data: as.numeric(rte)
Chi-squared = 65.637, df = 16, p-value = 5.722e-08
```

```
LM d'ordre : 17
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: as.numeric(rte)
Chi-squared = 65.799, df = 17, p-value = 1.126e-07
```

```
LM d'ordre : 18
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: as.numeric(rte)
Chi-squared = 65.98, df = 18, p-value = 2.137e-07
```

```
LM d'ordre : 19
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: as.numeric(rte)
Chi-squared = 66.155, df = 19, p-value = 3.959e-07
```

```
LM d'ordre : 20
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: as.numeric(rte)
Chi-squared = 65.924, df = 20, p-value = 8.308e-07
```

5.2 Résultats des tests d'ARCH sur les résidus du GARCH(1,1)

```
LM d'ordre : 2
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 4.2704, df = 2, p-value = 0.1182
```

```
LM d'ordre : 3
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 7.1903, df = 3, p-value = 0.06607
```

```
LM d'ordre : 4
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 7.3295, df = 4, p-value = 0.1195
```

```

LM d'ordre : 5

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 7.346, df = 5, p-value = 0.1961

LM d'ordre : 6

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 7.7374, df = 6, p-value = 0.258

LM d'ordre : 7

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 7.9346, df = 7, p-value = 0.3384

LM d'ordre : 8

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 9.819, df = 8, p-value = 0.278

LM d'ordre : 9

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 9.9135, df = 9, p-value = 0.3575

LM d'ordre : 10

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 10.125, df = 10, p-value = 0.4296

LM d'ordre : 11

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 10.1, df = 11, p-value = 0.5214

LM d'ordre : 12

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 10.575, df = 12, p-value = 0.5657

```

```

LM d'ordre : 13

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 12.041, df = 13, p-value = 0.5243

LM d'ordre : 14

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 12.284, df = 14, p-value = 0.5835

LM d'ordre : 15

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 37.513, df = 15, p-value = 0.001064

LM d'ordre : 16

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 40.791, df = 16, p-value = 0.0005961

LM d'ordre : 17

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 40.947, df = 17, p-value = 0.0009497

LM d'ordre : 18

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 40.891, df = 18, p-value = 0.001576

LM d'ordre : 19

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 40.97, df = 19, p-value = 0.002436

LM d'ordre : 20

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 40.979, df = 20, p-value = 0.003748

```

5.3 Résultats des tests d'ARCH sur les résidus du GARCH(1,15)

```
> ArchTest(volat$res,lag=1)

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.18173, df = 1, p-value = 0.6699

> for (i in seq(2, 20)) {
+   volati <- ArchTest(volat$res, lag = i)
+   cat("LM d'ordre :", i, "\n")
+   print(volati)
+ }
LM d'ordre : 2

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.207, df = 2, p-value = 0.9017

LM d'ordre : 3

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.943, df = 3, p-value = 0.815

LM d'ordre : 4

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 1.1051, df = 4, p-value = 0.8935

LM d'ordre : 5

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 2.0367, df = 5, p-value = 0.844

LM d'ordre : 6

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 2.7942, df = 6, p-value = 0.8342

LM d'ordre : 7

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 4.5922, df = 7, p-value = 0.7096
```

```

LM d'ordre : 8

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 6.0053, df = 8, p-value = 0.6466

LM d'ordre : 9

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 6.023, df = 9, p-value = 0.7376

LM d'ordre : 10

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 7.426, df = 10, p-value = 0.6847

LM d'ordre : 11

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 8.0501, df = 11, p-value = 0.7088

LM d'ordre : 12

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 9.3396, df = 12, p-value = 0.6737

LM d'ordre : 13

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 11.033, df = 13, p-value = 0.6081

LM d'ordre : 14

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 12.342, df = 14, p-value = 0.5788

LM d'ordre : 15

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 12.538, df = 15, p-value = 0.6379

```

```

LM d'ordre : 16

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 17.488, df = 16, p-value = 0.3547

LM d'ordre : 17

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 17.404, df = 17, p-value = 0.4273

LM d'ordre : 18

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 17.52, df = 18, p-value = 0.4877

LM d'ordre : 19

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 17.843, df = 19, p-value = 0.533

LM d'ordre : 20

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 17.907, df = 20, p-value = 0.5935

```

5.4 Dickey-Fuller

```

> summary(ur.df(rte,type= "trend",lags=0))

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.554e-03  6.123e-04   2.538  0.0112 *
z.lag.1      -1.024e+00  2.123e-02 -48.219  <2e-16 ***
tt           -8.637e-07  4.749e-07  -1.819   0.0691 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Value of test-statistic is: -48.2194 775.0444 1162.565

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12

```

```
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34
```

5.4.1 test de RU quand il existe une constante mais pas de tendance dans la série

```
> summary(ur.df(rte,type= "drift",lags=0))
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0005894  0.0003060   1.926   0.0542 .
z.lag.1      -1.0223615  0.0212283 -48.160   <2e-16 ***
---

```

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

```
Value of test-statistic is: -48.1604 1159.712
```

Critical values for test statistics:

```
      1pct  5pct 10pct
tau2 -3.43 -2.86 -2.57
phi1  6.43  4.59  3.78
```

6 Annexe 2 : Comparaison avec rtt

Dans cette partie, les propriétés testées sur les données d'entraînements ont été testée sur les données rtt.

Le tableau 2 suivant résume la comparaison entre les propriétés des rendements *rte* et *rtt*.

Propriété	rte	rtt
Asymétrie perte/gain	Skewness significative positive	Skewness significative positive
Queues de distribution	Epaisse : kurtosis = 6.5	Epaisse : kurtosis = 11.8
Auto-corrélation des rendements	Auto-corrélation aux ordres 4 et 5	Auto-corrélation à tous les ordres
Auto-corrélation des rendements au carré	Auto-corrélation à tous les ordres	Auto-corrélation à tous les ordres
Modèle ARMA	MA(4)	MA(6)
Espérance des aléas du ARMA nulle ?	Oui	Oui
Cluster de volatilité	Effet ARCH :	Effet ARCH :
	hétéoscédasticité conditionnelle	hétéoscédasticité conditionnelle
Queues épaisses conditionnelles	GARCH(1,1) et reste	GARCH(1,10) et reste
	d'hétéoscédasticité conditionnelle	d'hétéoscédasticité conditionnelle
	kurtosis = 5.2	kurtosis = 6.4
Effet de levier	Non	Non
Saisonnalité	Non	Non
Stationnarité	Oui	Oui

Table 2: Tableau des propriétés pour *rte* et *rtt*

6.1 Code et Output R avec rtt

```
# On test H0 : E(rte)=μ=0 avec la statistique t
> rbar<-mean(rtt)
> rbar
[1] -1.064648e-05
```

```
> s=sd(rtt)
> rbar/(s/sqrt(T))
[1] -0.01838307
```

La moyenne des rendements rtt n'est pas exactement centrée en zéro.

```
###-----
###      1 : Asymétrie Perte/Gain    (skewness)
###-----
> agostino.test(rtt) #par défaut bilatéral # Test d'Agostino
```

```
D'Agostino skewness test
data:  rtt
skew = -0.51124, z = -6.80588, p-value = 1.004e-11
alternative hypothesis: data have a skewness
```

```
###-----
###      2 : Queues de distribution épaisse  (kurtosis)
###-----
```

```
> # Test d'Anscombe modifié
> anscombe.test(rtt) #par défaut bilatéral
```

```
Anscombe-Glynn kurtosis test
```



```
data: rtt
kurt = 11.849, z = 14.209, p-value < 2.2e-16
alternative hypothesis: kurtosis is not equal to 3
```

```
###-----
###      3 : Autocorrélations fortes et faibles
###-----
```

```
> for (i in 1:25 ) {
+   pvaluesrt[i] = Box.test(rtt,lag=i,type="Ljung-Box")$p.value
+   pvaluesrt2[i] = Box.test(rtt^2,lag=i,type="Ljung-Box")$p.value
+ }
```

```
> pvaluesrt2
[1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

```
> pvaluesrt
[1] 1.533123e-03 9.945937e-05 2.752391e-04 4.090560e-04 1.020584e-03 3.406961e-05 8.190344e-05
[8] 1.774128e-04 1.678517e-04 2.103411e-04 4.095997e-04 6.051949e-04 1.058250e-03 1.306383e-03
[15] 1.793203e-03 1.013055e-03 1.083448e-03 9.542810e-04 7.424916e-04 1.090356e-03 2.896792e-04
[22] 4.686990e-04 6.550966e-04 3.047804e-04 4.791161e-04
```

Il semble il y avoir de l'autocorrélation à tous les ordres pour les rendements et pour les rendements au carré.

```
> eacf(rtt)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x o o o x o o o o o o o o
1 x o o o o x o o o o o o o o
2 x x o o o x o o o o o o o o
3 x x x o o x o o o o o o o o
4 o x o x o x o o o o o o o o
5 o x x x x x o o o o o o o o
6 x x x x x o x o o o o o o o
7 x x x x x o o o o o o o o o
```

Un modèle MA(6) est estimé ici.

```
> reg<-Arima(rtt, order=c(0,0,6))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ma1	-0.08638435	0.02904735	-2.9739	0.002940 **
ma2	0.08841171	0.02933759	3.0136	0.002582 **
ma3	-0.02016595	0.02937330	-0.6865	0.492373
ma4	-0.03162616	0.02879005	-1.0985	0.271982
ma5	0.00123445	0.02988846	0.0413	0.967055
ma6	-0.09477461	0.02903820	-3.2638	0.001099 **
intercept	-0.00001319	0.00049104	-0.0269	0.978569

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

```
> reg<-Arima(rtt, order=c(0,0,6),fixed=c(NA, NA,0,NA,0,NA,NA))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ma1	-8.4588e-02	2.8904e-02	-2.9265	0.0034280	**
ma2	8.6579e-02	2.9127e-02	2.9725	0.0029543	**
ma4	-3.3084e-02	2.8387e-02	-1.1655	0.2438345	
ma6	-9.5693e-02	2.8735e-02	-3.3302	0.0008678	***
intercept	-1.2818e-05	5.0054e-04	-0.0256	0.9795700	

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

```
> reg<-Arima(rtt, order=c(0,0,6),fixed=c(NA, NA,0,0,0,NA,NA))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ma1	-8.3608e-02	2.8816e-02	-2.9014	0.003715	**
ma2	8.9403e-02	3.0038e-02	2.9763	0.002918	**
ma6	-9.2102e-02	2.8726e-02	-3.2062	0.001345	**
intercept	-1.1835e-05	5.2388e-04	-0.0226	0.981976	

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

```
> reg<-Arima(rtt, order=c(0,0,6),fixed=c(NA, NA,0,0,0,NA,0))
> coeftest(reg)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ma1	-0.083605	0.028816	-2.9013	0.003716	**
ma2	0.089395	0.030038	2.9761	0.002920	**
ma6	-0.092102	0.028726	-3.2062	0.001345	**

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

Les coefficients 1, 2 et 6 sont significatifs.

```
> residu<-reg$res
> t.test(residu)
```

One Sample t-test
data: residu
t = -0.018192, df = 1175, p-value = 0.9855
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-0.001133588 0.001112760
sample estimates:
mean of x
-1.041424e-05

L'espérance des aléas est nulle.

```
> residuv=(residu-mean(residu))/sd(residu)
> K<-20
> tmp<-rep(0,K)
> for(i in 1:K){
```

```

+   tmp[i]<-Box.test(residuv,lag=i,type="Ljung-Box")$p.value
+ }
> tmp
[1] 0.9577419 0.9905299 0.8830718 0.7563695 0.8627158 0.9283830 0.9647884 0.9740436 0.9229285
[10] 0.8982059 0.9360029 0.9447176 0.9645245 0.9677944 0.9787927 0.8885970 0.8137158 0.7567382
[19] 0.5341711 0.5911561

```

Toutes les valeurs sont ≥ 0.05 , donc les résidus ne sont pas auto-corrélés.

```

###-----
###      4 : Clusters de volatilité
###-----

```

```

> for (i in seq(1, 20)) {
+   LMi <- ArchTest(as.numeric(rte), lag = i)
+   cat("LM d'ordre :", i, "\n")
+   print(LMi)
+ }

LM d'ordre : 1
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 1.4081, df = 1, p-value = 0.2354

LM d'ordre : 2
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 30.827, df = 2, p-value = 2.023e-07

LM d'ordre : 3
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 30.881, df = 3, p-value = 9.004e-07

LM d'ordre : 4
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 33.007, df = 4, p-value = 1.19e-06

LM d'ordre : 5
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 33.806, df = 5, p-value = 2.603e-06

LM d'ordre : 6
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 37.479, df = 6, p-value = 1.42e-06

LM d'ordre : 7
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 37.463, df = 7, p-value = 3.832e-06

LM d'ordre : 8
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)

```

Chi-squared = 37.809, df = 8, p-value = 8.165e-06

LM d'ordre : 9
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 38.008, df = 9, p-value = 1.737e-05

LM d'ordre : 10
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 37.963, df = 10, p-value = 3.852e-05

LM d'ordre : 11
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 37.982, df = 11, p-value = 7.879e-05

LM d'ordre : 12
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 38.447, df = 12, p-value = 0.0001298

LM d'ordre : 13
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 38.833, df = 13, p-value = 0.0002126

LM d'ordre : 14
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 39.931, df = 14, p-value = 0.0002615

LM d'ordre : 15
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 62.211, df = 15, p-value = 1.047e-07

LM d'ordre : 16
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 65.636, df = 16, p-value = 5.724e-08

LM d'ordre : 17
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 65.798, df = 17, p-value = 1.126e-07

LM d'ordre : 18
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 65.979, df = 18, p-value = 2.137e-07

LM d'ordre : 19
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 66.154, df = 19, p-value = 3.96e-07

```
LM d'ordre : 20
ARCH LM-test; Null hypothesis: no ARCH effects
data: as.numeric(rte)
Chi-squared = 65.924, df = 20, p-value = 8.309e-07
```

Hormis pour le premier test, toutes les p-valeurs sont ≤ 0.05 . On peut rejeter l'hypothèse nulle en faveur de l'hypothèse alternative. Cela indique la présence d'effets ARCH dans les résidus, montrant ainsi une hétéroscédasticité conditionnelle.

```
###-----
###      5 : Queues épaisses conditionnelles
###-----

> volat<-garch(residuv,order=c(1,1))
> summary(volat)

Call:
garch(x = residuv, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
      Min       1Q   Median       3Q      Max
-6.442643 -0.534504 -0.005721  0.529595  5.291608

Coefficient(s):
      Estimate Std. Error  t value Pr(>|t|)
a0 8.260e-01   6.467e-02   12.77  <2e-16 ***
a1 1.896e-01   1.740e-02   10.90  <2e-16 ***
b1 1.302e-14   6.501e-02    0.00      1
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
Jarque Bera Test

data: Residuals
X-squared = 791.1, df = 2, p-value < 2.2e-16
```

Box-Ljung test

```
data: Squared.Residuals
X-squared = 1.1725, df = 1, p-value = 0.2789
```

Le coefficient b1 est ici non-significatif. On change de modèle pour trouver un modèle GARCH avec des coefficients significatifs.

```
###-----
###      5 : Queues épaisses conditionnelles
###-----
```

```
Les combinaisons essayées précédemment :
#volat<-garch(residuv,order=c(1,1))
```

```
#volat<-garch(residuv,order=c(1,2))
#volat<-garch(residuv,order=c(1,3))
#volat<-garch(residuv,order=c(2,1))
#volat<-garch(residuv,order=c(3,1))
# volat<-garch(residuv,order=c(10,1))
# volat<-garch(residuv,order=c(1,9))
# volat<-garch(residuv,order=c(1,11))
#volat<-garch(residuv,order=c(0,1))
#volat<-garch(residuv,order=c(0,2))
```

Finalement, j'ai opté pour :

```
>volat<-garch(residuv,order=c(1,10))
>summary(volat)
```

Call:

```
garch(x = residuv, order = c(1, 10))
```

Model:

```
GARCH(1,10)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-4.82493	-0.58425	-0.01483	0.55241	5.69776

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t)	
a0	4.537e-01	1.044e-01	4.346	1.39e-05	***
a1	1.624e-01	2.834e-02	5.729	1.01e-08	***
a2	2.017e-02	3.992e-02	0.505	0.613432	
a3	3.015e-02	2.947e-02	1.023	0.306179	
a4	1.197e-02	2.054e-02	0.582	0.560246	
a5	4.880e-02	2.764e-02	1.765	0.077514	.
a6	2.216e-03	2.342e-02	0.095	0.924628	
a7	8.373e-02	2.330e-02	3.593	0.000326	***
a8	3.796e-14	3.135e-02	0.000	1.000000	
a9	1.501e-02	2.152e-02	0.697	0.485644	
a10	9.039e-02	1.667e-02	5.421	5.93e-08	***
b1	3.591e-02	2.009e-01	0.179	0.858148	

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

Diagnostic Tests:

Jarque Bera Test

data: Residuals

X-squared = 567.73, df = 2, p-value < 2.2e-16

Box-Ljung test

data: Squared.Residuals

X-squared = 5.2992e-05, df = 1, p-value = 0.9942

```
> ArchTest(volat$res,lag=1)
```

```

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 5.2785e-05, df = 1, p-value = 0.9942

> for (i in seq(2, 20)) {
+   volati <- ArchTest(volat$res, lag = i)
+   cat("LM d'ordre :", i, "\n")
+   print(volati)
+ }
LM d'ordre : 2

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.0039475, df = 2, p-value = 0.998

LM d'ordre : 3

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.064474, df = 3, p-value = 0.9957

LM d'ordre : 4

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.20504, df = 4, p-value = 0.9951

LM d'ordre : 5

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.29834, df = 5, p-value = 0.9977

LM d'ordre : 6

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.30414, df = 6, p-value = 0.9995

LM d'ordre : 7

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.54742, df = 7, p-value = 0.9993

LM d'ordre : 8

ARCH LM-test; Null hypothesis: no ARCH effects

```

```

data: volat$res
Chi-squared = 0.50618, df = 8, p-value = 0.9999

LM d'ordre : 9

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 0.53761, df = 9, p-value = 1

LM d'ordre : 10

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 3.3773, df = 10, p-value = 0.9711

LM d'ordre : 11

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 3.9916, df = 11, p-value = 0.9702

LM d'ordre : 12

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 4.0235, df = 12, p-value = 0.983

LM d'ordre : 13

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 4.1682, df = 13, p-value = 0.9893

LM d'ordre : 14

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 4.1974, df = 14, p-value = 0.9942

LM d'ordre : 15

ARCH LM-test; Null hypothesis: no ARCH effects

data: volat$res
Chi-squared = 4.3427, df = 15, p-value = 0.9964

LM d'ordre : 16

ARCH LM-test; Null hypothesis: no ARCH effects

```



```
data: volat$res
Chi-squared = 4.3554, df = 16, p-value = 0.9981
```

```
LM d'ordre : 17
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 4.7573, df = 17, p-value = 0.9984
```

```
LM d'ordre : 18
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 5.8128, df = 18, p-value = 0.9969
```

```
LM d'ordre : 19
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 5.7602, df = 19, p-value = 0.9984
```

```
LM d'ordre : 20
```

```
ARCH LM-test; Null hypothesis: no ARCH effects
```

```
data: volat$res
Chi-squared = 5.7742, df = 20, p-value = 0.9992
```

Ici toutes les p-valeurs sont supérieures à 0.05, donc toute l'hétéroscédasticité conditionnelle à été prise en compte.

```
> anscombe.test(volat$res)
```

```
Anscombe-Glynn kurtosis test
```

```
data: volat$res
kurt = 6.4147, z = 9.9236, p-value < 2.2e-16
alternative hypothesis: kurtosis is not equal to 3
```

Le kurtosis de 6.4 indique que les valeurs extrêmes sont sur-représentées parmis les résidus comparé à loi normale.

```
###-----
###      6 : Effet de levier
###-----
```

```
> sig<-rep(0,T)
> for(t in 1:T)
+ {
+   sig[t]<-sqrt(sum(rtt[t-22]-(sum(rtt[t-22]/22)))^2/22)
+ }
> sigma=sig[24:T]*100
> plot(log(pt[24:length(rtt)]),type='l',col=2,axes=F,xlab="", ylab="",lwd=3)
> axis(2,at=seq(1,3.5,by=0.25)) #axe de gauche
> par(new=T)
> plot(sigma, col="grey",type='l',axes = F,xlab="", ylab="")
> axis(4,at=seq(0,4,by=0.25))#axe de droite
> legend("topleft", c("log(pt)","sigma"),col = c(2, 1),lty=c(1,1))
```

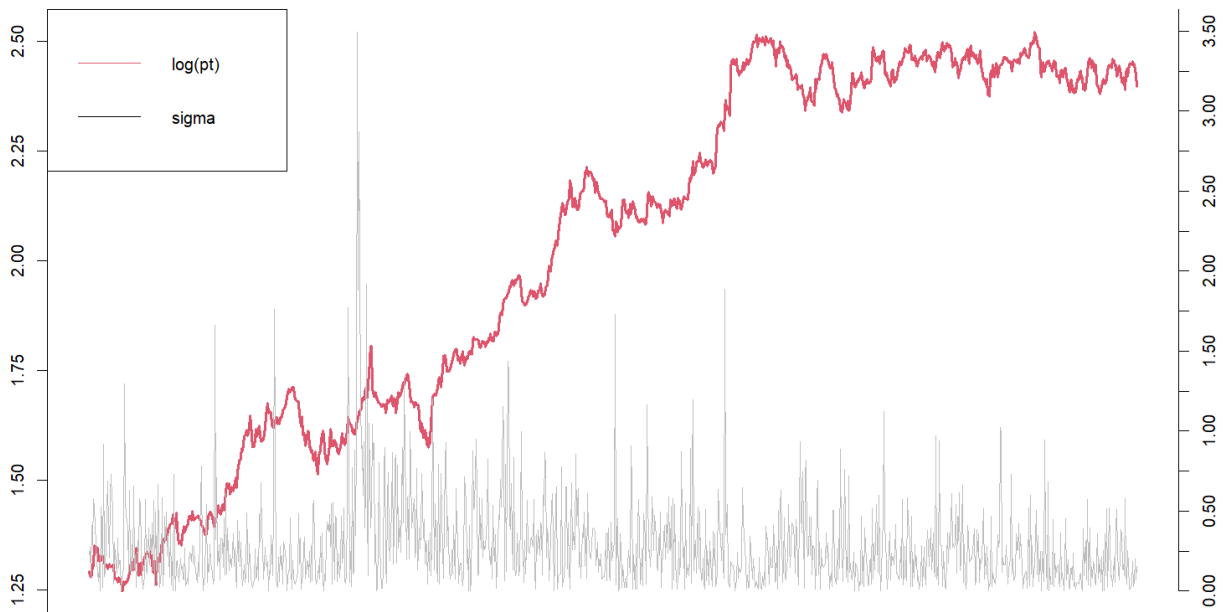


Figure 5: Effet de levier d'AmBev sur les données *rtt*: relation entre le logarithme et l'écart-type des rendements

On remarque l'absence d'effet de levier, car les volatilités les plus importantes ne sont pas visibles pendant des phases de baisse du cours.

```
###-----
###      7 : La saisonnalité
###-----

### EFFET WEEK-END

jour=format(dates[1:T], format = "%A")
> tableaures <- data.frame(matrix(NA,ncol=5,nrow=4))
> colnames(tableaures) <- c("lundi","mardi","mercredi","jeudi","vendredi")
> rownames(tableaures) <- c("moyenne en %","écart-type annuel en %","skewness","kurtosis")
>
> # Données du mardi : remplissage du tableau
> rtmar<-as.numeric(rtt[jour=="mardi"])
> mardi<-mean(rtmar) #moyenne journaliere
> tableaures[1,2] <- mardi*100 #moyenne journaliere en %
> tableaures[2,2] <- sd(rtmar)*100*sqrt(252) #ecart-type annualise en %
> tableaures[3,2] <- skewness(rtmar)
> tableaures[4,2] <- kurtosis(rtmar)
>
> rtmer<-as.numeric(rtt[jour=="mercredi"])
> mer<-mean(rtmer)
> tableaures[1,3] <- mer*100
```

```

> tableaures[2,3] <- sd(rtmer)*100*sqrt(252)
> tableaures[3,3] <- skewness(rtmer)
> tableaures[4,3] <- kurtosis(rtmer)
>
> rtjeu<-as.numeric(rtt[jour=="jeudi"])
> jeudi<-mean(rtjeu)
> tableaures[1,4] <- jeudi*100
> tableaures[2,4] <- sd(rtjeu)*100*sqrt(252)
> tableaures[3,4] <- skewness(rtjeu)
> tableaures[4,4] <- kurtosis(rtjeu)
>
> rtven<-as.numeric(rtt[jour=="vendredi"])
> ven<-mean(rtven)
> tableaures[1,5] <- ven*100
> tableaures[2,5] <- sd(rtven)*100*sqrt(252)
> tableaures[3,5] <- skewness(rtven)
> tableaures[4,5] <- kurtosis(rtven)
>
> rtlun<-as.numeric(rtt[jour=="lundi"])
> lundi<-mean(rtlun)
> tableaures[1,1] <- lundi*100
> tableaures[2,1] <- sd(rtlun)*100*sqrt(252)
> tableaures[3,1] <- skewness(rtlun)
> tableaures[4,1] <- kurtosis(rtlun)
>
> tableaures

```

	lundi	mardi	mercredi	jeudi	vendredi
moyenne en %	0.1030654	-0.1536909	-0.007554822	-0.1220013	0.1754332
écart-type annuel en %	35.0693542	27.8083241	27.988059532	34.1369009	31.7682229
skewness	-0.6860862	-0.1724195	0.391936469	-1.7219205	0.3722583
kurtosis	9.1136852	3.6973350	1.734367286	16.8561168	2.6140701

On remarque ici que les rendements sont plus élevées les lundis et vendredis et sont plus faibles les mardis. La variance (écart-type) est sensiblement la même tous les jours de la semaine. Il n'y a pas de saisonnalité hebdomadaire.

Effet Janvier

```

> monthplot(rtt, ylab="rendement",main="", cex.main=1,col.base=2,lwd.base=3)

```

On voit sur la figure 6 qu'il n'y a pas de saisonnalité mensuelle marquante sur les données *rtt*, malgré des pics de rendements négatifs en septembre et novembre.

```

###-----
###      8 : Stationnarité
###-----

```

```

> library(urca)
> summary(ur.df(rtt,type= "trend",lags=0))

```

```

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

```

```

Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt)

```

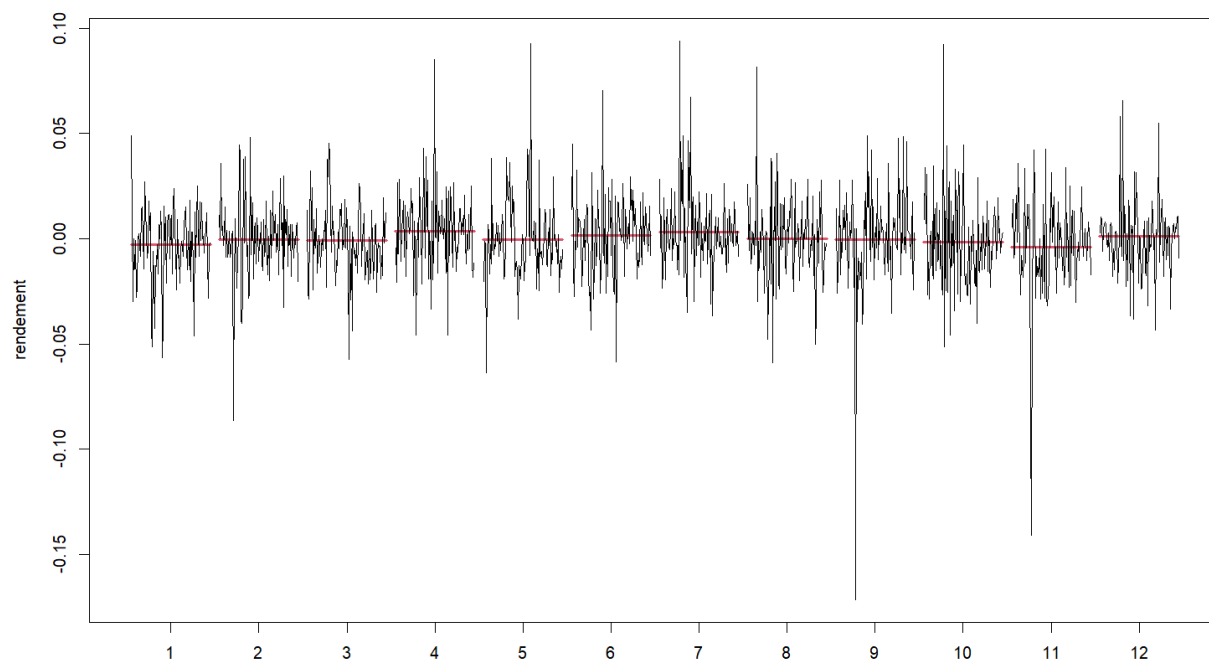


Figure 6: Rendement logarithmique mensuelle d'AmBev sur *rtt*

Residuals:

	Min	1Q	Median	3Q	Max
	-0.176241	-0.010070	-0.000093	0.010115	0.096957

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.275e-04	1.153e-03	0.197	0.844
z.lag.1	-1.092e+00	2.901e-02	-37.651	<2e-16 ***
tt	-4.762e-07	1.699e-06	-0.280	0.779

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

Residual standard error: 0.01975 on 1172 degrees of freedom

Multiple R-squared: 0.5474, Adjusted R-squared: 0.5466

F-statistic: 708.8 on 2 and 1172 DF, p-value: < 2.2e-16

Value of test-statistic is: -37.651 472.5387 708.8043

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.96	-3.41	-3.12
phi2	6.09	4.68	4.03
phi3	8.27	6.25	5.34

Le coefficient β_1 n'est pas significatif, car la p-valeur associé à *tt* est 0.77 > 0.05. Le processus générateur de *rtt* ne peut pas être TS.

```
plot(ur.df(rtt,lag=0,type="trend"))
> # On passe au modèle contenant une constante mais pas de tendance.
> summary(ur.df(rtt,type= "drift",lags=0))
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression drift

```
Call:
lm(formula = z.diff ~ z.lag.1 + 1)
```

```
Residuals:
      Min       1Q   Median       3Q      Max
-0.176098 -0.010073 -0.000132  0.010102  0.097088
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0000525  0.0005759  -0.091    0.927
z.lag.1      -1.0922884  0.0290002  -37.665 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.01974 on 1173 degrees of freedom
Multiple R-squared:  0.5474, Adjusted R-squared:  0.547
F-statistic: 1419 on 1 and 1173 DF, p-value: < 2.2e-16
```

Value of test-statistic is: -37.6649 709.3259

```
Critical values for test statistics:
      1pct  5pct 10pct
tau2  -3.43 -2.86 -2.57
phi1   6.43  4.59  3.78
```

Ici on voit que le coefficient d'interception, qui décrit β_0 n'est pas significatif, donc il faut faire le test de RU quand il n'existe ni tendance ni constante.

```
> summary(ur.df(rtt,type= "none",lags=0))
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression none

```
Call:
lm(formula = z.diff ~ z.lag.1 - 1)
```

```
Residuals:
      Min       1Q   Median       3Q      Max
-0.176150 -0.010126 -0.000185  0.010049  0.097036
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1 -1.092229    0.02899   -37.68 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

Residual standard error: 0.01973 on 1174 degrees of freedom

Multiple R-squared: 0.5474, Adjusted R-squared: 0.547

F-statistic: 1420 on 1 and 1174 DF, p-value: < 2.2e-16

Value of test-statistic is: -37.6808

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

On trouve $t = -37.68 < -1.95$. donc on rejette H_0 donc le PGD n'est pas DS. On en conclut que le PGD est stationnaire. Cette conclusion n'est valide que si les aléas de la régression de Dickey et Fuller ne sont pas auto-corrélés.

```
> plot(ur.df(rtt,lag=0,type="none"))
```

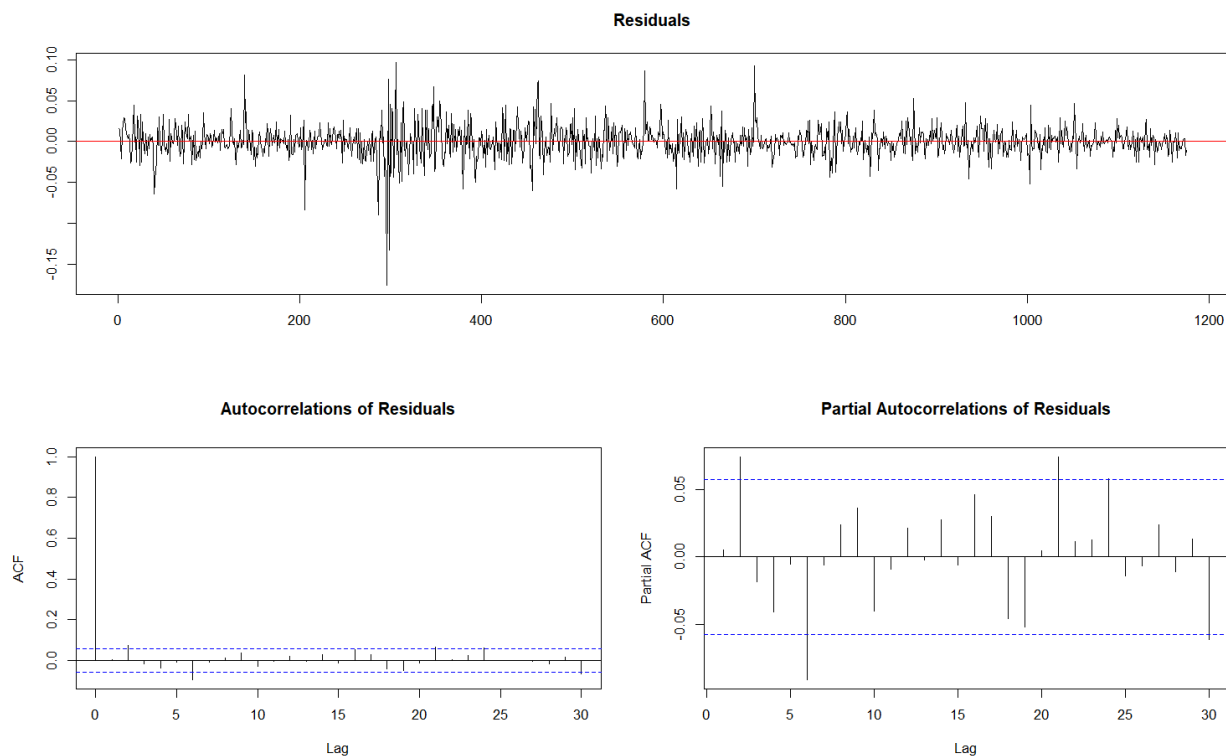


Figure 7: Corrélogramme des résidus de la régression DF "none"

On voit sur la figure 7 qu'il y a de l'autocorrélation partielle dans les résidus de rtt aux ordres 6, 22 et 30. Donc la conclusion précédente n'est pas valide et on doit effectuer un test de Diskey-Fuller Augmenté.

```
> ### test de Dickey Fuller Augmenté (ADF)
>
> # On introduit k variables explicatives
> # On determine k avec la partie entière de la formule de Schwert
> Schwert<-as.integer(12*(T/100)^(0.25))
> Schwert
```

```

[1] 22
>
> summary(ur.df(rtt,type= "none",lags=Schwert))

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.176290 -0.010680 -0.000186  0.009576  0.095415

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.039325   0.155540  -6.682 3.69e-11 ***
z.diff.lag1  -0.056020   0.152420  -0.368  0.7133
z.diff.lag2   0.031493   0.149152   0.211  0.8328
z.diff.lag3   0.036112   0.145236   0.249  0.8037
z.diff.lag4  -0.007889   0.141255  -0.056  0.9555
z.diff.lag5  -0.027086   0.137604  -0.197  0.8440
z.diff.lag6  -0.118188   0.134225  -0.881  0.3788
z.diff.lag7  -0.132947   0.130456  -1.019  0.3084
z.diff.lag8  -0.103568   0.126145  -0.821  0.4118
z.diff.lag9  -0.065667   0.121612  -0.540  0.5893
z.diff.lag10 -0.100693   0.116749  -0.862  0.3886
z.diff.lag11 -0.113205   0.111777  -1.013  0.3114
z.diff.lag12 -0.100448   0.106470  -0.943  0.3457
z.diff.lag13 -0.100365   0.100928  -0.994  0.3202
z.diff.lag14 -0.076670   0.095389  -0.804  0.4217
z.diff.lag15 -0.084128   0.089135  -0.944  0.3455
z.diff.lag16 -0.040404   0.082185  -0.492  0.6231
z.diff.lag17  0.002771   0.074860   0.037  0.9705
z.diff.lag18 -0.034939   0.067970  -0.514  0.6073
z.diff.lag19 -0.096779   0.060469  -1.600  0.1098
z.diff.lag20 -0.101149   0.052694  -1.920  0.0552 .
z.diff.lag21 -0.027842   0.043655  -0.638  0.5237
z.diff.lag22 -0.009289   0.029487  -0.315  0.7528
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01957 on 1130 degrees of freedom
Multiple R-squared:  0.5654, Adjusted R-squared:  0.5566
F-statistic: 63.92 on 23 and 1130 DF, p-value: < 2.2e-16

Value of test-statistic is: -6.6821

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62

```

On diminue la valeur de lag jusqu'à avoir tous les gammas avec une valeur de la statistique t calculée supérieure en valeur absolue à 1.6.

```
>summary(ur.df(rtt,type= "none",lags=20))
>summary(ur.df(rtt,type= "none",lags=10))
>summary(ur.df(rtt,type= "none",lags=7))
>summary(ur.df(rtt,type= "none",lags=5))
>summary(ur.df(rtt,type= "none",lags=4))
>summary(ur.df(rtt,type= "none",lags=3))
>summary(ur.df(rtt,type= "none",lags=2))

> summary(ur.df(rtt,type= "none",lags=1))

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.174578 -0.010063 -0.000319  0.009892  0.098338

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.01047    0.04300  -23.497  <2e-16 ***
z.diff.lag   -0.07636    0.02905   -2.628   0.0087 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01969 on 1172 degrees of freedom
Multiple R-squared:  0.5496, Adjusted R-squared:  0.5489
F-statistic: 715.2 on 2 and 1172 DF,  p-value: < 2.2e-16

Value of test-statistic is: -23.497

Critical values for test statistics:
    1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

Pour $k=1$, toutes les valeurs t sont supérieures en valeur absolue à 1.6. La valeur du test statistique est $-23.497 < -1.95$ donc le test ADF permet de conclure à la stationnarité du PGD qui a généré *rtt*.