Exploration des indicateurs comptables des sociétés non-financières françaises depuis 1950

Jolibert Grégoire & Combe Oscar

Université Paris-Dauphine, Décembre 2023

Introduction :

Ce mémoire a pour but d'étudier les relations existantes, depuis 1950, entre les principaux indicateurs comptables des sociétés non-financières françaises. Rappelons que les sociétés non financières constituent l'ensemble des unités institutionnelles qui sont des producteurs marchands dont la fonction principale consiste à produire des biens et des services non financiers.

Au départ, nous les étudierons distinctement dans le but d'observer leurs caractéristiques propres. Il s'agira ensuite de les confronter afin d'analyser leurs corrélations dans le temps.

Enfin, ce choix d'étude nous est apparu pertinent de part ses perspectives Statistiques et Comptables, deux matières essentielles au champ économique.
Il était en effet important pour nous d'opter pour un sujet et des données qui puissent favoriser des interprétations économiques intéréssantes donnant un sens plus concret à l'étude.

Les indicateurs comptables utilisés dans notre base de données**:
- Taux de marge.
- Taux d'investissement.
- Taux d'autofinancement.

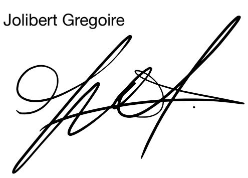
** *Source : INSEE.*
** *Champ : France, sociétés non financières (hors entreprises individuelles)*

Sommaire :

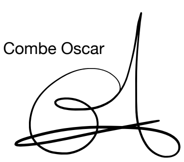
"Nous déclarons sur l'honneur que ce mémoire a été écrit de notre main, sans aide extérieure non autorisée, qu'il n'a pas été présenté auparavant pour évaluation et qu'il n'a jamais été publié, dans sa totalité ou en partie. Toutes parties, groupes de mots ou idées, aussi limités soient-ils, y compris des tableaux, graphiques, cartes etc. qui sont empruntés ou qui font référence à d'autres sources bibliographiques sont présentés comme tels, sans exception aucune."

Rendu le 5 décembre 2023,
à M. Lada, Professeur à l'Université Paris-Dauphine

Jolibert Gregoire



Combe Oscar



Question I : Description des données

Question 1.a) : Définition des variables

- La population est l'ensemble des sociétés en France.
- L'unité statistiques est représentée, dans cette étude, par les sociétés non financières.
- Les variables :
  - Date (allant de 1950 à 2022), c'est une variable qualitative ordinale exprimée en année.
  - $Taux\ de\ marge = \frac{Excédent\ Brut\ d'Exploitation}{Valeur\ Ajoutée\ brute}$, c'est une variable quantitative continue exprimée en pourcentage.
  - $Taux\ d'investissement = \frac{Formation\ Brute\ de\ Capital\ Fixe}{Valeur\ Ajoutée\ brute}$, c'est une variable quantitative continue exprimée en pourcentage.
  - $Taux\ d'autofinancement = \frac{Epargne\ Brute}{Formation\ Brute\ de\ Capital\ Fixe}$, c'est une variable quantitative continue exprimée en pourcentage.

Question 1.b) : Distribution des variables

- Pour notre première variable "date", de nature qualitative ordinale, il n'est pas utile de faire un tableau de fréquence et une représentation graphique associée.
  En effet, les écarts entre les années sont uniformes d'une unité.
- Pour nos prochaines variables "taux de marge, d'investissement et d'autofinancement", de nature quantitative continue, on utilisera un histogramme :

```
#Introduction au code
rm(list=ls())
library(readxl)
options(digits=4)
Base1 <- read_excel("Base_ProjetR.xlsx")
View(Base1)
anyNA(Base1)
```

```
## [1] FALSE
```

```
#Déclarartion des racourcis pour les différentes variables
tauxdemarge1 <- Base1$`Taux de marge1`
tauxdinvest2 <- Base1$`Taux d'investissement2`
tauxdautof3 <- Base1$`Taux d'autofinancement3`
```

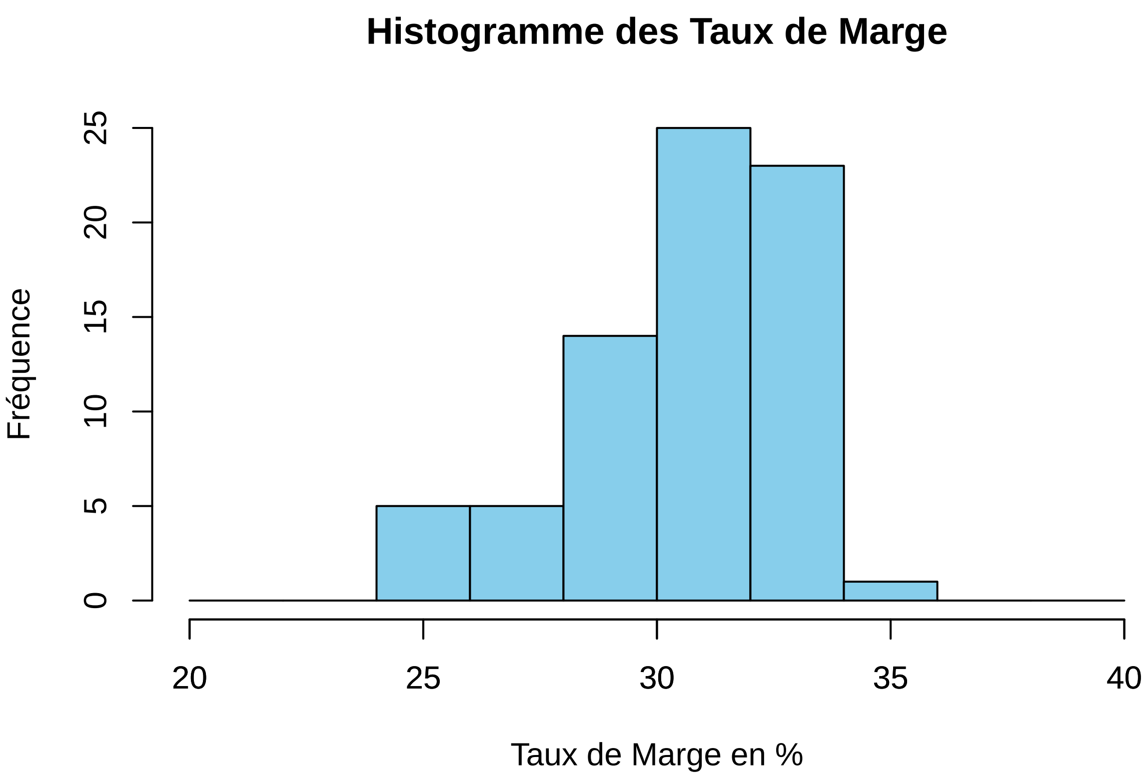
```
#tableaux de fréquence du taux de marge
borne_Marge<-seq(20,40,2)
Taux_de_marge<- cut(tauxdemarge1, borne_Marge)

#tableau de fréquence en numéraire
addmargins(table(Taux_de_marge))
```

```
## Taux_de_marge
## (20,22] (22,24] (24,26] (26,28] (28,30] (30,32] (32,34] (34,36] (36,38] (38,40]
##       0       0       5       5      14      25      23       1       0       0
##     Sum
##      73
```

```r
#tableau de féquence en %
formatC(addmargins(prop.table(table(Taux_de_marge))) * 100, format = "f", digits = 2)
```

```
## Taux_de_marge
##  (20,22]  (22,24]  (24,26]  (26,28]  (28,30]  (30,32]  (32,34]  (34,36]
##   "0.00"   "0.00"   "6.85"   "6.85"  "19.18"  "34.25"  "31.51"   "1.37"
##  (36,38]  (38,40]      Sum
##   "0.00"   "0.00" "100.00"
```

```r
#histogramme du taux de marge
hist(tauxdemarge1, breaks = borne_Marge,
col = "skyblue", border = "black",
main = "Histogramme des Taux de Marge", xlab = "Taux de Marge en %",
ylab = "Fréquence")
axis(1, at = seq(20, 40, 5))
```



Interprétation : On remarque ici une distribution assez déséquilibrée, en effet deux tiers des taux de marge se situent entre 30 et 34 pourcents. Nous pouvons suggérer que le taux de marge est relativement stable depuis 1950. Cette stabilité à travers le temps pourrait indiquer une gestion financière efficace des sociétés non financières française, évoquant ainsi une pérennité historique dans leur capacité à maintenir des niveaux de rentabilité constants, et par conséquent, à s'adapter aux changements du marché.

```r
#tableaux de fréquence du taux d'investissement
borne_Investissement<-seq(15,30,2)
Taux_dInvest<- cut(tauxdinvest2, borne_Investissement)

#tableau de fréquence en numéraire
addmargins(table(Taux_dInvest))
```

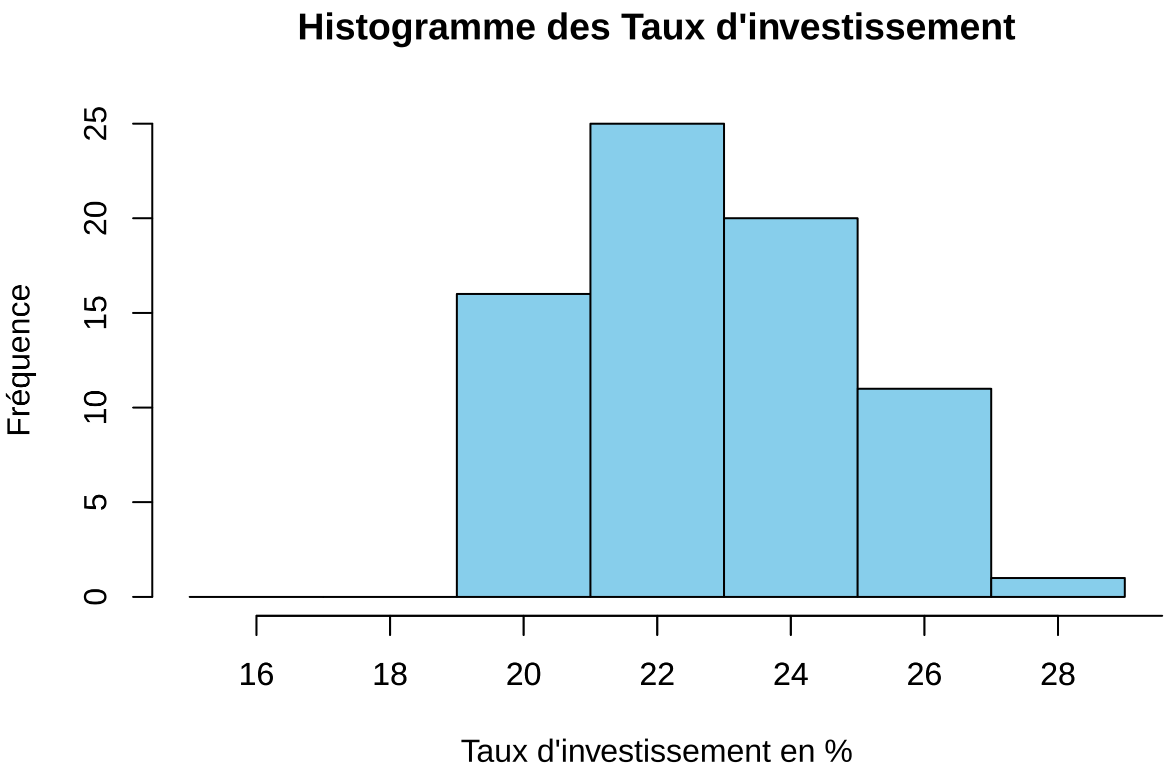
```
## Taux_dInvest
```

```
## (15,17] (17,19] (19,21] (21,23] (23,25] (25,27] (27,29]     Sum
##       0       0      16      25      20      11       1      73
```

```
#tableau de féquence en %
formatC(addmargins(prop.table(table(Taux_dInvest))) * 100, format = "f", digits = 2)
```

```
## Taux_dInvest
##  (15,17]  (17,19]  (19,21]  (21,23]  (23,25]  (25,27]  (27,29]      Sum
##   "0.00"   "0.00"  "21.92"  "34.25"  "27.40"  "15.07"   "1.37" "100.00"
```

```
#histogramme du taux d'investissement
hist(tauxdinvest2, breaks = borne_Investissement,
col = "skyblue", border = "black",
main = "Histogramme des Taux d'investissement", xlab = "Taux d'investissement en %",
ylab = "Fréquence")
axis(1, at = seq(16, 32, 2))
```



Interprétation : On remarque une légère prépondérance du taux d'investissement compris entre 21 et 23 pourcent, mais globalement les fréquences sont distribuées équitablement pour des taux allant de 19 à 27 pourcents. Ainsi, nous pouvons parler de variabilité des taux d'investissement depuis 1950. Cette variabilité globale des taux d'investissement entre 19 et 27 pourcents souligne l'adaptabilité des entreprises face aux changements économiques. Par exemple, une entreprise ayant initialement investi 19% de son chiffre d'affaires dans la recherche et le développement (R&D) pourrait, en réponse à des changements économiques favorables, augmenter ce taux à 27% pour renforcer son innovation et sa compétitivité sur le marché.

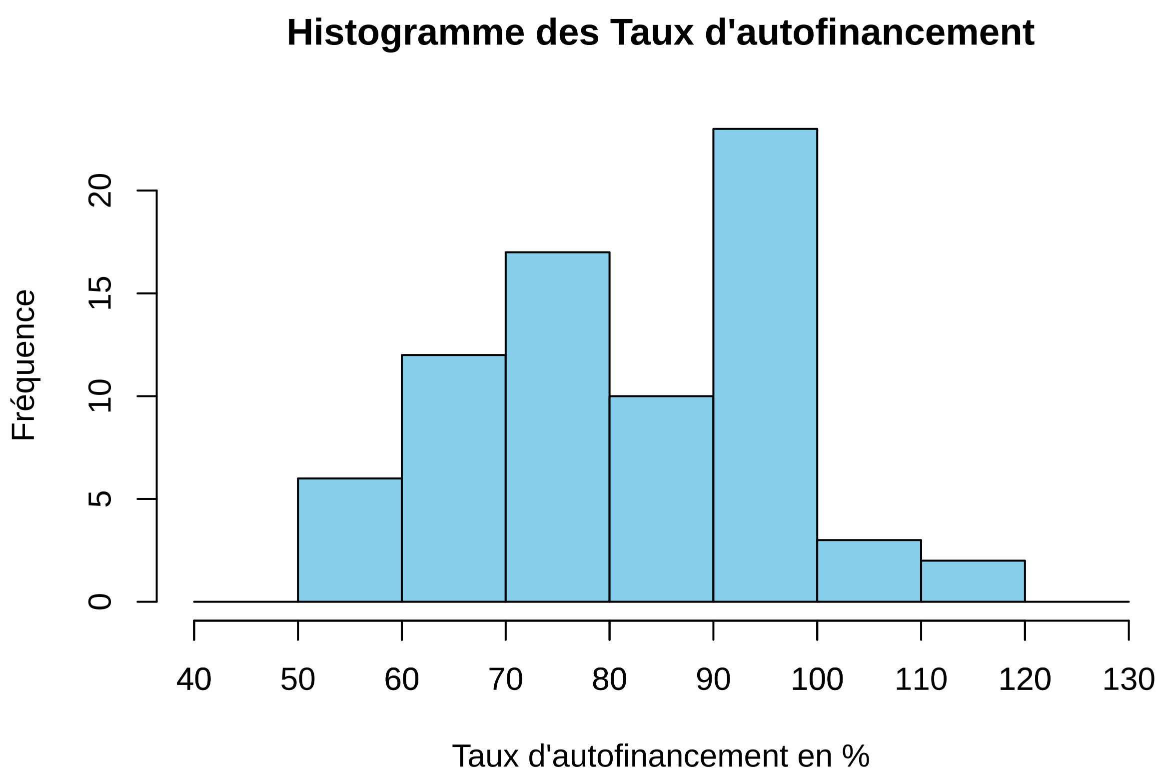
```r
#tableaux de fréquence du taux d'autofinancement
borne_Autof<-seq(40,130,10)
Taux_dAutof<- cut(tauxdautof3, borne_Autof)

#tableau de fréquence en numéraire
addmargins(table(Taux_dAutof))
```

```
## Taux_dAutof
##   (40,50]   (50,60]   (60,70]   (70,80]   (80,90]  (90,100] (100,110] (110,120]
##         0         6        12        17        10        23         3         2
## (120,130]       Sum
##         0        73
```

```r
#tableau de féquence en %
formatC(addmargins(prop.table(table(Taux_dAutof))) * 100, format = "f", digits = 2)
```

```
## Taux_dAutof
##   (40,50]   (50,60]   (60,70]   (70,80]   (80,90]  (90,100] (100,110] (110,120]
##    "0.00"    "8.22"   "16.44"   "23.29"   "13.70"   "31.51"    "4.11"    "2.74"
## (120,130]       Sum
##    "0.00"  "100.00"
```

```r
#histogramme du taux d'autofinancement
hist(tauxdautof3, breaks=borne_Autof,
col = "skyblue", border = "black",
main = "Histogramme des Taux d'autofinancement", xlab = "Taux d'autofinancement en %",
ylab = "Fréquence")
axis(1, at = seq(40, 130, 10))
```



Interprétation : Comme pour le taux d'investissement, on remarque une certaine hétérogénéité dans la distribution du taux d'autofinancement. Nous pouvons également parler de variabilité depuis 1950, malgré une prépondérance du taux compris entre 90 et 100 pourcents. Reprenons notre exemple précédent, si une entreprise effectue d'importants investissements en R&D pour stimuler l'innovation, elle peut connaître une baisse du taux d'autofinancement si les bénéfices générés par ces investissements tardent à se concrétiser. Le manque immédiat de retour sur investissement peut réduire la capacité de l'entreprise à autofinancer ses activités.

Conclusion :

Les statistiques descriptives jouent un rôle crucial dans l'analyse économique en résumant et décrivant les caractéristiques clés des données comptables, telles que le taux de marge, le taux d'investissement et le taux de marge. Elles fournissent une vue d'ensemble de la tendance centrale et de la dispersion de ces variables, permettant ainsi une compréhension approfondie de la situation économique à l'échelle globale. En l'absence de ces outils statistiques, il serait difficile de définir de manière significative l'analyse économique, surtout étant donné que les indicateurs comptables sont de nature micro-économique, se concentrant sur des entités spécifiques telles que les entreprises.

Question II : Estimation de la moyenne théorique

Dans cette *Question II*, nous chercherons à estimer la moyenne, plus que la variance, afin de l'utiliser pour la suite de notre projet (*cf. Question III*).

Question 2.a) : Estimation ponctuelle de la moyenne

Nous utiliserons la moyenne empirique qui est un estimateur convergent et sans biais de la moyenne théorique (LGN), rappelons sa formule : $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

```
#calcul des moyennes
Xbar_TauxdeMarge <- mean(tauxdemarge1)
Xbar_Tauxdinvest <- mean(tauxdinvest2)
Xbar_Tauxdautofi <- mean(tauxdautof3)

# Création d'un tableau
tableau_moyennes <- data.frame(
  Variable = c("Taux de marge", "Taux d'investissement", "Taux d'autofinancement"),
  Moyenne = format(c(Xbar_TauxdeMarge, Xbar_Tauxdinvest, Xbar_Tauxdautofi),digits=4)
  )

# Affichage des résultats sous forme d'un tableau
print(tableau_moyennes)
```

```
##                 Variable Moyenne
## 1          Taux de marge   30.51
## 2  Taux d'investissement   22.88
## 3 Taux d'autofinancement   81.64
```

Question 2.b) : Estimation de la moyenne par Intervalle de Confiance (IC) de niveau 90% & 95%, avec une variance inconnue

Dans cette *Question 2.b)*, nous réaliserons une estimation par intervalle de confiance seulement sur deux de nos variables afin d'éviter toute répétition.

Rédaction :

La moyenne empirique est un estimateur convergent et sans biais de la moyenne théorique, nous l'utiliserons donc pour nos constructions d'IC.
D'après la LGN, la variance empirique modifiée estime la variance théorique : $S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$
Ayant un échantillon de taille n > 30 et une variance inconnue, nous pouvons utiliser le TCL, et appliquer les théorèmes de Continuité puis de Slutsky, afin de centrer et réduire notre moyenne empirique.

Nous obtenons donc la fonction pivotale pour la moyenne (notée m), qui sera utilisée dans nos IC :

$$\frac{\overline{X} - m}{\frac{S'}{\sqrt{n}}} \xrightarrow{d} N(0,1)$$

On peut ainsi calculer les différents IC :

1) IC de niveau 90% pour le taux de marge

```
n1 <-length(tauxdemarge1)
scor1 <- sd(tauxdemarge1)
```

```
a1 <- Xbar_TauxdeMarge - qnorm(0.95)*scor1/sqrt(n1)
a2 <- Xbar_TauxdeMarge + qnorm(0.95)*scor1/sqrt(n1)
c(a1,a2)
```

```
## [1] 30.05 30.97
```

- IC de niveau 90% pour le taux de marge: [30,05 ; 30,97]

2) IC de niveau 95% pour le taux de marge

```
n1 <-length(tauxdemarge1)
scor1 <- sd(tauxdemarge1)
a1_bis <- Xbar_TauxdeMarge - qnorm(0.975)*scor1/sqrt(n1)
a2_bis <- Xbar_TauxdeMarge + qnorm(0.975)*scor1/sqrt(n1)
c(a1_bis,a2_bis)
```

```
## [1] 29.96 31.06
```

- IC de niveau 95% pour le taux de marge: [29,96 ; 31,06]

Interprétation du 1er {resp. 2ème} IC:

Il indique que nous sommes certains à 90% {resp.95%} que '*la moyenne théorique du taux de marge*' se situe dans l'intervalle [30,05 ; 30,97] {resp.[29,96 ; 31,06] }.
Dans ce cas, la valeur estimée de 30,51 est incluse dans cet intervalle, ce qui suggère que notre estimation est cohérente avec la vraie moyenne théorique.

3) IC de niveau 90% pour le taux d'investissement

```
n2 <- length(tauxdinvest2)
scor2 <- sd(tauxdinvest2)
b1 <- Xbar_Tauxdinvest - qnorm(0.95)*scor2/sqrt(n2)
b2 <- Xbar_Tauxdinvest + qnorm(0.95)*scor2/sqrt(n2)
c(b1,b2)
```

```
## [1] 22.51 23.25
```

- IC de niveau 90% pour le taux d'investissement: [22,51 ; 23,25]

4) IC de niveau 95% pour le taux d'investissement

```
n2 <- length(tauxdinvest2)
scor2 <- sd(tauxdinvest2)
b1_bis <- Xbar_Tauxdinvest - qnorm(0.975)*scor2/sqrt(n2)
b2_bis <- Xbar_Tauxdinvest + qnorm(0.975)*scor2/sqrt(n2)
c(b1_bis,b2_bis)
```

```
## [1] 22.44 23.32
```

- IC de niveau 95% pour le taux d'investissement: [22,44 ; 23,32]

Interprétation du 3ème {resp. 4ème} IC:

Il indique que nous sommes certains à 90% {resp.95%} que '*la moyenne théorique du taux d'investissement*' se situe dans l'intervalle [22,51 ; 23,25] {resp.[22,44 ; 23,32] }.
Dans ce cas, la valeur estimée de 22.88 est incluse dans cet intervalle, ce qui suggère que notre estimation est cohérente avec la vraie moyenne théorique.

Conclusion :

L'évaluation des intervalles de confiance sur des variables comptables telles que le taux de marge, le taux d'investissement ou encore le taux d'autofinancement, présente plusieurs avantages importants dans le domaine financier et comptable.

En fournissant un intervalle de confiance, on communique plus efficacement l'incertitude entourant les estimations. Cela aide les décideurs à comprendre que les valeurs rapportées ne sont pas des certitudes absolues, mais plutôt des estimations avec une certaine marge d'erreur.
Cela contribue à une prise de décision plus éclairée, car les gestionnaires peuvent améliorer leur évaluation des risques liés aux données financières.

Enfin, en comparant les intervalles de confiance entre différentes périodes ou secteurs, on peut déterminer si des différences observées sont statistiquement significatives. Cela permet d'identifier les variations importantes par rapport à la simple comparaison des valeurs ponctuelles.

Question III : Test de conformité de la moyenne théorique au seuil de 5% & 10%, avec une variance inconnue

Dans cette *Question III*, nous réaliserons les 3 types de tests de conformités à 5% & 10% de la manière suivante :

- unilatéral à droite (avec le taux de marge). D'après le TNP, la région critique s'écrit W tel que $(\overline{X} > c)$
- unilatéral à gauche (avec le taux d'investissemnt). D'après le TNP, la région critique s'écrit W tel que $(\overline{X} < c)$
- bilatéral (avec le taux d'autofinancement). D'après le TNP, la région critique s'écrit W tel que $(\overline{X} < c_1)$ *ou* $(\overline{X} > c_2)$

Remarque : la valeur c représente la valeur critique, elle est obtenue grâce à la Contrainte de Première Espèce (CPE)

Rédaction :

La moyenne empirique est un estimateur convergent et sans biais de la moyenne théorique, elle sera donc notre statistique de test.
Ayant un échantillon de taille n > 30 et une variance inconnue, nous pouvons utiliser le TCL, et appliquer les théorèmes de Continuité puis de Slutsky, afin de centrer et réduire notre statistique de test.

Nous obtenons alors U qui servira dans la CPE afin de trouver la valeur critique :

$$U = \frac{\overline{X} - m}{\frac{S'}{\sqrt{n}}} \xrightarrow{d} N(0, 1)$$

On peut ainsi calculer les différents tests :

1) Test sur la moyenne du taux de marge, au seuil æ = 10% i.e avec un IC de 90%.

```
#Hypothèse 0: mean(tauxdemarge1) = 30.51
#Hypothèse 1: mean(tauxdemarge1) > 30.51
#CPE: P(X>c|HO) = æ

HO <- 30.51
#Test unilatéral à droite
t.test(tauxdemarge1, mu = HO, alternative = 'greater',  conf.level = 0.90)
```

```
##
##  One Sample t-test
##
## data:  tauxdemarge1
## t = -0.0067, df = 72, p-value = 0.5
## alternative hypothesis: true mean is greater than 30.51
## 90 percent confidence interval:
##  30.15   Inf
## sample estimates:
## mean of x
##     30.51
```

- En sachant que notre æ = 0.10, et que notre p-value = 0.50, nous pouvons affirmer que HO n'est pas rejetée (car æ < p-value). Ainsi, nous sommes certains à 90% que notre vrai moyenne théorique se situe au dessus de 30.15.

2) Test sur la moyenne du taux de marge, au seuil æ = 5% i.e avec un IC de 95%.

```r
#Hypothèse 0: mean(tauxdemarge1) = 30.51
#Hypothèse 1: mean(tauxdemarge1) > 30.51
#CPE: P(X>c|HO) = æ

HO <- 30.51
#Test unilatéral à droite
t.test(tauxdemarge1, mu = HO, alternative = 'greater',  conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  tauxdemarge1
## t = -0.0067, df = 72, p-value = 0.5
## alternative hypothesis: true mean is greater than 30.51
## 95 percent confidence interval:
##  30.04   Inf
## sample estimates:
## mean of x
##     30.51
```

- En sachant que notre æ = 0.05, et que notre p-value = 0.50, nous pouvons affirmer que HO n'est pas rejetée (car æ < p-value). Ainsi, nous sommes certains à 95% que notre vrai moyenne théorique se situe au dessus de 30.04.

3) Test sur la moyenne du taux d'investissement, au seuil æ = 10% i.e avec un IC de 90%.

```r
#Hypothèse 0_bis: mean(tauxdinvest2) = 22.88
#Hypothèse 1_bis: mean(tauxdinvest2) < 22.88
#CPE: P(X<c|HO) = æ

HO_bis <- 22.88
#Test unilatéral à gauche
t.test(tauxdinvest2, mu = HO_bis, alternative = 'less',  conf.level = 0.90)
```

```
##
##  One Sample t-test
##
## data:  tauxdinvest2
## t = 0.0029, df = 72, p-value = 0.5
## alternative hypothesis: true mean is less than 22.88
## 90 percent confidence interval:
##   -Inf 23.17
## sample estimates:
## mean of x
##     22.88
```

- En sachant que notre æ = 0.10, et que notre p-value = 0.50, nous pouvons affirmer que HO_bis n'est pas rejetée (car æ < p-value). Ainsi, nous sommes certains à 90% que notre vrai moyenne théorique se situe en dessous de 23.17.

4) Test sur la moyenne du taux d'investissement, au seuil æ = 5% i.e avec un IC de 95%.

```r
#Hypothèse 0_bis: mean(tauxdinvest2) = 22.88
#Hypothèse 1_bis: mean(tauxdinvest2) < 22.88
#CPE: P(X<c|HO) = æ

HO_bis <- 22.88
#Test unilatéral à gauche
t.test(tauxdinvest2, mu = HO_bis, alternative = 'less',  conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  tauxdinvest2
## t = 0.0029, df = 72, p-value = 0.5
## alternative hypothesis: true mean is less than 22.88
## 95 percent confidence interval:
##   -Inf 23.26
## sample estimates:
## mean of x
##     22.88
```

- En sachant que notre æ = 0.05, et que notre p-value = 0.50, nous pouvons affirmer que HO_bis n'est pas rejetée (car æ < p-value). Ainsi, nous sommes certains à 95% que notre vrai moyenne théorique se situe en dessous de 23.26.

5) Test sur la moyenne du taux d'autofinancement, au seuil æ = 10% i.e avec un IC de 90%

```r
#Hypothèse 0_ter: mean(tauxdautof3) = 81.64
#Hypothèse 1_ter: mean(tauxdautof3) \neq 81.64 (i.e différent de 81.64)
#CPE: P(c1<X<c2|HO) = 1-æ

HO_ter <- 81.64
t.test(tauxdautof3, mu = HO_ter,  conf.level = 0.90) #Test bilatéral
```

```
##
##  One Sample t-test
##
## data:  tauxdautof3
## t = -0.001, df = 72, p-value = 1
## alternative hypothesis: true mean is not equal to 81.64
## 90 percent confidence interval:
##  78.67 84.61
## sample estimates:
## mean of x
##     81.64
```

- Ici, la p-value est égale à 1, cela nous indique un manque de preuves statistiques pour rejeter H0_ter. Ainsi, nous sommes certains à 90% que notre vraie moyenne théorique se situe entre 78.67 et 84.61.

6) Test sur la moyenne du taux d'autofinancement, au seuil æ = 5% i.e avec un IC de 95%.

```r
#Hypothèse 0_ter: mean(tauxdautof3) = 81.64
#Hypothèse 1_ter: mean(tauxdautof3) \neq 81.64 (i.e différent de 81.64)
#CPE: P(c1<X<c2|HO) = 1-æ
HO_ter <- 81.64
```

```r
t.test(tauxdautof3, mu = H0_ter,  conf.level = 0.95) #Test bilatéral
```

```
##
##  One Sample t-test
##
## data:  tauxdautof3
## t = -0.001, df = 72, p-value = 1
## alternative hypothesis: true mean is not equal to 81.64
## 95 percent confidence interval:
##  78.09 85.19
## sample estimates:
## mean of x
##     81.64
```

- Ici, la p-value est égale à 1, cela nous indique un manque de preuves statistiques pour rejeter H0_ter. Ainsi, nous sommes certains à 95% que notre vraie moyenne théorique se situe entre 78.09 et 85.19.

Conclusion :

Les tests d'hypothèses permettent de valider ou de réfuter des hypothèses spécifiques sur les variables financières. Par exemple, on pourrait tester si le taux d'autofinancement moyen d'une entreprise diffère de celui de l'industrie dans son ensemble.
Les résultats des tests d'hypothèses fournissent une base objective pour la prise de décision. Par exemple, si un test suggère que l'introduction d'une nouvelle stratégie a un impact significatif sur le taux d'investissement, cela peut influencer les décisions futures.

Question IV : Test de comparaison des moyennes théoriques de deux variables au seuil de 5% & 10%, avec des variances inconnues

Dans cette *Question IV*, nous comparerons la moyenne du taux de marge à celle du taux d'investissement puis à celle du taux d'autofinancement, afin d'étudier de possibles corrélations.

Posons les hypothèses : $\begin{cases} H_0 : m_1 - m_2 = 0 \\ H_1 : m_1 - m_2 \neq 0 \end{cases}$
D'après le TNP, la région critique s'écrit W tel que $(\overline{X_1} - \overline{X_2} < c_1)$ *ou* $(\overline{X_1} - \overline{X_2} > c_2)$.

Rédaction :

Rappelons que la moyenne empirique est un estimateur convergent et sans biais de la moyenne théorique.
Notre statistique de test s'écrira alors : $\overline{X_1} - \overline{X_2}$
Ayant 2 échantillons de taille n > 30 et des variances inconnues, nous pouvons utiliser le TCL, et appliquer les théorèmes de Continuité puis de Slutsky, afin de centrer et réduire notre statistique de test.

Nous obtenons alors U qui servira dans la CPE afin de trouver les valeurs critiques :

Cas où : $\sigma_1 = \sigma_2$

$$U = \frac{(\overline{X_1} - \overline{X_2}) - (m_1 - m_2)}{\sqrt{(\frac{1}{n_1} + \frac{1}{n_2}).S''^2}} \xrightarrow{d} N(0,1)$$

$avec,\ S''^2 = \frac{n_1.S_1^2 + n_2.S_2^2}{n_1 + n_2}$

Cas où : $\sigma_1 \neq \sigma_2$

$$U = \frac{(\overline{X_1} - \overline{X_2}) - (m_1 - m_2)}{\sqrt{(\frac{S_1'^2}{n_1} + \frac{S_2'^2}{n_2})}} \xrightarrow{d} N(0,1)$$

On peut ainsi calculer les différents tests :

Remarque : Au préalable, pour comparer deux moyennes, il conviendra de faire un test de comparaison des variances (des variables en questions) afin de vérifier si nous sommes, ou non, dans un cas avec variances égales.

1) Test de comparaison, au seuil æ=5%, des moyennes du taux de marge et du taux d'investissement

```
#Vérifions si les variances des deux variables sont égales, ou non, avec un test au seuil æ=5%

#Hypothèse 0 : var_tauxdemarge1 = var_tauxdinvest2
#Hypothèse 1 : var_tauxdemarge1 \neq var_tauxdinvest2

var.test(tauxdemarge1, tauxdinvest2, alternative="two.sided", conf.level = 0.95)
```

```
##
##  F test to compare two variances
##
## data:  tauxdemarge1 and tauxdinvest2
## F = 1.5, num df = 72, denom df = 72, p-value = 0.08
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.9572 2.4291
## sample estimates:
## ratio of variances
##              1.525
```

- La p-value est supérieur au seuil de æ=5% ainsi l'hypothèse 0 n'est pas rejetée. Nous supposerons donc, à ce seuil, que les variances sont égales.

```
#Test de comparaisons de moyennes avec les variances égales au seuil æ=5%

#Hypothèse 0 : mean(tauxdemarge1) = mean(tauxdinvest2)
#Hypothèse 1 : mean(tauxdemarge1) \neq mean(tauxdinvest2)
#CPE : P(c1<X1-X2<c2|HO) = 1-æ

t.test(tauxdemarge1, tauxdinvest2, alternative = "two.sided", var.equal = TRUE, conf.level = 0.95)
```

```
##
##  Two Sample t-test
##
## data:  tauxdemarge1 and tauxdinvest2
## t = 21, df = 144, p-value <2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  6.917 8.338
## sample estimates:
## mean of x mean of y
##     30.51     22.88
```

Interprétation plus bas

2) Test de comparaison, au seuil æ=10%, des moyennes du taux de marge et du taux d'investissement

```
#Vérifions si les variances des deux variables sont égales, ou non, avec un test au seuil æ=10%

#Hypothèse 0 : var_tauxdemarge1 = var_tauxdinvest2
#Hypothèse 1 : var_tauxdemarge1 \neq var_tauxdinvest2

var.test(tauxdemarge1, tauxdinvest2, alternative="two.sided", conf.level = 0.90)
```

```
##
##  F test to compare two variances
##
## data:  tauxdemarge1 and tauxdinvest2
## F = 1.5, num df = 72, denom df = 72, p-value = 0.08
## alternative hypothesis: true ratio of variances is not equal to 1
## 90 percent confidence interval:
##  1.032 2.253
## sample estimates:
## ratio of variances
##             1.525
```

- La p-value (=8%) est inférieure au seuil de æ=10% ainsi l'hypothèse 0 est rejetée. Nous supposerons donc, à ce seuil, que les variances sont différentes.

```
#Test de comparaisons de moyennes avec les variances différentes au seuil æ=10%

#Hypothèse 0 : mean(tauxdemarge1) = mean(tauxdinvest2)
#Hypothèse 1 : mean(tauxdemarge1) \neq mean(tauxdinvest2)
```

```r
#CPE : P(c1<X1-X2<c2|H0) = 1-æ

t.test(tauxdemarge1, tauxdinvest2, alternative = "two.sided", var.equal = FALSE, conf.level = 0.90)
```

```
##
##  Welch Two Sample t-test
##
## data:  tauxdemarge1 and tauxdinvest2
## t = 21, df = 138, p-value <2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
##  7.032 8.223
## sample estimates:
## mean of x mean of y
##     30.51     22.88
```

Interprétation du 1er {resp. 2ème} Test:

Ici la p-value est inférieure à æ=0,05 {resp. æ=10%}, donc l'hypothèse 0 est rejetée. C'est à dire que les taux de marge et d'investissement sont rarement égaux, nous sommes certains à 95% {resp. 90%} que l'écart entre ces deux moyennes se situe entre [6,917 ; 8,338] {resp. [7,032 ; 8,223]. Cette différence entre les moyennes de taux de marge et de taux d'investissement pourrait résulter de divers facteurs économiques, et de politique d'entreprise sur la période de 1950 à 2022.

3) Test de comparaison, au seuil æ=5%, des moyennes du taux de marge et du taux d'autofinancement

```r
#Vérifions si les variances des deux variables sont égales, ou non, avec un test au seuil æ=5%

#Hypothèse 0 : var_tauxdemarge1 = var_tauxdautof3
#Hypothèse 1 : var_tauxdemarge1 \neq var_tauxdautof3

var.test(tauxdemarge1, tauxdautof3, alternative="two.sided", conf.level = 0.95)
```

```
##
##  F test to compare two variances
##
## data:  tauxdemarge1 and tauxdautof3
## F = 0.025, num df = 72, denom df = 72, p-value <2e-16
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.01544 0.03919
## sample estimates:
## ratio of variances
##             0.0246
```

- La p-value est inférieure au seuil de æ=5% ainsi l'hypothèse 0 est rejetée. Nous supposerons donc, à ce seuil, que les variances sont différentes.

```r
#Test de comparaisons de moyennes avec les variances différentes au seuil æ=5%

#Hypothèse 0 : mean(tauxdemarge1) = mean(tauxdautof3)
#Hypothèse 1 : mean(tauxdemarge1) \neq mean(tauxdautof3)
#CPE : P(c1<X1-X2<c2|H0) = 1-æ
```

```r
t.test(tauxdemarge1, tauxdautof3, alternative = "two.sided", var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  tauxdemarge1 and tauxdautof3
## t = -28, df = 76, p-value <2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -54.72 -47.54
## sample estimates:
## mean of x mean of y
##     30.51     81.64
```

Interprétation plus bas

4) Test de comparaison, au seuil æ=10%, des moyennes du taux de marge et du taux d'autofinancement

```r
#Vérifions si les variances des deux variables sont égales, ou non, avec un test au seuil æ=10%

#Hypothèse 0 : var_tauxdemarge1 = var_tauxdautof3
#Hypothèse 1 : var_tauxdemarge1 \neq var_tauxdautof3

var.test(tauxdemarge1, tauxdautof3, alternative="two.sided", conf.level = 0.90)
```

```
##
##  F test to compare two variances
##
## data:  tauxdemarge1 and tauxdautof3
## F = 0.025, num df = 72, denom df = 72, p-value <2e-16
## alternative hypothesis: true ratio of variances is not equal to 1
## 90 percent confidence interval:
##  0.01665 0.03634
## sample estimates:
## ratio of variances
##             0.0246
```

- La p-value est inférieure au seuil de æ=10% ainsi l'hypothèse 0 est rejetée. Nous supposerons donc, à ce seuil, que les variances sont différentes.

```r
#Test de comparaisons de moyennes avec les variances différentes au seuil æ=10%

#Hypothèse 0 : mean(tauxdemarge1) = mean(tauxdautof3)
#Hypothèse 1 : mean(tauxdemarge1) \neq mean(tauxdautof3)
#CPE : P(c1<X1-X2<c2|H0) = 1-æ

t.test(tauxdemarge1, tauxdautof3, alternative = "two.sided", var.equal = FALSE, conf.level = 0.90)
```

```
##
##  Welch Two Sample t-test
##
## data:  tauxdemarge1 and tauxdautof3
## t = -28, df = 76, p-value <2e-16
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
##  -54.13 -48.13
## sample estimates:
## mean of x mean of y
##     30.51     81.64
```

Interprétation du 3ème {resp. 4ème} Test:

Ici la p-value est inférieure à æ=0,05 {resp. æ=10%}, donc l'hypothèse 0 est rejetée. C'est à dire que les taux de marge et d'autofinancement sont rarement égaux, nous sommes certains à 95% {resp. 90%} que l'écart en valeur absolue entre ces deux moyennes se situe entre [47.54 ; 54.72] {resp. [48.13 ; 54.13]. La disparité entre ces moyennes peut être attribuée à des choix fluctuants dans la gestion des coûts ou des bénéfices, mais aussi au niveau des politiques de financement.

Conclusion

Il n'est pas étonnant de constater que ces taux soient globalement inégaux, en effet ils résultent de facteurs et de politiques d'entreprises bien distinctes. Il intéragissent entre eux car une politique favorisant un de ces taux va nécessairement influer sur un autre mais sans pour autant qu'ils soient égaux, ils n'évoluent d'ailleurs pas forcément dans le même sens.
Une entreprise peut être rentable (avoir un taux de marge élevé) mais ne pas générer suffisamment de flux de trésorerie pour financer ses propres investissements, et inversement.

Question V : Liens entre le taux d'investissement et le taux d'autofinancement

Dans cette *Dernière Question* nous étudierons le lien entre deux ces deux indicateurs, nous étude incluera :

- des tableaux de contingences (empiriques & théoriques)
- un tableau des profils-lignes
- un test d'indépendance du χ^2 au seuil de 5% & 10%.

Question 5.a) : Tableaux de contingences

Tableau de contingence empirique :

```r
# Création des séquences
borne_Investissement <- seq(15, 30, 5)
borne_Autof <- seq(40, 130, 30)

# Découpage des données en catégories
Taux_dInvest_cut <- cut(tauxdinvest2, borne_Investissement)
Taux_dAutof_cut <- cut(tauxdautof3, borne_Autof)

# Création du tableau
TC_emp_obs <- addmargins(table(Taux_dAutof_cut, Taux_dInvest_cut))

#Nomination des axes & valeurs
colnames(TC_emp_obs) <- c('[15,20]','[20,25]','[25,30]','Total')
rownames(TC_emp_obs) <- c('[40,70]','[70,100]','[100,130]','Total')

#Affichage du tableau de contingence empirique
print(TC_emp_obs)
```

```
##                  Taux_dInvest_cut
## Taux_dAutof_cut [15,20] [20,25] [25,30] Total
##       [40,70]         0      14       4    18
##       [70,100]        1      41       8    50
##       [100,130]       2       3       0     5
##       Total           3      58      12    73
```

- Nous obtenons les *effectifs marginaux* qui nous servirons pour le tableau de contingence théorique. Mais nous obtenons également les effectifs observés pour les différentes paires de modalités, ceux-ci seront utilisés lors du test d'indépendance du χ^2.

- Prenons l'exemple de l'effectif égal à 41.
  Il renvoie au fait qu'il y'a 41 fois, depuis 1950, des taux d'investissement compris entre 20% et 25% qui s'associent à des taux d'autofinancement compris entre 70% et 100%.

Tableau de contingence théorique :

Les effectifs théoriques T_{ij} sont obtenus de la manière suivante:

$T_{ij} = \frac{n_{\bullet i}.n_{j\bullet}}{n}$; avec $n_{i\bullet}$ et $n_{\bullet j}$ les $effectifs$ $marginaux$

```r
# Création des séquences
borne_Investissement <- seq(15, 30, 5)
borne_Autof <- seq(40, 130, 30)
```

```r
# Découpage des données en catégories
Taux_dInvest_cut <- cut(tauxdinvest2, borne_Investissement)
Taux_dAutof_cut <- cut(tauxdautof3, borne_Autof)

# Déclaration des effectif marginaux empiriques
row_totals <- margin.table(TC_emp_obs, 1)
col_totals <- margin.table(TC_emp_obs, 2)
total <- sum(TC_emp_obs)

# Création du tableau de contingence théorique
TC_emp_theorique <- outer(row_totals, col_totals) / total

# Affichage du tableau de contingence théorique
print(TC_emp_theorique)
```

```
##                Taux_dInvest_cut
## Taux_dAutof_cut [15,20] [20,25] [25,30] Total
##       [40,70]    0.7397  14.301  2.9589    18
##       [70,100]   2.0548  39.726  8.2192    50
##       [100,130]  0.2055   3.973  0.8219     5
##       Total      3.0000  58.000 12.0000    73
```

- Nous obtenons les effectifs théoriques pour les différentes paires de modalités, ceux-ci seront aussi utilisés lors du test d'indépendance du χ^2.

Question 5.b) : Tableau de Profils-Lignes

```r
#Création de tableau
TPL <- addmargins(prop.table(addmargins(table(Taux_dAutof_cut, Taux_dInvest_cut),1),1),2)

colnames(TPL) <- c('Taux faible','Taux moyen','Taux fort','Total')
rownames(TPL) <- c('Taux faible','Taux moyen','Taux fort','Total')

#Affichage du tableau de Profils-Lignes
TPL
```

```
##                Taux_dInvest_cut
## Taux_dAutof_cut Taux faible Taux moyen Taux fort  Total
##     Taux faible      0.0000     0.7778    0.2222 1.0000
##     Taux moyen       0.0200     0.8200    0.1600 1.0000
##     Taux fort        0.4000     0.6000    0.0000 1.0000
##     Total            0.0411     0.7945    0.1644 1.0000
```

Interprétation :

On constate grâce à ce tableau qu'un taux d'autofinancement jugé comme "faible" implique pas un taux d'investissement faible, cela souligne la diversité des moyens de financement des entreprises (emprunt,augmentation de capital...) autres que leur propre trésorerie.
On note également que les entreprises qui investissent le plus ne sont pas forcément celles avec le taux d'autofinancement le plus élevé, celles-ci préférant peut-être d'autres canaux.
Par exemple une entreprise se trouvant dans une bonne conjecture économique et investissant massivement va pouvoir se tourner aisément vers l'emprunt bancaire.

***Question 5.c) : Test d'indépendance du χ^2* entre le taux d'investissement et le taux d'autofinancement**

Rédaction :

Dans ce test nous mettons en relation les effectifs observés et les effectifs théoriques, afin de vérifier si les taux d'autofinancement et d'investissement sont réellement corrélés dans le temps.

Posons les hypothèses : $\begin{cases} H_0 : Nos\ deux\ taux\ sont\ indépendants\ /non\ corrélés \\ H_1 : Nos\ deux\ taux\ sont\ dépendants\ /corrélés \end{cases}$

La statistique de test est la suivante :

$$\Delta = \sum_{i=1}^{k} \sum_{j=1}^{z} \frac{(n_{ij} - T_{ij})^2}{T_{ij}} \xrightarrow{d} \chi^2_{(k-1)(z-1)};\ avec\ T_{ij} > 4\ pour\ tout\ i, j$$

Hypothèse : Nous supposerons par simplification $T_{ij} > 4$

```r
#Hypothèse 0: Le taux d'autofinancement et d'investissement sont indépendants
#Hypothèse 1: Le taux d'autofinancement et d'investissement sont dépendants
#Rejet de H0 au seuil æ : si delta > c; ou si p-value > æ

#Test du chi-2
test_chi2 <- chisq.test(TC_emp_obs, TC_emp_theorique, correct=FALSE)
```

```
## Warning in chisq.test(TC_emp_obs, TC_emp_theorique, correct = FALSE):
## Chi-squared approximation may be incorrect
```

```r
test_chi2
```

```
##
##  Pearson's Chi-squared test
##
## data:  TC_emp_obs
## X-squared = 18, df = 9, p-value = 0.03
```

- Le message "Warning" signifie que le test de χ^2 doit remplir la contition citée plus haut. *Si ce n'est pas le cas il est possible que l'approximation du test soit "incorrecte".*

Interprétation :

Ici la p-value est inférieure à æ=5% (donc < æ=10%), ainsi l'hypothèse O est rejetée, ce qui nous permet d'affirmer que nos taux d'investissement et d'autofinancement sont corrélés dans le temps.
Reprenons notre exemple énoncé en début d'étude sur l'investissement en Recherche et Développement (R&D) qui stipulait une possible corrélation entre une augmentation du taux d'investissement, et une baisse du taux d'autofinancement du fait que le retour sur investissement pouvait tarder. Grâce à ce test nous comprenons qu'une relation existe entre ces deux indicateurs comptables.

Conclusion

Même si le taux d'autofinancement et d'investissement ne sont pas parfaitement corrélés comme nous l'avons vu précédemment, une explication de la non-indépendance entre le taux d'investissement et le taux d'autofinancement pourrait être que les entreprises ou les entités étudiées ajustent leur taux d'investissement en fonction de leur capacité d'autofinancement. Par exemple, si les entreprises ont une capacité d'autofinancement élevée, elles peuvent être plus enclines à investir d'avantage.

Des liens sont également présents pour nos deux autres combinaisons de variables, en effectuant un test d'indépendance entre le taux de marge et le taux d'investissement on aurait sûrement trouvé une dépendance. En général, une entreprise avec un taux de marge élevé peut avoir plus de fonds disponibles après le paiement de ses coûts d'exploitation et de ses charges financières. Si cette entreprise choisit de réinvestir une part significative de ces bénéfices dans des projets d'investissement, le taux d'investissement pourrait être élevé.

De même, l'augmentation du taux de marge pourrait conduire à plus de fonds et plus de facilité à s'autofinancer. Aussi, une entreprise peut avoir un taux de marge élevé, mais si elle distribue la majorité de ses bénéfices sous forme de dividendes, l'autofinancement pourrait être limité. Dans tous les cas on observerait ainsi une dépendance entre ces deux variables.

Conclusion de l'étude :

Notre étude a plongé dans la dynamique financière des entreprises non-financières françaises depuis 1950, cherchant à comprendre les relations entre le taux de marge, d'investissement et d'autofinancement. Au-delà des chiffres, elle offre un aperçu des liens subtils qui guident les choix stratégiques des entreprises.

Perspectives Statistiques:

Les résultats soulignent une corrélation notable entre le taux de marge, d'investissement et d'autofinancement. Chaque variation dans l'un résonne dans les autres, offrant un panorama statistique de l'évolution de ces indicateurs au fil du temps. Cette cohérence met en lumière l'importance de ces variables dans les décisions financières des entreprises.

Perspectives Comptables:

D'un point de vue comptable, chaque pourcentage dans l'autofinancement laisse une empreinte distincte dans les choix d'investissement, influençant directement la marge. Cette perspective comptable offre une compréhension tangible de l'impact concret de chaque décision financière.

En résumé, notre étude offre une analyse éclairante de la trajectoire financière de ces entreprises. Derrière les ratios, se dessine un récit financier où le taux de marge, d'investissement et d'autofinancement interagissent pour façonner l'évolution économique. Ces conclusions apportent un éclairage précieux pour comprendre comment ces entreprises ont navigué dans le monde financier au fil des décennies, soulignant l'importance cruciale de ces indicateurs dans leurs choix stratégiques.

Cher lecteur, nous tenions à vous exprimer notre profonde reconnaissance pour avoir consacré votre temps à explorer notre projet. Nous espérons sincèrement que cette immersion dans nos travaux a su captiver votre intérêt et répondre à vos attentes.

Jolibert Grégoire & Combe Oscar