

III) Statistiques

Capsule 13. ANOVA

Masoumeh Sajedi

Biostatisticienne, Unité de recherche clinique appliquée (URCA)

Contents

1	Introduction	3
2	Objectif et l'hypothèse nulle/alternative	3
3	Base de données	4
3.1	Importation et préparation des données	4
4	ANOVA à un facteur	6
4.1	Effectuer l'ANOVA et interpréter les résultats	7
4.2	Tests Post-Hoc	8
4.2.1	Test HSD de Tukey	8
4.2.2	Test t par paires	9
4.3	Hypothèses de validité	10
4.3.1	Indépendance	10
4.3.2	Normalité	10
4.3.3	Homogénéité de la variance	11
5	ANOVA à deux facteur	12
5.1	Effectuer l'ANOVA à deux facteurs	14
5.2	Test Post-Hoc	15
6	Exercice	15
7	Références	18

Contents

1 Introduction

ANOVA (ANALYsis Of VARiance) est un test statistique permettant de déterminer si deux ou plusieurs moyennes de population sont différentes. En d'autres termes, il est utilisé pour comparer deux ou plusieurs groupes pour voir s'ils sont significativement différents.

Notez que le test t de Student est utilisé pour comparer les moyennes entre deux groupes, et que l'ANOVA à deux groupes retourne la même conclusion qu'un test t.

Il existe plusieurs versions de l'ANOVA (par exemple, à un facteur, à deux facteurs, mixte, à mesures répétées, etc.). Dans ce document, nous présentons l'ANOVA à un facteur et à deux facteurs .

Dans l'ANOVA à un facteur, les moyennes se rapportent aux différentes modalités d'une seule variable indépendante catégorielle (ou facteur).

Celle à deux facteurs est utilisée pour tester l'effet combiné de deux facteurs sur une même variable dépendante.

Bien que l'ANOVA soit utilisée pour faire des inférences sur les moyennes de différents groupes, la méthode est appelée "analyse de la variance". Il est appelé ainsi car il compare la variabilité observée entre les moyennes des différents groupes (la variance *inter-groupe*) et la variance au sein de chaque groupe (la variance *intra-groupe*).

Les deux variances sont comparées l'une à l'autre en prenant le rapport

$$\frac{\text{variance intergroup}}{\text{variance intragroup}}$$

puis en comparant ce rapport à un seuil de la distribution de probabilité de Fisher (un seuil basé sur un niveau de signification spécifique, généralement 5 %).

Dans ce document, nous présentons la méthode ANOVA d'un point de vue pratique, afin d'aborder les points suivants:

- Objectif de l'ANOVA et l'hypothèse nulle et alternative
- Effectuer l'ANOVA dans R
- Interpréter les résultats
- Les hypothèses de l'ANOVA et comment les vérifier
- Réaliser des tests post-hoc et visualiser les résultats

2 Objectif et l'hypothèse nulle/alternative

L'ANOVA est utilisée pour déterminer si les mesures sont similaires à travers différentes modalités d'une variable catégorielle. Elle est également utilisée pour comparer l'impact de différentes modalités d'une variable catégorielle sur une variable quantitative continue.

Pour un facteur avec a modalités, les hypothèses nulle et alternative d'une ANOVA sont:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_a$$
$$H_1 : \mu_i \neq \mu_j \quad \text{pour au moins une paire}(i, j)$$

3 Base de données

Nous utilisons la base de données de l'étude:

Durocher, Jill, 2019, "A double-blind, randomized controlled trial on oxytocin route for prevention of post-partum hemorrhage in Argentina", <https://doi.org/10.7910/DVN/MDZRKU>, Harvard Dataverse, V1.

L'ensemble de données contient les données de 543 patients. Le base de données contient 148 variables. Nous nous concentrons uniquement sur les 4 variables: la perte de sang totale à l'arrêt du saignement actif (*perdida_sangre_total*), l'âge de la femme (*edad*), l'âge gestationnel (*edadgest*), et groupe d'étude (1:IV infusion et 2:IM Injection). Nous prenons un sous-ensemble de données pour les 480 participants randomisées. Nous ajoutons une variable catégorielle pour âge de la femme (moins de 19 ans, entre 20 et 30 ans et plus de 30 ans), et âge gestationnel (moins de 38 semaines, entre 38 et 40 semaines et plus de 40 semaines).

Nous allons utiliser les tests d'ANOVA pour nous aider à répondre aux questions suivant:

1. La perte de sang totale est-elle différente entre les 3 groupes d'âge de femme?
2. La perte de sang totale est-elle différente entre les 3 groupes d'âge de femme et cela dépend-il de leur groupe d'étude?

Pour répondre à la première question, nous utilisons une ANOVA à un facteur. Un test ANOVA à deux facteurs peut répondre à la deuxième question.

3.1 Importation et préparation des données

Dans le script ci-dessous, nous avons le code pour lire les données, modifier les noms des variables et ajouter une variable catégorielle pour âge de la femme et .

```
library(haven)
library(tidyverse)

dataset = read_sav("Gynuity_Argentina_oxytocin_IVIM_PPH_Prevention_DB.sav")

dat=dataset%>%
  select(edad,studygrp_AR,perdida_sangre_total,edadgest)%>%
  filter(studygrp_AR %in% c("1","2"))

#modification des noms des variables
colnames(dat)=c("age_femme","groupe_étude","perte_sang_totale","age_gestationnel")

# l'ajout d'une variable catégorielle pour âge de la femme
attach(dat)
age_femme_grp= case_when(age_femme>= 30~ '>=30',
  age_femme > 19 & age_femme < 35 ~ '20-30',
  age_femme <= 19 ~ '<=19')

age_gestationnel_grp= case_when(age_gestationnel>= 40~ '>=40',
  age_gestationnel > 38 & age_gestationnel < 40 ~ '38-40',
  age_gestationnel <= 38 ~ '<=38')

dat$age_femme_grp = factor(age_femme_grp,
  levels=c("<=19","20-30",
  ">=30"))
dat$age_gestationnel_grp=factor(age_gestationnel_grp,
  levels=c("<=38","38-40",">=40"))
dat$groupe_étude=as.factor(dat$groupe_étude)
```


4 ANOVA à un facteur

Nous voulons savoir s'il existe une différence significative entre les perte de sange totale moyens des femme dans les 3 groupes d'âge.

Ici, le facteur est le variable `age_femme` qui contient 3 modalités ou groupes.

Ci-dessous, nous avons quelques statistiques descriptives de base de notre ensemble de données.

```
group_by(dat, groupe_étude) %>%
  summarise(
    count = n(),
    mean = mean(perte_sang_totale, na.rm = TRUE),
    sd = sd(perte_sang_totale, na.rm = TRUE)
  )

## # A tibble: 2 x 4
##   groupe_étude count  mean   sd
##   <fct>         <int> <dbl> <dbl>
## 1 1             239  363.  322.
## 2 2             241  406.  344.

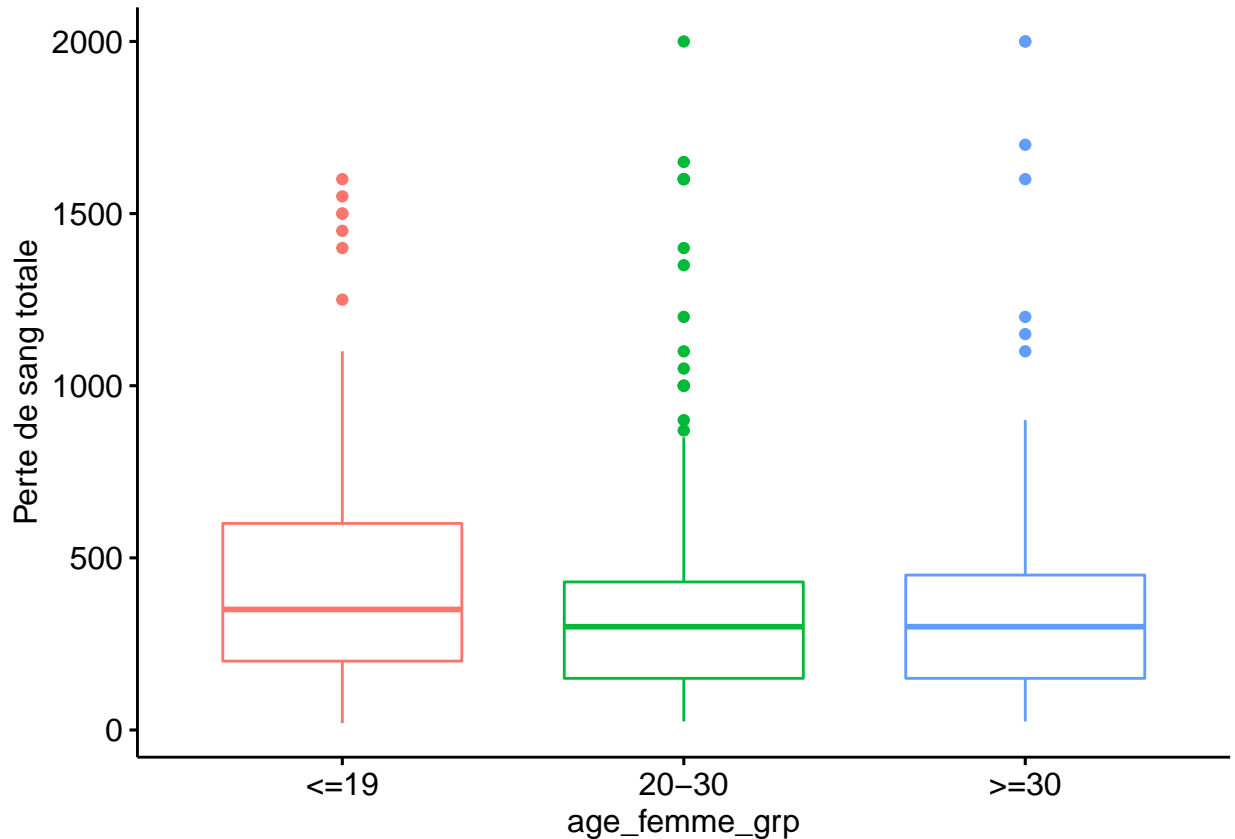
group_by(dat, age_femme_grp) %>%
  summarise(
    count = n(),
    mean = mean(perte_sang_totale, na.rm = TRUE),
    sd = sd(perte_sang_totale, na.rm = TRUE)
  )

## # A tibble: 3 x 4
##   age_femme_grp count  mean   sd
##   <fct>         <int> <dbl> <dbl>
## 1 <=19           106  458.  361.
## 2 20-30          292  352.  294.
## 3 >=30           82  407.  411.

# Visualisation des données

library("ggpubr")

ggboxplot(dat, x = "age_femme_grp", y = "perte_sang_totale",
  color = "age_femme_grp",
  ylab = "Perte de sang totale",
  xlab="age_femme_grp") +theme(legend.position = "none")
```



Il y a respectivement 106, 292, et 82 femmes de l'âge moins de 19 ans, entre 20 et 30 ans et plus de 30 ans.

Il y a respectivement 239, et 241 femmes dans le groupe 1 (IV infusion) et 2 (IM injection).

La moyenne de la *Perte de sange totale* est la plus faible pour les femme entre 20 et 30 ans et la plus élevée pour les femmes de moins de 19 ans.

Les boxplot montrent que, du moins pour notre échantillon, les femmes de moins de 19 ans semblent avoir *Perte de sange totale* le plus élevée, et les femmes entre 20 et 30 ans ont le *Perte de sange totale* le moins élevé.

Les boxplot et les statistiques descriptives ne suffisent cependant pas à conclure que la *Perte de sange totale* est significativement différentes dans les 3 groupes d'âge.

4.1 Effectuer l'ANOVA et interpréter les résultats

La fonction R `aov()` peut être utilisée pour ajuster un modèle d'analyse de la variance. La fonction `summary.aov()` est utilisée pour résumer ce modèle.

```
res_aov <- aov(perte_sang_totale ~ age_femme_grp, data = dat)
summary(res_aov)

##           Df    Sum Sq Mean Sq F value Pr(>F)
## age_femme_grp  2    911524  455762   4.146 0.0164 *
## Residuals    474 52104892  109926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 3 observations deleted due to missingness
```

Comme la valeur de p est inférieure à 0,05, nous rejetons l'hypothèse nulle que toutes les moyennes sont égales. Par conséquent, nous pouvons conclure qu'au moins une des catégories est différente des autres. Cependant, les résultats d'une ANOVA, ne nous disent pas quel(s) groupe(s) est(sont) différent(s) des autres. Pour tester cela, nous devons utiliser d'autres types de tests, appelés tests post-hoc ou des tests de comparaison par paires multiples. Dans la session suivante, deux de ces tests sont présentés.

Dans le cas où la valeur de p est supérieur au niveau de signification de 0,05, nous ne pouvons pas rejeter l'hypothèse nulle que toutes les moyennes sont égales.

4.2 Tests Post-Hoc

Comme le résultat du test ANOVA est significatif, à fin de déterminer si la différence moyenne entre des paires spécifiques de groupes est statistiquement significative, il est possible d'effectuer une comparaison multiple par paires entre les moyennes des groupes.

Nous présentons deux des tests post hoc les plus courants: le test HSD de Tukey (Tukey Honest Significant Differences) et le test t par paires.

4.2.1 Test HSD de Tukey

Le test HSD de Tukey, utilisé pour comparer tous les groupes entre eux (donc toutes les comparaisons possibles de 2 groupes). La fonction `TukeyHSD()` prend l'ANOVA ajustée comme argument.

```
TukeyHSD(res_aov)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = perte_sang_totale ~ age_femme_grp, data = dat)
##
## $age_femme_grp
##          diff          lwr          upr          p adj
## 20-30-<=19 -105.84045 -194.58205 -17.09886 0.0145065
## >=30-<=19  -51.29101 -166.56732  63.98531 0.5481602
## >=30-20-30  54.54945  -43.37773 152.47662 0.3904644
```

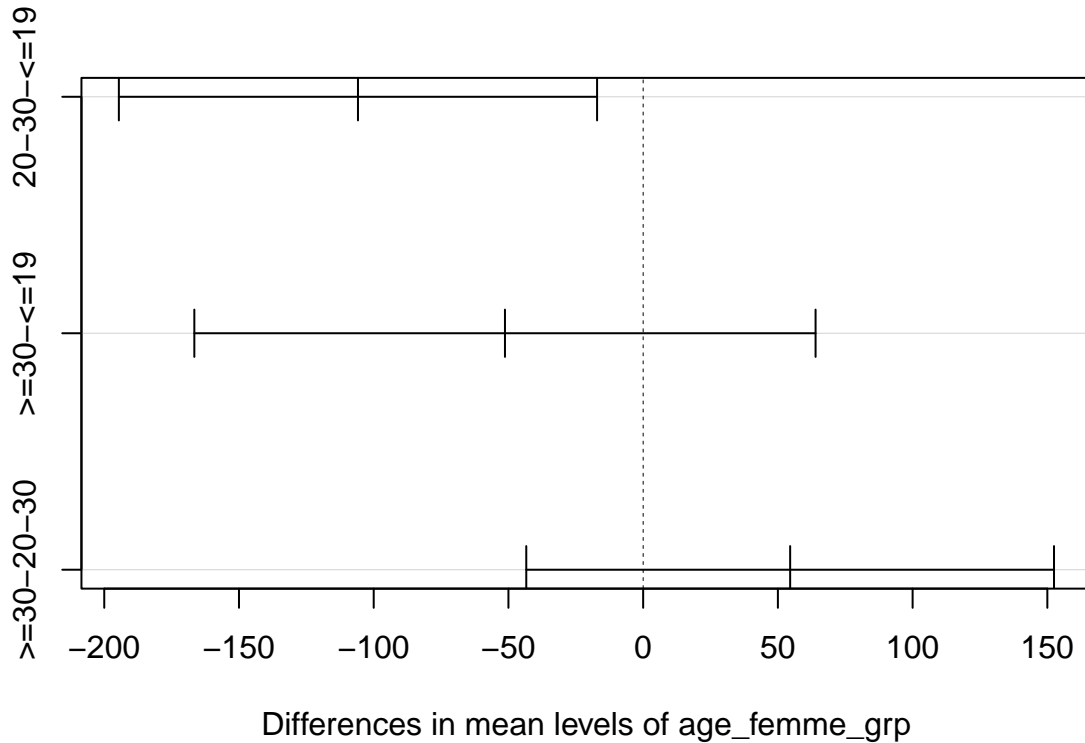
- *diff*: signifie la différence entre les moyennes des deux groupes. Ici on compare les difference entre les femmes âgées
 1. entre 20 et 30 ans vs moins de 19 ans
 2. plus que 30 ans vs moins de 19 ans
 3. plus que 30 ans vs entre 20 et 30 ans
- *lwr*, *upr*: signifie la limite inférieure et la limite supérieure de l'intervalle de confiance à 95% (par défaut)
- *p adj*: est la valeur de p après ajustement pour les comparaisons multiples.

Il ressort de la sortie que seule la différence entre *20-30 ans* et *<=19 ans* est significative avec une valeur de p ajustée de 0.014.

Les résultats du test post-hoc peuvent être visualisés en utilisant de la fonction `plot()`:

```
plot(TukeyHSD(res_aov))
```


95% family-wise confidence level



Nous observons que l'intervalle de confiance correspondant à la comparaison des modalités *20-30 ans* et *<=19 ans* ne franchit pas la ligne zéro, ce qui indique que ces deux modalités sont significativement différents.

4.2.2 Test t par paires

La fonction `pairwise.t.test()` peut également être utilisée pour calculer des comparaisons par paires entre les modalités avec des corrections pour les tests multiples.

```
pairwise.t.test(dat$perte_sang_totale, dat$age_femme_grp,  
               p.adjust.method = "bonferroni")
```

```
##  
## Pairwise comparisons using t tests with pooled SD  
##  
## data: dat$perte_sang_totale and dat$age_femme_grp  
##  
##      <=19  20-30  
## 20-30 0.016 -  
## >=30 0.888 0.573  
##  
## P value adjustment method: bonferroni
```

Les valeurs P sont ajustées à l'aide de la méthode de correction des tests multiples de Bonferroni.

4.3 Hypothèses de validité

Afin de pouvoir interpréter les résultats d'un test ANOVA, certaines hypothèses doivent être satisfaites. Lorsqu'une ou plusieurs hypothèses ne sont pas satisfaites, bien qu'il soit techniquement possible de réaliser ces tests, il serait erroné d'interpréter les résultats et de se fier aux conclusions.

Il y a trois hypothèses à vérifier:

- L'indépendance: les données doivent être indépendantes entre les groupes et au sein de chaque groupe. Si les mesures ont été collectées sur les mêmes individus, l'ANOVA à mesures répétées doit être utilisée afin de prendre en compte la dépendance entre les échantillons.
- La normalité des résidus: Dans le cas de petits échantillons, les résidus doivent suivre approximativement une distribution normale. La normalité n'est pas requise lorsque le nombre d'observations dans chaque groupe est important (généralement $n \geq 30$).
- Homogénéité de la variance: les variances des différents groupes doivent être égales. Si l'hypothèse d'égalité des variances est rejetée, l'ANOVA de Welch (`oneway.test` dans R) peut être utilisée.

Notez qu'une alternative non paramétrique à l'ANOVA à un facteur est le test de somme des rangs de Kruskal-Wallis (`kruskal.test` dans R), qui peut être utilisé lorsque les hypothèses ANOVA ne sont pas satisfaites.

4.3.1 Indépendance

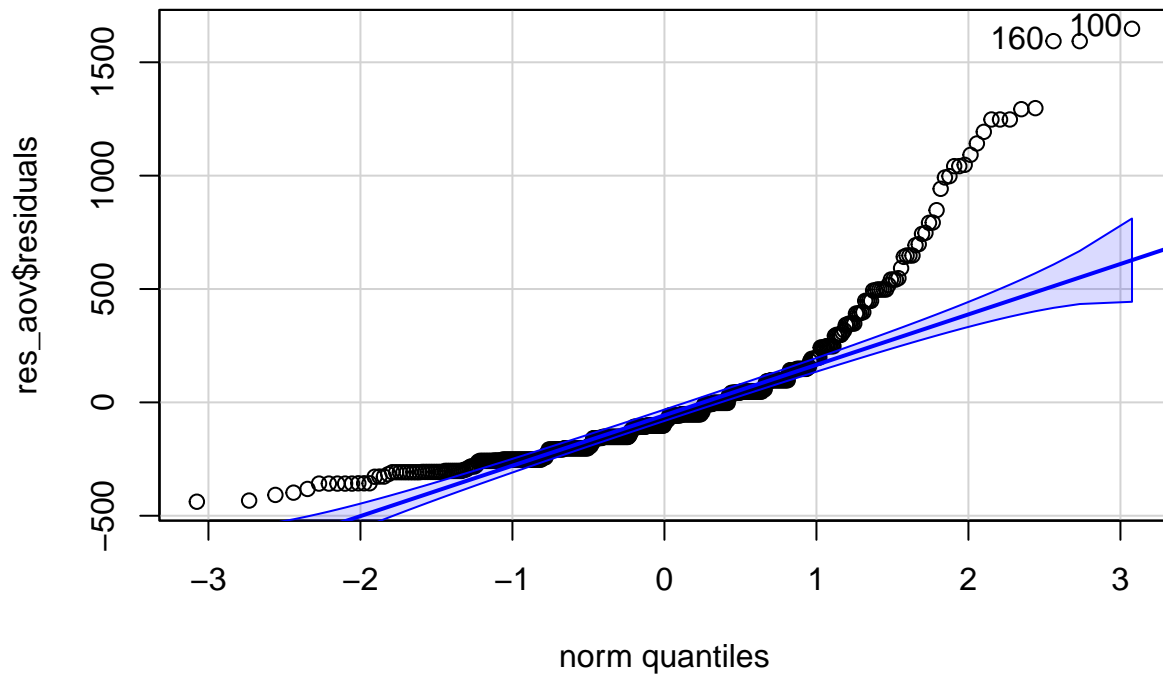
Les mesures sont indépendantes entre les groupes d'âges et nous ne sommes pas en présence de mesures répétées pour chaque sujet.

4.3.2 Normalité

L'hypothèse de normalité peut être testée visuellement (un histogramme, un QQ-plot), et/ou formellement (un test de normalité tel que le test de Shapiro-Wilk ou de Kolmogorov-Smirnov).

Étant donné que le nombre d'observations dans chaque groupe est grand, nous n'avons pas besoin de vérifier la normalité. Normalement, on testerait directement l'homogénéité des variances sans tester la normalité.

```
#plot(res_aov, which = 2)
# autre option
library(car)
qqPlot(res_aov$residuals)
```



```
## 100 160
## 100 159
# test Shapiro-Wilk test
shapiro.test(res_aov$residuals)
```

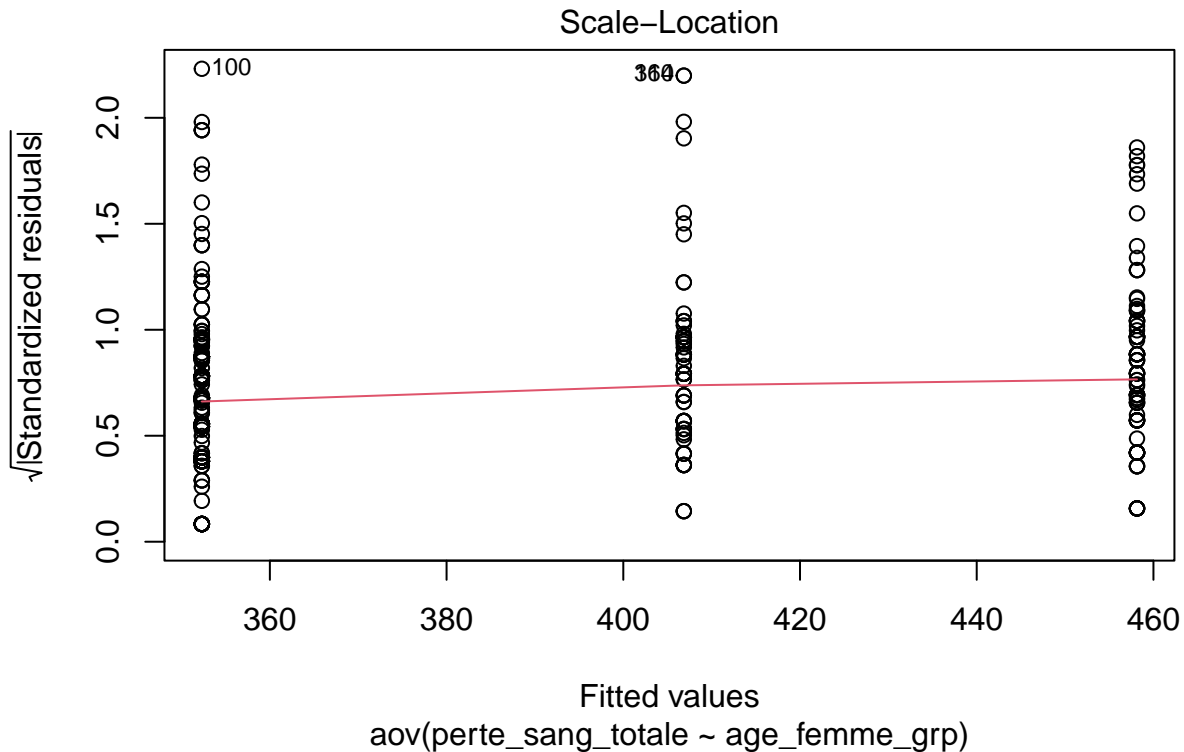
```
##
## Shapiro-Wilk normality test
##
## data: res_aov$residuals
## W = 0.78786, p-value < 2.2e-16
```

La p-value est inférieure à un niveau alpha de 0.05, alors l'hypothèse nulle que les résidus sont normalement distribués est rejetée. Cependant, comme mentionné ci-dessus, étant donné que la taille de notre échantillon est grande, nous n'avons pas besoin de tester la normalité.

4.3.3 Homogénéité de la variance

Le diagramme des valeurs résiduelles en fonction des valeurs ajustées peut être utilisé pour vérifier l'homogénéité des variances.

```
plot(res_aov, which = 3)
```



Il n'y a pas de relations évidentes entre les valeurs résiduelles et les valeurs ajustées (la moyenne de chaque groupe). On peut donc supposer l'homogénéité des variances.

Il est également possible d'utiliser le test de Levene pour vérifier l'homogénéité des variances.

```
library(car)
leveneTest(perte_sang_totale ~ age_femme_grp, data = dat)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 2  2.5734 0.07734 .
##      474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

La valeur de p étant supérieure au seuil de signification de 0,05, nous ne rejetons pas l'hypothèse nulle, nous ne pouvons donc pas rejeter l'hypothèse que les variances entre les groupes d'âge sont égales.

5 ANOVA à deux facteurs

Le test ANOVA à deux facteurs est utilisé pour évaluer simultanément l'effet de deux variables sur une variable de réponse.

Les hypothèses nulles sont

- Il n'y a pas de différence dans les moyennes des deux facteurs
- Il n'y a pas d'interaction entre les facteurs

L'ANOVA à deux facteurs, suppose que les observations sont normalement distribuées et des données dans les différents groupes ont des variances égales. Pour tester ces hypothèses, nous suivons la procédure qui a été expliquée pour celle à un facteur.

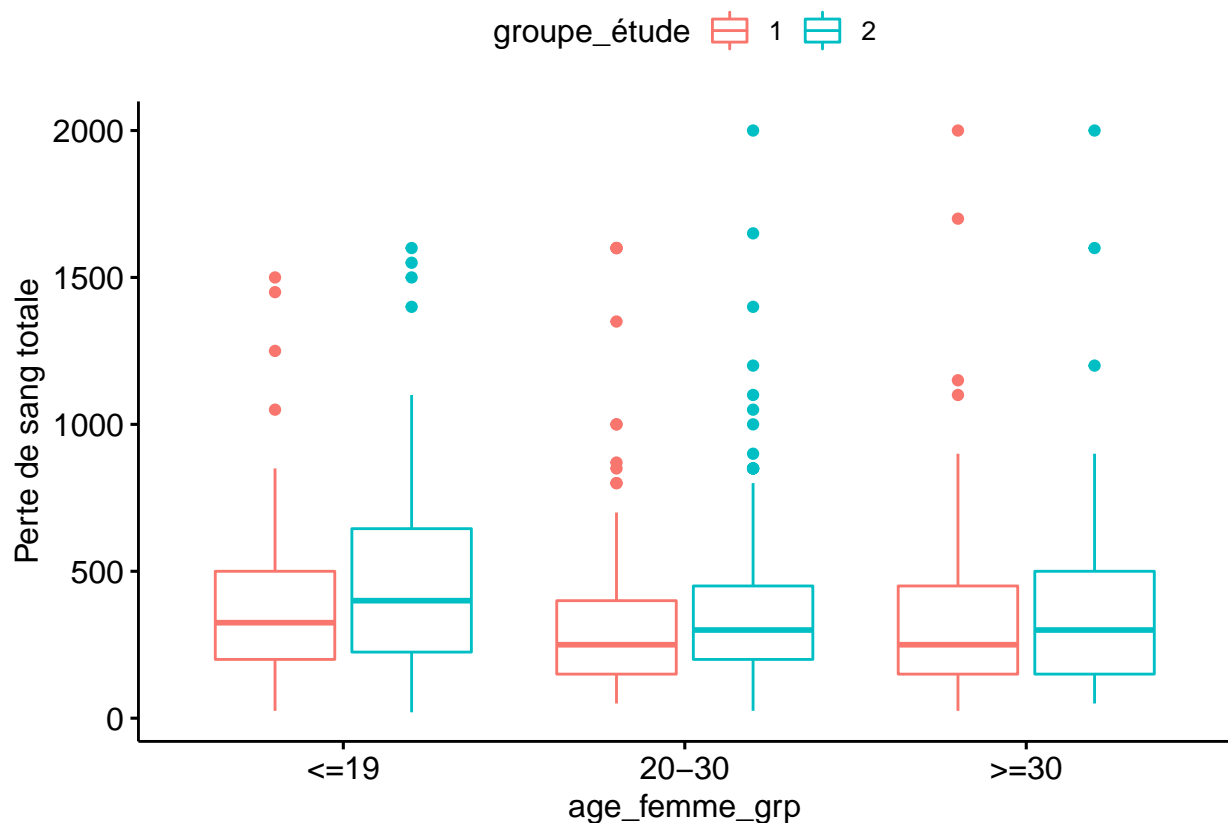
Dans notre exemple, les facteurs sont *groupe_étude* et *age_femme_grp* et ils contiennent respectivement 2 et 3 modalités.

Ci-dessous, nous avons quelques statistiques descriptives.

```
#Visualisation des données
```

```
library("ggpubr")
```

```
ggboxplot(dat,x = "age_femme_grp", y = "perte_sang_totale",
           color = "groupe_étude",
           ylab = "Perte de sang totale",
           xlab="age_femme_grp")
```



```
# Calcul de la moyenne et l'écart-type par groupes
```

```
group_by(dat, groupe_étude, age_femme_grp) %>%
  summarise(
    count = n(),
    mean = mean(perte_sang_totale, na.rm = TRUE),
    sd = sd(perte_sang_totale, na.rm = TRUE)
  )
```

```
## # A tibble: 6 x 5
```

```
## # Groups:   groupe_étude [2]
```

```
##   groupe_étude age_femme_grp count  mean    sd
```

```
## <fct>      <fct>      <int> <dbl> <dbl>
## 1 1         <=19         50  415.  329.
## 2 1         20-30        149 338.  286.
## 3 1         >=30         40  393.  426.
## 4 2         <=19         56  498.  386.
## 5 2         20-30        143 367.  303.
## 6 2         >=30         42  420.  400.
```

```
#Génération de tables de fréquences:
table(dat$groupe_étude, dat$age_femme_grp)
```

```
##
##      <=19 20-30 >=30
## 1      50  149   40
## 2      56  143   42
```

Le nombre de sujets dans chaque groupe n'est pas égale. Nous sommes en présence d'un plan déséquilibré.

5.1 Effectuer l'ANOVA à deux facteurs

Il existe trois façons différentes d'exécuter une ANOVA dans un plan déséquilibré. Ils sont connus sous le nom de sommes de carrés de type I, de type II et de type III. Les trois méthodes donnent le même résultat lorsque le plan est équilibré.

Pour les plans équilibrés, la fonction `summary.aov()` est utilisée pour résumer l'analyse du modèle de variance. La fonction `Anova()` peut être utilisée pour calculer le test ANOVA à deux facteurs pour les plans déséquilibrés.

```
res_aov2 <- aov(perte_sang_totale ~ age_femme_grp + groupe_étude,
               data = dat)
Anova(res_aov2, type = "III")
```

```
## Anova Table (Type III tests)
##
## Response: perte_sang_totale
##              Sum Sq Df F value Pr(>F)
## (Intercept) 16140698  1 147.0758 < 2e-16 ***
## age_femme_grp  888690  2   4.0489 0.01805 *
## groupe_étude  195928  1   1.7853 0.18214
## Residuals    51908964 473
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nous pouvons conclure que la groupe d'âge est statistiquement significatif.

Le modèle ajusté ci-dessus est appelé modèle additif. Il fait l'hypothèse que les deux variables factorielles sont indépendantes. Si nous pensons que ces deux variables peuvent interagir nous effectuons un test à deux facteurs avec effet d'interaction.

```
res_aov3 <- aov(perte_sang_totale ~ age_femme_grp * groupe_étude,
               data = dat)
Anova(res_aov3, type = "III")
```

```
## Anova Table (Type III tests)
##
## Response: perte_sang_totale
##              Sum Sq Df F value Pr(>F)
## (Intercept)  8598805  1 78.1128 <2e-16 ***
```

```
## age_femme_grp          262107  2  1.1905 0.3050
## groupe_étude          180150  1  1.6365 0.2014
## age_femme_grp:groupe_étude  60371  2  0.2742 0.7603
## Residuals              51848593 471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nous observerons que l'interaction n'est pas significative, dans ce cas il faut utiliser le modèle additif.

Une interaction significative implique que la valeur moyenne de la variable réponse pour chaque modalité d'un facteur change selon la modalité de l'autre facteur.

5.2 Test Post-Hoc

Nous pouvons effectuer un test post-hoc pour déterminer si la différence moyenne entre des paires spécifiques des groupes est statistiquement significative. Ci-dessous, nous voyons comment obtenir les résultats du test HSD de Tukey pour une variable spécifique.

```
TukeyHSD(res_aov2, which = "age_femme_grp")

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = perte_sang_totale ~ age_femme_grp + groupe_étude, data = dat)
##
## $age_femme_grp
##          diff          lwr          upr          p adj
## 20-30-<=19 -105.84045 -194.50921 -17.17169 0.0144085
## >=30-<=19  -51.29101 -166.47271  63.89070 0.5476173
## >=30-20-30  54.54945  -43.29736 152.39626 0.3898614
```

Il ressort de la sortie que seule la différence entre *20-30 ans* et *<=19 ans* est significative avec une valeur de p ajustée de 0,014.

6 Exercice

En utilisant le test ANOVA, répondez à la question suivante:

1. La perte de sang totale est-elle différente entre les 3 groupes d'âge gestationnel?

```
group_by(dat, age_gestationnel_grp) %>% summarise(
  count = n(),
  mean = mean(perte_sang_totale, na.rm = TRUE),
  sd = sd(perte_sang_totale, na.rm = TRUE))
```

```
## # A tibble: 3 x 4
##   age_gestationnel_grp count  mean    sd
##   <fct>                <int> <dbl> <dbl>
## 1 <=38                  122  341.  317.
## 2 38-40                 226  383.  325.
## 3 >=40                 132  428.  358.
```

La moyenne de la Perte de sange totale est la plus faible pour pour les accouchements avant 38 semaines de gestation et la plus élevée pour les accouchements après 40 semaines de gestation.

Afin de voir si les différences observées sont statistiquement significatives, nous effectuons le test d'ANOVA.

```
res_aov_exc <- aov(perte_sang_totale ~ age_gestationnel_grp, data = dat)
summary(res_aov_exc)
```

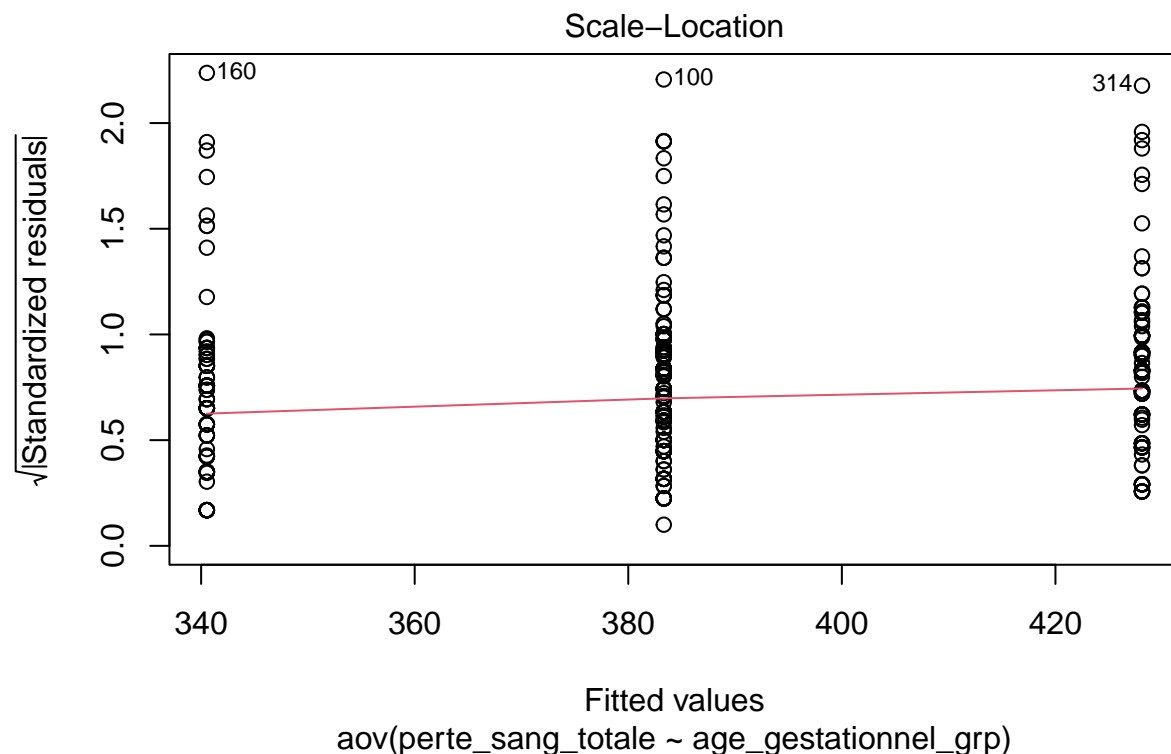
```
##                Df    Sum Sq Mean Sq F value Pr(>F)
## age_gestationnel_grp  2    485102  242551    2.189  0.113
## Residuals           474  52531314  110826
## 3 observations deleted due to missingness
```

La valeur de p est supérieure à 0,05 ($p=0,113$), nous ne pouvons pas rejeter l'hypothèse nulle que toutes les moyennes sont égales.

Nous vérifions la validité des hypothèses:

- Indépendance: Les mesures sont indépendantes entre les groupes d'âges gestationnel et nous ne sommes pas en présence de mesures répétées pour chaque sujet.
- Normalité: Étant donné que le nombre d'observations dans chaque groupe est grand, nous n'avons pas besoin de vérifier la normalité.
- Homogénéité de la variance

```
plot(res_aov_exc, which = 3) #Homogénéité de la variance
```



```
leveneTest(perte_sang_totale ~ age_gestationnel_grp, data = dat)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  2  1.0431 0.3532
##      474
```


Il n'y a pas de relations évidentes entre les valeurs résiduelles et les valeurs ajustées (la moyenne de chaque groupe). En plus, la valeur de p du test de Levene est supérieure au seuil de signification de 0,05. Alors, nous ne pouvons pas rejeter l'hypothèse que les variances entre les groupes d'âge sont égales.

7 Références

1. **Base de données:** Durocher, Jill, 2019, “A double-blind, randomized controlled trial on oxytocin route for prevention of postpartum hemorrhage in Argentina”, Harvard Dataverse, V1, <https://doi.org/10.7910/DVN/MDZRKU>, UNF:6:cFZwNd9CHxjepreGPi0ZFg==
2. **Packages tidyverse:** <https://cran.r-project.org/web/packages/tidyverse/index.html>
3. **Package car:** Fox J, Weisberg S (2019). An R Companion to Applied Regression, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
4. **Package ggplot2:** <https://cran.r-project.org/web/packages/ggpubr/index.html>
5. **aov RDocumentation:** <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/aov>
6. **TukeyHSD RDocumentation:** <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/TukeyHSD>