

R301. Inferencia base de datos Galaad

Julio David Ruiz Mendoza*

jueves 03 de junio de 2021, 18:31

Índice

1	Lectura y preparación de datos	2
2	Inferencia	3
2.1	Peso	3
2.1.1	Sexo	3
2.1.2	Semana de gestación	8
2.1.3	Peso al nacer	12
2.1.4	Grupo	16
2.2	Días de hospitalización	22
2.2.1	Sexo	22
2.2.2	Semana de gestación	23
2.2.3	Peso al nacer	24
2.2.4	Grupo	26
3	Recomendaciones	27
4	Agradecimientos	27
5	Referencias y bibliografía	27

*Estudiante Física UM, julioruizm14@gmail.es

1 Lectura y preparación de datos

Trabajamos con el fichero: `database.csv`
 Proveniente del fichero: Variables osteopenia COMPLETO Original.xlsx
 N° de observaciones: 106
 N° de variables: 31

Vista de datos:

```
##      Sexo      SG Peso.al.nacer      Num.Grupo Hospital VST1 NTx1.crea
## 1  Mujer [34,36) [1000,1500) Movilizaciones Arrixaca 1622 7057.305
## 2  Hombre [28,32) [1500,2500) Masaje Arrixaca 1644 3962.891
## 3  Mujer [34,36) [2500,3000] Movilizaciones Arrixaca 1617 3451.997
## 4  Mujer [32,34) [1500,2500) Terapia Arrixaca 1649 3712.299
## 5  Mujer [34,36) [1500,2500) Movilizaciones Arrixaca 1582 15691.385
## 6  Mujer [32,34) [1500,2500) Terapia Arrixaca 1628 NA
##      NTxSerum_1 OC1 BC1 Peso.1 Talla.1 PC1 VST2 NTx2.crea Peso.2 Talla.2 PC2 VST3
## 1      NA NA NA 1560 41.0 32 1567 7057.305 2140 47 34.5 1585
## 2      NA NA NA 1460 42.0 29 1610 1400.721 1950 42 30.0 1610
## 3      NA NA NA 2480 48.0 33 1605 3528.538 2980 51 33.0 1605
## 4      NA NA NA 1419 42.0 29 1624 5451.739 1803 44 33.0 1631
## 5      NA NA NA 2130 42.5 33 1572 4628.934 2770 48 35.0 1572
## 6      NA NA NA 2120 44.4 30 1620 NA NA NA NA 1620
##      NTx3.crea NTxSerum_3 OC3 BC3 Peso.3 Talla.3 PC.3 Peso.alta Talla.alta PC.alta
## 1 7057.305 NA NA NA 2660 47.0 35.5 2555 47.0 35.5
## 2 3854.461 NA NA NA 2520 42.5 32.5 2740 45.5 34.0
## 3 3528.538 NA NA NA NA NA 2980 51.0 33.0
## 4 7010.474 NA NA NA 2410 47.0 35.0 2410 47.0 35.0
## 5 4628.934 NA NA NA 3200 NA NA 3220 NA NA
## 6 NA NA NA NA NA NA 2415 48.0 30.5
##      Días.de.tratamiento Días.de.Hospitalización
## 1 18 38
## 2 20 49
## 3 7 20
## 4 20 37
## 5 20 45
## 6 4 12
```

2 Inferencia

Se trata de ver si hay diferencias significativas en las distantes variables (numéricas) a lo largo del tiempo y respecto a las diferentes variables sociodemográficas. Dado que tenemos datos independientes y medidas repetidas, se trata de un estudio mixto para el tiempo y cada una de las variables sociodemográficas, por lo que haremos uso de un ANOVA Mixto cuando sea posible. En caso, de no cumplirse los supuestos necesarios, aplicaremos un test ANOVA robusto mixto.

2.1 Peso

Veremos el peso en todos los tiempos, pero como nos quedamos con menos datos (28 individuos), posteriormente lo veremos sin el último tiempo (alta).

2.1.1 Sexo

Comprobamos los supuestos.

```
## NORMALIDAD:
```

```
## Hombre
```

```
## Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.1[dd$Sexo == "Hombre"]
```

```
## W = 0.79092, p-value = 0.005328
```

```
## -----
```

```
## Mujer
```

```
## Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.1[dd$Sexo == "Mujer"]
```

```
## W = 0.91744, p-value = 0.176
```

```
## -----
```

```
## Hombre
```

```
## Peso.2
```

```

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.2[dd$Sexo == "Hombre"]
## W = 0.87045, p-value = 0.05304

## -----

## Mujer

## Peso.2

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.2[dd$Sexo == "Mujer"]
## W = 0.93679, p-value = 0.3437

## -----

## Hombre

## Peso.3

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.3[dd$Sexo == "Hombre"]
## W = 0.93019, p-value = 0.3428

## -----

## Mujer

## Peso.3

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.3[dd$Sexo == "Mujer"]
## W = 0.95599, p-value = 0.6232

## -----

## Hombre

## Peso.alta

```

```

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.alta[dd$Sexo == "Hombre"]
## W = 0.94745, p-value = 0.5601

## -----

## Mujer

## Peso.alta

##
## Shapiro-Wilk normality test
##
## data: dd$Peso.alta[dd$Sexo == "Mujer"]
## W = 0.95658, p-value = 0.6333

## HOMOCESTIDAD

##

## Peso.1

##
## Fligner-Killeen test of homogeneity of variances
##
## data: Peso.1 by Sexo
## Fligner-Killeen:med chi-squared = 0.0094223, df = 1, p-value = 0.9227

## -----

## Peso.2

##
## Fligner-Killeen test of homogeneity of variances
##
## data: Peso.2 by Sexo
## Fligner-Killeen:med chi-squared = 0.057409, df = 1, p-value = 0.8106

## -----

## Peso.3

##
## Fligner-Killeen test of homogeneity of variances
##
## data: Peso.3 by Sexo
## Fligner-Killeen:med chi-squared = 1.1282e-05, df = 1, p-value = 0.9973

```

```
## -----

## Peso.alta

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.alta by Sexo
## Fligner-Killeen:med chi-squared = 0.23146, df = 1, p-value = 0.6304

Nos falta normalidad para uno de los cruce y tenemos homocedasticidad. Dado que el ANOVA es robusto
frente a la falta de algo de normalidad, procedemos a hacer el ANOVA mixto.

## Warning: Converting "id" to factor for ANOVA.

## Warning: Data is unbalanced (unequal N per group). Make sure you specified a
## well-considered value for the type argument to ezANOVA().

## $ANOVA
##      Effect DFn DFd      SSn      SSd      F      p p<.05
## 1 (Intercept)   1  26 585964688.46 18044659 844.2986839 2.373304e-21  *
## 2      Sexo     1  26   163820.25 18044659   0.2360436 6.311478e-01
## 3     tiempo    3  78 21859606.21 2938253 193.4312081 4.997964e-36  *
## 4 Sexo:tiempo    3  78   18821.28 2938253   0.1665457 9.186239e-01
##      ges
## 1 0.9654287920
## 2 0.0077468352
## 3 0.5102315938
## 4 0.0008961774
##
## $'Mauchly's Test for Sphericity'
##      Effect      W      p p<.05
## 3     tiempo 0.3183394 3.245822e-05  *
## 4 Sexo:tiempo 0.3183394 3.245822e-05  *
##
## $'Sphericity Corrections'
##      Effect      GGe      p[GG] p[GG]<.05      HFe      p[HF] p[HF]<.05
## 3     tiempo 0.6062208 8.972252e-23  * 0.6493146 3.164267e-24  *
## 4 Sexo:tiempo 0.6062208 8.272977e-01  * 0.6493146 8.416558e-01
```

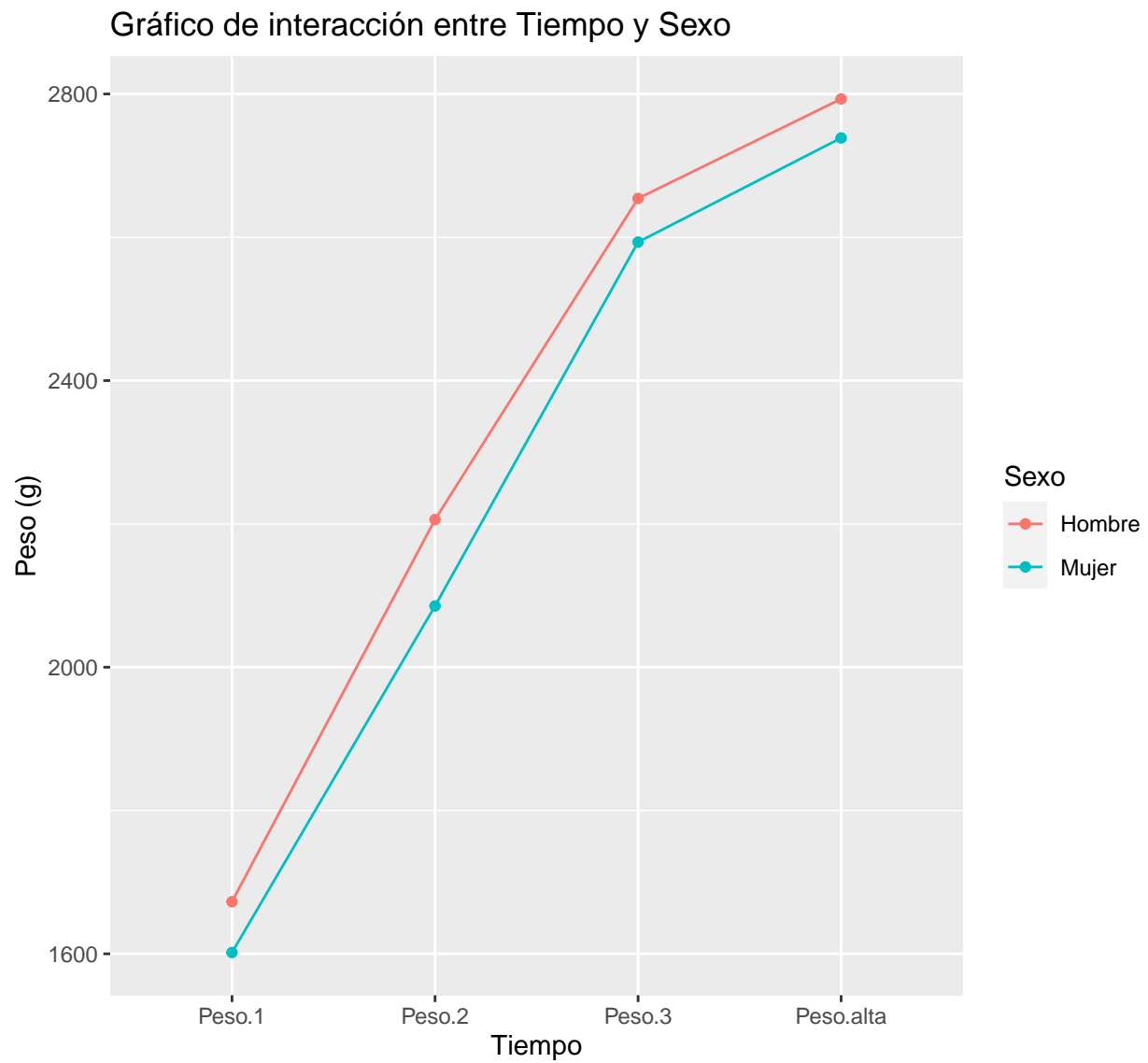
No se cumple la esfericidad por lo que nos tenemos que fijar en la corrección Greenhouse-Geisser y comprobamos que se encuentran diferencias para el tiempo ($\eta^2 = 0.5097636$), pero no para el sexo ni para la interacción. Donde eta cuadrado es el tamaño del efecto cuyos rangos vienen dados como sigue:

Efecto	Pequeño	Mediano	Grande
η^2	0.01	0.06	0.14

Veamos ahora donde se encuentran estas diferencias respecto al tiempo.

```
##
## Pairwise comparisons using Wilcoxon signed rank test with continuity correction
##
## data: ddRe$valor and ddRe$tiempo
##
##      Peso.1  Peso.2  Peso.3
## Peso.2  2.4e-05 -      -
## Peso.3  2.4e-05 2.4e-05 -
## Peso.alta 2.4e-05 2.4e-05 0.0013
##
## P value adjustment method: bonferroni
```

Se encuentran diferencias entre todos los tiempos. Veamos un gráfico para la interacción.



2.1.2 Semana de gestación

En este caso, al quedarnos con aquellos individuos para los que disponemos de datos del peso en todos los tiempo, para el nivel [32,34) quedan 2 individuos y para el nivel [34,36) 3 individuos, por lo que uniremos estos 2 niveles en un único nivel [32,36).

Comprobamos los supuestos.

```
## NORMALIDAD:
```

```
##
```

```
## [28,32)    Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  dd$Peso.1[dd$SG == "[28,32)"]
```

```
## W = 0.8446, p-value = 0.002175
```

```
## -----
```

```
## [32,36)    Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  dd$Peso.1[dd$SG == "[32,36)"]
```

```
## W = 0.73558, p-value = 0.02177
```

```
## -----
```

```
## [28,32)    Peso.2
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  dd$Peso.2[dd$SG == "[28,32)"]
```

```
## W = 0.91004, p-value = 0.04093
```

```
## -----
```

```
## [32,36)    Peso.2
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  dd$Peso.2[dd$SG == "[32,36)"]
```

```
## W = 0.85708, p-value = 0.2179
```

```
## -----
```



```

## [28,32)    Peso.3

##
##  Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$SG == "[28,32)"]
## W = 0.956, p-value = 0.3876

## -----

## [32,36)    Peso.3

##
##  Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$SG == "[32,36)"]
## W = 0.96311, p-value = 0.8294

## -----

## [28,32)    Peso.alta

##
##  Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$SG == "[28,32)"]
## W = 0.96536, p-value = 0.5795

## -----

## [32,36)    Peso.alta

##
##  Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$SG == "[32,36)"]
## W = 0.87215, p-value = 0.2753

## HOMOCESTIDAD

##

## Peso.1

##
##  Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.1 by SG
## Fligner-Killeen:med chi-squared = 0.68242, df = 1, p-value = 0.4088

```

```

## -----

## Peso.2

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.2 by SG
## Fligner-Killeen:med chi-squared = 0.47032, df = 1, p-value = 0.4928

## -----

## Peso.3

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.3 by SG
## Fligner-Killeen:med chi-squared = 0.34312, df = 1, p-value = 0.558

## -----

## Peso.alta

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.alta by SG
## Fligner-Killeen:med chi-squared = 0.22637, df = 1, p-value = 0.6342

Nos falta normalidad para algún cruce y tenemos homocedasticidad. Dado que el ANOVA es robusto frente
a la falta de algo de normalidad, procedemos a hacer el ANOVA mixto.

## Warning: Converting "id" to factor for ANOVA.

## Warning: Data is unbalanced (unequal N per group). Make sure you specified a
## well-considered value for the type argument to ezANOVA().

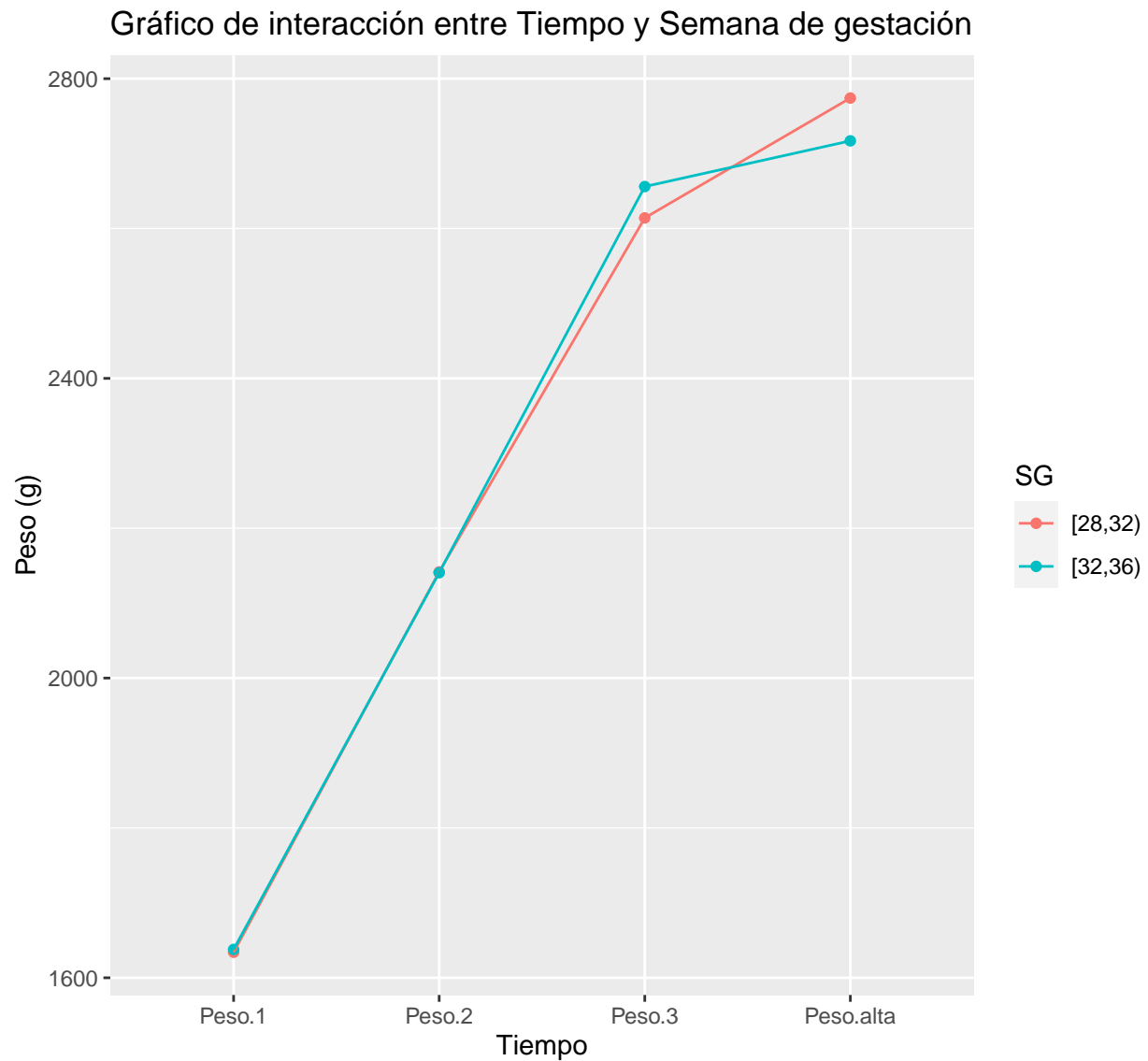
## $ANOVA
##      Effect DFn DFd      SSn      SSd      F      p p<.05
## 1 (Intercept)   1  26 3.444392e+08 18208316 4.918313e+02 2.045754e-18 *
## 2             SG   1  26 1.630125e+02 18208316 2.327686e-04 9.879438e-01
## 3      tiempo   3  78 1.275816e+07  2936569 1.129591e+02 2.628794e-28 *
## 4  SG:tiempo   3  78 2.050519e+04  2936569 1.815503e-01 9.086265e-01
##           ges
## 1 9.421614e-01
## 2 7.709252e-06
## 3 3.763131e-01
## 4 9.688077e-04
##
## $'Mauchly's Test for Sphericity'

```

```
##      Effect      W      p p<.05
## 3      tiempo 0.3174725 3.148976e-05 *
## 4 SG:tiempo 0.3174725 3.148976e-05 *
##
## $'Sphericity Corrections'
##      Effect      GGe      p[GG] p[GG]<.05      HFe      p[HF] p[HF]<.05
## 3      tiempo 0.6041106 5.033060e-18 * 0.6467891 3.901853e-19 *
## 4 SG:tiempo 0.6041106 8.136662e-01      0.6467891 8.282019e-01
```

No se cumple la esfericidad por lo que nos tenemos que fijar en la corrección *Greenhouse-Geisser* y comprobamos que se encuentran diferencias para el tiempo ($\eta^2 = 0.3763131$), pero no para la semana de gestación ni para la interacción.

Veamos un gráfico para la interacción.



2.1.3 Peso al nacer

En este caso, al quedarnos con aquellos individuos para los que disponemos de datos del peso en todos los tiempo, para el nivel [2500,3000] queda 1 individuo así que lo uniremos con el nivel [15000,2500) en un único nivel, que será [1500,3000]. Comprobamos los supuestos.

```
## NORMALIDAD:
```

```
##
```

```
## [1000,1500)    Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.1[dd$Peso.al.nacer == "[1e+03,1.5e+03)"]
```

```
## W = 0.95595, p-value = 0.6904
```

```
## -----
```

```
## [1500, 3000]    Peso.1
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.1[dd$Peso.al.nacer == "[1.5e+03,3e+03]"]
```

```
## W = 0.84073, p-value = 0.01287
```

```
## -----
```

```
## [1000,1500)    Peso.2
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.2[dd$Peso.al.nacer == "[1e+03,1.5e+03)"]
```

```
## W = 0.98463, p-value = 0.9948
```

```
## -----
```

```
## [1500, 3000]    Peso.2
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: dd$Peso.2[dd$Peso.al.nacer == "[1.5e+03,3e+03]"]
```

```
## W = 0.88635, p-value = 0.05906
```

```
## -----
```

```

## [1000,1500)    Peso.3

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$Peso.al.nacer == "[1e+03,1.5e+03)"]
## W = 0.98612, p-value = 0.9971

## -----

## [1500, 3000]    Peso.3

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$Peso.al.nacer == "[1.5e+03,3e+03]"]
## W = 0.84413, p-value = 0.01436

## -----

## [1000,1500)    Peso.alta

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$Peso.al.nacer == "[1e+03,1.5e+03)"]
## W = 0.94754, p-value = 0.5615

## -----

## [1500, 3000]    Peso.alta

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$Peso.al.nacer == "[1.5e+03,3e+03]"]
## W = 0.933, p-value = 0.3024

## HOMOCESTIDAD

##

## Peso.1

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.1 by Peso.al.nacer
## Fligner-Killeen:med chi-squared = 2.7267, df = 1, p-value = 0.09869

```

```

## -----

## Peso.2

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.2 by Peso.al.nacer
## Fligner-Killeen:med chi-squared = 1.1586, df = 1, p-value = 0.2818

## -----

## Peso.3

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.3 by Peso.al.nacer
## Fligner-Killeen:med chi-squared = 0.061963, df = 1, p-value = 0.8034

## -----

## Peso.alta

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.alta by Peso.al.nacer
## Fligner-Killeen:med chi-squared = 0.44626, df = 1, p-value = 0.5041

Nos falta normalidad para un par de cruces y tenemos homocedasticidad. Dado que el ANOVA es robusto
frente a la falta de algo de normalidad, procedemos a hacer el ANOVA mixto.

## Warning: Converting "id" to factor for ANOVA.

## Warning: Data is unbalanced (unequal N per group). Make sure you specified a
## well-considered value for the type argument to ezANOVA().

## $ANOVA
##           Effect DFn DFd          SSn          SSd          F          p
## 1      (Intercept)   1  26 576834381.44 13164481 1139.254500 5.321186e-23
## 2      Peso.al.nacer   1  26  5043998.58 13164481    9.961955 4.014442e-03
## 3           tiempo    3  78 21740875.86  2860146  197.634222 2.387650e-36
## 4  Peso.al.nacer:tiempo  3  78   96927.71  2860146   0.881116 4.546786e-01
##  p<.05           ges
## 1      * 0.972970594
## 2      * 0.239408051
## 3      * 0.575680826
## 4      0.006012306
##
## $'Mauchly's Test for Sphericity'

```

```

##          Effect          W          p p<.05
## 3          tiempo 0.3145026 2.836614e-05      *
## 4 Peso.al.nacer:tiempo 0.3145026 2.836614e-05      *
##
## $'Sphericity Corrections'
##          Effect          GGe          p[GG] p[GG]<.05          HFe          p[HF]
## 3          tiempo 0.591759 1.779365e-22      * 0.6320331 7.578136e-24
## 4 Peso.al.nacer:tiempo 0.591759 4.099370e-01      0.6320331 4.157498e-01
## p[HF]<.05
## 3          *
## 4

```

No se cumple la esfericidad por lo que nos tenemos que fijar en la corrección *Greenhouse-Geisser* y comprobamos que se encuentran diferencias para el tiempo ($\eta^2 = 0.5756808$) y para el peso al nacer ($\eta^2 = 0.2394081$), pero no para la interacción. Veamos para qué tiempos tenemos las diferencias en el peso al nacer.

```

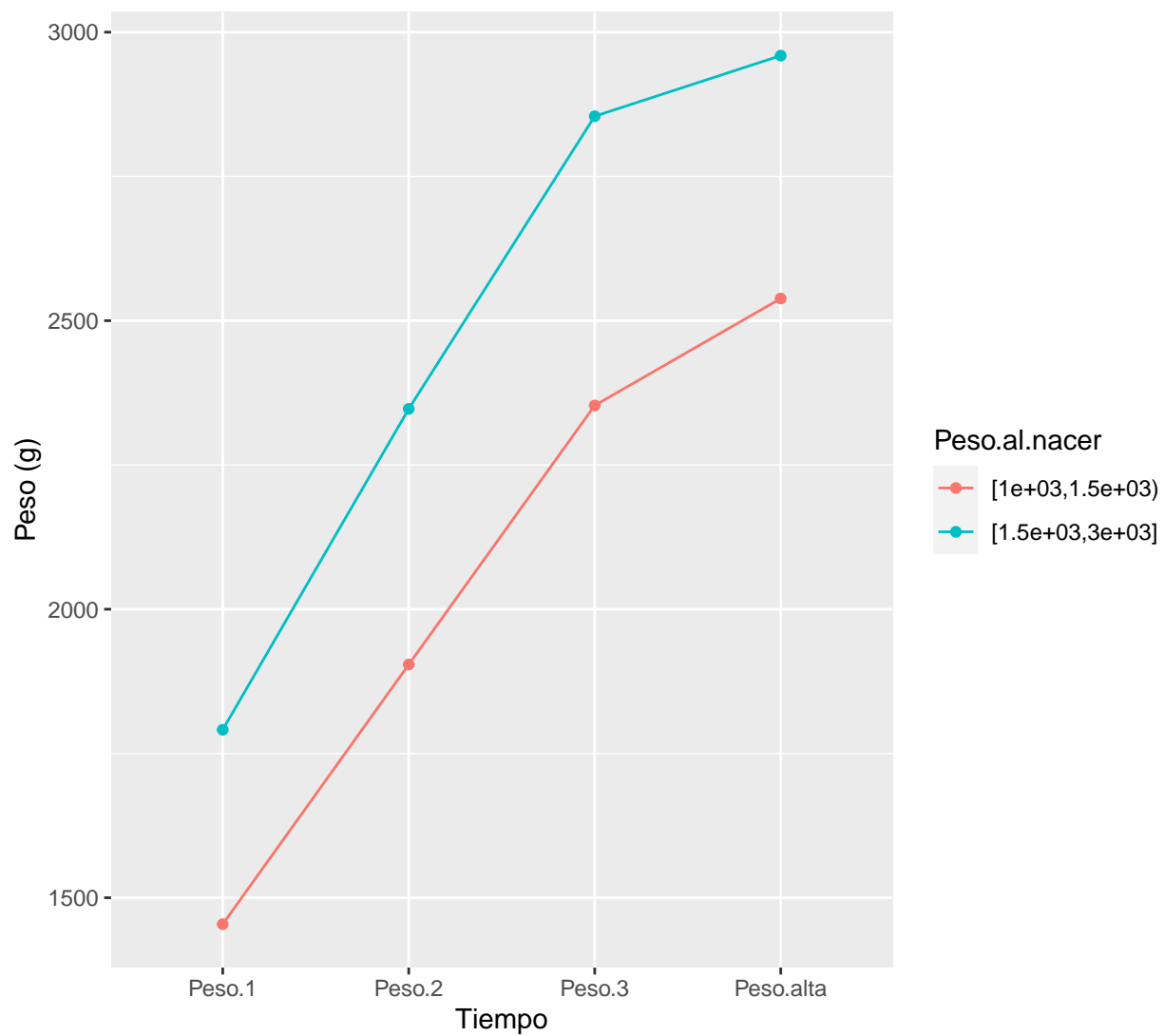
##
## Two Sample t-test
##
## data:  Peso.1 by Peso.al.nacer
## t = -3.0467, df = 26, p-value = 0.005252
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -564.1987 -109.6064
## sample estimates:
## mean in group [1e+03,1.5e+03) mean in group [1.5e+03,3e+03]
##          1454.231          1791.133
##
##
## Two Sample t-test
##
## data:  Peso.2 by Peso.al.nacer
## t = -2.9718, df = 26, p-value = 0.006301
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -749.4888 -136.6035
## sample estimates:
## mean in group [1e+03,1.5e+03) mean in group [1.5e+03,3e+03]
##          1904.154          2347.200
##
##
## Two Sample t-test
##
## data:  Peso.3 by Peso.al.nacer
## t = -3.1483, df = 26, p-value = 0.004094
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -828.5227 -173.9901
## sample estimates:
## mean in group [1e+03,1.5e+03) mean in group [1.5e+03,3e+03]
##          2353.077          2854.333

```

```
##
## Two Sample t-test
##
## data:  Peso.alta by Peso.al.nacer
## t = -2.4844, df = 26, p-value = 0.01974
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -769.08423  -72.65936
## sample estimates:
## mean in group [1e+03,1.5e+03) mean in group [1.5e+03,3e+03]
##                2538.462                2959.333
```

Encontramos diferencias en todos los tiempos. Veamos un gráfico para la interacción.

Gráfico de interacción entre Tiempo y Semana de gestación



2.1.4 Grupo

Comprobamos los supuestos.


```

## NORMALIDAD:

##

## Terapia    Peso.1

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.1[dd$Num.Grupo == "Terapia"]
## W = 0.93053, p-value = 0.4531

## -----

## Movilizaciones    Peso.1

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.1[dd$Num.Grupo == "Movilizaciones"]
## W = 0.95592, p-value = 0.7386

## -----

## Masaje    Peso.1

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.1[dd$Num.Grupo == "Masaje"]
## W = 0.9691, p-value = 0.8909

## -----

## Terapia    Peso.2

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.2[dd$Num.Grupo == "Terapia"]
## W = 0.95864, p-value = 0.7702

## -----

## Movilizaciones    Peso.2

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.2[dd$Num.Grupo == "Movilizaciones"]
## W = 0.97542, p-value = 0.9361

```

```

## -----

## Masaje    Peso.2

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.2[dd$Num.Grupo == "Masaje"]
## W = 0.93358, p-value = 0.5492

## -----

## Terapia    Peso.3

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$Num.Grupo == "Terapia"]
## W = 0.90152, p-value = 0.2276

## -----

## Movilizaciones    Peso.3

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$Num.Grupo == "Movilizaciones"]
## W = 0.97635, p-value = 0.9427

## -----

## Masaje    Peso.3

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.3[dd$Num.Grupo == "Masaje"]
## W = 0.87887, p-value = 0.1837

## -----

## Terapia    Peso.alta

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$Num.Grupo == "Terapia"]
## W = 0.95088, p-value = 0.6789

```

```

## -----

## Movilizaciones    Peso.alta

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$Num.Grupo == "Movilizaciones"]
## W = 0.99132, p-value = 0.9982

## -----

## Masaje    Peso.alta

##
## Shapiro-Wilk normality test
##
## data:  dd$Peso.alta[dd$Num.Grupo == "Masaje"]
## W = 0.95229, p-value = 0.7343

## HOMOCESTIDAD

##

## Peso.1

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.1 by Num.Grupo
## Fligner-Killeen:med chi-squared = 12.824, df = 2, p-value = 0.001642

## -----

## Peso.2

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.2 by Num.Grupo
## Fligner-Killeen:med chi-squared = 13.807, df = 2, p-value = 0.001005

## -----

## Peso.3

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  Peso.3 by Num.Grupo
## Fligner-Killeen:med chi-squared = 2.8532, df = 2, p-value = 0.2401

```

```
## -----
```

```
## Peso.alta
```

```
##
```

```
## Fligner-Killeen test of homogeneity of variances
```

```
##
```

```
## data:  Peso.alta by Num.Grupo
```

```
## Fligner-Killeen:med chi-squared = 4.427, df = 2, p-value = 0.1093
```

Tenemos normalidad pero no homocedasticidad y no podemos realizar un ANOVA mixto, pasamos a hacer un ANOVA robusto mixto.

```
## Call:
```

```
## tsplit(formula = valor ~ Num.Grupo * tiempo, id = id, data = dt)
```

```
##
```

```
##           value df1    df2 p.value
```

```
## Num.Grupo      8.0751  2  9.0965  0.0096
```

```
## tiempo        138.8501  3 10.4888  0.0000
```

```
## Num.Grupo:tiempo  0.9008  6  9.3373  0.5325
```

Se encuentran diferencias para el tiempo ($\eta^2 = 0.5865396$) y para el grupo ($\eta^2 = 0.2337761$), pero no para la interacción. Ahora buscamos dónde se encuentran las diferencias entre los grupos para los distintos tiempos.

```
## Tiempo: Peso.1
```

```
##
```

```
## Pairwise comparisons using Wilcoxon rank sum test with continuity correction
```

```
##
```

```
## data:  dt$valor[dt$tiempo == "Peso.1"] and dt$Num.Grupo[dt$tiempo == "Peso.1"]
```

```
##
```

```
##           Terapia Movilizaciones
```

```
## Movilizaciones 0.148  -
```

```
## Masaje        0.079  1.000
```

```
##
```

```
## P value adjustment method: bonferroni
```

```
## Tiempo: Peso.2
```

```
##
```

```
## Pairwise comparisons using Wilcoxon rank sum test with continuity correction
```

```
##
```

```
## data:  dt$valor[dt$tiempo == "Peso.2"] and dt$Num.Grupo[dt$tiempo == "Peso.2"]
```

```
##
```

```
##           Terapia Movilizaciones
```

```
## Movilizaciones 0.31  -
```

```
## Masaje        0.02  1.00
```

```
##
```

```
## P value adjustment method: bonferroni
```

```
## Tiempo: Peso.3
```

```

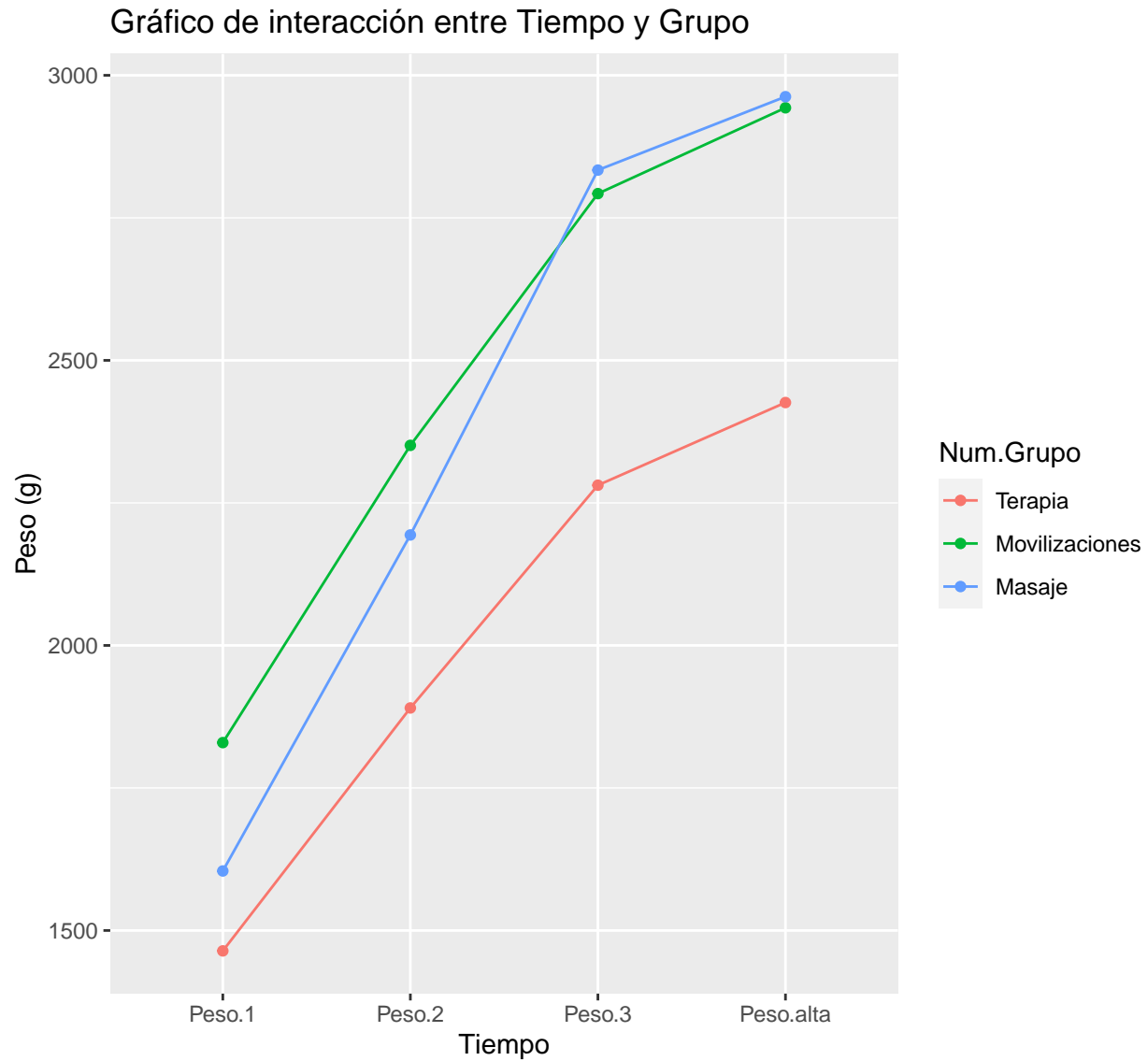
##
## Pairwise comparisons using t tests with pooled SD
##
## data:  dt$valor[dt$tiempo == "Peso.3"] and dt$Num.Grupo[dt$tiempo == "Peso.3"]
##
##               Terapia Movilizaciones
## Movilizaciones 0.038      -
## Masaje         0.034    1.000
##
## P value adjustment method: bonferroni

## Tiempo: Peso.alta

##
## Pairwise comparisons using t tests with pooled SD
##
## data:  dt$valor[dt$tiempo == "Peso.alta"] and dt$Num.Grupo[dt$tiempo == "Peso.alta"]
##
##               Terapia Movilizaciones
## Movilizaciones 0.039      -
## Masaje         0.044    1.000
##
## P value adjustment method: bonferroni

```

Encontramos diferencias entre Terapia y Masaje en los 4 tiempos y entre Terapia y Movilizaciones para los tiempos 1 y 2. Veamos un gráfico para la interacción.



2.2 Días de hospitalización

En este caso, como se trata de un único tiempo, haremos un estudio respecto cada una de las variables sociodegráficas, por lo que aplicaremos un t-test independiente o un ANOVA de una vía dependiente de los niveles del factor.

2.2.1 Sexo

En este caso, como sexo tiene dos niveles aplicaremos un t-test independiente. Comprobamos los supuestos

NORMALIDAD:

```
##
## Shapiro-Wilk normality test
##
```

```
## data: dd$Días.de.Hospitalización[dd$Sexo == "Hombre"]
## W = 0.7962, p-value = 1.305e-06

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Sexo == "Mujer"]
## W = 0.79398, p-value = 1.239e-05

## HOMOCEASTICIDAD

##
## Fligner-Killeen test of homogeneity of variances
##
## data: Días.de.Hospitalización by Sexo
## Fligner-Killeen:med chi-squared = 0.0073887, df = 1, p-value = 0.9315
```

No tenemos normalidad pero sí homocedasticidad, por lo que no podemos aplicar el t-test. Aplicamos en este caso la *U de Mann-Whitney*, que es la alternativa no paramétrica a la falta de normalidad.

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: Días.de.Hospitalización by Sexo
## W = 793.5, p-value = 0.6326
## alternative hypothesis: true location shift is not equal to 0
```

No se encuentran diferencias para el sexo.

2.2.2 Semana de gestación

En este caso como para el nivel [24,28) solo hay 1 dato para cada tiempo, lo uniremos con el nivel [28,32) en un único nivel, que será [24,32). Nos quedamos así con 3 niveles en la variable, por lo que aplicaremos un ANOVA de una vía. Comprobamos los supuestos.

```
## NORMALIDAD:

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$SG == "[24,32)"]
## W = 0.75984, p-value = 5.077e-08

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$SG == "[32,34)"]
## W = 0.88952, p-value = 0.01093
```

```
## -----
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: dd$Días.de.Hospitalización[dd$SG == "[34,36)"]  
## W = 0.88097, p-value = 0.3428
```

```
## HOMOCEASTICIDAD
```

```
##  
## Fligner-Killeen test of homogeneity of variances  
##  
## data: Días.de.Hospitalización by SG  
## Fligner-Killeen:med chi-squared = 2.037, df = 2, p-value = 0.3611
```

No tenemos normalidad pero sí homocedasticidad, por lo que no podemos aplicar el ANOVA. Aplicamos el test de Welch, que es la alternativa no paramétrica.

```
##           Df Sum Sq Mean Sq F value Pr(>F)  
## SG           2   4885   2442.7    6.846 0.0018 **  
## Residuals    80   28543    356.8  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
## 23 observations deleted due to missingness
```

Se encuentran diferencias para la semana de gestación. Vemos entre qué niveles se encuentran esas diferencias.

```
##  
## Pairwise comparisons using t tests with pooled SD  
##  
## data: dd$Días.de.Hospitalización and dd$SG  
##  
##           [24,32) [32,34)  
## [32,34) 0.0013 -  
## [34,36) 0.9502 1.0000  
##  
## P value adjustment method: bonferroni
```

Se encuentran diferencias entre las semanas [24,32) y [32,34).

2.2.3 Peso al nacer

En este caso para el nivel [2500,3000] quedan 2 individuos así que lo uniremos con el nivel [15000,2500) en un único nivel, que será [1500,3000]. Nos quedamos así con 3 niveles en la variable, por lo que aplicaremos un ANOVA de una vía. Comprobamos los supuestos.

```
## NORMALIDAD:
```



```
##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Peso.al.nacer == "[0,1e+03)"]
## W = 0.93052, p-value = 0.5841

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Peso.al.nacer == "[1e+03,1.5e+03)"]
## W = 0.7092, p-value = 3.029e-07

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Peso.al.nacer == "[1.5e+03,3e+03)"]
## W = 0.97665, p-value = 0.5671
```

HOMOCEASTICIDAD

```
##
## Fligner-Killeen test of homogeneity of variances
##
## data: Días.de.Hospitalización by Peso.al.nacer
## Fligner-Killeen:med chi-squared = 0.78519, df = 2, p-value = 0.6753
```

Nos falta normalidad para uno de los niveles y tenemos homocedasticidad. Dado que el ANOVA es robusto frente a la falta de algo de normalidad, procedemos a hacer el ANOVA de una vía.

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Peso.al.nacer  2   7063    3532   10.72 7.53e-05 ***
## Residuals    80  26365     330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 23 observations deleted due to missingness
```

Se encuentran diferencias para el peso al nacer. Vemos entre qué niveles se encuentran esas diferencias.

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: dd$Días.de.Hospitalización and dd$Peso.al.nacer
##
##           [0,1e+03) [1e+03,1.5e+03)
## [1e+03,1.5e+03) 0.16771 -
## [1.5e+03,3e+03) 0.00088 0.00211
##
## P value adjustment method: bonferroni
```

2.2.4 Grupo

Como tenemos 3 niveles en la variable, aplicaremos un ANOVA de una vía. Comprobamos los supuestos.

```
## NORMALIDAD:

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Num.Grupo == "Terapia"]
## W = 0.78609, p-value = 4.751e-05

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Num.Grupo == "Movilizaciones"]
## W = 0.75099, p-value = 3.84e-05

## -----

##
## Shapiro-Wilk normality test
##
## data: dd$Días.de.Hospitalización[dd$Num.Grupo == "Masaje"]
## W = 0.90249, p-value = 0.01125

## HOMOCEASTICIDAD

##
## Fligner-Killeen test of homogeneity of variances
##
## data: Días.de.Hospitalización by Num.Grupo
## Fligner-Killeen:med chi-squared = 0.6627, df = 2, p-value = 0.718
```

No tenemos normalidad ni homocedasticidad, por lo que no podemos aplicar el ANOVA. Aplicamos el test de Welch, que es la alternativa no paramétrica.

```
## Call:
## welchADF.test(formula = Días.de.Hospitalización ~ Num.Grupo,
## data = dd)
##
## WJ statistic Numerator DF Denominator DF Pr(>WJ)
## Num.Grupo      0.5091          2          49.54 0.6041
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No se encuentran diferencias para el grupo.

3 Recomendaciones

Les recordamos que dependiendo de la normativa de escritura utilizada en sus investigaciones, deberán escribir de la manera más adecuada los resultados obtenidos en este informe (por ejemplo, en la rama de Ciencias Sociales, siguiendo normativa APA, un p-valor de 0.025, será un p-valor de .025).

4 Agradecimientos

Sección de Apoyo Estadístico (SAE), Área Científica y de Investigación (ACTI), Universidad de Murcia (www.um.es/web/acti).

Análisis realizados con **R** versión 4.0.3 (R Core Team 2020).

5 Referencias y bibliografía

- Dag, O., Dolgun, A., & Konar, N. M. (2018). onewaytests: An R Package for One-Way Tests in Independent Groups Designs. *The R Journal*, 10(1), 175–199. <https://journal.r-project.org/archive/2018/0D0A%20%20%20RJ-2018-022/index.html>
- Dag, O., Dolgun, A., Konar, N. M., Weerahandi, S., & Ananda, M. (2021). *Onewaytests: One-way tests in independent groups designs*. <https://CRAN.R-project.org/package=onewaytests>
- DataCamp, P. by. (n.d.). *RDocumentation*. <https://www.rdocumentation.org/>
- Lawrence, M. A. (2016). *Ez: Easy analysis and visualization of factorial experiments*. <http://github.com/mike-lawrence/ez>
- Luque-Calvo, P. L. (2017). *Escribir un trabajo fin de estudios con r markdown*. Disponible en <http://destio.us.es/calvo>.
- Mair, P., & Wilcox, R. (2020). Robust Statistical Methods in R Using the WRS2 Package. *Behavior Research Methods*, 52, 464–488.
- Mair, P., & Wilcox, R. (2021). *WRS2: A collection of robust statistical methods*. <https://r-forge.r-project.org/projects/psychor/>
- Pablo J. Villacorta. (2017). The welchADF package for robust hypothesis testing in unbalanced multivariate mixed models with heteroscedastic and non-normal data. *The R Journal*, 9(2), 309–328. <https://doi.org/10.32614/RJ-2017-049>
- R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- team, O. (2021). *Latex documentation*. Overleaf team. <https://es.overleaf.com/learn>
- Villacorta, P. J. (2019). *WelchADF: Welch-james statistic for robust hypothesis testing under heterocedasticity and non-normality*. <http://decsai.ugr.es/~pjvi/r-packages.html%3E>
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1–20. <http://www.jstatsoft.org/v21/i12/>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Wickham, H. (2020). *Reshape2: Flexibly reshape data: A reboot of the reshape package*. <https://github.com/hadley/reshape>
- Wickham, H. (2021). *Tidyr: Tidy messy data*. <https://CRAN.R-project.org/package=tidyr>

- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2020). *Ggplot2: Create elegant data visualisations using the grammar of graphics*. <https://CRAN.R-project.org/package=ggplot2>
- Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr>
- Yihui Xie, E. R., Christophe Dervieux. (2021). *R markdown cookbook*. <https://bookdown.org/yihui/rmarkdown-cookbook/>

Índice de tablas

Índice de figuras