

1 Gestió de les dades

Llegim les dades `pesosIndividuales.xlsx`

Posem les variables `Box` i `Treat` com a factors.

Creem les següents variables d'interès amb l'objectiu de controlar l'efecte individual i el pes al naixement:

- Guany de pes:

$$BW_{41-28} = BW_{41} - BW_{28}$$

- Index de l'increment de pes:

$$Index_{41-28} = 100 * \frac{BW_{41}}{BW_{28}}$$

- Taxa de l'increment de pes:

$$Taxa_{41-28} = 100 * \frac{BW_{41} - BW_{28}}{BW_{28}}$$

Caldrà tenir en compte que els valors petits d'aquestes variables poden provindre d'individus amb pesos inicialment grans (o també petits).

1.1 Tractament de les dades faltants

1.1.1 Eliminació de les dades faltants

Si concluïm en que la distribució de les dades faltants es pot considerar aleatòria i que n'hi ha poques podem omitir aquestes dades i treure-les de la base de dades sabent que no esbiaixarà l'anàlisi. Encara que cal tenir en compte que hi haurà una petita pèrdua d'informació.

2 Comparació gràfica

Usem el package `patchwork` per a fer layouts de `ggplot2`.

Amb les llibreries `ggplot2`, `devtools` i `easyGgplot2` podrem obtenir ggplots d'una manera considerablement senzilla.

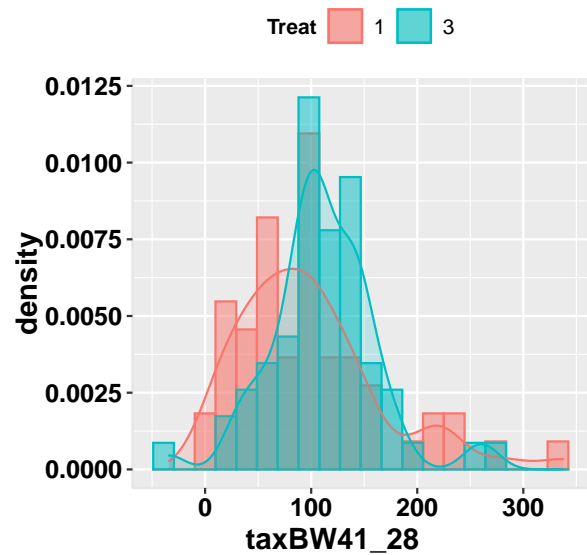
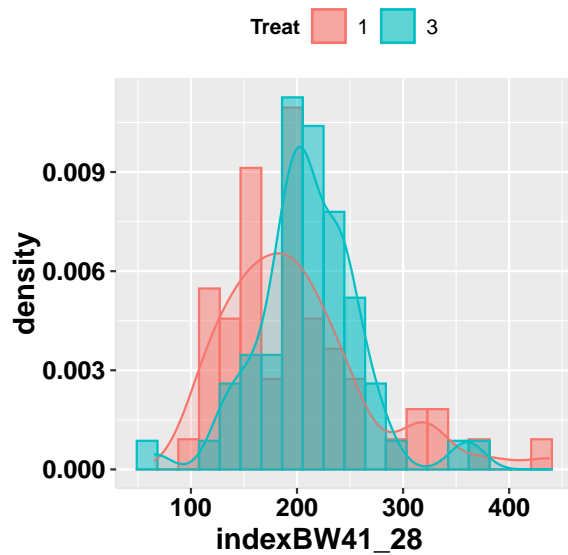
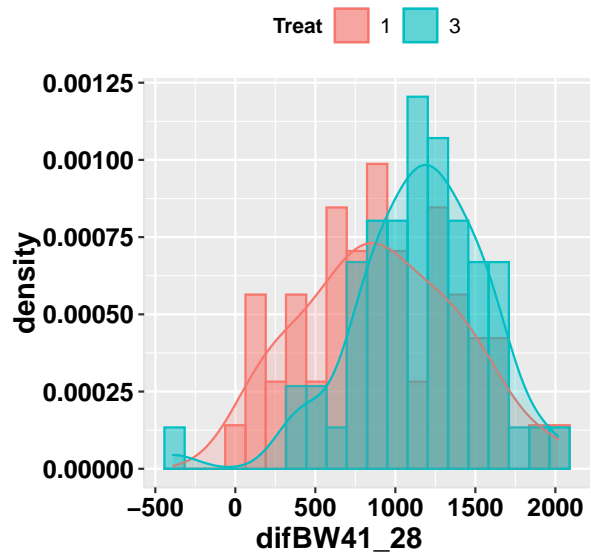
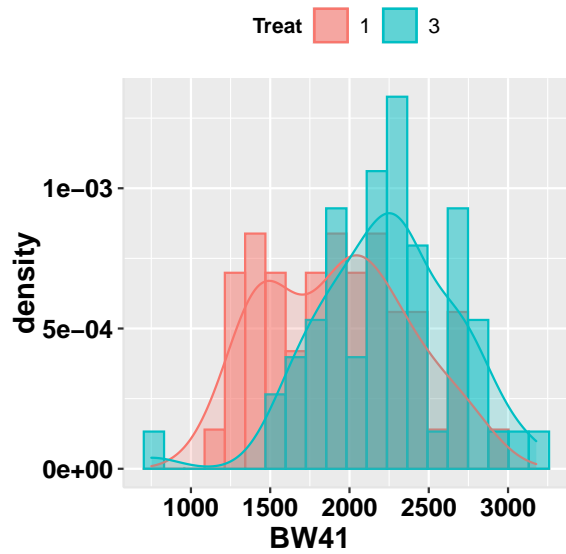
Creem una funció per a simplificar l'obtenció dels histogrames per a comparar tractaments:

```
histComparatiu(BD, Tract1, Tract2, TipusGraf=c(1,2), NombreBins, var)
```

Els arguments de la funció són:

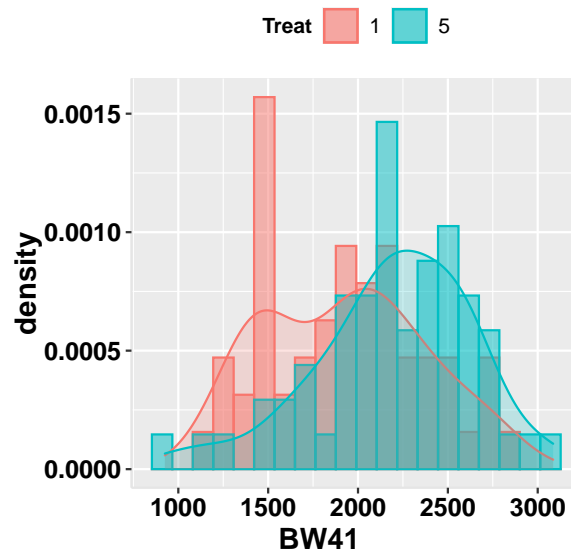
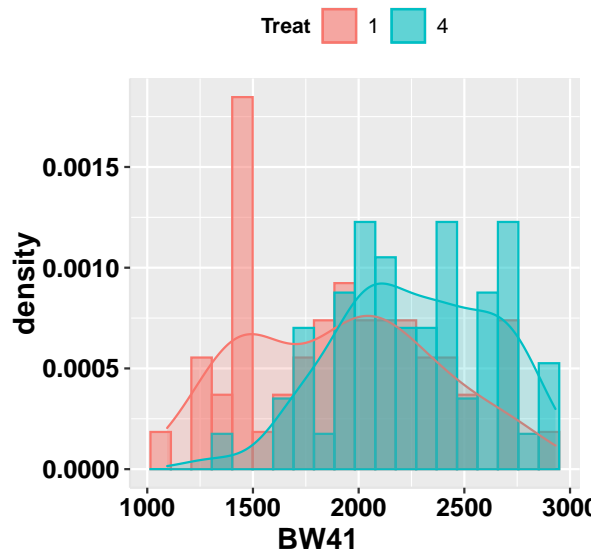
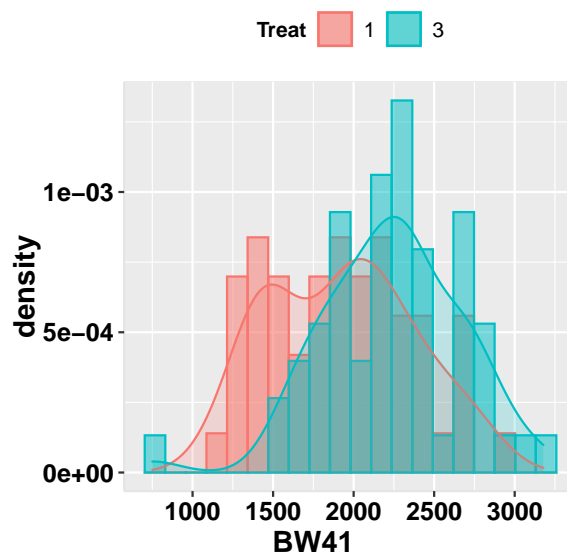
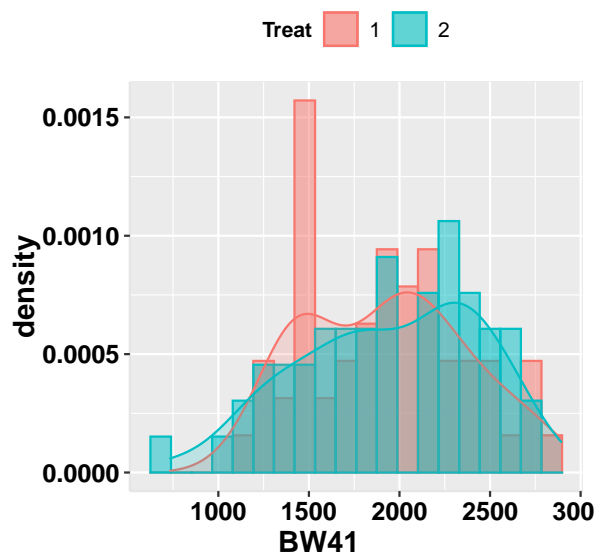
- `BD` la base de dades.
- `Tract1` i `Tract2` són els tractaments a comparar.
- `TipusGraf` si és 1 representem els histogrames solapant-se al mateix eix de les y, si val 2 representem els histogrames un a sobre de l'altre però en diferents eix y.
- `NombreBins` és el nombre de "caixes" en que estarà dividit l'histograma de cadascun dels tractaments, el valor predeterminat és 10.
- `var` és la variable sobre la que es faràn els histogrames.

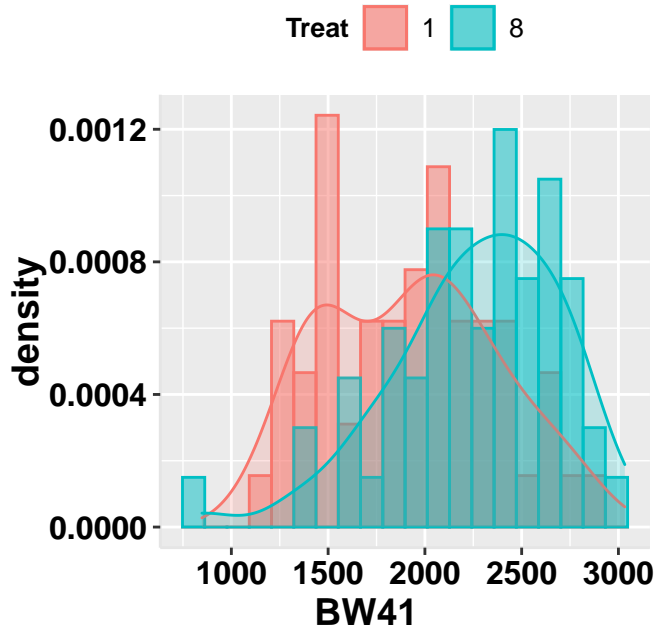
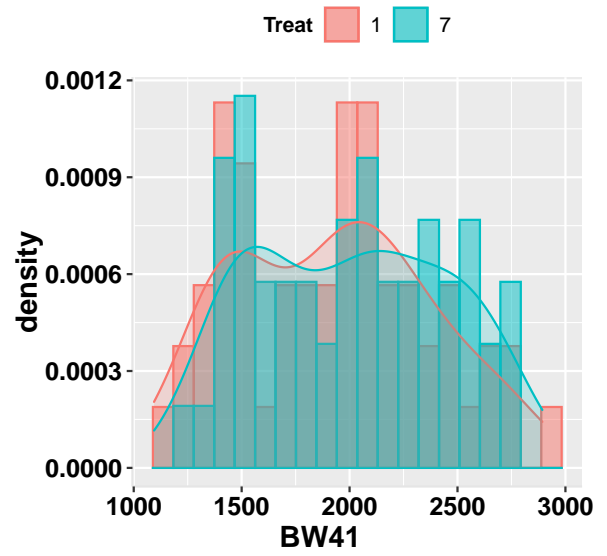
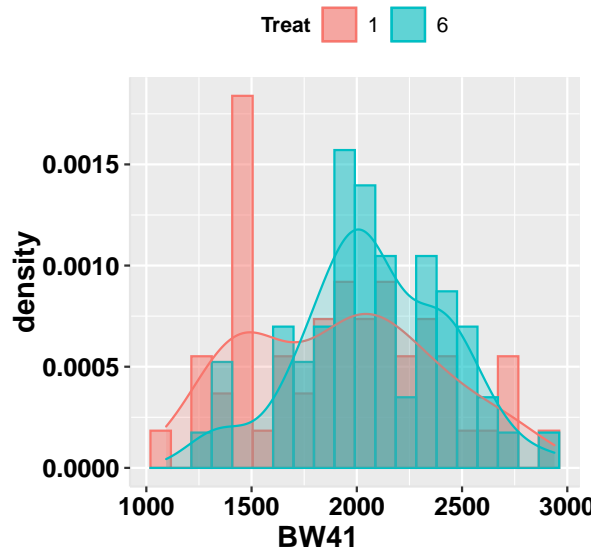
2.1 Histogrammes de les noves variables:



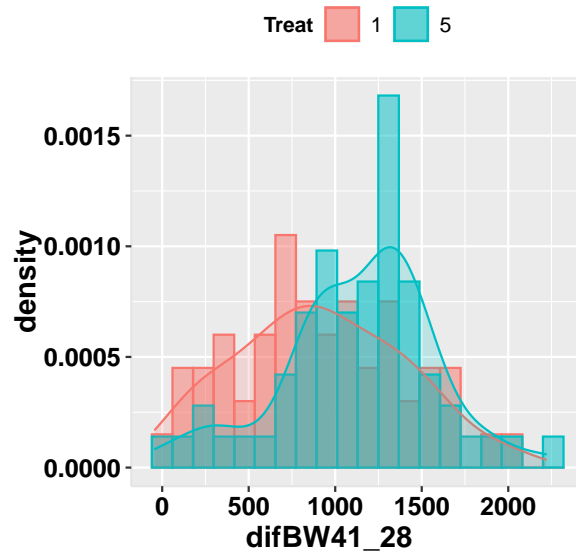
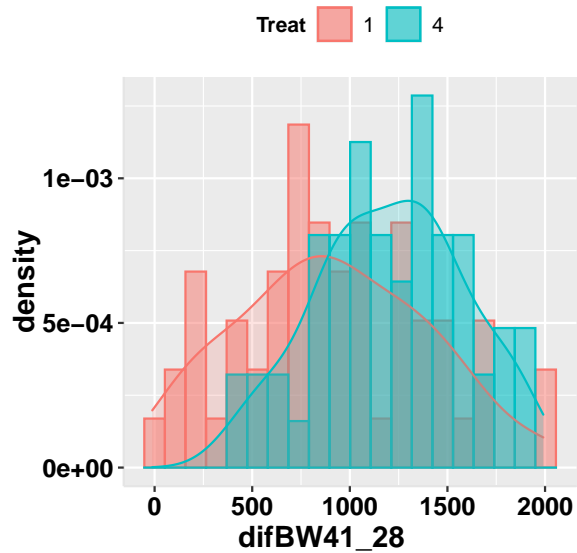
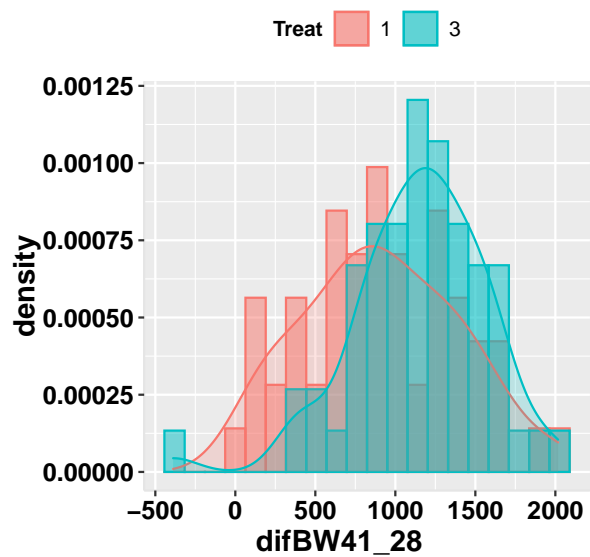
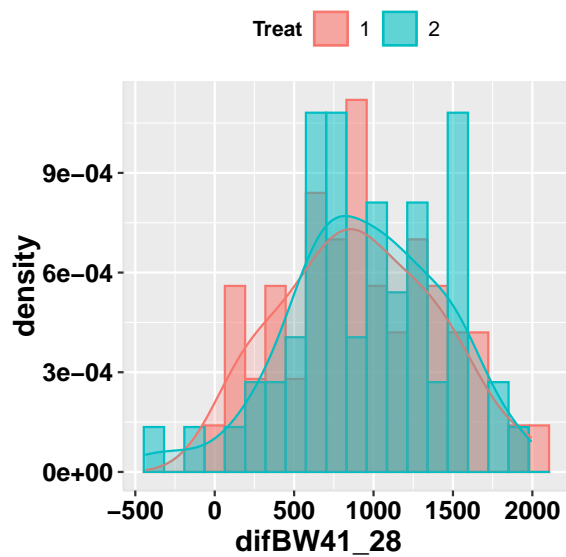
2.2 Comparació de tots els tractaments amb el control

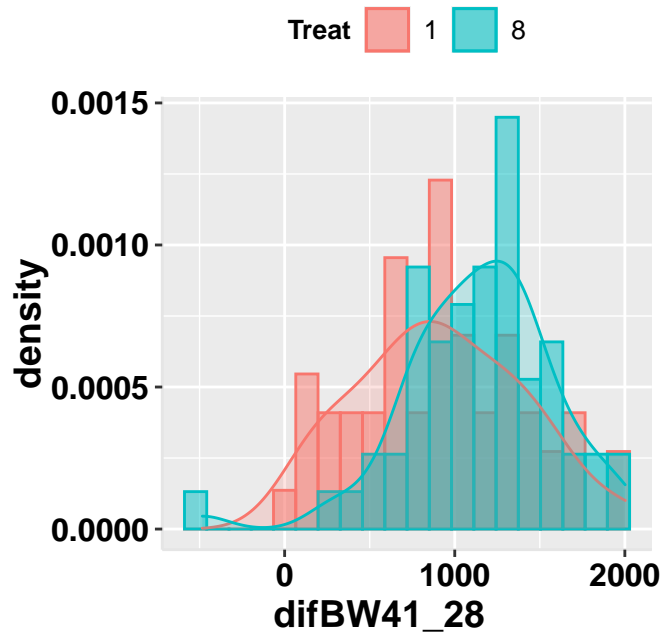
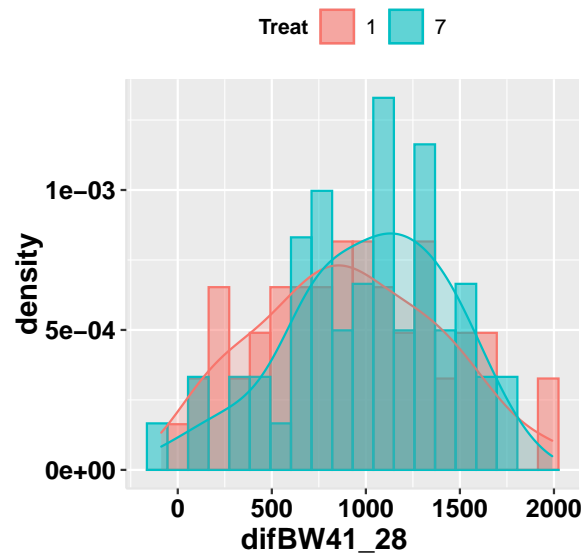
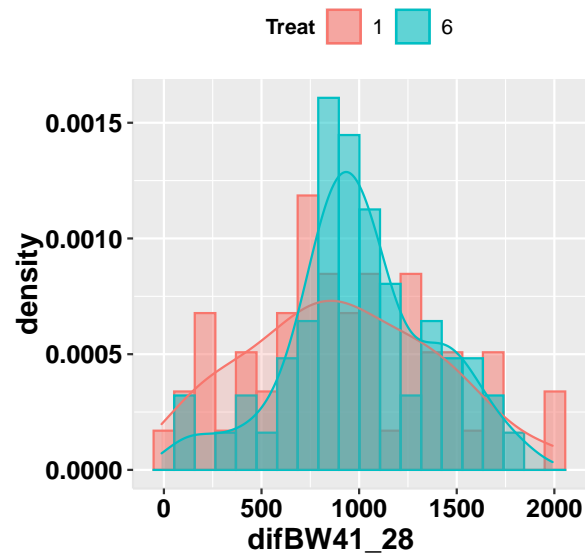
2.2.1 Variable BW41





2.2.2 Variable difBW41₂₈





3 Two-Sample Rank Test to detect a shift in a proportion of the "treated" population

Test bi-mostrat per detectar un canvi positiu en una proporció de la població (tractament) comparada a una altra (control).

```
quantileTest(x, y, alternative = "greater", target.quantile = 0.5, target.r = NULL, exact.p = TRUE)
```

- **x**: Vector numèric d'observacions del grup tractament.
- **y**: Vector numèric d'observacions del grup control.
- **alternative**: Tipus d'hipòtesi alternativa.
 - "greater": La cua dreta del grup tractament desplaçada cap a la dreta de la cua dreta del grup control.
 - "less": La cua esquerra del grup tractament desplaçada cap a la esquerra de la cua esquerra del grup control.
- **target.quantile**: Quantil utilitzat com a punt de tall inferior per a la prova. A causa de la naturalesa discreta dels quantils empírics, el límit superior dels possibles quantils empírics sovint difereix del valor de target.quantile.

H_1 : La porció ϵ de la distribució per al grup de tractament (la distribució de X) es desplaça cap a la dreta de la distribució per al grup de referència (la distribució de Y).

3.1 Resultats del test comparant Tractament vs. control

Apliquem el test al quantil 0.2, de manera que un 20% de la població és menor o igual al valor del quantil. I a la variable diferència dels pesos individuals a la setmana 28 i a la setmana 41.

Veiem un exemple de la sortida de RStudio i una taula resumint els resultats per a tots els tractaments respecte el control.

```
##
## Quantile Test
##
## data: subset(pesInd, Treat == 2)$difBW41_28subset(pesInd, Treat == 1)$difBW41_28
## k (# x obs of r largest) = 94, r = 181, m =
## 115.00000, n = 111.00000, quantile.ub = 0.20264,
## p-value = 0.3206
## alternative hypothesis: true e is 0
```

Tractaments	Quantil 20	P-value quantileTest
1	503	0.32064
2	573.8	
1	503	0.00325
3	819	
1	503	4e-05
4	883.6	
1	503	9e-05
5	865.8	
1	503	0.00555
6	745.2	
1	503	0.29542
7	609	
1	503	0.01499
8	766	

Taula 1: Taula d'anàlisi de quantiles per a la variable guany de pes $BW_{41_{28}}$

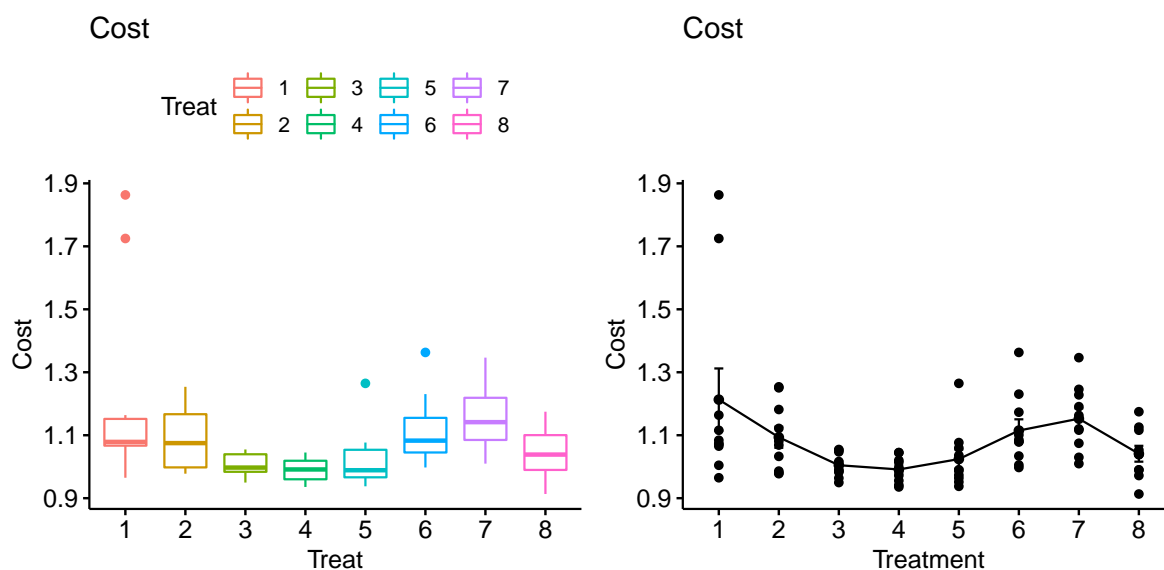
4 Costos

Usarem la variable `Coste_per` PV del fitxer `CalculoIndices.xlsx` modificat.

La variable s'ha calculat de la següent forma:

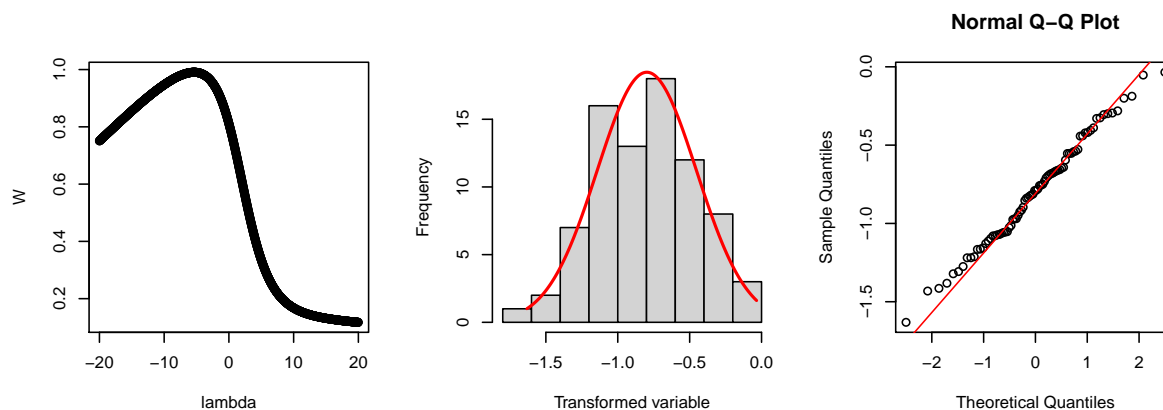
$$\text{Cost/kg PV (€/kg)} = (\text{cost pinso [€/kg]} * \text{consum pinso [kg]}) / \text{guany pes [kg]}$$

4.1 Exploració de la variable Cost

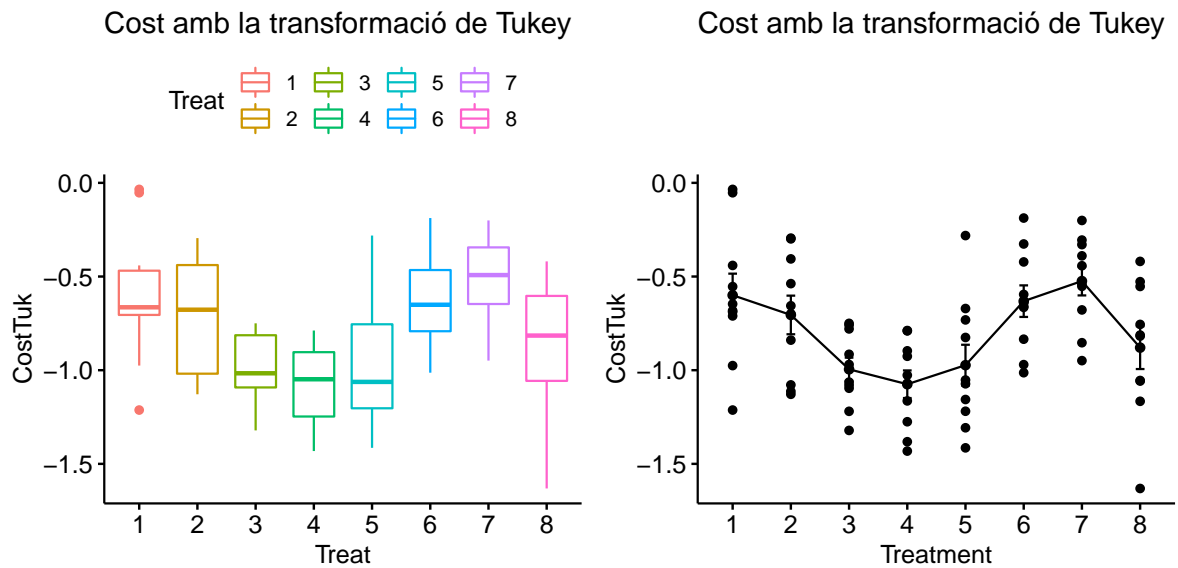


Per tal de millorar la normalitat del costos apliquem una transformació de Tukey de la llibreria `rcompanion`:

```
##  
##      lambda      W Shapiro.p.value  
## 585    -5.4 0.9906          0.8288  
##  
## if (lambda > 0){TRANS = x ^ lambda}  
## if (lambda == 0){TRANS = log(x)}  
## if (lambda < 0){TRANS = -1 * x ^ lambda}
```



El programa selecciona una lambda de -5.4 de manera que com $\lambda < 0$ la transformació aplicada als costos serà $costTuk = -1 * Cost^\lambda$



4.2 Anàlisi dels resultats

Amb les dades normalitzades ajustem un model ANOVA i realitzem els següents tests:

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Treat      7  3.035   0.4336    4.881 0.000152 ***
## Residuals  72  6.397   0.0888
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tukey multiple pairwise-comparisons: Crea un conjunt d'interval de confiança per les diferències de les mitjanes per als diferents tractaments.

```
##           diff      lwr      upr      p adj
## 2-1 -0.10534095 -0.5214742  0.31079227 0.99313670
## 3-1 -0.39553259 -0.8116658  0.02060063 0.07431879
## 4-1 -0.47522197 -0.8913552 -0.05908875 0.01427889
## 5-1 -0.37332879 -0.7894620  0.04280442 0.11097158
## 6-1 -0.03179923 -0.4479324  0.38433399 0.99999757
## 7-1  0.07552814 -0.3406051  0.49166136 0.99915669
## 8-1 -0.27980476 -0.6959380  0.13632846 0.42544395
```

Si observem el p-valor ajustat per comparacions múltiples observem que els tractaments més significatius respecte del control són els tractaments 4,3 i 5, en ordre de significació.

Pairwise t-test: Calcula comparacions per parelles entre els diferent tractaments amb correccions per a proves múltiples.

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data:  costos$CostTuk and costos$Treat
##
##      1      2      3      4      5      6      7
## 2 0.5691 -      -      -      -      -      -
## 3 0.0190 0.0764 -      -      -      -      -
## 4 0.0067 0.0246 0.6277 -      -      -      -
## 5 0.0246 0.0963 0.8682 0.5691 -      -      -
## 6 0.8422 0.6277 0.0248 0.0078 0.0318 -      -
## 7 0.6277 0.2949 0.0067 0.0027 0.0078 0.5691 -
## 8 0.0847 0.3030 0.5691 0.2572 0.5907 0.1249 0.0265
##
## P value adjustment method: BH
```

Cal destacar que les diferències en els p-valors per a les comparacions dos a dos es deuen a que el t.test i les comparacions de Tukey utilitzen metodologies diferents per a obtenir el p-valor.

Referències

- [1] <https://www.rdocumentation.org/packages/ggplot2/versions/3.3.2>
- [2] <https://www.rdocumentation.org/packages/devtools/versions/2.3.2>
- [3] <https://github.com/kassambara/easyGgplot2>
- [4] <https://www.rdocumentation.org/packages/patchwork/versions/1.1.0>
- [5] <https://www.rdocumentation.org/packages/EnvStats/versions/2.4.0>
- [6] <https://www.rdocumentation.org/packages/EnvStats/versions/2.3.1/topics/quantileTest>
<https://www.jstor.org/stable/2532001?seq=1&cid=pdf-reference>
- [7] <https://www.rdocumentation.org/packages/xtable/versions/1.8-4>
- [8] <https://www.rdocumentation.org/packages/rcompanion/versions/2.3.26>
- [9] <https://www.rdocumentation.org/packages/ggpubr/versions/0.4.0>
- [10] <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/aov>
- [11] <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/TukeyHSD>
- [12] <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/pairwise.t.test>