

Analisis poblaciones de trucha en León

Santiago Fraga Sáenz

13/12/2023

Contents

1. ANÁLISIS Y OBSERVACIÓN PREVIA DE LOS DATOS PUBLICADOS POR LA JUNTA DE CASTILLA Y LEÓN	1
2. ANÁLISIS DE LA VARIABLE BIOMASA	2
2.1 VISUALIZACIÓN DE LOS DATOS	3
2.2 MEDIAS Y DESVIACIONES DE LA VARIABLE BIOMASA	4
2.3 MODELO PARA LA COMPARACIÓN DE LA BIOMASA	5
2. ANÁLISIS DE LA VARIABLE DENSIDAD	16
2.1 VISUALIZACIÓN DE LOS DATOS	16
2.2 MEDIAS Y DESVIACIONES DE LA VARIABLE DENSIDAD	18
2.3 MODELO PARA LA COMPARACIÓN DE LA DENSIDAD	19
3. ANÁLISIS DE LA VARIABLE PESO MEDIO	30
3.1 VISUALIZACIÓN DE LOS DATOS	30
3.2 MEDIAS Y DESVIACIONES DE LA VARIABLE PESO MEDIO	32
3.3 MODELO PARA LA COMPARACIÓN DEL PESO MEDIO	33
3. DISCUSIÓN	44

1. ANÁLISIS Y OBSERVACIÓN PREVIA DE LOS DATOS PUBLICADOS POR LA JUNTA DE CASTILLA Y LEÓN

Preparamos los datos. Antes de comenzar con el análisis mostramos la estructura de los datos una vez tabulados. Estos datos se han extraído de los informes publicados por la Junta de Castilla Y León correspondientes con las redes de vigilancia y seguimineto link. Se ha creado un data frame para la provincia de León que comprende las siguientes variables:

##	Estacion	Provincia	Gestion	Nivel	año	Biomasa
##	LE-AREC-3: 8	Leon:1064	AREC: 24	1:104	2014 :133	Min. : 0.01
##	LE-AREC-7: 8		CCM :120	2:960	2015 :133	1st Qu.: 2.90
##	LE001 : 8		CM : 96		2016 :133	Median : 5.59
##	LE002 : 8		CSM : 24		2017 :133	Mean : 7.33
##	LE003 : 8		EDS : 8		2018 :133	3rd Qu.:10.58
##	LE004 : 8		L :720		2019 :133	Max. :38.12
##	(Other) :1016		V : 72		(Other):266	NA's :671
##	Densidad	Peso_medio				
##	Min. :0.0000	Min. : 4.50				
##	1st Qu.:0.0800	1st Qu.:22.51				
##	Median :0.1500	Median :36.75				

```
## Mean      :0.2204    Mean      : Inf
## 3rd Qu.   :0.3200    3rd Qu.   :54.66
## Max.      :2.0400    Max.      : Inf
## NA's      :673       NA's      :673
```

Las variables Estación, Provincia, Gestión, Nivel y año son de tipo factor y el resto numéricas. Dentro de la gestión diferenciamos los siguientes niveles en función del tramo en el que se situó la estación de muestreo:

L= Tramo libre

CCM= Coto con muerte mixto

V= Vedado

CM= coto con muerte

AREC= Aguas en régimen especial controlado

ARE= Aguas en regimen especial

CSM= Coto sin muerte

Los datos de las variables métricas de las poblaciones de trucha son:

Biomasa; medida en gr/m2

Densidad; medido en indv/m2

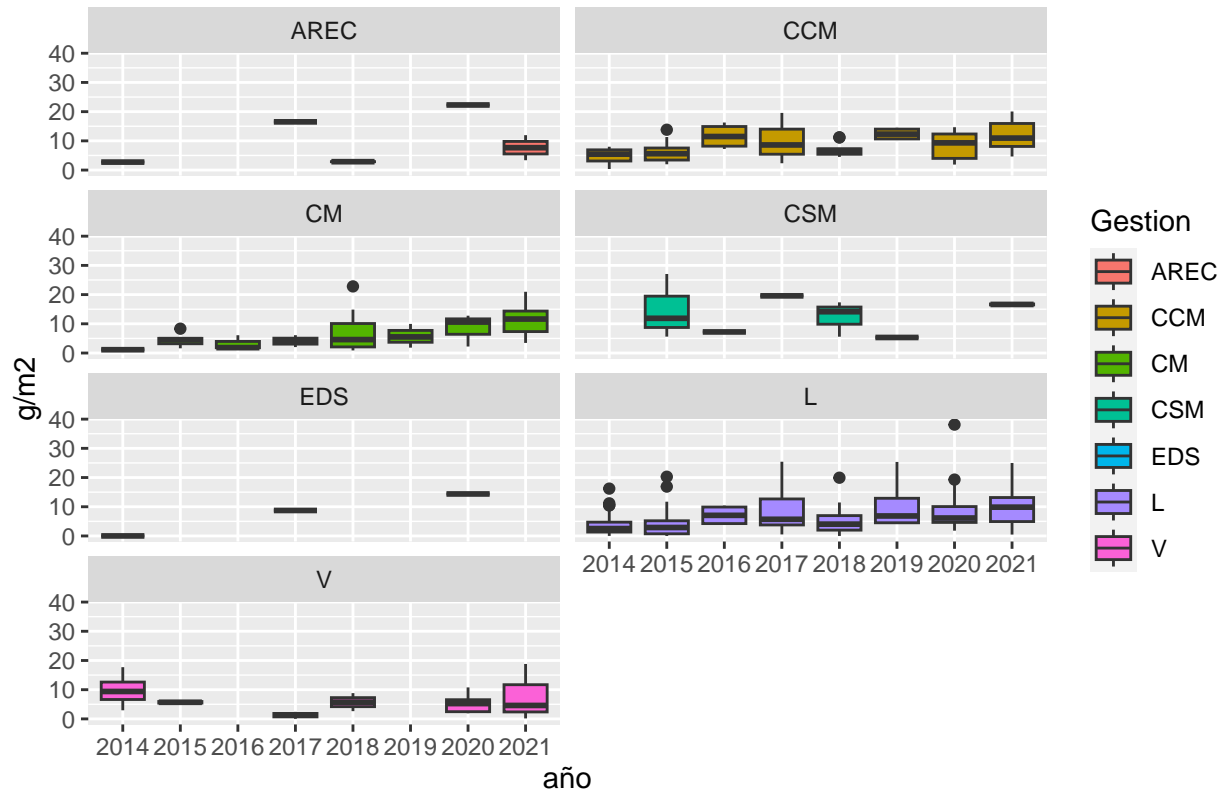
Peso medio: Resultado de dividir la biomasa entre la densidad, medida en gr/indv

2. ANÁLISIS DE LA VARIABLE BIOMASA

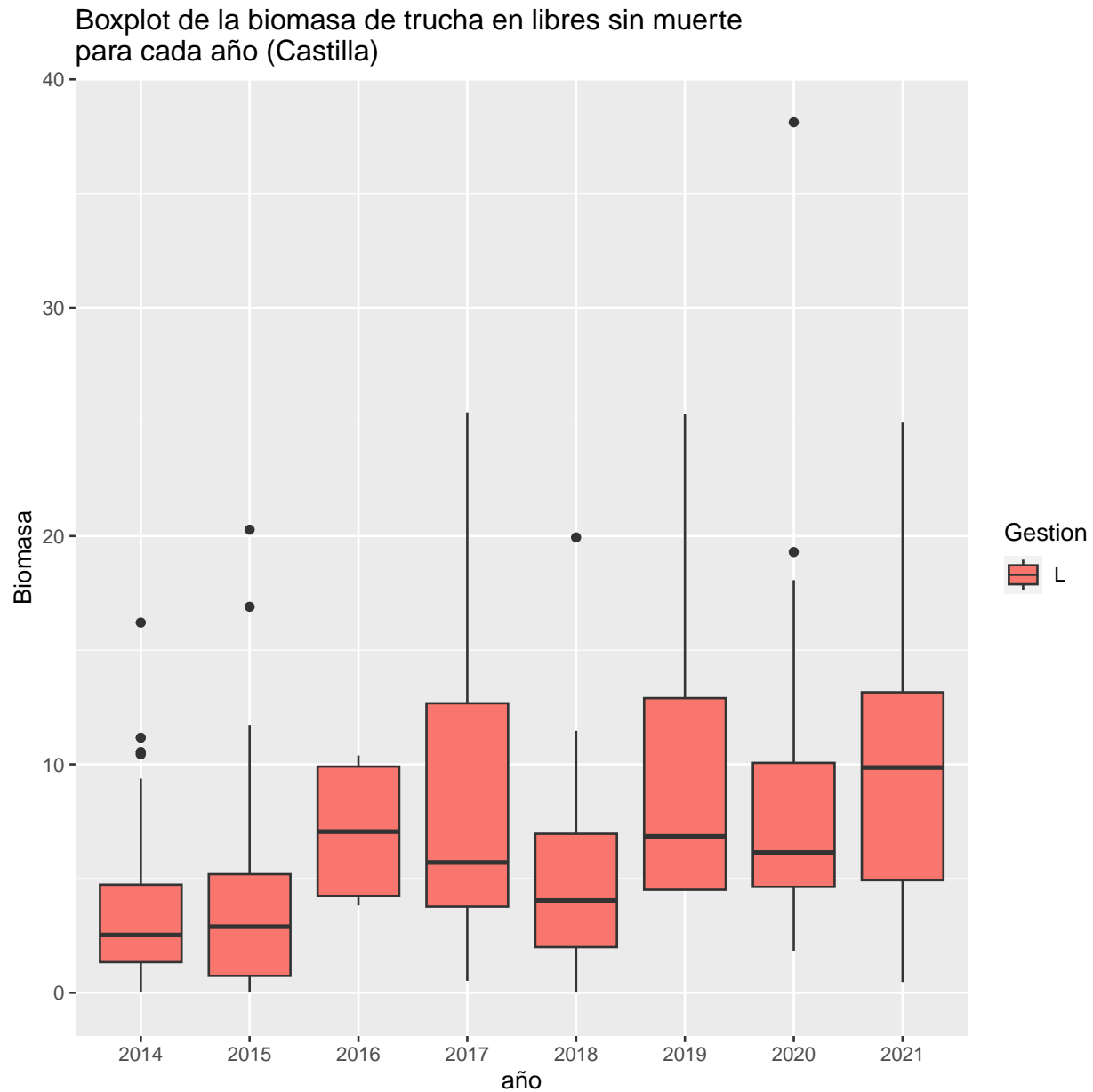
Para dar respuesta a la pregunta que se plantea en esta investigación que no es otra que ver si el cambio normativo introducido por la ley de pesca de 2013 ha supuesto cambios detectables en las variables poblacionales de trucha, tenemos que definir como, en función de los datos existentes, vamos a comprobarlo. Debido a que los datos publicados solo exponen la biomasa y la densidad como únicas variables obtenidas para cada estación, se propone definir un análisis de comparación entre grupos considerando el año 2014 como año de comparación ya que los efectos del cambio de gestión derivados de la nueva ley aún serían incipientes. De los modelos de gestión existentes el posible gran cambio se produciría en los tramos libres ya que pasarían a ser gestionados en el totalidad de las aguas trucheras como libres sin muerte. Observamos en primer lugar la distribución de las variables poblacionales, como variables respuesta, respecto a los grupos anuales y a los sistemas de gestión.

2.1 VISUALIZACIÓN DE LOS DATOS

Boxplot de la biomasa de trucha para cada año y tipo de gestión en León



Representamos las variables para los tramos libres, donde es previsible que se puedan observar con los efectos que se quieren comprobar, ya que en ellos la gestión ha pasado de poder sacrificarse las capturas que superasen la talla mínima sin un control de acceso del número de pescadores a ser tramos donde el sacrificio está prohibido, y todo este se produce por primera vez en el año 2014. Por ello este año es el año de referencia con el que se compararán el resto de años para observar si hay cambios significativos en las variables medidas.



Se observa la existencia de posibles outliers que para la comprobación de las comparaciones y realizar las comprobaciones de hipótesis deberá ser tenida en cuenta para elegir la técnica estadística. Determinamos a continuación los valores medios para cada año para los tramos de pesca libre, recogiendo igualmente las desviaciones estándar de cada grupo.

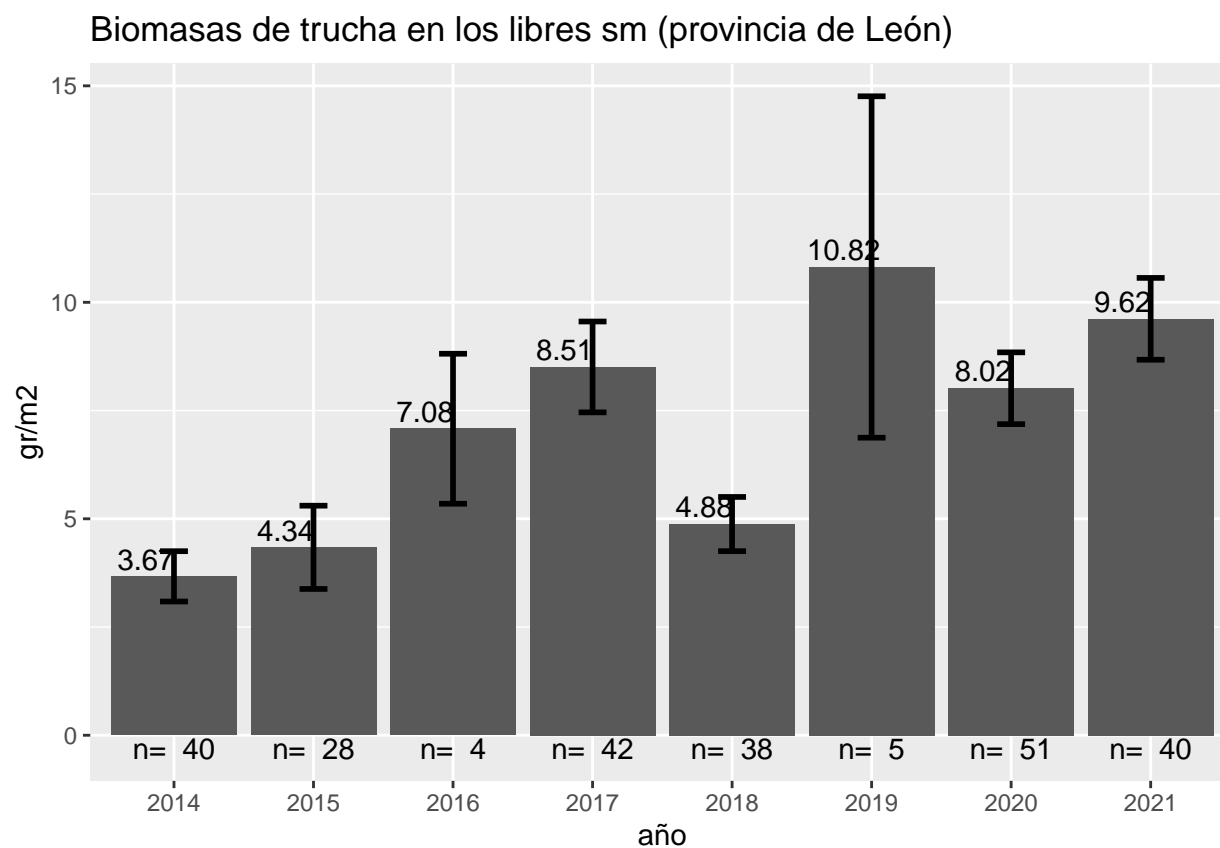
2.2 MEDIAS Y DESVIACIONES DE LA VARIABLE BIOMASA

Para los tramos libres de León, comprobamos los valores medios de biomasa de trucha (en gr/m²) en los diferentes años, así como los valores de la desviación estándar.

Table 1: Valores medios y desviaciones estandar

	2014	2015	2016	2017	2018	2019	2020	2021
Media	3.670250	4.340357	7.080000	8.507143	4.878421	10.816000	8.015294	9.618250
Desv	3.675736	5.088801	3.464265	6.800465	3.858286	8.816594	5.917669	5.978341

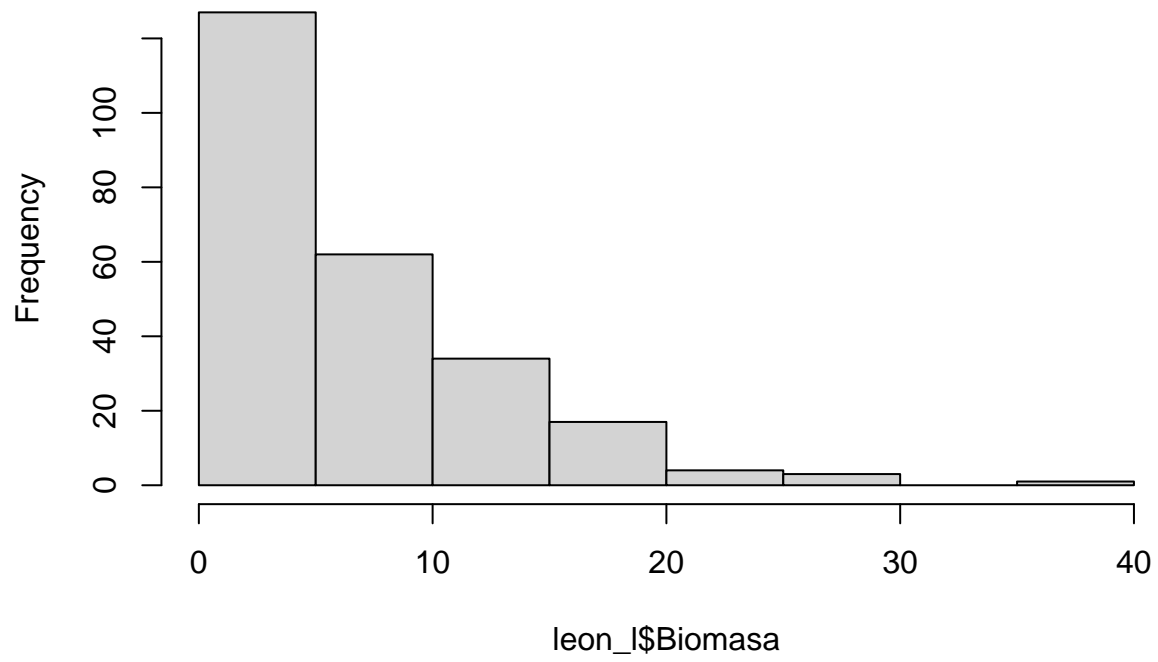
Se observa un aumento de los valores medios respecto a 2014. Representamos la distribución mediante histograma de la variable biomasa

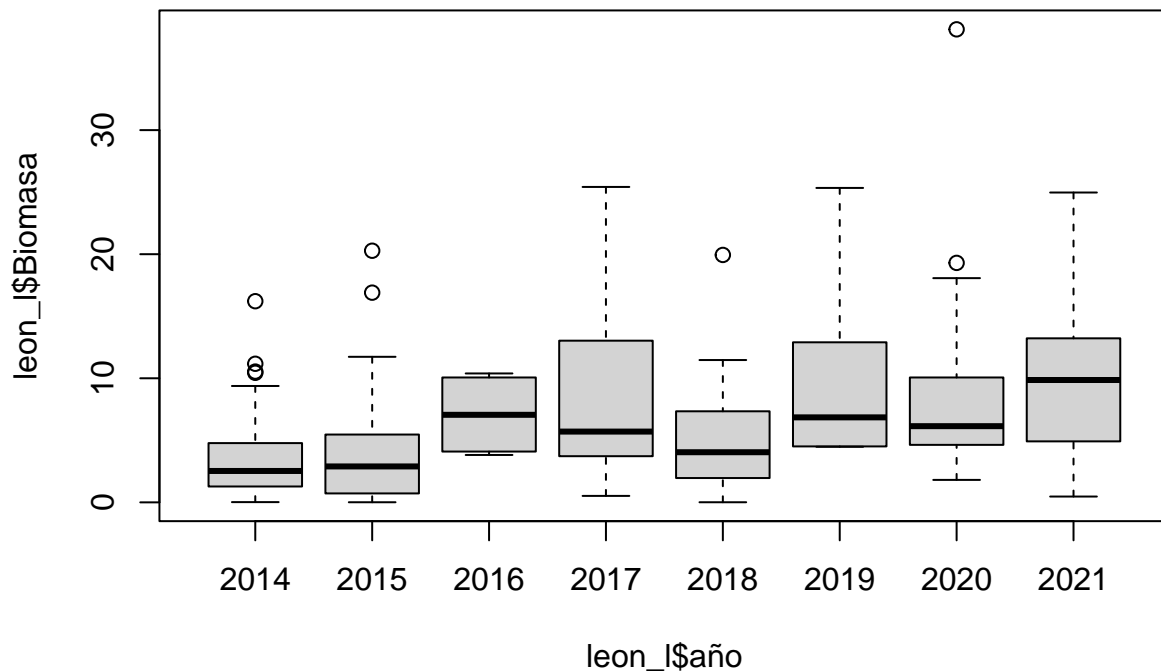


2.3 MODELO PARA LA COMPARACIÓN DE LA BIOMASA

Para la comparación de los grupos optamos en primer lugar por observar los datos y ver los supuestos para adaptarlo a un modelo lineal. Comprobamos la normalidad y la homocedasticidad. Primero los visualizamos

Histogram of leon_I\$Biomasa





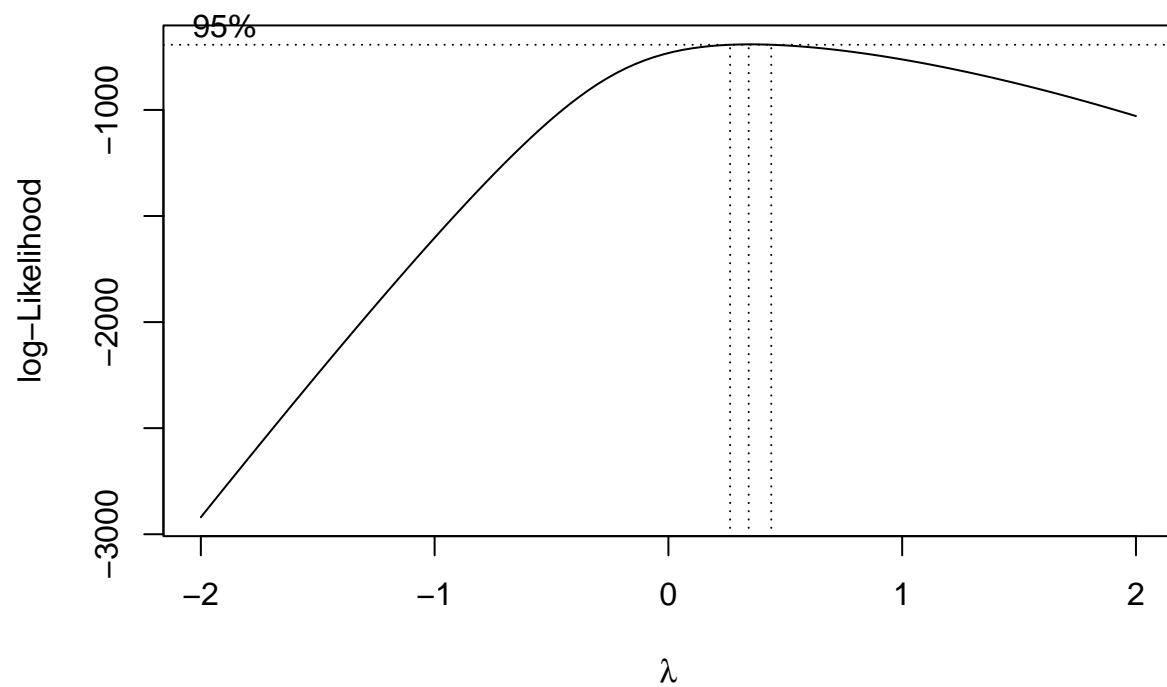
Comprobamos la normalidad y homocedasticidad aunque es visible que no parecen cumplirse ninguno de los 2 criterios

```
##
##  Shapiro-Wilk normality test
##
## data:  leon_tb_l$Biomasa
## W = 0.86414, p-value = 5.19e-14
```

No se puede aceptar la la distribución normal de los datos

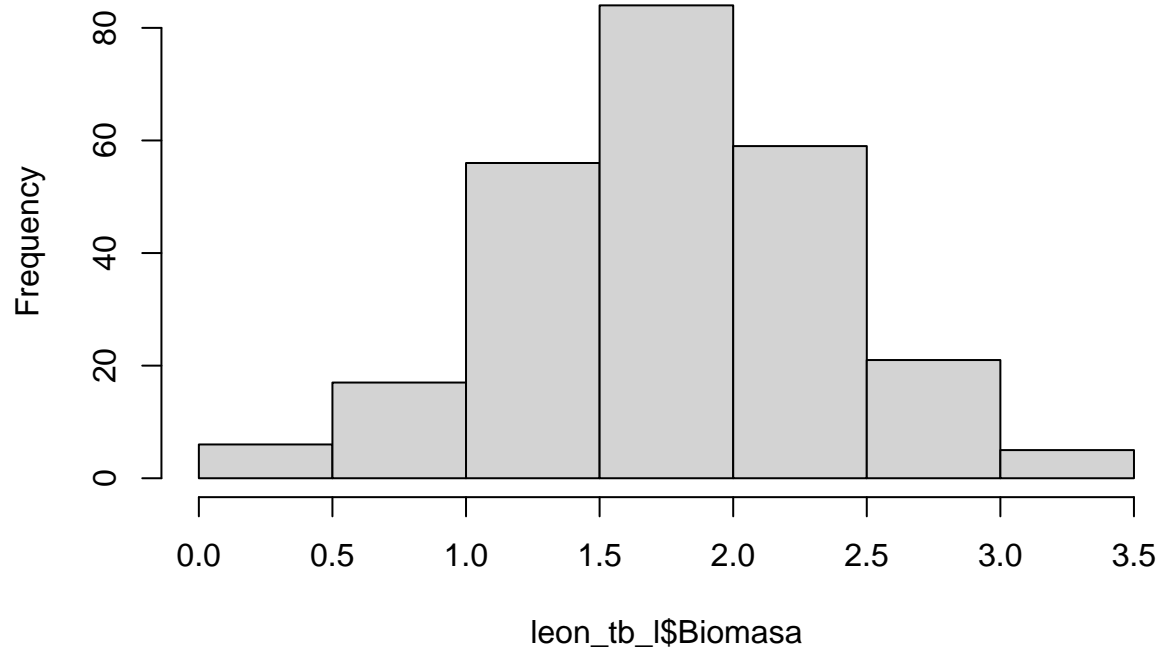
```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  7  2.4056 0.02137 *
##      240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tampoco se puede aceptar la homocedasticidad de los datos de biomasa para los diferentes años. Por ello es preciso transformar los datos y ver si con ello podemos asumir un modelo lineal. Para ello determinaremos los valores de λ mediante la aplicación de una transformación de Box Cox

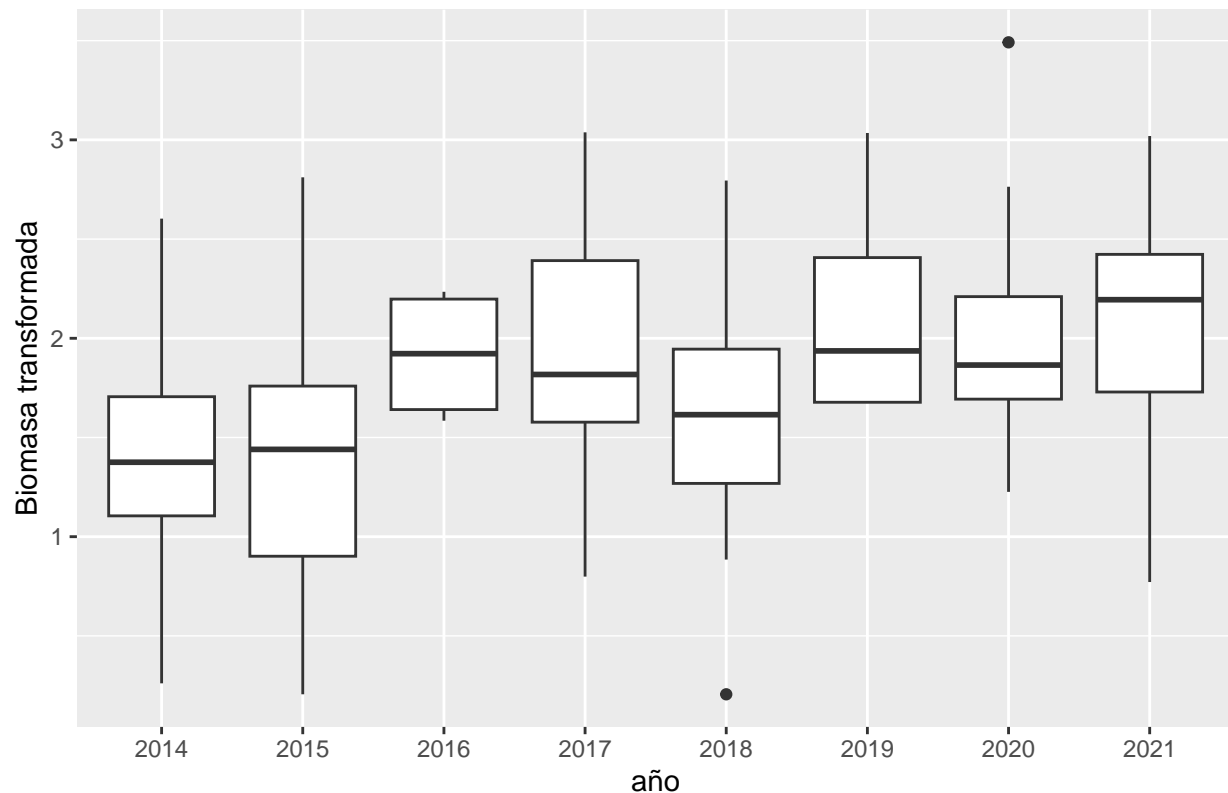


El valor de λ obtenido es 0.3434343 es el valor de la potencia usada en la transformación de los datos.
Una vez transformado los datos visualizamos nuevamente

Histogram of leon_tb_I\$Biomasa



Boxplot de las biomásas transformadas



Visualmente ahora parece que los datos pueden cumplir a los criterios de normalidad y homocedasticidad, pese a ello realizamos los test.

```
##
##  Shapiro-Wilk normality test
##
## data:  leon_tb_l$Biomasa
## W = 0.99583, p-value = 0.7467
```

Se acepta la la distribución normal de los datos

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  7  1.5997 0.1361
##      240
```

Se acepta la homocedasticidad

Planteamos ahora un modelo lineal mixto. Esto es conveniente en nuestro caso ya que podemos establecer una parte aleatoria del modelo con los datos anidados por la variable Estación. Como término fijo utilizaríamos la variable año, ya que se trata del factor principal que nos explica precisamente si existen diferencias en los valores de la biomasa considerando el año 2014 como el año que se produce el cambio de política de gestión al establecer los libres como libres sin muerte. Para la construcción del modelo seguimos el procedimiento que recomienda Zuur et al. (2009) y que se secuenciaría en fases:

1. Estructura aleatoria óptima. Usando un modelo saturado (beyond optimal model), se determina la estructura óptima del componente aleatorio, la cual no debe contener información que esté en la componente fija. Debemos:

- construir un modelo saturado.

- comparar modelos con distinta estructura aleatoria, mediante máxima verosimilitud restringida (REML).

2. Estructura fija óptima. Una vez encontramos la estructura aleatoria óptima, podemos encontrar la estructura fija óptima. Comparamos los modelos anidados mediante máxima verosimilitud (ML), manteniendo la misma estructura aleatoria.

3. Ajuste del modelo final con REML.

```
library(nlme)
####PROCEDIMIENTO DE ZUUR PARA LA SELECCION DEL MODELO
##1. Determinamos los efectos variables y para ello saturamos los efectos fijos
leon_tb_1<-leon_tb_1[!is.na(leon_tb_1$Biomasa),]
m1a<-gls(Biomasa~1+año,data=leon_tb_1,method="REML")
m1b<-lme(Biomasa~1+año,random=~1|Estacion,data=leon_tb_1,method="REML")
anova(m1a,m1b)#Efectivamente al año parece afectar a la VD, seleccionamos como variable aleatoria la es

##      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
## m1a      1   9 416.4606 447.7864 -199.2303
## m1b      2  10 347.5091 382.3155 -163.7545 1 vs 2 70.95153 <.0001

##2. Seleccionamos las variables de la estructura fija.
m1c<-lme(Biomasa~1,random=~1|Estacion,data=leon_tb_1,method="ML")
m1d<-lme(Biomasa~1+año,random=~1|Estacion,data=leon_tb_1,method="ML")
anova(m1c,m1d)#El modelo m1d es el elegido

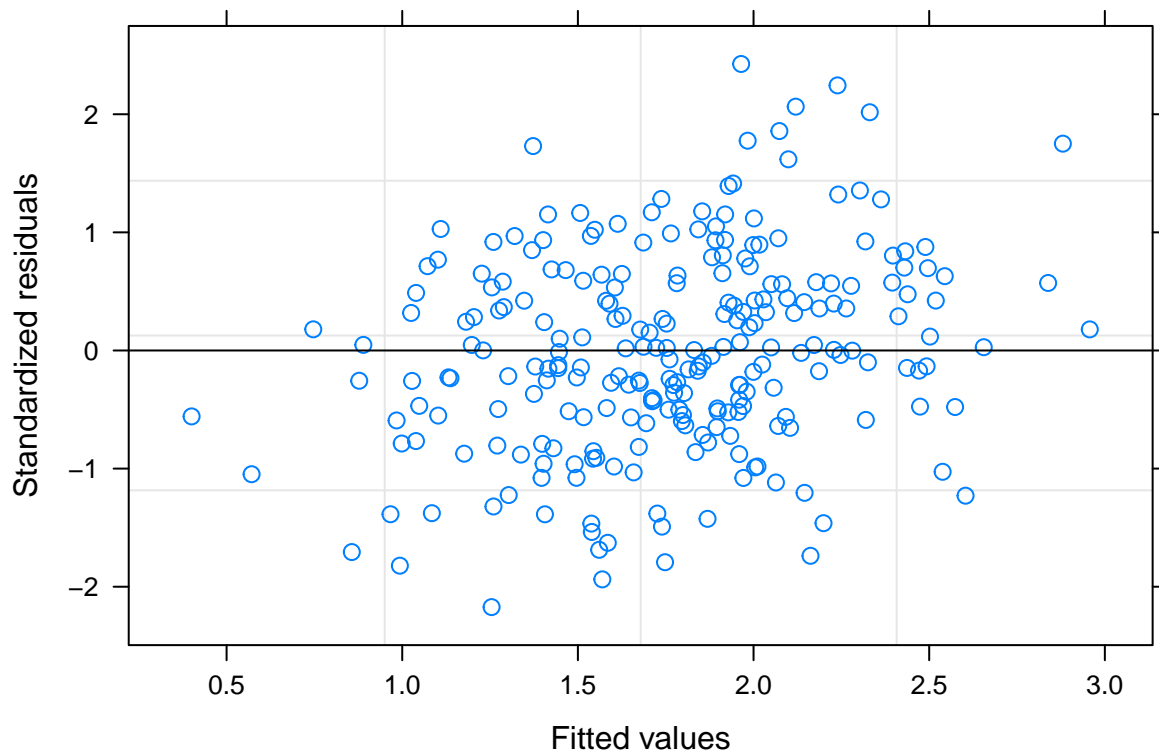
##      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
## m1c      1   3 406.8656 417.4058 -200.4328
## m1d      2  10 323.3767 358.5110 -151.6883 1 vs 2 97.4889 <.0001

##3. Ajuste del modelo final con el método REML
m1d_final<-lme(Biomasa~1+año,random=~1|Estacion,method="REML",data=leon_tb_1)
summary(m1d_final)#Diferencias entre 2017 y 2016 respecto de 2014

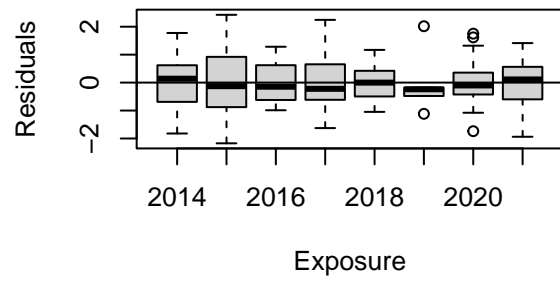
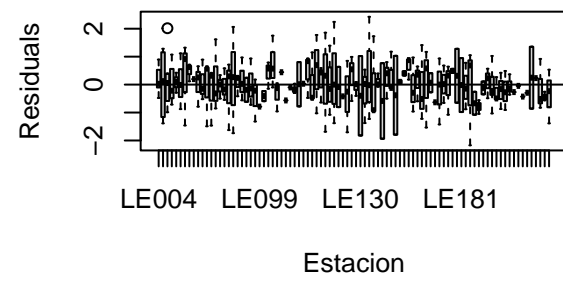
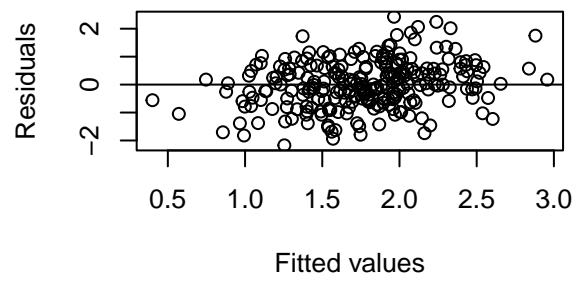
## Linear mixed-effects model fit by REML
## Data: leon_tb_1
##      AIC      BIC    logLik
## 347.5091 382.3155 -163.7545
##
## Random effects:
## Formula: ~1 | Estacion
##      (Intercept) Residual
## StdDev: 0.3922224 0.3489627
##
## Fixed effects: Biomasa ~ 1 + año
##      Value Std.Error DF t-value p-value
## (Intercept) 1.3458764 0.07720062 151 17.433492 0.0000
## año2015 0.1004071 0.11249997 151 0.892508 0.3735
## año2016 0.3313852 0.19989252 151 1.657817 0.0994
## año2017 0.5129883 0.07799500 151 6.577195 0.0000
## año2018 0.2711809 0.10599031 151 2.558545 0.0115
## año2019 0.6574566 0.18210715 151 3.610273 0.0004
## año2020 0.5550115 0.07737549 151 7.172963 0.0000
## año2021 0.7391176 0.10509690 151 7.032725 0.0000
## Correlation:
##      (Intr) año2015 año2016 año2017 año2018 año2019 año2020
## año2015 -0.613
## año2016 -0.283 0.269
```

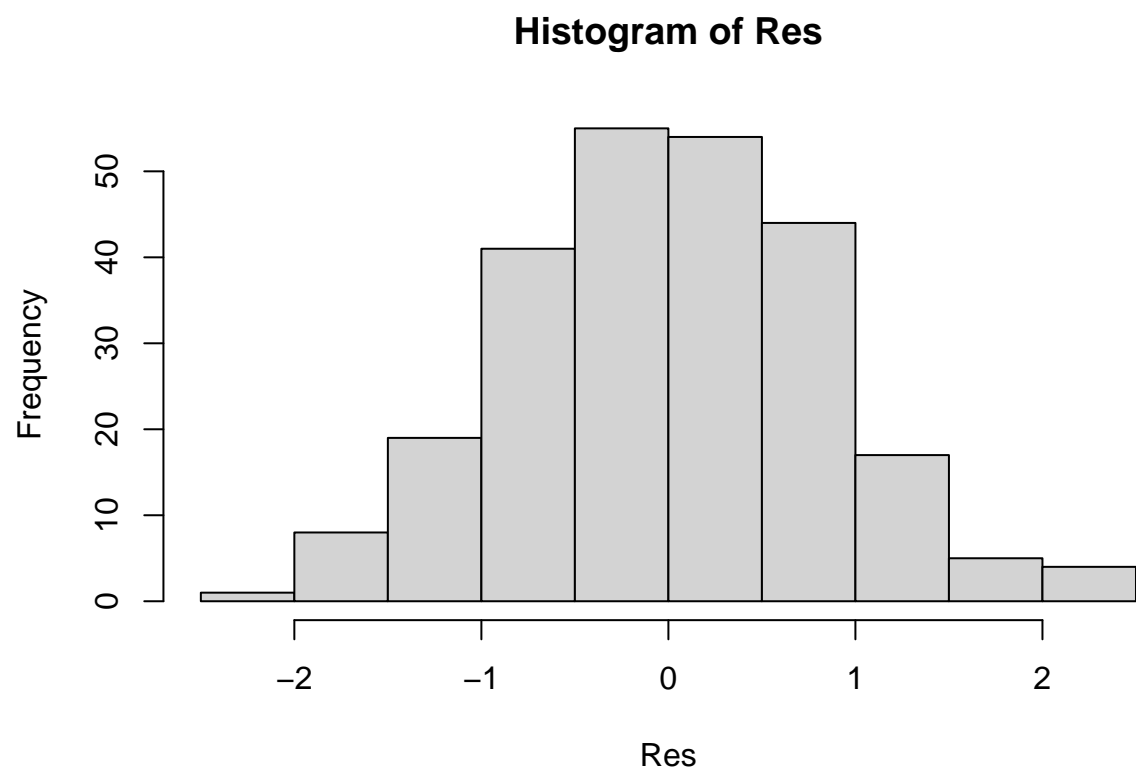
```
## año2017 -0.531  0.379  0.225
## año2018 -0.654  0.665  0.286  0.403
## año2019 -0.315  0.302  0.239  0.250  0.319
## año2020 -0.588  0.411  0.235  0.525  0.436  0.262
## año2021 -0.661  0.670  0.287  0.406  0.716  0.320  0.440
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.1725640501 -0.5476011726 -0.0002408912  0.5722837173  2.4257130817
##
## Number of Observations: 248
## Number of Groups: 90
```

El modelo final lo utilizamos para contrastar los diferentes niveles de la variable predictora, en este caso los diferentes años mediante comparaciones múltiples mediante el test de Tukey. Comprobamos el modelo de forma gráfica

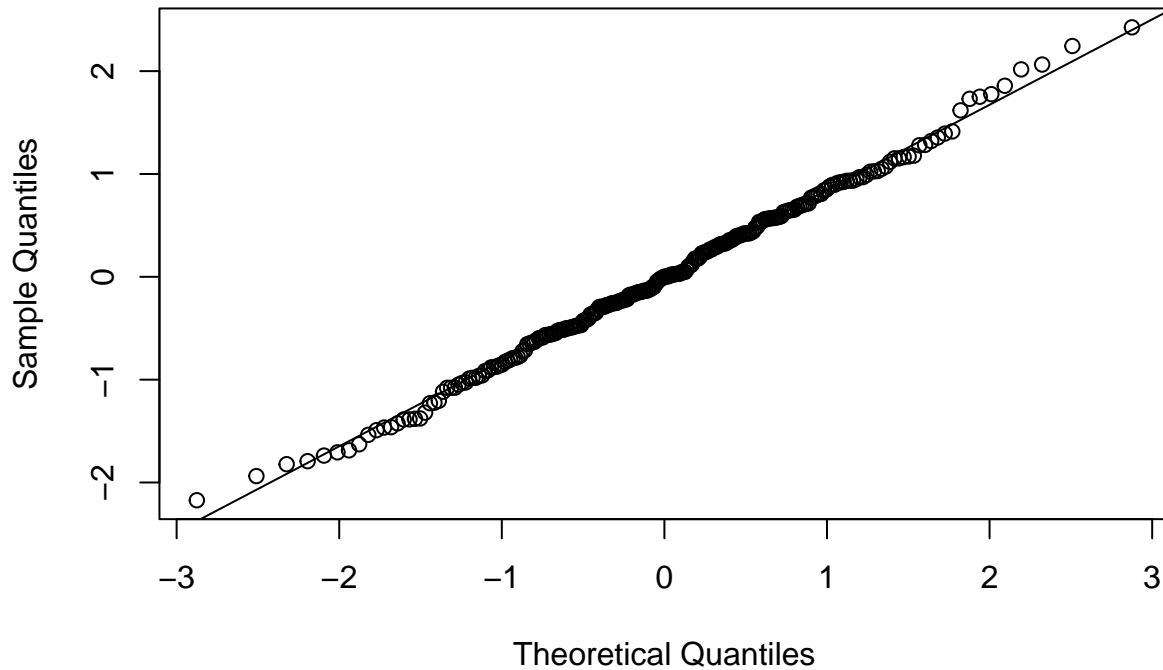


Residuals vs. fitted





Normal Q-Q Plot



```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme.formula(fixed = Biomasa ~ 1 + año, data = leon_tb_1, random = ~1 |
## Estacion, method = "REML")
##
## Linear Hypotheses:
```

	Estimate	Std. Error	z value	Pr(> z)
2015 - 2014 == 0	0.10041	0.11250	0.893	0.98401
2016 - 2014 == 0	0.33139	0.19989	1.658	0.68240
2017 - 2014 == 0	0.51299	0.07799	6.577	< 0.001 ***
2018 - 2014 == 0	0.27118	0.10599	2.559	0.14957
2019 - 2014 == 0	0.65746	0.18211	3.610	0.00609 **
2020 - 2014 == 0	0.55501	0.07738	7.173	< 0.001 ***
2021 - 2014 == 0	0.73912	0.10510	7.033	< 0.001 ***
2016 - 2015 == 0	0.23098	0.20123	1.148	0.93647
2017 - 2015 == 0	0.41258	0.10992	3.753	0.00356 **
2018 - 2015 == 0	0.17077	0.08960	1.906	0.50963
2019 - 2015 == 0	0.55705	0.18283	3.047	0.04040 *
2020 - 2015 == 0	0.45460	0.10721	4.240	< 0.001 ***
2021 - 2015 == 0	0.63871	0.08860	7.209	< 0.001 ***
2017 - 2016 == 0	0.18160	0.19758	0.919	0.98107
2018 - 2016 == 0	-0.06020	0.19766	-0.305	0.99999

```
## 2019 - 2016 == 0 0.32607 0.23599 1.382 0.84594
## 2020 - 2016 == 0 0.22363 0.19663 1.137 0.93935
## 2021 - 2016 == 0 0.40773 0.19735 2.066 0.40051
## 2018 - 2017 == 0 -0.24181 0.10322 -2.343 0.24077
## 2019 - 2017 == 0 0.14447 0.17931 0.806 0.99128
## 2020 - 2017 == 0 0.04202 0.07575 0.555 0.99917
## 2021 - 2017 == 0 0.22613 0.10233 2.210 0.31231
## 2019 - 2018 == 0 0.38628 0.17912 2.157 0.34367
## 2020 - 2018 == 0 0.28383 0.10031 2.829 0.07498 .
## 2021 - 2018 == 0 0.46794 0.07951 5.885 < 0.001 ***
## 2020 - 2019 == 0 -0.10245 0.17825 -0.575 0.99895
## 2021 - 2019 == 0 0.08166 0.17877 0.457 0.99977
## 2021 - 2020 == 0 0.18411 0.09938 1.853 0.54705
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Podemos ver en la tabla de contrastes que para los datos de biomasa existen diferencias significativas de los datos del año 2014 respecto a los de 2017 en adelante, si exceptuamos el año 2018, año que parece producirse una reducción en la biomasa.

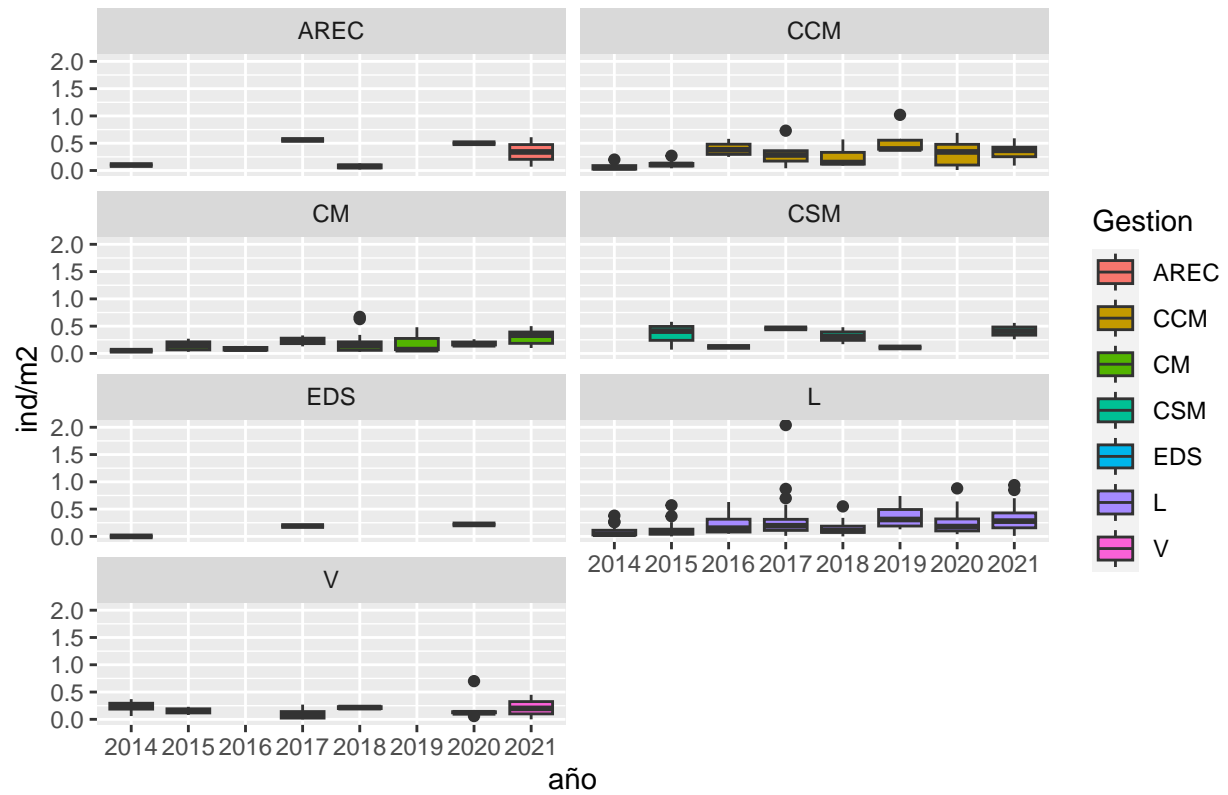
2. ANÁLISIS DE LA VARIABLE DENSIDAD

Analizaremos la siguiente variable poblacional que corresponde con densidad de trucha en ind/m². Como en el caso de la biomasa, interesa conocer la existencia de diferencias significativas derivadas de los diferentes inventarios para ver si el cambio de políticas de gestión que se produce en el año 2013 y que tiene como primer año de implementación el 2014, ha supuesto un cambio en las variables de las poblaciones de trucha. Como en el caso de la biomasa, se escogieron los libres por ser los que más se notaría las medidas implementadas al declararse como de pesca sin muerte. Observamos en primer lugar la distribución de las variables poblacionales, como variables respuesta, respecto a los grupos anuales y a los sistemas de gestión.

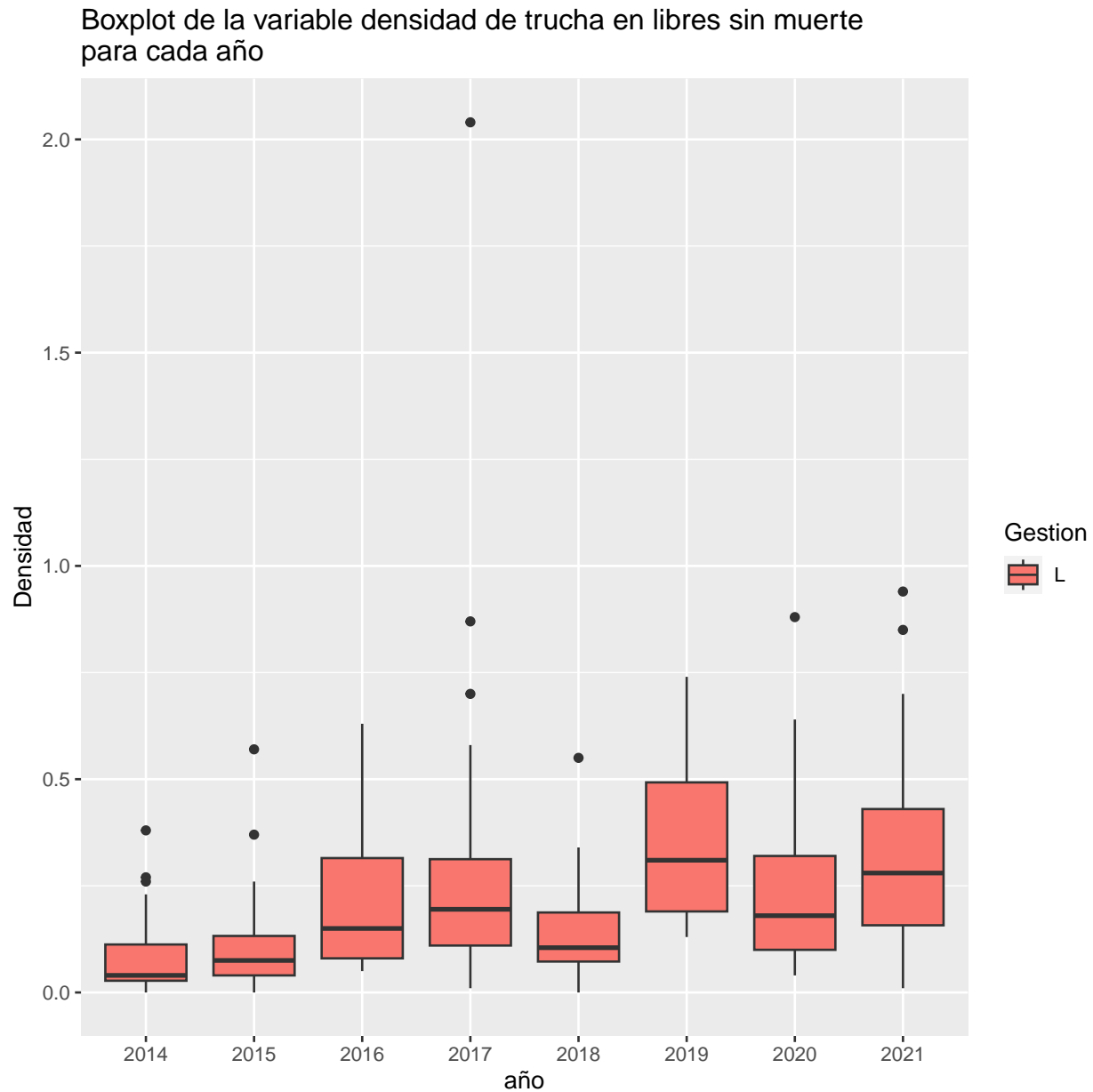
2.1 VISUALIZACIÓN DE LOS DATOS

```
## Estacion Provincia Gestion Nivel año Biomasa
## LE004 : 8 Leon:720 L:720 1: 40 2014 : 90 Min. : 0.010
## LE009 : 8 2:680 2015 : 90 1st Qu.: 2.627
## LE010 : 8 2016 : 90 Median : 4.925
## LE070 : 8 2017 : 90 Mean : 6.802
## LE072 : 8 2018 : 90 3rd Qu.: 9.750
## LE073 : 8 2019 : 90 Max. :38.120
## (Other):672 (Other):180 NA's :472
## Densidad Peso_medio
## Min. :0.0000 Min. : 4.50
## 1st Qu.:0.0600 1st Qu.:21.82
## Median :0.1400 Median :37.00
## Mean :0.2054 Mean : Inf
## 3rd Qu.:0.2650 3rd Qu.:54.82
## Max. :2.0400 Max. : Inf
## NA's :473 NA's :473
```


Boxplot de la densidad de trucha para cada año y tipo de gestión en León



Representamos las variables para los tramos libres, donde es previsible que se puedan observar con los efectos que se quieren comprobar.



Se observa la existencia de posibles outliers que para la comprobación de las comparaciones y realizar las comprobaciones de hipótesis deberá ser tenida en cuenta para elegir la técnica estadística. Determinamos a continuación los valores medios para cada año para los tramos de pesca libre, recogiendo igualmente las desviaciones estándar de cada grupo.

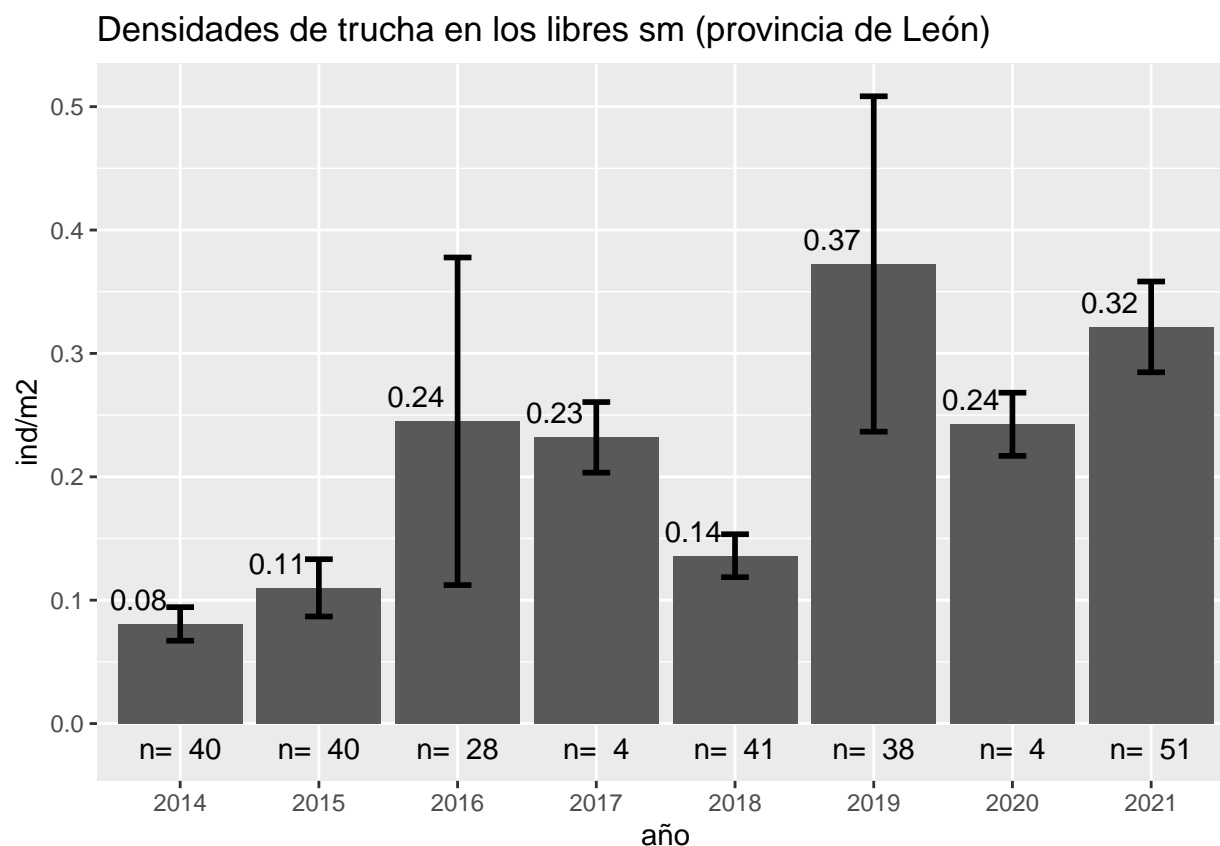
2.2 MEDIAS Y DESVIACIONES DE LA VARIABLE DENSIDAD

Para los tramos libres de León, comprobamos los valores medios de biomasa de trucha (en ind/m²) en los diferentes años, así como los valores de la desviación estándar.

Table 2: Valores medios y desviaciones estandar

	2014	2015	2016	2017	2018	2019	2020	2021
Media	0.0807500	0.1100000	0.2450000	0.2319512	0.1360526	0.3725000	0.2425490	0.3215000
Desv	0.0860944	0.1228971	0.2655184	0.1832378	0.1073146	0.2718302	0.1828753	0.2326828

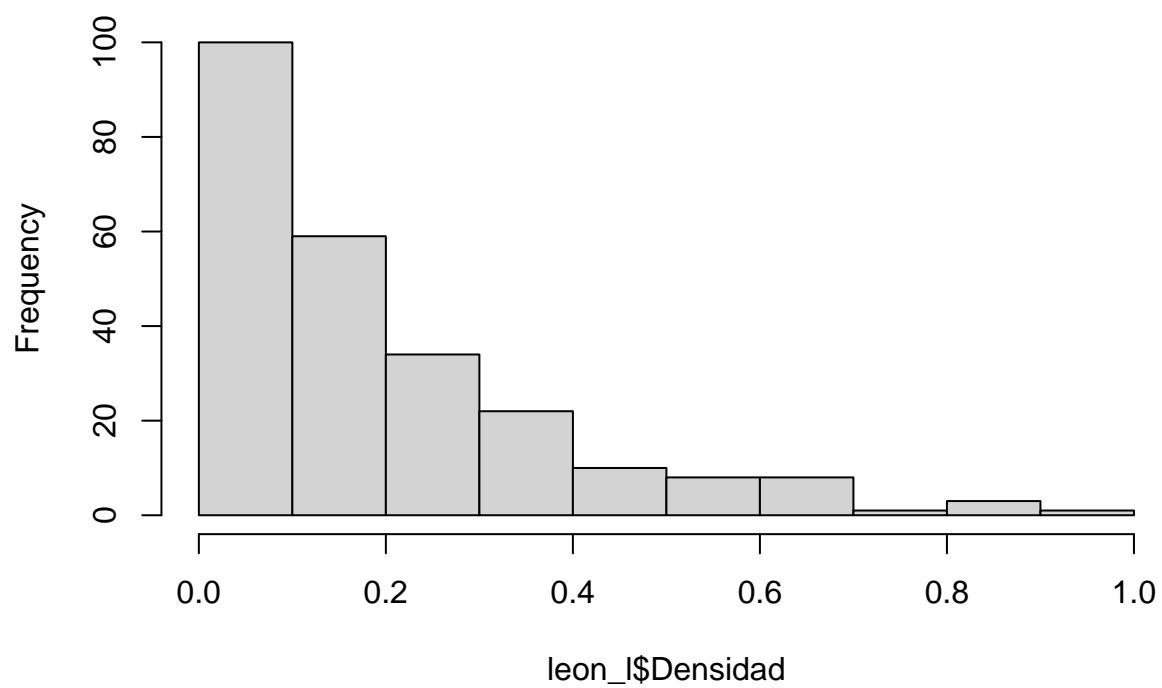
Se observa un aumento de los valores medios respecto a 2014. Representamos la distribución mediante histograma de la variable densidad

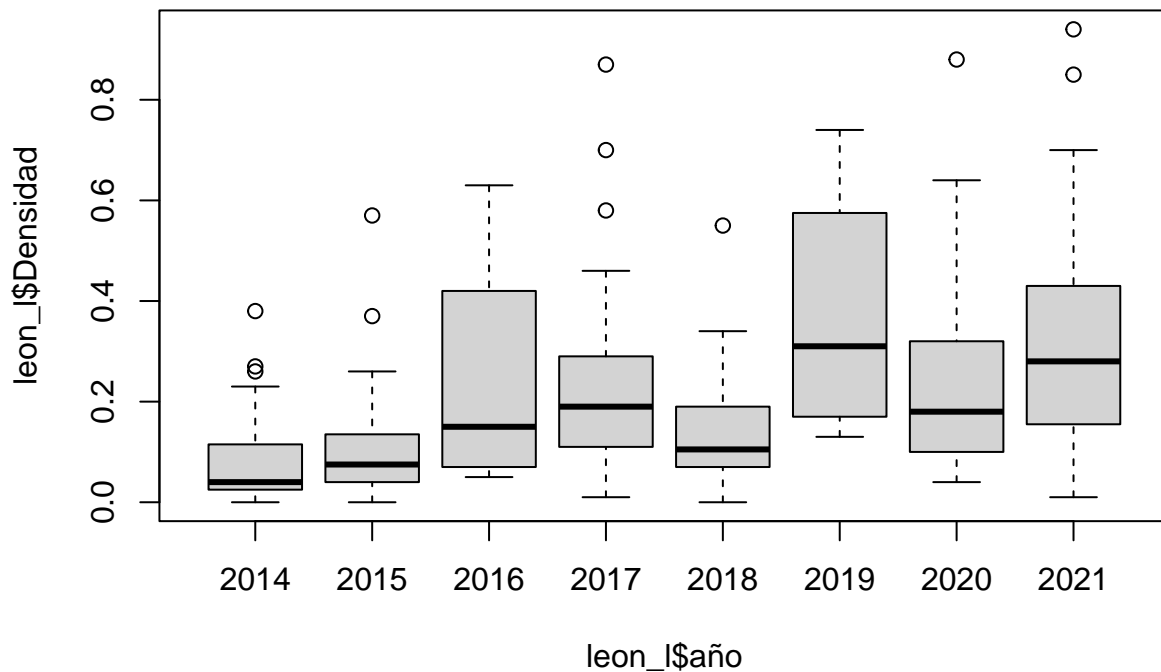


2.3 MODELO PARA LA COMPARACIÓN DE LA DENSIDAD

Para la comparación de los grupos optamos en primer lugar por observar los datos y ver los supuestos para adaptarlo a un modelo lineal. Comprobamos la normalidad y la homocedasticidad. Primero los visualizamos

Histogram of leon_I\$Densidad





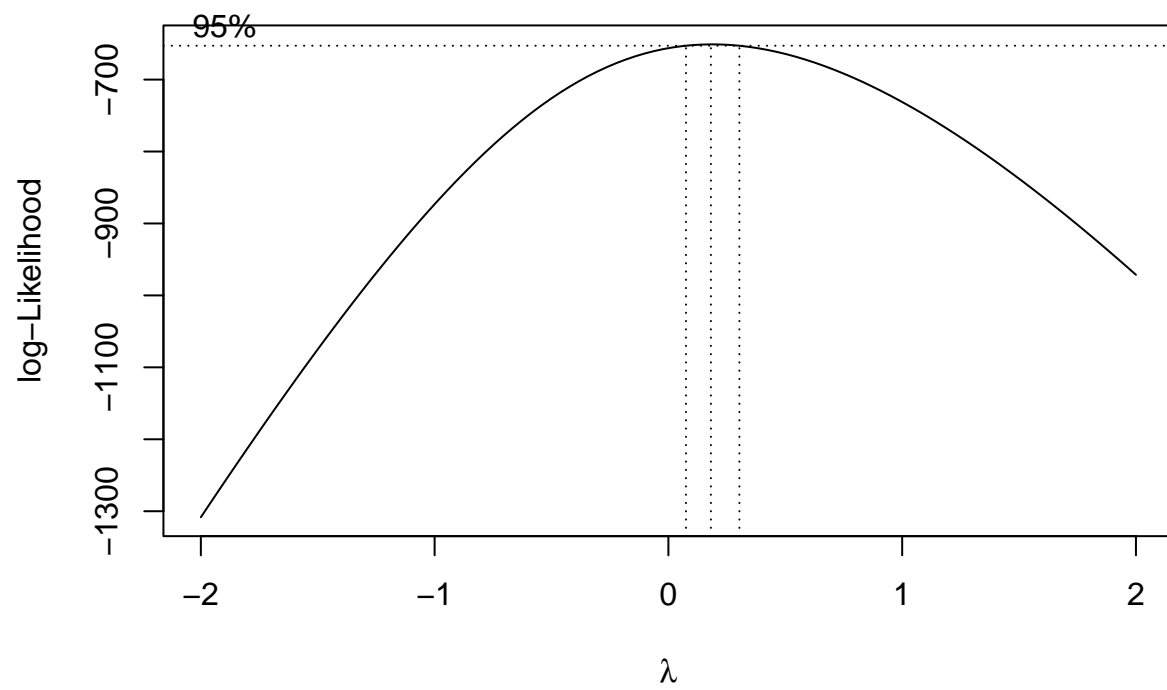
Comprobamos la normalidad y homocedasticidad aunque es visible que no parecen cumplirse ninguno de los 2 criterios

```
##
##  Shapiro-Wilk normality test
##
## data:  leon_l$Densidad
## W = 0.83854, p-value = 2.69e-15
```

No se puede aceptar la la distribución normal de los datos

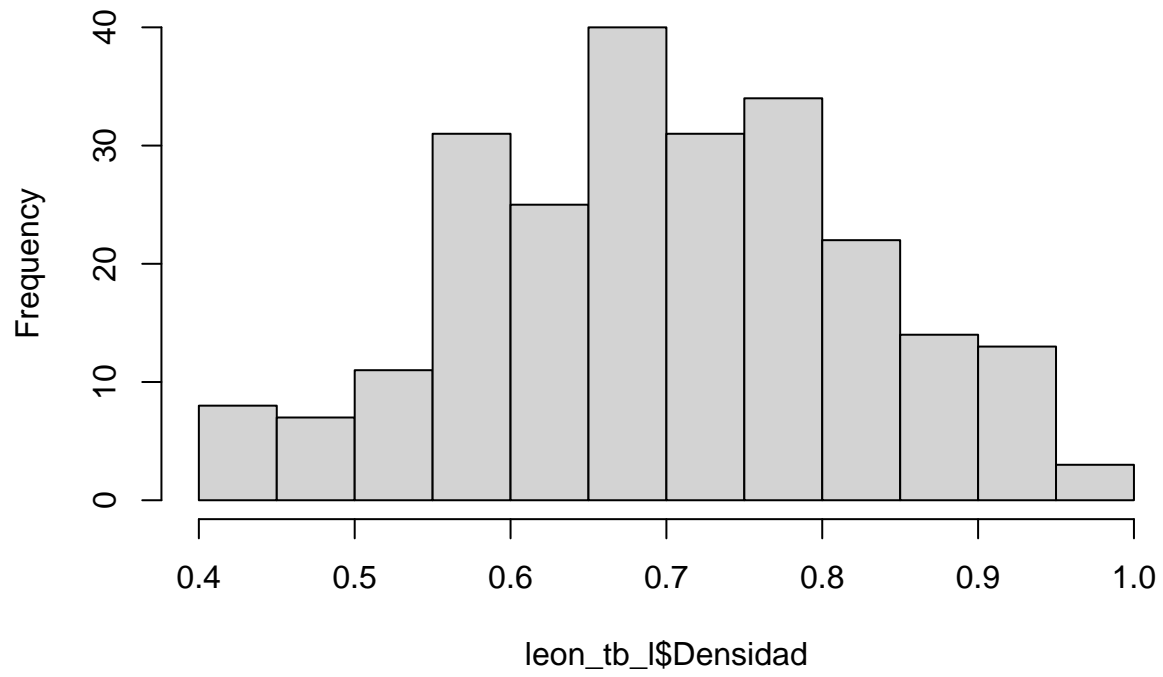
```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value    Pr(>F)
## group  7  5.2825 1.259e-05 ***
##      238
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

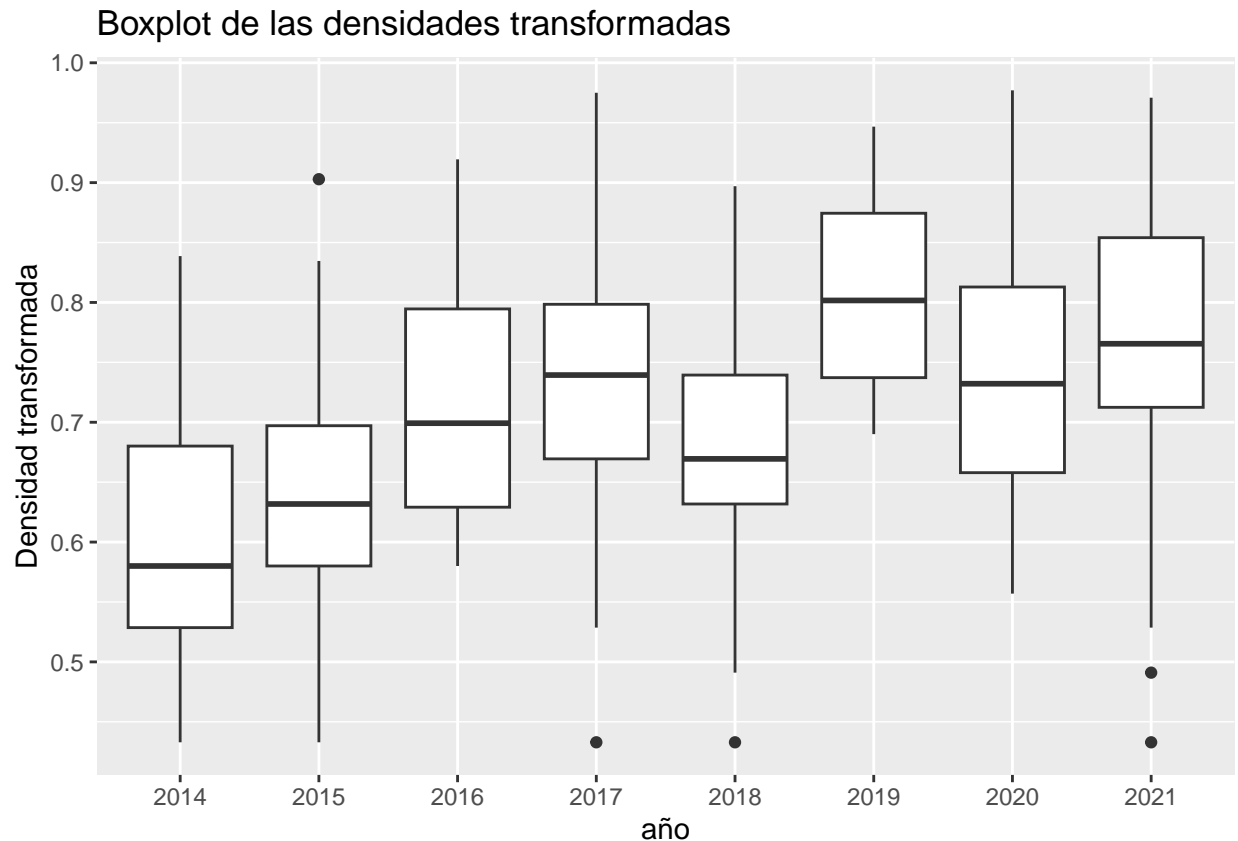
Tampoco se puede aceptar la homocedasticidad de los datos de densidad para los diferentes años. Por ello es preciso transformar los datos y ver si con ello podemos asumir un modelo lineal. Para ello determinaremos los valores de λ mediante la aplicación de una transformación de Box Cox



El valor de λ obtenido es 0.1818182 es el valor de la potencia usada en la transformación de los datos.
Una vez transformado los datos visualizamos nuevamente

Histogram of leon_tb_I\$Densidad





Visualmente ahora parece que los datos pueden cumplir a los criterios de normalidad y homocedasticidad, pese a ello realizamos los test.

```
##
##  Shapiro-Wilk normality test
##
## data:  leon_tb_1$Densidad
## W = 0.98914, p-value = 0.06927
```

Se acepta la la distribución normal de los datos

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  7  0.6972 0.6744
##      231
```

Se acepta la homocedasticidad

Planteamos ahora un modelo lineal mixto. Esto es conveniente en nuestro caso ya que podemos establecer una parte aleatoria del modelo con los datos anidados por la variable Estación. Como término fijo utilizaríamos la variable año, ya que se trata del factor principal que nos explica precisamente si existen diferencias en los valores de la densidad considerando el año 2014 como el año que se produce el cambio de política de gestión al establecer los libres como libres sin muerte. Para la construcción del modelo seguimos el procedimiento que recomienda Zuur et al. (2009) y que se secuenciaría en fases:

1. Estructura aleatoria óptima. Usando un modelo saturado (beyond optimal model), se determina la estructura óptima del componente aleatorio, la cual no debe contener información que esté en la componente fija. Debemos:

- construir un modelo saturado.

- comparar modelos con distinta estructura aleatoria, mediante máxima verosimilitud restringida (REML).

2. Estructura fija óptima. Una vez encontramos la estructura aleatoria óptima, podemos encontrar la estructura fija óptima. Comparamos los modelos anidados mediante máxima verosimilitud (ML), manteniendo la misma estructura aleatoria.

3. Ajuste del modelo final con REML.

```
library(nlme)
####PROCEDIMIENTO DE ZUUR PARA LA SELECCION DEL MODELO
##1. Determinamos los efectos variables y para ello saturamos los efectos fijos
leon_tb_1<-leon_tb_1[!is.na(leon_tb_1$Densidad),]
m1a<-glS(Densidad~1+año,data=leon_tb_1,method="REML")
m1b<-lme(Densidad~1+año,random=~1|Estacion,data=leon_tb_1,method="REML")
anova(m1a,m1b)#Efectivamente al año parece afectar a la VD, seleccionamos como variable aleatoria la es

##      Model df      AIC      BIC   logLik   Test  L.Ratio p-value
## m1a      1  9 -321.1165 -290.1348 169.5583
## m1b      2 10 -362.0582 -327.6340 191.0291 1 vs 2 42.94169 <.0001

##2. Seleccionamos las variables de la estructura fija.
m1c<-lme(Densidad~1,random=~1|Estacion,data=leon_tb_1,method="ML")
m1d<-lme(Densidad~1+año,random=~1|Estacion,data=leon_tb_1,method="ML")
anova(m1c,m1d)#El modelo m1d es el elegido

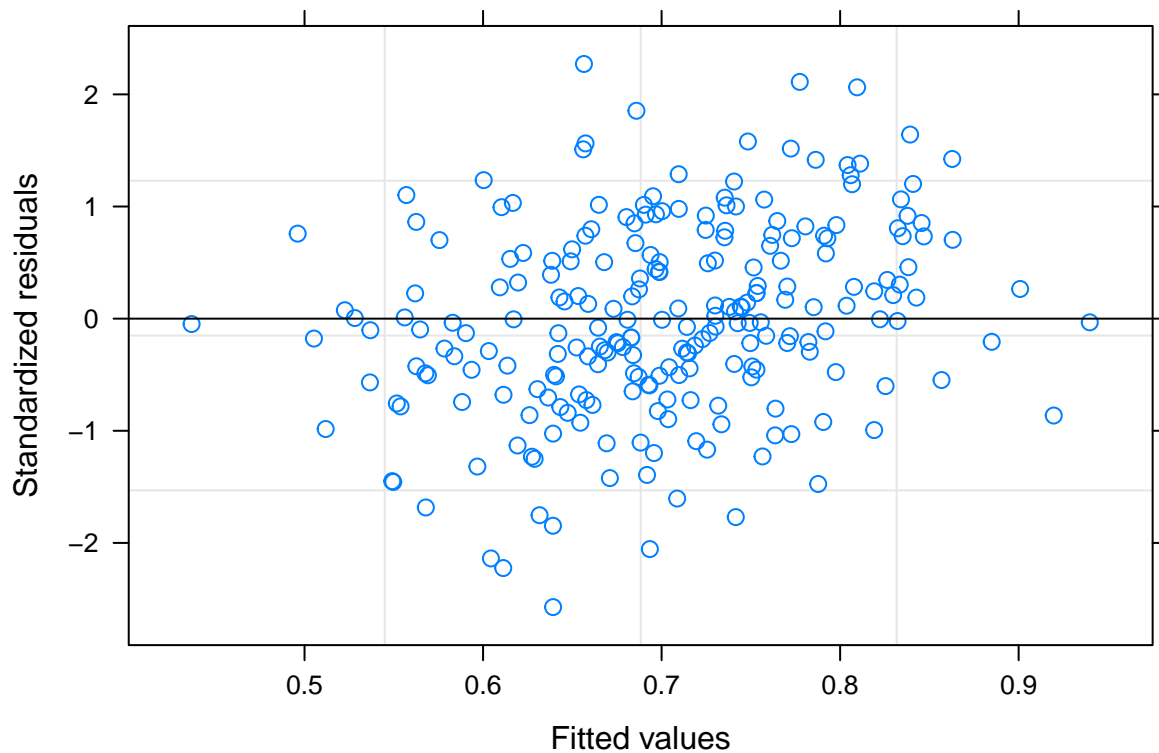
##      Model df      AIC      BIC   logLik   Test  L.Ratio p-value
## m1c      1  3 -328.2504 -317.8210 167.1252
## m1d      2 10 -409.7011 -374.9365 214.8506 1 vs 2 95.45076 <.0001

##3. Ajuste del modelo final con el método REML
m1d_final<-lme(Densidad~1+año,random=~1|Estacion,method="REML",data=leon_tb_1)
summary(m1d_final)#Diferencias entre 2017 y 2016 respecto de 2014

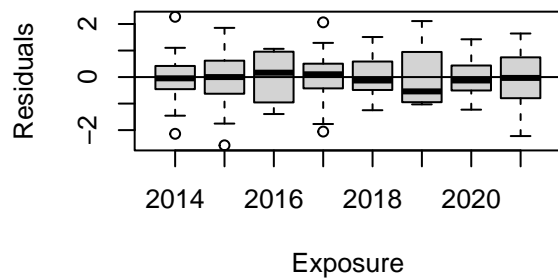
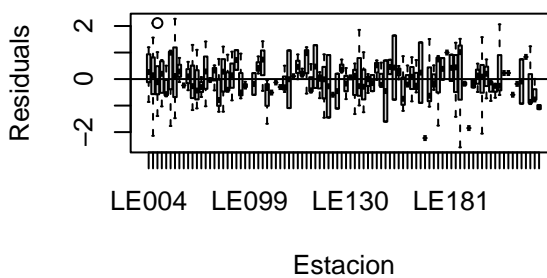
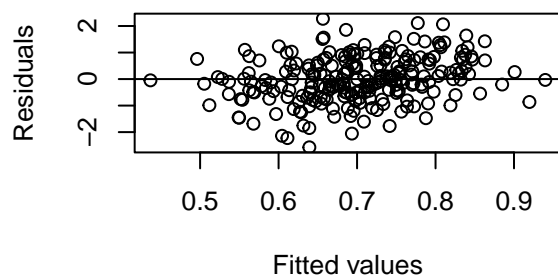
## Linear mixed-effects model fit by REML
## Data: leon_tb_1
##      AIC      BIC   logLik
## -362.0582 -327.634 191.0291
##
## Random effects:
## Formula: ~1 | Estacion
##      (Intercept)  Residual
## StdDev:  0.07692056 0.08021259
##
## Fixed effects:  Densidad ~ 1 + año
##      Value Std.Error DF  t-value p-value
## (Intercept) 0.5892415 0.01722955 143 34.19947  0.0000
## año2015      0.0529923 0.02553027 143  2.07567  0.0397
## año2016      0.0804010 0.04584639 143  1.75370  0.0816
## año2017      0.1312110 0.01836909 143  7.14303  0.0000
## año2018      0.0786468 0.02369991 143  3.31844  0.0011
## año2019      0.1659635 0.04584639 143  3.61999  0.0004
## año2020      0.1471381 0.01811142 143  8.12405  0.0000
## año2021      0.1861422 0.02347420 143  7.92965  0.0000
## Correlation:
##      (Intr) año2015 año2016 año2017 año2018 año2019 año2020
## año2015 -0.622
## año2016 -0.294  0.261
```

```
## año2017 -0.575  0.400  0.232
## año2018 -0.675  0.631  0.280  0.431
## año2019 -0.294  0.261  0.235  0.232  0.280
## año2020 -0.637  0.434  0.244  0.541  0.469  0.244
## año2021 -0.683  0.637  0.281  0.435  0.688  0.281  0.473
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.5724777 -0.5170112 -0.0324419  0.6336792  2.2712043
##
## Number of Observations: 239
## Number of Groups: 89
```

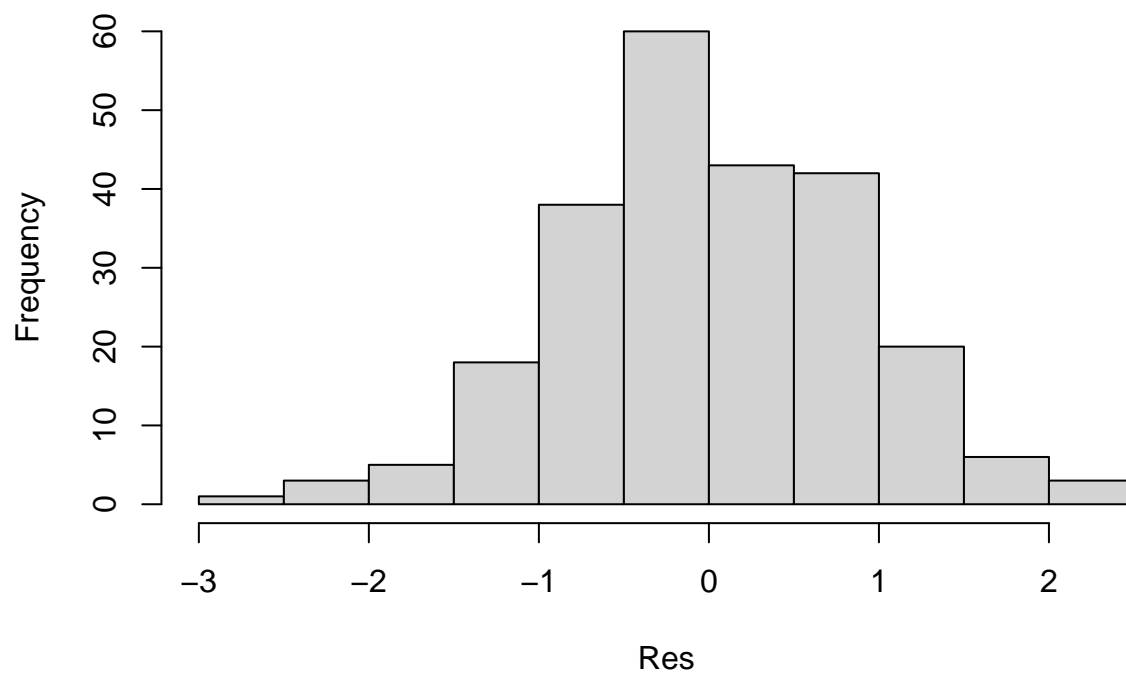
El modelo final lo utilizamos para contrastar los diferentes niveles de la variable predictora, en este caso los diferentes años mediante comparaciones múltiples mediante el test de Tukey. Comprobamos el modelo de forma gráfica



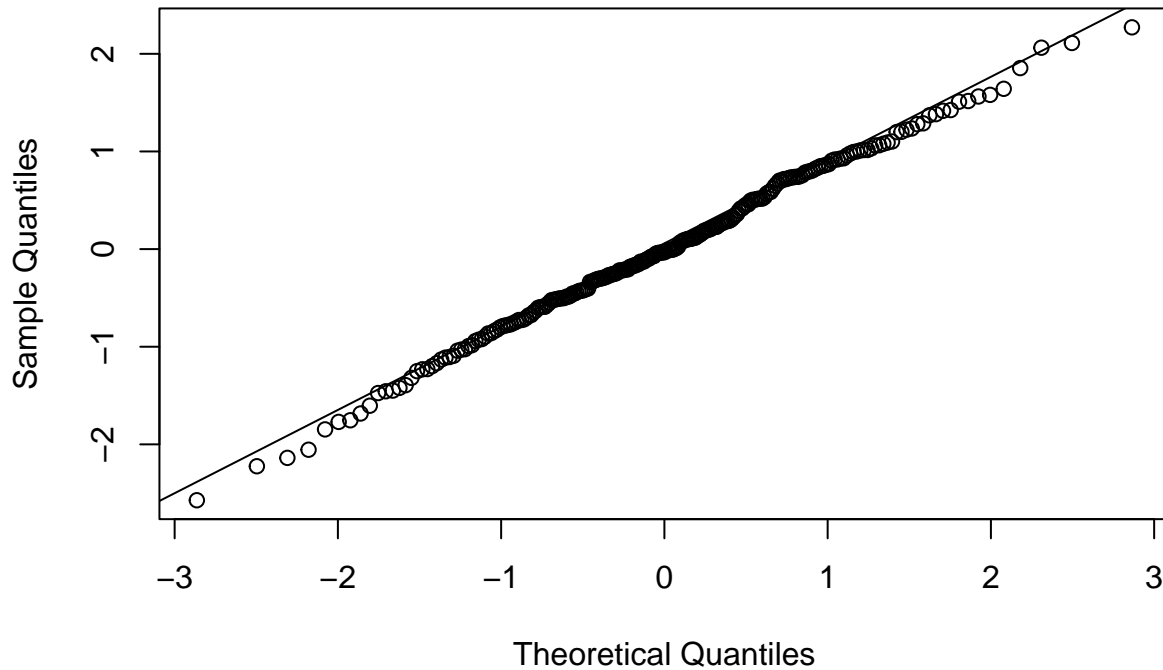
Residuals vs. fitted



Histogram of Res



Normal Q-Q Plot



```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme.formula(fixed = Densidad ~ 1 + año, data = leon_tb_1, random = ~1 |
## Estacion, method = "REML")
##
## Linear Hypotheses:
```

	Estimate	Std. Error	z value	Pr(> z)
2015 - 2014 == 0	0.052992	0.025530	2.076	0.39478
2016 - 2014 == 0	0.080401	0.045846	1.754	0.61737
2017 - 2014 == 0	0.131211	0.018369	7.143	< 0.001 ***
2018 - 2014 == 0	0.078647	0.023700	3.318	0.01701 *
2019 - 2014 == 0	0.165963	0.045846	3.620	0.00591 **
2020 - 2014 == 0	0.147138	0.018111	8.124	< 0.001 ***
2021 - 2014 == 0	0.186142	0.023474	7.930	< 0.001 ***
2016 - 2015 == 0	0.027409	0.046280	0.592	0.99874
2017 - 2015 == 0	0.078219	0.024782	3.156	0.02872 *
2018 - 2015 == 0	0.025654	0.021198	1.210	0.91734
2019 - 2015 == 0	0.112971	0.046280	2.441	0.19568
2020 - 2015 == 0	0.094146	0.024043	3.916	0.00188 **
2021 - 2015 == 0	0.133150	0.020954	6.354	< 0.001 ***
2017 - 2016 == 0	0.050810	0.045254	1.123	0.94346
2018 - 2016 == 0	-0.001754	0.045325	-0.039	1.00000

```
## 2019 - 2016 == 0 0.085562 0.056719 1.509 0.77769
## 2020 - 2016 == 0 0.066737 0.044988 1.483 0.79234
## 2021 - 2016 == 0 0.105741 0.045245 2.337 0.24391
## 2018 - 2017 == 0 -0.052564 0.022891 -2.296 0.26454
## 2019 - 2017 == 0 0.034752 0.045254 0.768 0.99351
## 2020 - 2017 == 0 0.015927 0.017484 0.911 0.98208
## 2021 - 2017 == 0 0.054931 0.022663 2.424 0.20322
## 2019 - 2018 == 0 0.087317 0.045325 1.926 0.49588
## 2020 - 2018 == 0 0.068491 0.022078 3.102 0.03420 *
## 2021 - 2018 == 0 0.107495 0.018624 5.772 < 0.001 ***
## 2020 - 2019 == 0 -0.018825 0.044988 -0.418 0.99987
## 2021 - 2019 == 0 0.020179 0.045245 0.446 0.99980
## 2021 - 2020 == 0 0.039004 0.021838 1.786 0.59452
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Podemos ver en la tabla de contrastes que para los datos de biomasa existen diferencias significativas de los datos del año 2014 respecto a los de 2017 en adelante.

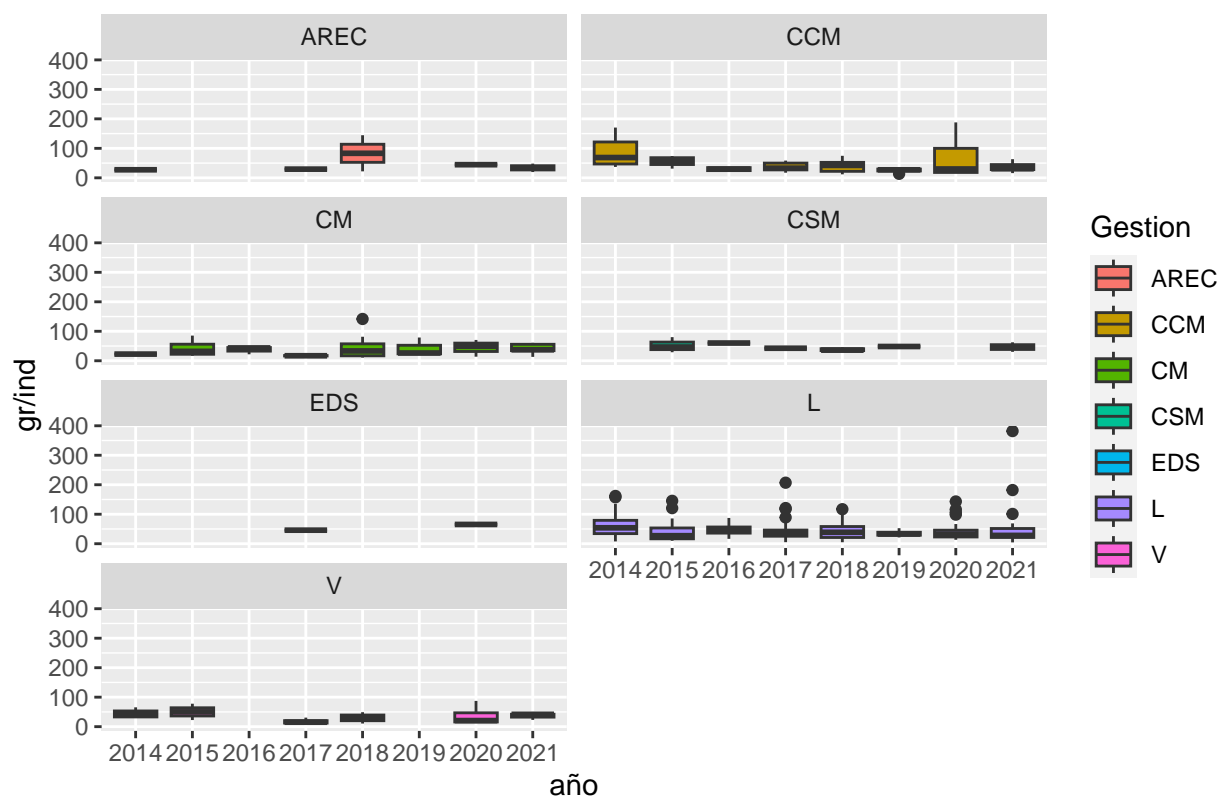
3. ANÁLISIS DE LA VARIABLE PESO MEDIO

Otra variable determinada de las 2 anteriores por simple cociente entre la biomasa y la densidad es el peso medio (gr/ind). Con ella podemos comprobar si desde 2014 ha habido variaciones significativas en la misma, también centrándonos en los tramos libres, desde la entrada de la normativa que exige la pesca sin muerte.

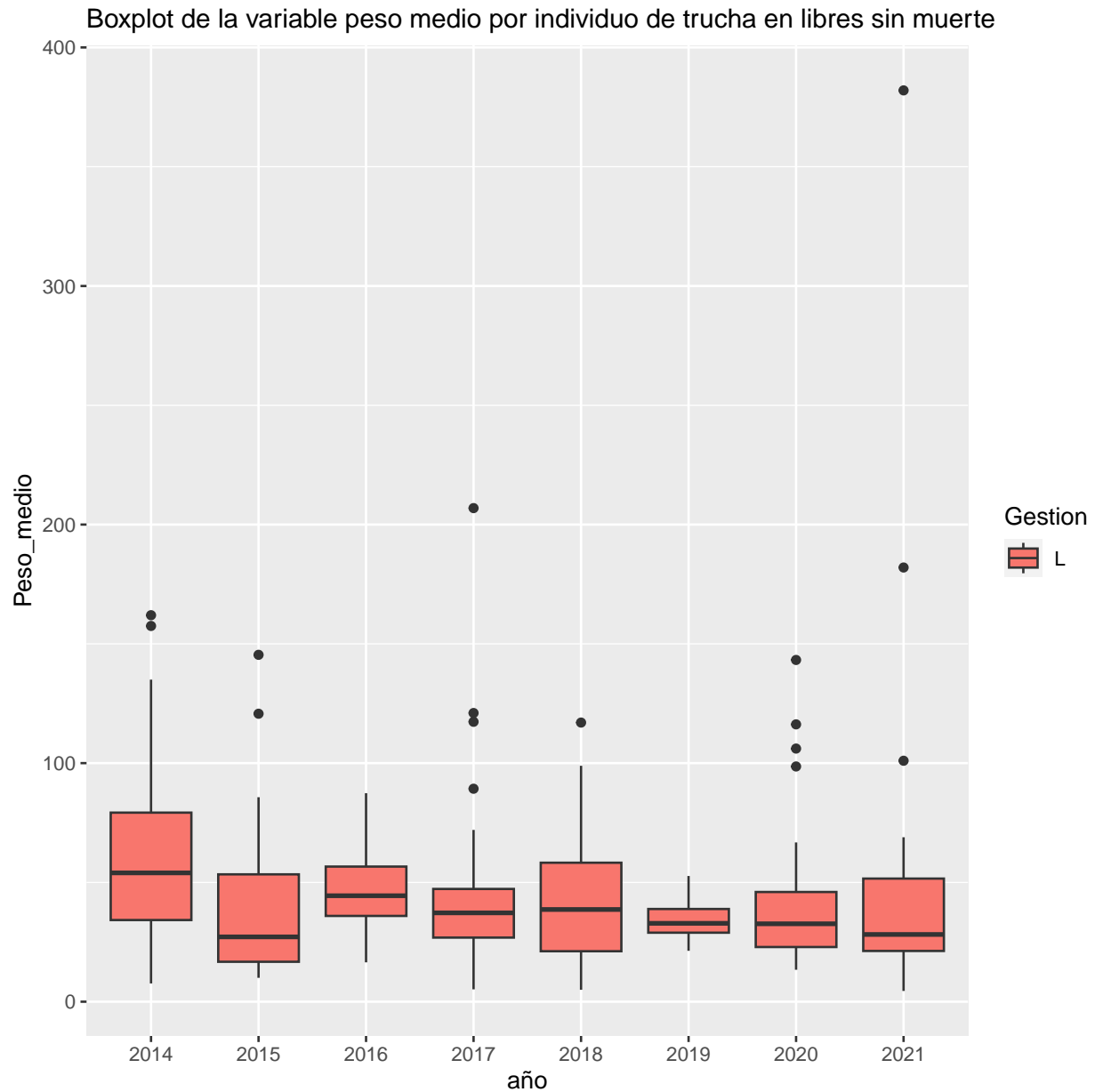
3.1 VISUALIZACIÓN DE LOS DATOS

```
## Estacion Provincia Gestion Nivel año Biomasa
## LE009 : 8 Leon:245 L:245 1: 34 2020 :51 Min. : 0.010
## LE010 : 8 2:211 2017 :41 1st Qu.: 2.630
## LE004 : 7 2021 :40 Median : 4.930
## LE070 : 6 2014 :39 Mean : 6.791
## LE189 : 5 2018 :38 3rd Qu.: 9.740
## LE072 : 3 2015 :28 Max. :38.120
## (Other):208 (Other): 8
## Densidad Peso_medio
## Min. :0.0000 Min. : 4.50
## 1st Qu.:0.0600 1st Qu.: 22.04
## Median :0.1400 Median : 37.00
## Mean :0.1987 Mean : 45.90
## 3rd Qu.:0.2600 3rd Qu.: 54.81
## Max. :0.9400 Max. :382.00
##
```

Boxplot del peso en gr/ind para cada año y tipo de gestión en León



Representamos las variables para los tramos libres, donde es previsible que se puedan observar con los efectos que se quieren comprobar por la introducción de la modalidad de sin muerte de forma general.



Se observa la existencia de outliers que para la comprobación de las comparaciones y realizar las comprobaciones de hipótesis deberá ser tenida en cuenta para elegir la técnica estadística. Determinamos a continuación los valores medios para cada año para los tramos de pesca libre, recogiéndose igualmente las desviaciones estándar de cada grupo.

3.2 MEDIAS Y DESVIACIONES DE LA VARIABLE PESO MEDIO

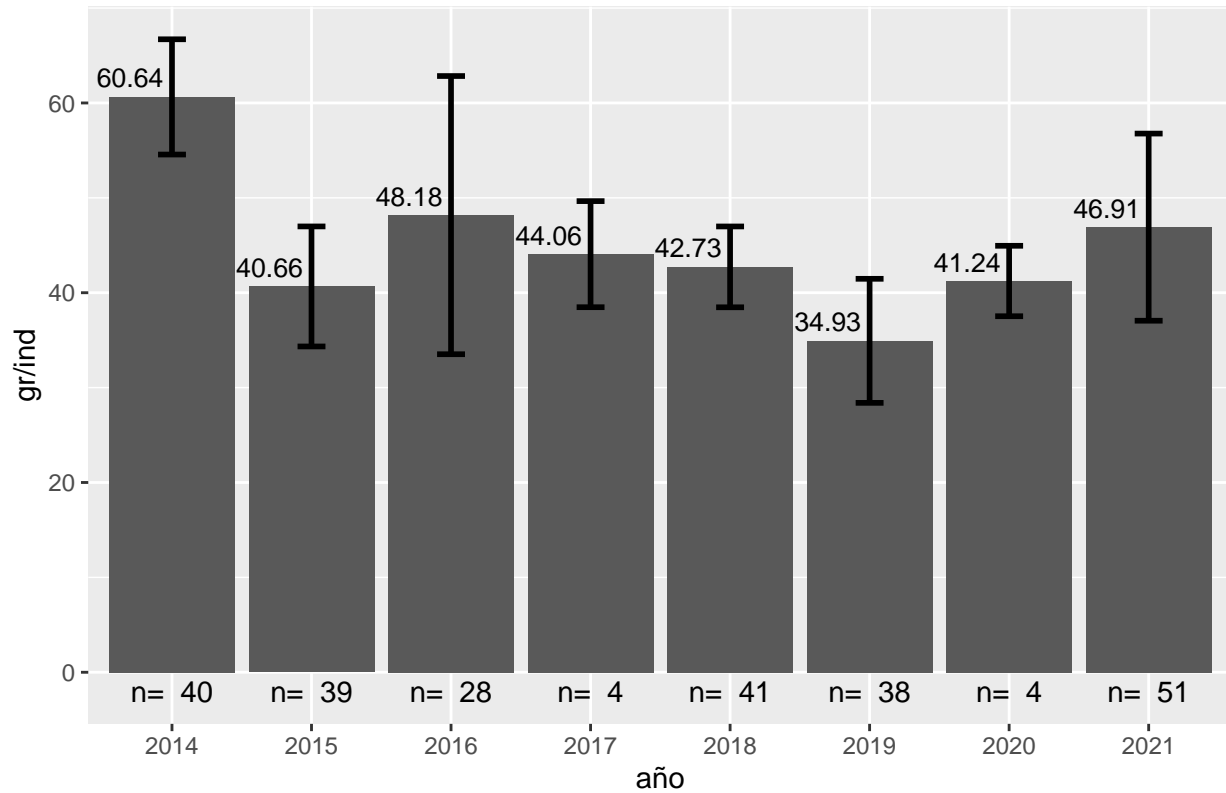
Para los tramos libres de León, comprobamos los valores medios del peso medio de trucha (en gr/ind) en los diferentes años, así como los valores de la desviación estándar.

Table 3: Valores medios y desviaciones estandar

	2014	2015	2016	2017	2018	2019	2020	2021
Media	60.63667	40.66071	48.17750	44.06390	42.72500	34.93000	41.23627	46.91050
Desv	37.93280	33.46672	29.31809	35.79717	26.25097	13.07535	26.50332	62.37319

Se observa un aumento de los valores medios respecto a 2014. Representamos la distribución mediante histograma de la variable densidad

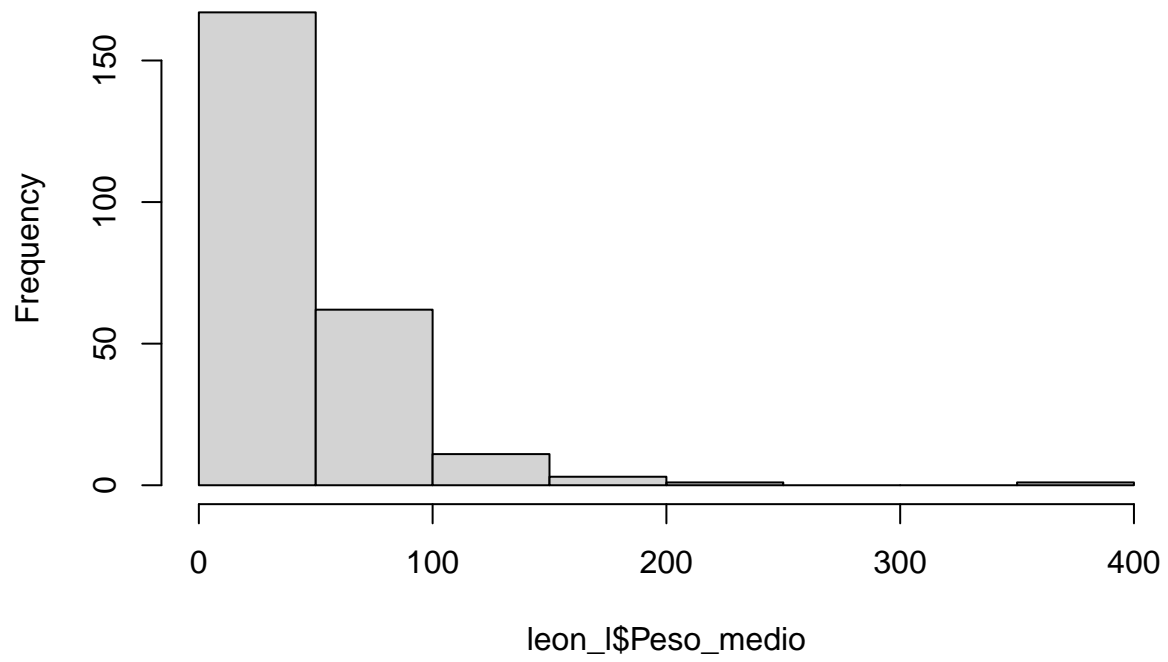
Peso medio de trucha en los libres sm (provincia de León)

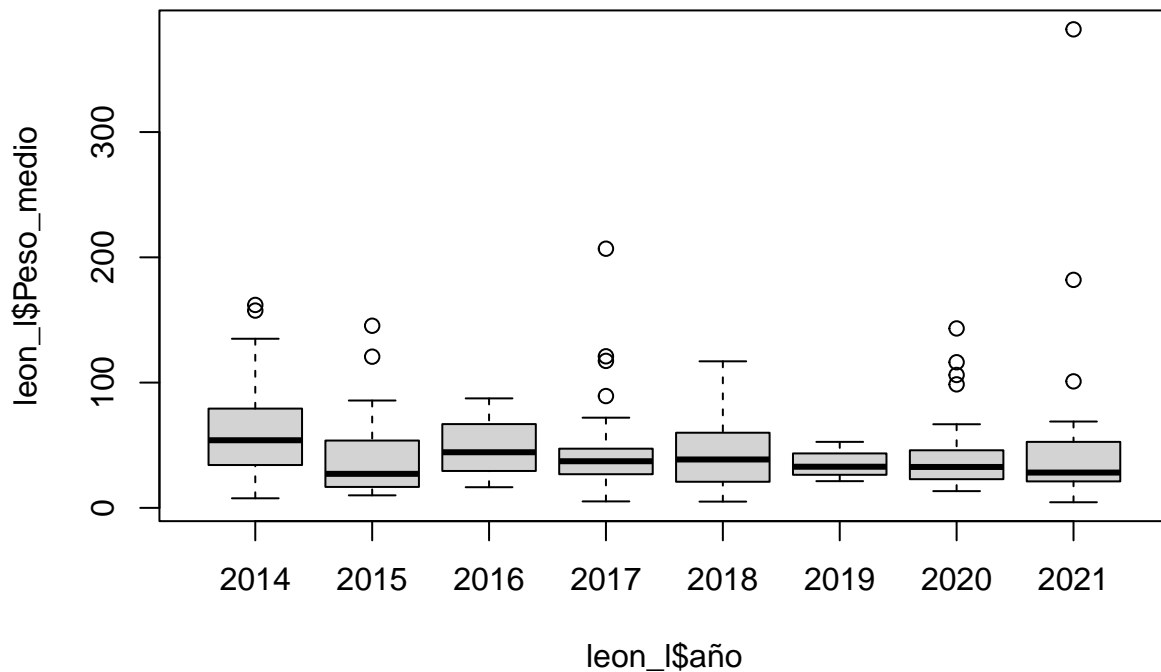


2.3 MODELO PARA LA COMPARACIÓN DEL PESO MEDIO

Para la comparación de los grupos optamos en primer lugar por observar los datos y ver los supuestos para adaptarlo a un modelo lineal. Comprobamos la normalidad y la homocedasticidad. Primero los visualizamos

Histogram of leon_I\$Peso_medio





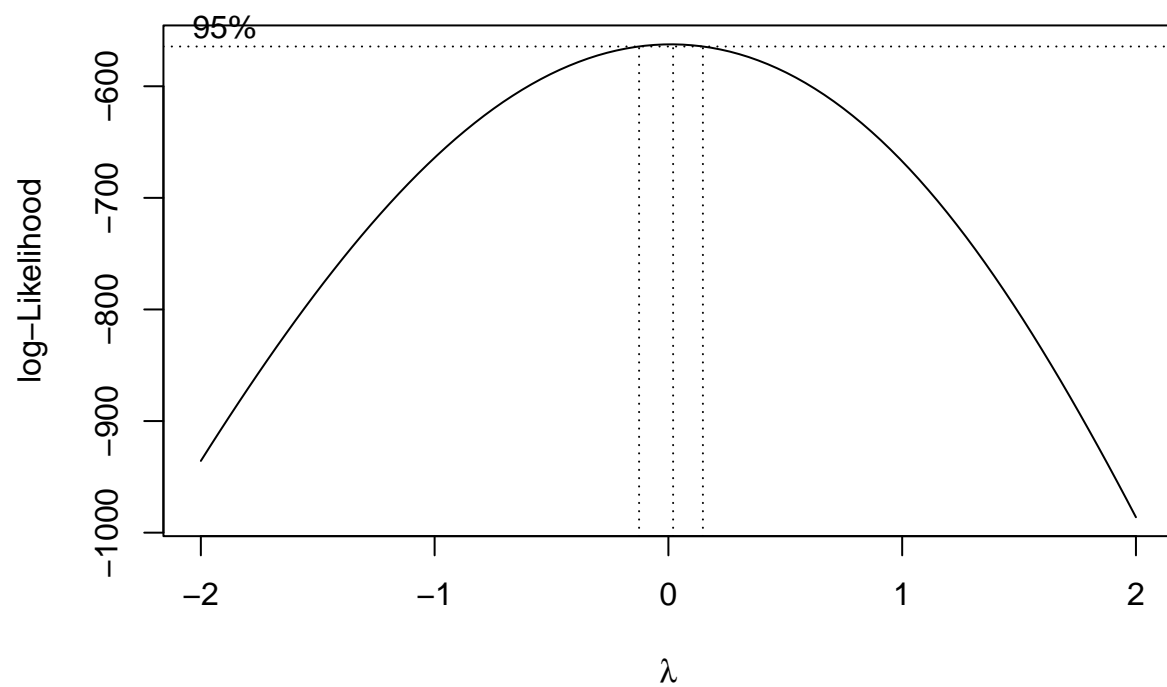
Comprobamos la normalidad y homocedasticidad aunque es visible que no parecen cumplirse ninguno de los 2 criterios

```
##
## Shapiro-Wilk normality test
##
## data: leon_l$Densidad
## W = 0.83839, p-value = 2.86e-15
```

No se puede aceptar la la distribución normal de los datos

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value    Pr(>F)
## group  7  5.2073 1.541e-05 ***
##      237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tampoco se puede aceptar la homocedasticidad de los datos de densidad para los diferentes años. Por ello es preciso transformar los datos y ver si con ello podemos asumir un modelo lineal. Para ello determinaremos los valores de λ mediante la aplicación de una transformación de Box Cox

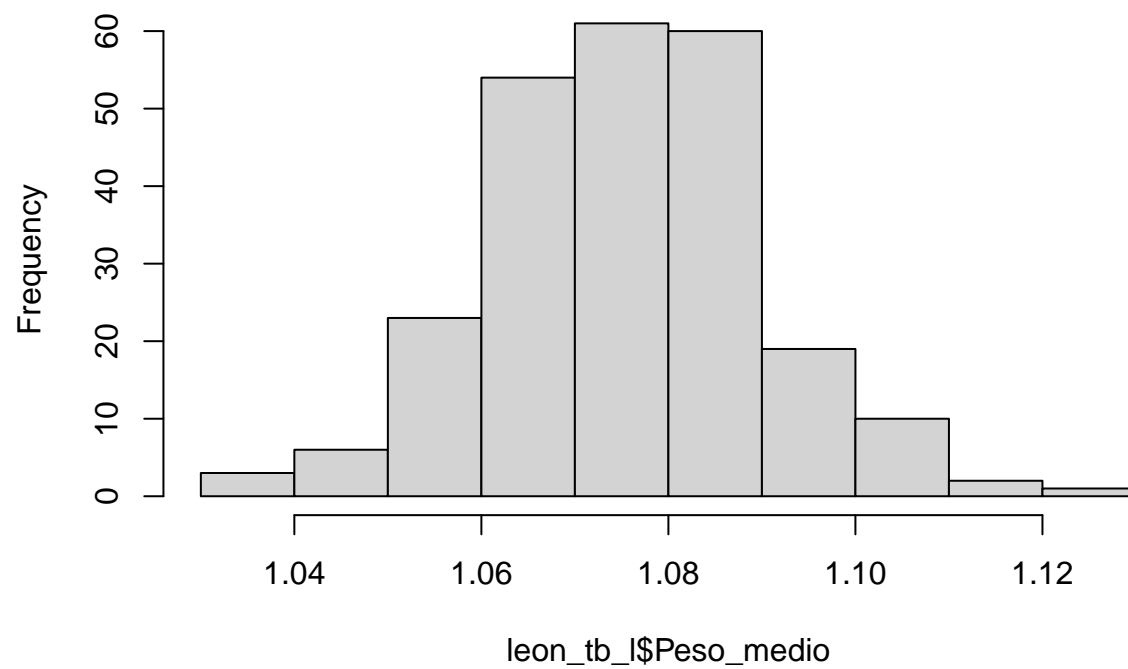


El valor de λ obtenido es 0.020202 es el valor de la potencia usada en la transformación de los datos.

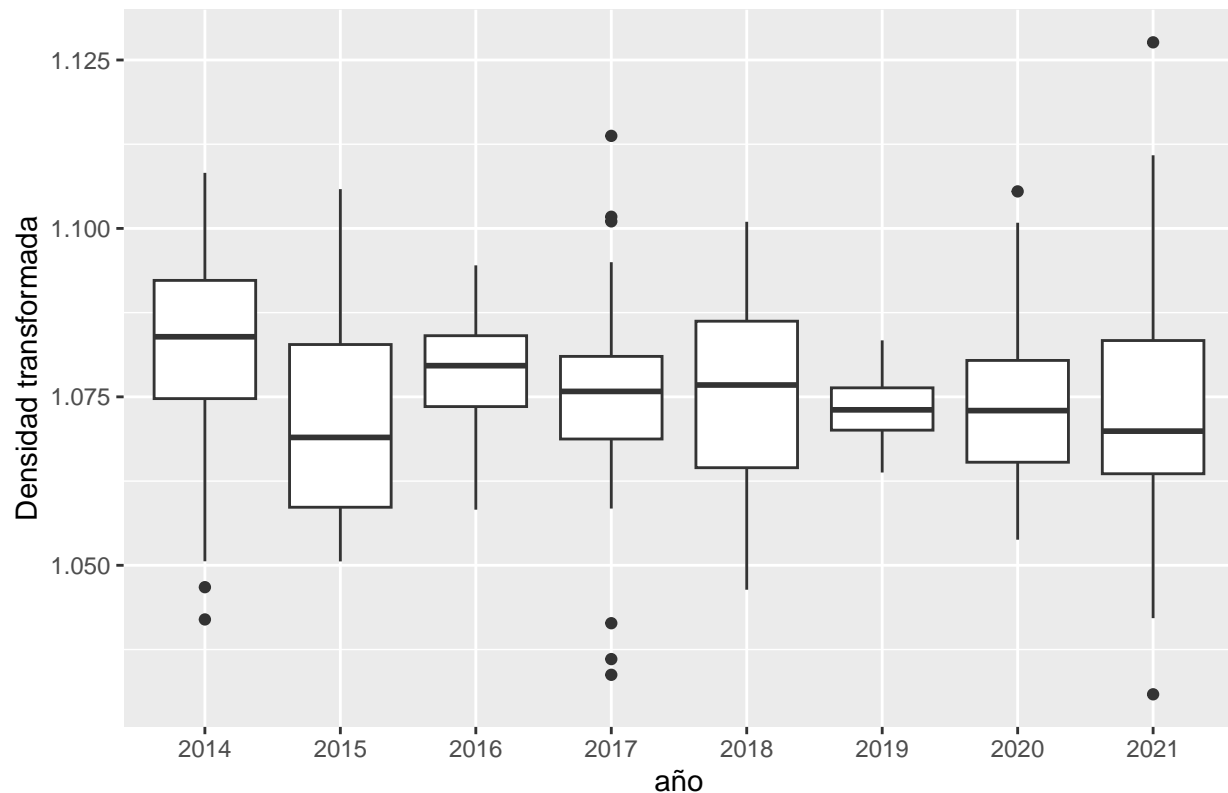
Transformamos los datos

Una vez transformado los datos visualizamos nuevamente

Histogram of leon_tb_I\$Peso_medio



Boxplot de las densidades transformadas



Visualmente ahora parece que los datos pueden cumplir a los criterios de normalidad y homocedasticidad, pese a ello realizamos los test.

```
##
## Shapiro-Wilk normality test
##
## data: leon_tb_1$Peso_medio
## W = 0.9941, p-value = 0.4743
```

Se acepta la la distribución normal de los datos

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  7  0.5712  0.779
##      231
```

Se acepta la homocedasticidad

Planteamos ahora un modelo lineal mixto. Esto es conveniente en nuestro caso ya que podemos establecer una parte aleatoria del modelo con los datos anidados por la variable Estación. Como término fijo utilizaríamos la variable año, ya que se trata del factor principal que nos explica precisamente si existen diferencias en los valores de la densidad considerando el año 2014 como el año que se produce el cambio de política de gestión al establecer los libres como libres sin muerte. Para la construcción del modelo seguimos el procedimiento que recomienda Zuur et al. (2009) y que se secuenciaría en fases: 1. Estructura aleatoria óptima. Usando un modelo saturado (beyond optimal model), se determina la estructura óptima del componente aleatorio, la cual no debe contener información que esté en la componente fija. Debemos:

- construir un modelo saturado.
- comparar modelos con distinta estructura aleatoria, mediante máxima verosimilitud restringida (REML).

2. Estructura fija óptima. Una vez encontramos la estructura aleatoria óptima, podemos encontrar la estructura fija óptima. Comparamos los modelos anidados mediante máxima verosimilitud (ML), manteniendo la misma estructura aleatoria.

3. Ajuste del modelo final con REML.

```
library(nlme)
####PROCEDIMIENTO DE ZUUR PARA LA SELECCION DEL MODELO
##1. Determinamos los efectos variables y para ello saturamos los efectos fijos
leon_tb_1<-leon_tb_1[!is.na(leon_tb_1$Peso_medio),]
m1a<-glS(Peso_medio~1+año,data=leon_tb_1,method="REML")
m1b<-lme(Peso_medio~1+año,random=~1|Estacion,data=leon_tb_1,method="REML")
anova(m1a,m1b)#Efectivamente al año parece afectar a la VD, seleccionamos como variable aleatoria la es

##      Model df      AIC      BIC  logLik  Test  L.Ratio p-value
## m1a      1  9 -1250.487 -1219.505 634.2435
## m1b      2 10 -1286.370 -1251.946 653.1851 1 vs 2 37.88316 <.0001

##2. Seleccionamos las variables de la estructura fija.
m1c<-lme(Peso_medio~1,random=~1|Estacion,data=leon_tb_1,method="ML")
m1d<-lme(Peso_medio~1+año,random=~1|Estacion,data=leon_tb_1,method="ML")
anova(m1c,m1d)#El modelo m1d es el elegido

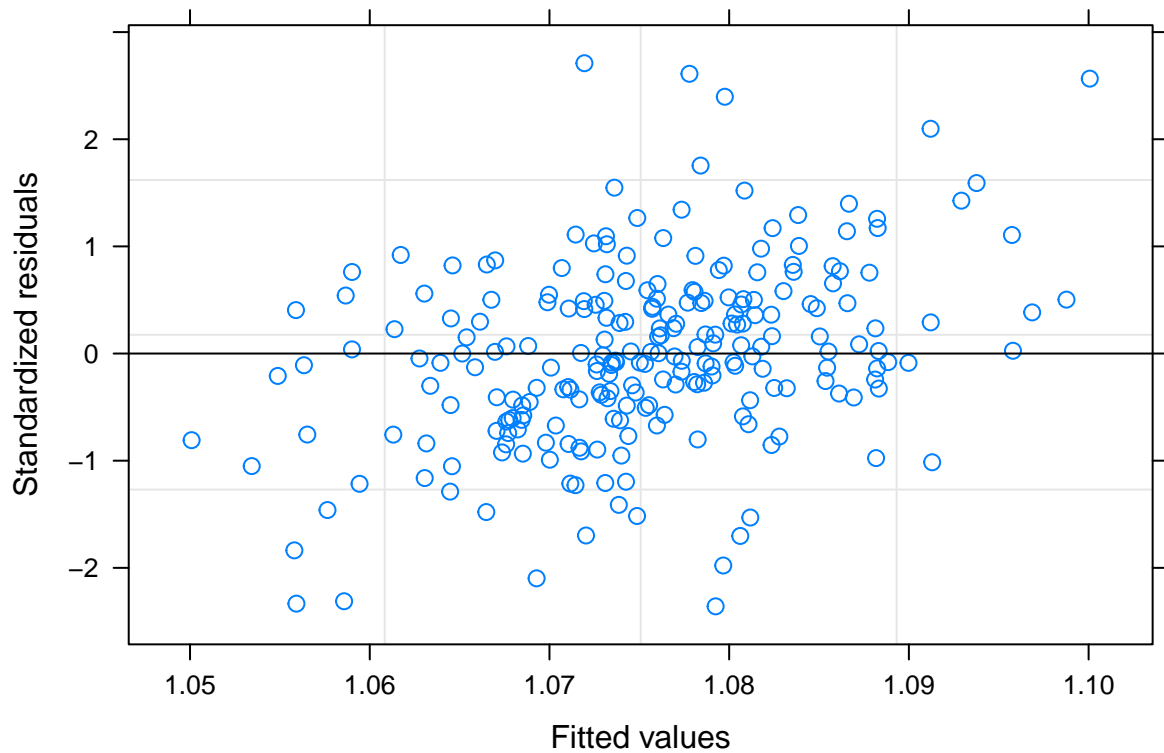
##      Model df      AIC      BIC  logLik  Test  L.Ratio p-value
## m1c      1  3 -1363.442 -1353.013 684.7212
## m1d      2 10 -1366.083 -1331.319 693.0417 1 vs 2 16.64103 0.0199

##3. Ajuste del modelo final con el método REML
m1d_final<-lme(Peso_medio~1+año,random=~1|Estacion,method="REML",data=leon_tb_1)
summary(m1d_final)#Diferencias entre 2017 y 2016 respecto de 2014

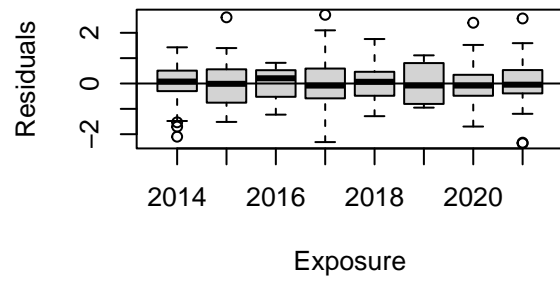
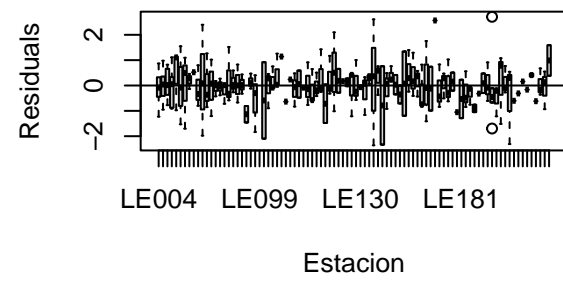
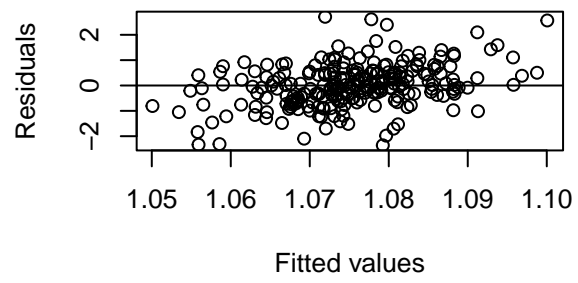
## Linear mixed-effects model fit by REML
## Data: leon_tb_1
##      AIC      BIC  logLik
## -1286.37 -1251.946 653.1851
##
## Random effects:
## Formula: ~1 | Estacion
##      (Intercept)  Residual
## StdDev:  0.01070344 0.01074035
##
## Fixed effects:  Peso_medio ~ 1 + año
##      Value Std.Error DF t-value p-value
## (Intercept) 1.0818501 0.002338837 143 462.5591 0.0000
## año2015     -0.0106075 0.003451569 143 -3.0732 0.0025
## año2016     -0.0030930 0.006149433 143 -0.5030 0.6158
## año2017     -0.0075608 0.002460199 143 -3.0733 0.0025
## año2018     -0.0060582 0.003208584 143 -1.8881 0.0610
## año2019     -0.0077743 0.006149433 143 -1.2642 0.2082
## año2020     -0.0074722 0.002428914 143 -3.0763 0.0025
## año2021     -0.0091514 0.003178765 143 -2.8789 0.0046
## Correlation:
##      (Intr) añ2015 añ2016 añ2017 añ2018 añ2019 añ2020
## año2015 -0.621
## año2016 -0.293 0.264
## año2017 -0.568 0.397 0.233
## año2018 -0.672 0.638 0.283 0.427
```

```
## año2019 -0.293  0.264  0.237  0.233  0.283
## año2020 -0.630  0.432  0.245  0.540  0.465  0.245
## año2021 -0.680  0.644  0.284  0.431  0.695  0.284  0.469
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.35920539 -0.48285040 -0.01712622  0.49199585  2.70961349
##
## Number of Observations: 239
## Number of Groups: 89
```

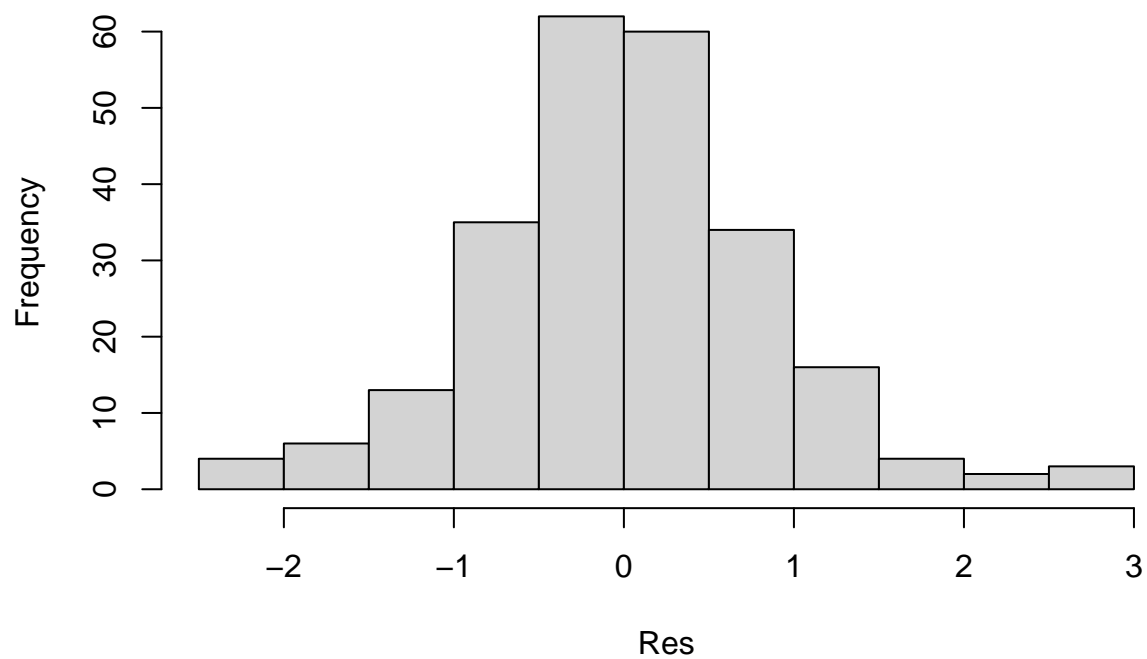
El modelo final lo utilizamos para contrastar los diferentes niveles de la variable predictora, en este caso los diferentes años mediante comparaciones múltiples mediante el test de Tukey. Comprobamos el modelo de forma gráfica



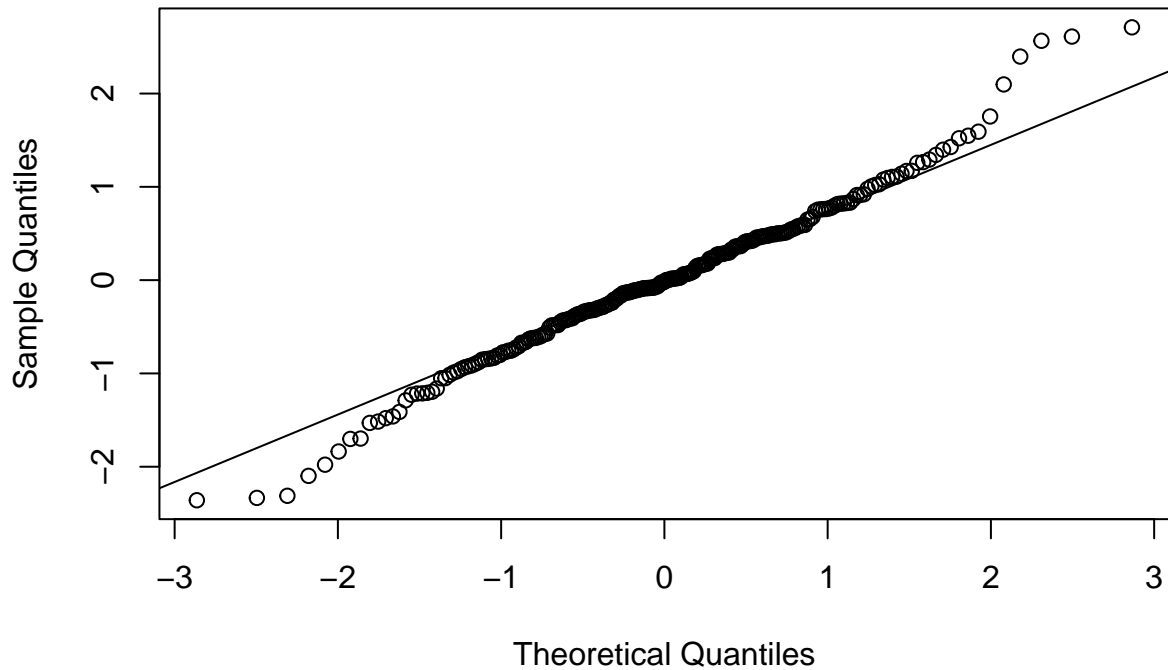
Residuals vs. fitted



Histogram of Res



Normal Q-Q Plot



```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lme.formula(fixed = Peso_medio ~ 1 + año, data = leon_tb_1,
## random = ~1 | Estacion, method = "REML")
##
## Linear Hypotheses:
```

	Estimate	Std. Error	z value	Pr(> z)	
2015 - 2014 == 0	-1.061e-02	3.452e-03	-3.073	0.0373	*
2016 - 2014 == 0	-3.093e-03	6.149e-03	-0.503	0.9996	
2017 - 2014 == 0	-7.561e-03	2.460e-03	-3.073	0.0370	*
2018 - 2014 == 0	-6.058e-03	3.209e-03	-1.888	0.5224	
2019 - 2014 == 0	-7.774e-03	6.149e-03	-1.264	0.8977	
2020 - 2014 == 0	-7.472e-03	2.429e-03	-3.076	0.0369	*
2021 - 2014 == 0	-9.151e-03	3.179e-03	-2.879	0.0654	.
2016 - 2015 == 0	7.514e-03	6.205e-03	1.211	0.9171	
2017 - 2015 == 0	3.047e-03	3.349e-03	0.910	0.9822	
2018 - 2015 == 0	4.549e-03	2.841e-03	1.601	0.7200	
2019 - 2015 == 0	2.833e-03	6.205e-03	0.457	0.9998	
2020 - 2015 == 0	3.135e-03	3.252e-03	0.964	0.9752	
2021 - 2015 == 0	1.456e-03	2.808e-03	0.519	0.9995	
2017 - 2016 == 0	-4.468e-03	6.069e-03	-0.736	0.9950	
2018 - 2016 == 0	-2.965e-03	6.077e-03	-0.488	0.9996	

```
## 2019 - 2016 == 0 -4.681e-03 7.595e-03 -0.616 0.9984
## 2020 - 2016 == 0 -4.379e-03 6.034e-03 -0.726 0.9954
## 2021 - 2016 == 0 -6.058e-03 6.067e-03 -0.999 0.9698
## 2018 - 2017 == 0 1.503e-03 3.098e-03 0.485 0.9997
## 2019 - 2017 == 0 -2.135e-04 6.069e-03 -0.035 1.0000
## 2020 - 2017 == 0 8.865e-05 2.344e-03 0.038 1.0000
## 2021 - 2017 == 0 -1.591e-03 3.068e-03 -0.518 0.9995
## 2019 - 2018 == 0 -1.716e-03 6.077e-03 -0.282 1.0000
## 2020 - 2018 == 0 -1.414e-03 2.991e-03 -0.473 0.9997
## 2021 - 2018 == 0 -3.093e-03 2.495e-03 -1.240 0.9068
## 2020 - 2019 == 0 3.021e-04 6.034e-03 0.050 1.0000
## 2021 - 2019 == 0 -1.377e-03 6.067e-03 -0.227 1.0000
## 2021 - 2020 == 0 -1.679e-03 2.959e-03 -0.567 0.9990
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Podemos ver en la tabla de contrastes que para los datos de peso medio existen diferencias significativas de los datos del año 2014, solo con respecto a 2015. No obstante los datos parecen indicar una reducción ligera del peso medio probablemente debido a que la pesca sin muerte en los tramos libres ha supuesto un aumento apreciable de biomasa pero aún más las densidades con lo que el peso medio puede mostrar una ligera reducción.

3. DISCUSIÓN

De los resultados obtenidos para la provincia de León que es la provincia con mayor número de estaciones de muestreo, se observa que tanto para las variables biomasa como para la densidad esta aumenta en todos los años en comparación con el año base (2014) aunque los efectos son significativos para los años 2017, 2019, 2020 y 2021 en el caso de la biomasa de trucha y para los años 2017, 2018, 2019, 2020 y 2021 en el caso de la densidad.

Debido a la falta de datos ambientales, tales como factores hidrológicos, composición y estructura del hábitat que caractericen las estaciones es imposible profundizar en parámetros explicativos. Igualmente al no ser publicados las estructuras de las poblaciones de trucha inventariadas tampoco conocemos los parámetros poblacionales principales que nos ayuden a comprender la situación en profundidad. Mortalidad, crecimiento o producción, parámetros éstos relevantes para una mejor comprensión de las dinámicas poblacionales no es posible determinarlas al no haber sido publicadas las existencias por cohortes de las poblaciones inventariadas. No obstante el análisis de densidades, biomasa y en menor medida el peso medio, entre los diferentes años desde el 2014, parecen claros al mostrar los datos un importante incremento en los valores de la biomasa de trucha y sobre todo en las densidades, que han visto multiplicado por 3 los valores de 2014 y por 2 en el caso de la biomasa. La variable peso medio, sufre una ligera disminución ya que siendo determinada como el cociente entre la biomasa unitaria y la densidad unitaria, y debido a que el aumento de la densidad ha sido superior al aumento en biomasa, este valor ha disminuido aunque no de forma significativa.

Para analizar con más profundidad estos datos sería relevante determinar para las estaciones de muestreo la estructura de cohortes, parámetros de crecimiento y mortalidad. Con ello se podría determinar con más precisión el alcance de estos valores de peso medio.

Sería importante recabar datos anteriores a 2014, ya que pese a que se ha tomado como año de comparación para comprobar los efectos de la declaración de los libres como libres sin muerte, debido a que los muestreos se realizaron al final de la temporada de pesca, es posible que en el año 2014 también se hayan notado los efectos de la entrada en vigor de la ley, y se hayan reclutado en las tallas pescables individuos que no han sido objeto de extracción, y por ello se hayan atenuado las diferencias entre los grupos.