Regresiones tamaño y tipo

D. Daza

junio 13, 2023 | 11:05:12 | CDT

Table of Contents

[1. Modelos 2](#_Toc137550940)

[1.1 Cargas vs Tipo 2](#_Toc137550941)

[1.2 Cargas vs Tamaño 5](#_Toc137550942)

[2. Comprobación de supuestos 13](#_Toc137550943)

[2.1 Normalidad de residuos 13](#_Toc137550944)

[2.2 Independencia de los errores 15](#_Toc137550945)

[2.3 Linealidad de las relaciones entre las variables 16](#_Toc137550946)

[2.4 Homocedasticidad 21](#_Toc137550947)

# 1. Modelos

## 1.1 Cargas vs Tipo

#FORMAL GENERAL  
mfo1 <- lm(p\_formal ~ ca\_xe, data = base)  
  
summary(mfo1)

Call:  
lm(formula = p\_formal ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.037209 -0.016190 -0.003996 0.015013 0.042512   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 6.414e-02 1.737e-02 3.693 0.000662 \*\*\*  
ca\_xe -6.528e-06 6.249e-06 -1.045 0.302463   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.02093 on 40 degrees of freedom  
Multiple R-squared: 0.02656, Adjusted R-squared: 0.002221   
F-statistic: 1.091 on 1 and 40 DF, p-value: 0.3025

#FORMAL CONTROLES  
mfo2 <- lm(p\_formal ~ ca\_xe +  
 p\_informal + pobreza,  
 data = base)  
  
summary(mfo2)

Call:  
lm(formula = p\_formal ~ ca\_xe + p\_informal + pobreza, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.0072104 -0.0022978 0.0002455 0.0017084 0.0068145   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.084e-01 9.622e-03 32.051 <2e-16 \*\*\*  
ca\_xe -3.748e-07 1.118e-06 -0.335 0.7394   
p\_informal -3.109e-01 1.296e-02 -23.993 <2e-16 \*\*\*  
pobreza 1.435e-04 7.826e-05 1.834 0.0745 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.003686 on 38 degrees of freedom  
Multiple R-squared: 0.9713, Adjusted R-squared: 0.9691   
F-statistic: 429.2 on 3 and 38 DF, p-value: < 2.2e-16

vif(mfo2)

ca\_xe p\_informal pobreza   
 1.033467 2.513975 2.527391

# Obtiene un data frame de los resultados del modelo  
mfo2\_tidy <- broom::tidy(mfo2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mfo2\_table <- kable(mfo2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y empresas formales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mfo2\_table

Cargas administrativas y empresas formales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.3084124 | 0.0096224 | 32.0513776 | 0.0000000 |
| ca\_xe | -0.0000004 | 0.0000011 | -0.3351417 | 0.7393613 |
| p\_informal | -0.3108978 | 0.0129576 | -23.9934975 | 0.0000000 |
| pobreza | 0.0001435 | 0.0000783 | 1.8339891 | 0.0744958 |
|  |  |  |  |  |

#INFORMAL GENERAL  
mi1 <- lm(p\_informal ~ ca\_xe, data = base)  
  
summary(mi1)

Call:  
lm(formula = p\_informal ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.14932 -0.04771 0.01975 0.05436 0.11798   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 7.952e-01 5.840e-02 13.617 <2e-16 \*\*\*  
ca\_xe 2.160e-05 2.101e-05 1.028 0.31   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.07039 on 40 degrees of freedom  
Multiple R-squared: 0.02576, Adjusted R-squared: 0.0014   
F-statistic: 1.057 on 1 and 40 DF, p-value: 0.31

#INFORMAL CONTROLES  
mi2 <- lm(p\_informal ~ ca\_xe +  
 p\_formal + pobreza +   
 densidadpobkm2,  
 data = base)  
  
summary(mi2)

Call:  
lm(formula = p\_informal ~ ca\_xe + p\_formal + pobreza + densidadpobkm2,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.025578 -0.006810 0.001692 0.006231 0.020496   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 9.598e-01 1.660e-02 57.823 < 2e-16 \*\*\*  
ca\_xe -7.886e-08 3.409e-06 -0.023 0.981670   
p\_formal -2.916e+00 1.336e-01 -21.825 < 2e-16 \*\*\*  
pobreza 8.542e-04 2.305e-04 3.707 0.000684 \*\*\*  
densidadpobkm2 6.274e-07 3.379e-07 1.857 0.071279 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.01113 on 37 degrees of freedom  
Multiple R-squared: 0.9775, Adjusted R-squared: 0.975   
F-statistic: 401.3 on 4 and 37 DF, p-value: < 2.2e-16

vif(mi2)

ca\_xe p\_formal pobreza densidadpobkm2   
 1.053370 2.595009 2.403518 1.214474

# Obtiene un data frame de los resultados del modelo  
mi2\_tidy <- broom::tidy(mi2)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mi2\_table <- kable(mi2\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y empresas informales",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mi2\_table

Cargas administrativas y empresas informales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.9597792 | 0.0165985 | 57.8233962 | 0.0000000 |
| ca\_xe | -0.0000001 | 0.0000034 | -0.0231313 | 0.9816699 |
| p\_formal | -2.9157559 | 0.1335983 | -21.8247905 | 0.0000000 |
| pobreza | 0.0008542 | 0.0002305 | 3.7066607 | 0.0006843 |
| densidadpobkm2 | 0.0000006 | 0.0000003 | 1.8570489 | 0.0712786 |

## 1.2 Cargas vs Tamaño

#Modelo 1.1 Tamaño vs Cargas  
  
  
##########MICROEMPRESAS  
  
# Regresión lineal para microempresas   
micro\_reg11 <- lm(p\_micro ~ ca\_xe, data = base)  
  
summary(micro\_reg11)

Call:  
lm(formula = p\_micro ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.155137 -0.023870 0.003991 0.044157 0.091479   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 7.852e-01 4.615e-02 17.014 <2e-16 \*\*\*  
ca\_xe 2.735e-05 1.660e-05 1.647 0.107   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.05563 on 40 degrees of freedom  
Multiple R-squared: 0.06354, Adjusted R-squared: 0.04013   
F-statistic: 2.714 on 1 and 40 DF, p-value: 0.1073

# Regresión controles microempresas  
micro\_reg12 <- lm(p\_micro ~ ca\_xe +  
 escolaridadprom +  
 p\_informal +   
 razondeingreso,  
 data = base)  
  
summary(micro\_reg12)

Call:  
lm(formula = p\_micro ~ ca\_xe + escolaridadprom + p\_informal +   
 razondeingreso, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.051438 -0.006305 0.000230 0.013974 0.033296   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 4.446e-01 1.041e-01 4.272 0.00013 \*\*\*  
ca\_xe 1.091e-05 5.774e-06 1.889 0.06669 .   
escolaridadprom -1.474e-02 5.106e-03 -2.887 0.00646 \*\*   
p\_informal 5.479e-01 4.959e-02 11.047 2.84e-13 \*\*\*  
razondeingreso 5.640e-01 3.320e-01 1.699 0.09779 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.01903 on 37 degrees of freedom  
Multiple R-squared: 0.8986, Adjusted R-squared: 0.8876   
F-statistic: 81.96 on 4 and 37 DF, p-value: < 2.2e-16

vif(micro\_reg12)

ca\_xe escolaridadprom p\_informal razondeingreso   
 1.033083 2.659174 1.380897 2.788041

# Obtiene un data frame de los resultados del modelo  
micro12\_tidy <- broom::tidy(micro\_reg12)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
micro12\_table <- kable(micro12\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y microempresas",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
micro12\_table

Cargas administrativas y microempresas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.4446401 | 0.1040848 | 4.271903 | 0.0001300 |
| ca\_xe | 0.0000109 | 0.0000058 | 1.889395 | 0.0666947 |
| escolaridadprom | -0.0147404 | 0.0051057 | -2.887035 | 0.0064582 |
| p\_informal | 0.5478540 | 0.0495922 | 11.047181 | 0.0000000 |
| razondeingreso | 0.5639740 | 0.3320225 | 1.698602 | 0.0977897 |

##########PEQUEÑAS EMPRESAS  
  
# Regresión lineal para pequeñas empresas  
pequena\_reg11 <- lm(p\_pequeña ~ ca\_xe, data = base)  
  
summary(pequena\_reg11)

Call:  
lm(formula = p\_pequeña ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.043338 -0.018508 -0.001663 0.014954 0.063697   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 9.790e-02 2.056e-02 4.761 2.53e-05 \*\*\*  
ca\_xe -8.683e-06 7.397e-06 -1.174 0.247   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.02478 on 40 degrees of freedom  
Multiple R-squared: 0.0333, Adjusted R-squared: 0.009135   
F-statistic: 1.378 on 1 and 40 DF, p-value: 0.2474

# Regresión controles pequeñas empresas  
pequena\_reg12 <- lm(p\_pequeña ~ ca\_xe +  
 p\_informal +   
 razondeingreso,  
 data = base)  
  
summary(pequena\_reg12)

Call:  
lm(formula = p\_pequeña ~ ca\_xe + p\_informal + razondeingreso,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.012994 -0.003878 -0.001330 0.004317 0.018710   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.793e-01 1.486e-02 25.520 < 2e-16 \*\*\*  
ca\_xe -1.523e-06 2.178e-06 -0.699 0.48882   
p\_informal -2.931e-01 1.865e-02 -15.719 < 2e-16 \*\*\*  
razondeingreso -3.567e-01 8.700e-02 -4.099 0.00021 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.007204 on 38 degrees of freedom  
Multiple R-squared: 0.9224, Adjusted R-squared: 0.9163   
F-statistic: 150.6 on 3 and 38 DF, p-value: < 2.2e-16

vif(pequena\_reg12)

ca\_xe p\_informal razondeingreso   
 1.026438 1.362679 1.336405

# Obtiene un data frame de los resultados del modelo  
pequena12\_tidy <- broom::tidy(pequena\_reg12)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
pequena12\_table <- kable(pequena12\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y pequeñas empresas",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
pequena12\_table

Cargas administrativas y pequeñas empresas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.3793013 | 0.0148627 | 25.5204149 | 0.0000000 |
| ca\_xe | -0.0000015 | 0.0000022 | -0.6989775 | 0.4888225 |
| p\_informal | -0.2931010 | 0.0186458 | -15.7194009 | 0.0000000 |
| razondeingreso | -0.3566722 | 0.0870039 | -4.0994954 | 0.0002100 |

########## MEDIANAS EMPRESAS  
  
# Regresión lineal para medianas empresas  
mediana\_reg11 <- lm(p\_mediana ~ ca\_xe, data = base)  
  
summary(mediana\_reg11)

Call:  
lm(formula = p\_mediana ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.03746 -0.01870 -0.00242 0.01101 0.06911   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 8.608e-02 2.008e-02 4.287 0.000111 \*\*\*  
ca\_xe -1.219e-05 7.224e-06 -1.688 0.099190 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.0242 on 40 degrees of freedom  
Multiple R-squared: 0.0665, Adjusted R-squared: 0.04316   
F-statistic: 2.849 on 1 and 40 DF, p-value: 0.09919

# Regresión controles medianas empresas  
mediana\_reg12 <- lm(p\_mediana ~ ca\_xe +  
 delitos\_2019 + escolaridadprom +  
 p\_informal +   
 razondeingreso,  
 data = base)  
  
summary(mediana\_reg12)

Call:  
lm(formula = p\_mediana ~ ca\_xe + delitos\_2019 + escolaridadprom +   
 p\_informal + razondeingreso, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.012548 -0.006846 -0.001323 0.005187 0.021428   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.034e-01 4.773e-02 4.262 0.00014 \*\*\*  
ca\_xe -3.430e-06 2.831e-06 -1.212 0.23350   
delitos\_2019 3.977e-07 2.348e-07 1.694 0.09897 .   
escolaridadprom 7.434e-03 2.341e-03 3.176 0.00306 \*\*   
p\_informal -2.139e-01 2.306e-02 -9.273 4.49e-11 \*\*\*  
razondeingreso -3.118e-01 1.561e-01 -1.998 0.05336 .   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.008718 on 36 degrees of freedom  
Multiple R-squared: 0.891, Adjusted R-squared: 0.8758   
F-statistic: 58.84 on 5 and 36 DF, p-value: 2.553e-16

vif(mediana\_reg12)

ca\_xe delitos\_2019 escolaridadprom p\_informal razondeingreso   
 1.183500 1.245920 2.664173 1.423716 2.936954

# Obtiene un data frame de los resultados del modelo  
mediana12\_tidy <- broom::tidy(mediana\_reg12)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
mediana12\_table <- kable(mediana12\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y medianas empresas",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
mediana12\_table

Cargas administrativas y medianas empresas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | 0.2034021 | 0.0477286 | 4.261642 | 0.0001395 |
| ca\_xe | -0.0000034 | 0.0000028 | -1.211753 | 0.2335007 |
| delitos\_2019 | 0.0000004 | 0.0000002 | 1.693662 | 0.0989665 |
| escolaridadprom | 0.0074341 | 0.0023408 | 3.175856 | 0.0030596 |
| p\_informal | -0.2138695 | 0.0230646 | -9.272630 | 0.0000000 |
| razondeingreso | -0.3118043 | 0.1560875 | -1.997625 | 0.0533570 |

########## GRANDES EMPRESAS  
  
# Regresión lineal para grandes empresas  
grande\_reg11 <- lm(p\_grande ~ ca\_xe, data = base)  
  
summary(grande\_reg11)

Call:  
lm(formula = p\_grande ~ ca\_xe, data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.010678 -0.006283 -0.002383 0.004160 0.037717   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 3.079e-02 7.701e-03 3.999 0.000267 \*\*\*  
ca\_xe -6.477e-06 2.770e-06 -2.338 0.024489 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.009282 on 40 degrees of freedom  
Multiple R-squared: 0.1202, Adjusted R-squared: 0.09821   
F-statistic: 5.465 on 1 and 40 DF, p-value: 0.02449

# Regresión controles grandes empresas  
grande\_reg12 <- lm(p\_grande ~ ca\_xe +  
 escolaridadprom +  
 p\_informal,  
 data = base)  
  
summary(grande\_reg12)

Call:  
lm(formula = p\_grande ~ ca\_xe + escolaridadprom + p\_informal,   
 data = base)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.0083590 -0.0035838 -0.0008574 0.0012527 0.0229328   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -1.809e-03 2.301e-02 -0.079 0.9378   
ca\_xe -4.397e-06 1.914e-06 -2.297 0.0272 \*   
escolaridadprom 5.492e-03 1.174e-03 4.680 3.59e-05 \*\*\*  
p\_informal -3.936e-02 1.591e-02 -2.474 0.0179 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.00632 on 38 degrees of freedom  
Multiple R-squared: 0.6125, Adjusted R-squared: 0.5819   
F-statistic: 20.02 on 3 and 38 DF, p-value: 5.985e-08

vif(grande\_reg12)

ca\_xe escolaridadprom p\_informal   
 1.029694 1.274635 1.288439

# Obtiene un data frame de los resultados del modelo  
grande12\_tidy <- broom::tidy(grande\_reg12)  
  
# Convierte el data frame de los resultados del modelo en una tabla  
grande12\_table <- kable(grande12\_tidy, format = "pandoc",   
 caption = "Cargas administrativas y grandes empresas",   
 col.names = c("Variable", "Estimado", "Error estándar", "Estadístico t", "Valor p"))  
  
# Imprimir la tabla  
grande12\_table

Cargas administrativas y grandes empresas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimado | Error estándar | Estadístico t | Valor p |
| (Intercept) | -0.0018085 | 0.0230142 | -0.0785827 | 0.9377769 |
| ca\_xe | -0.0000044 | 0.0000019 | -2.2973516 | 0.0271984 |
| escolaridadprom | 0.0054924 | 0.0011737 | 4.6796305 | 0.0000359 |
| p\_informal | -0.0393562 | 0.0159052 | -2.4744168 | 0.0179244 |
|  |  |  |  |  |

# 2. Comprobación de supuestos

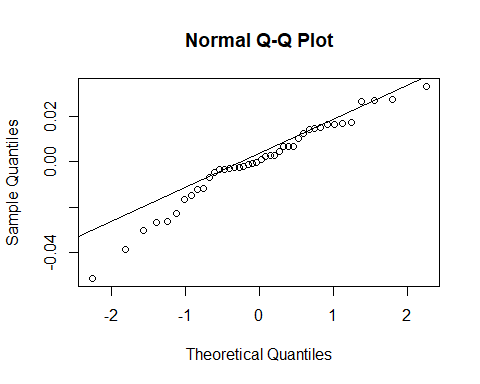
micro\_reg12

grande\_reg12

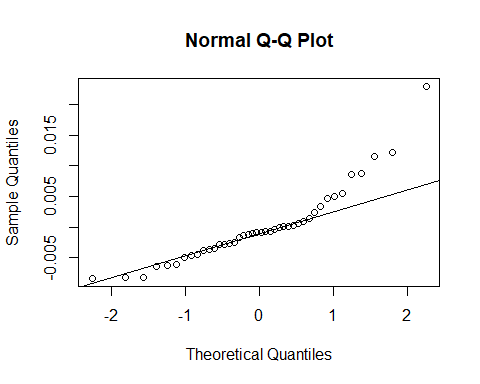
## 2.1 Normalidad de residuos

Normalidad de los residuos: Puedes usar un gráfico Q-Q de los residuos para verificar la normalidad. Los puntos deben caer aproximadamente en una línea recta si los residuos están normalmente distribuidos.

#MODELO MICROEMPRESAS  
# Generar residuos  
residuos1 <- residuals(micro\_reg12)  
  
# Crear gráfico Q-Q  
qqnorm(residuos1)  
qqline(residuos1)



#MODELO GRANDES EMPRESAS  
# Generar residuos  
residuos2 <- residuals(grande\_reg12)  
  
# Crear gráfico Q-Q  
qqnorm(residuos2)  
qqline(residuos2)



#Para usar la prueba de Shapiro-Wilk para verificar la normalidad, generalmente mirarías el valor p. Si el valor p es menor que 0.05, entonces tienes evidencia para rechazar la hipótesis nula y puedes concluir que tus datos probablemente no provienen de una distribución normal. Si el valor p es mayor que 0.05, entonces no tienes suficiente evidencia para rechazar la hipótesis nula y puedes concluir que tus datos pueden provenir de una distribución normal.  
  
shapiro.test(residuos1)

Shapiro-Wilk normality test  
  
data: residuos1  
W = 0.96282, p-value = 0.1862

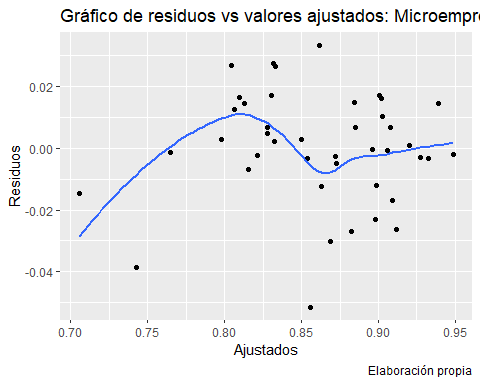
shapiro.test(residuos2)

Shapiro-Wilk normality test  
  
data: residuos2  
W = 0.87696, p-value = 0.0003155

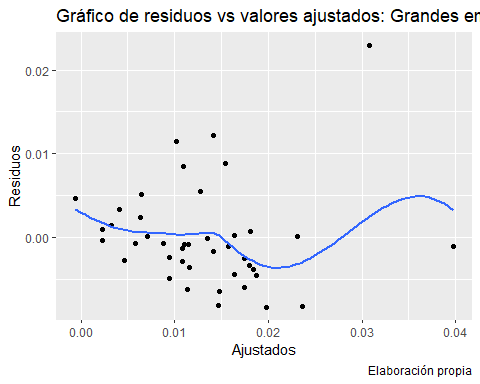
## 2.2 Independencia de los errores

Un gráfico de los residuos en función de los valores ajustados o el orden temporal puede ayudar a identificar si existe una correlación entre los errores.

#MODELO MICROEMPRESAS  
# Crear un gráfico de residuos vs valores ajustados  
ggplot(data = data.frame(Residuos = residuos1, Ajustados = fitted(micro\_reg12)), aes(x = Ajustados, y = Residuos)) +  
 geom\_point() +  
 geom\_smooth(se = FALSE) +  
 labs(title = "Gráfico de residuos vs valores ajustados: Microempresas",  
 caption = "Elaboración propia")



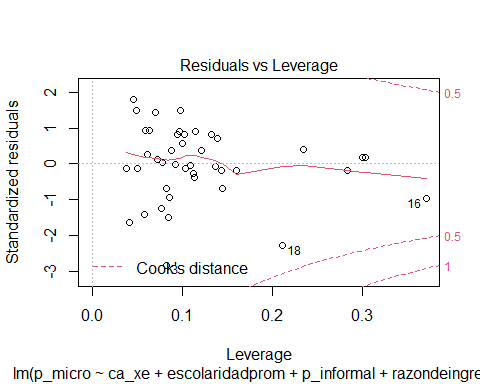
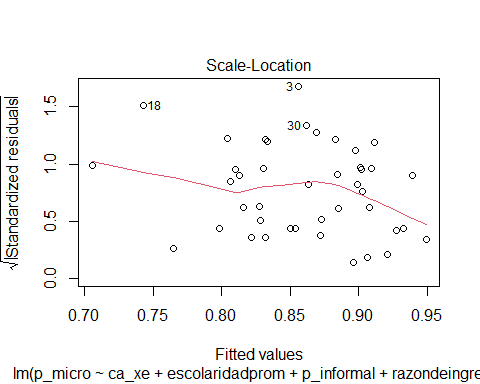
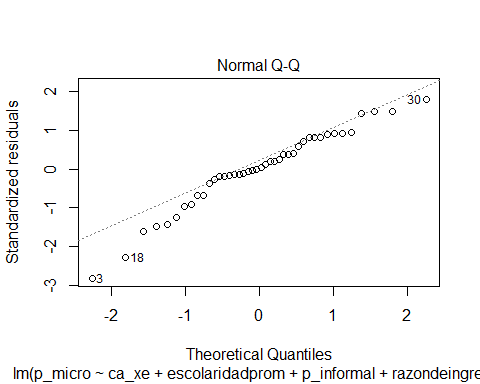
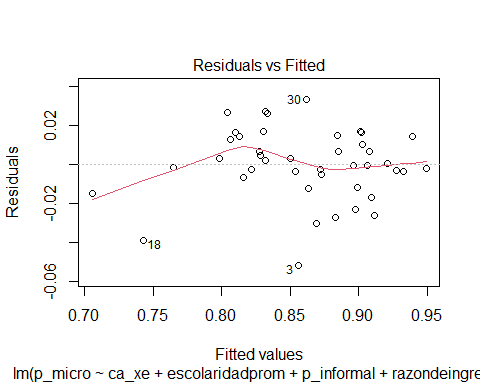
#MODELO GRANDES EMPRESAS  
# Crear un gráfico de residuos vs valores ajustados  
ggplot(data = data.frame(Residuos = residuos2, Ajustados = fitted(grande\_reg12)), aes(x = Ajustados, y = Residuos)) +  
 geom\_point() +  
 geom\_smooth(se = FALSE) +  
 labs(title = "Gráfico de residuos vs valores ajustados: Grandes empresas",  
 caption = "Elaboración propia")



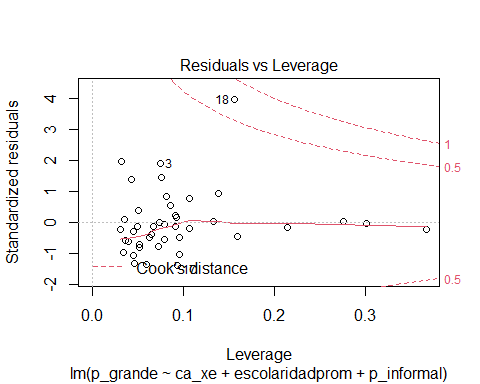
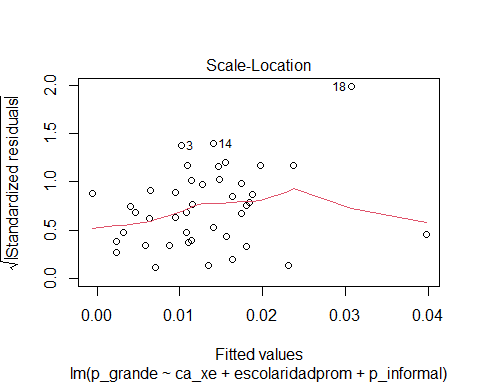
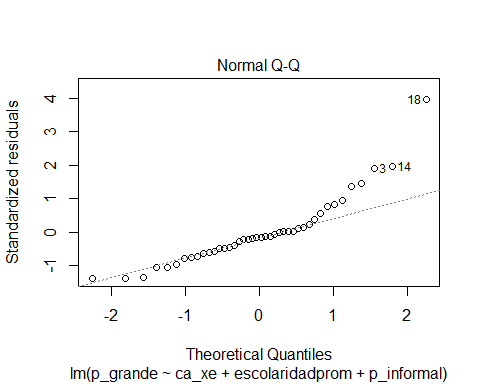
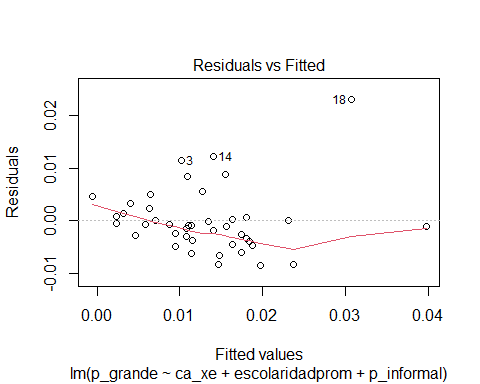
## 2.3 Linealidad de las relaciones entre las variables

Para verificar la linealidad, puedes hacer gráficos de residuos contra cada variable predictora.

# Gráfico de residuos contra cada variable predictora  
  
plot(micro\_reg12)



plot(grande\_reg12)



## 2.4 Homocedasticidad

Para comprobar la homocedasticidad (varianza constante de los errores), puedes hacer un gráfico de residuos contra valores ajustados. La dispersión de los puntos debe ser más o menos constante a lo largo del rango de valores ajustados.

#También puedes usar la prueba de Breusch-Pagan para comprobar la homocedasticidad. Si el valor p es significativo (por lo general, p < 0.05), entonces tienes evidencia de heterocedasticidad.  
  
library(lmtest)  
  
# Prueba de Breusch-Pagan Micro empresas  
bptest(micro\_reg12)

studentized Breusch-Pagan test  
  
data: micro\_reg12  
BP = 2.4197, df = 4, p-value = 0.6591

# Prueba de Breusch-Pagan Grandes empresas  
bptest(grande\_reg12)

studentized Breusch-Pagan test  
  
data: grande\_reg12  
BP = 6.1674, df = 3, p-value = 0.1037