

Taller 2 Regresión

1. Ajuste de modelo de regresión lineal múltiple

Considere las siguientes variables:

Y_i : i-ésima observación de la variable respuesta ‘Longitud de permanencia’ (DPERM).
 X_{i1} : i-ésima observación de la variable predictoria ‘Edad’ (EDAD).
 X_{i2} : i-ésima observación de la variable predictoria ‘Riesgo de infección’ (RINF).
 X_{i3} : i-ésima observación de la variable predictoria ‘Razón de rutina de cultivos’ (RRC).
 X_{i4} : i-ésima observación de la variable predictoria ‘Razón de rutina de rayos X del pecho’ (RRX).
 X_{i5} : i-ésima observación de la variable predictoria ‘Número de camas’ (NCAMAS).
 X_{i6} : i-ésima observación de la variable predictoria ‘Censo promedio diario’ (PDP).
 X_{i7} : i-ésima observación de la variable predictoria ‘Número de enfermeras’ (NENFERM).
 X_{i8} : i-ésima observación de la variable predictoria ‘Facilidades y servicios disponibles’ (FSD).

Se observa que se tienen 90 observaciones y $k = 8$ variables regresoras.

Se asume que el modelo de regresión lineal múltiple tiene la siguiente forma:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \beta_8 X_{i8} + E_i, \quad E_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2), i = 1, \dots, 90$$

Se ajusta un modelo de regresión lineal multiple:

Se escribe la ecuación ajustada:

$$\hat{Y}_i = -0.2084 + 0.1043X_{i1} + 0.3352X_{i2} + 0.0287X_{i3} + 0.0209X_{i4} - 0.0106X_{i5} + 0.0223X_{i6} - 0.006X_{i7} + 0.0041X_{i8}, i = 1, \dots, 90.$$

Se muestra la tabla de parámetros ajustados:

```
kable(summary(modelo1)$coefficients)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.2084670	1.9450831	-0.1071764	0.9149141
EDAD	0.1043806	0.0340139	3.0687638	0.0029235
RINF	0.3352223	0.1555968	2.1544291	0.0341775
RRC	0.0287063	0.0177746	1.6150188	0.1101955
RRX	0.0209817	0.0085074	2.4662932	0.0157625
NCAMAS	-0.0106992	0.0044559	-2.4011396	0.0186350
PDP	0.0223642	0.0048806	4.5822850	0.0000165
NENFERM	-0.0060256	0.0029319	-2.0552029	0.0430819
FSD	0.0041605	0.0195476	0.2128384	0.8319878

Se calcula la tabla ANOVA del modelo:

```
kable(myAnova(modelo1))
```

	Sum_of_Squares	DF	Mean_Square	F_Value	P_value
Model	217.568	8	27.19606	14.6536	5.0418e-13
Error	150.330	81	1.85592		

Con un p-value casi igual a cero, se concluye que al menos una de las covariable es significativa para explicar la variabilidad de la longitud de permanencia.

Del resumen del modelo se obtiene que el valor de R^2 es 0.5914, es decir que un 59.14% de la variabilidad total de la longitud de permanencia es explicada por el modelo. Se opina que este porcentaje de la variabilidad tan “bajo” puede deberse a que alguna de las covariables no sea significativa, o no es adecuado suponer que existe una relación lineal entre la longitud de permanencia y las covariables numéricas presentándose carencia de ajuste.

2. Coeficientes estandarizados:

Se calculan los coeficientes de un modelo de regresión lineal multiple con las variables estandarizadas, ordenados de menor a mayor:

Table 3: Tabla de coeficientes Estandarizados

	Coef.Std
FSD	0.0275649
RRC	0.1514452
RRX	0.2074015
RINF	0.2254678
EDAD	0.2354129
NENFERM	0.3839389
NCAMAS	0.9777531
PDP	1.6801779

De la tabla se concluye que la covariable que “más” aporta al modelo cuando los datos se encuentran estandarizados es el “Censo promedio diario”, indicando que un aumento unitario en el “Censo promedio diario” estandarizado aumentaría en 1.68 unidades en promedio la longitud de permanencia estandarizada, dado que el resto de covariables están en el modelo.

Además del “Censo promedio diario”, se encontró que dos de las covariables que estandarizadas más aportan al modelo son el número de camas y el número de enfermeras, sugiriendo que las primeras tres covariables que estandarizadas más aportan al modelo se relacionan con el tamaño de los hospitales.

3. Significancia individual de los parámetros del modelo:

Cada una de las pruebas t para la significancia individual de los parámetros del modelo tienen la siguiente forma:

$$H_0 : \beta_j = 0 \text{ vs. } H_1 : \beta_j \neq 0, j = 1, \dots, 8.$$

$$T_{0j} = \frac{\hat{\beta}_j}{s.e(\hat{\beta}_j)} \sim t_{81}, j = 1, \dots, 8.$$

$$p\text{-value}_j = P(|t_{81}| > |T_{0j}|).$$

Se crea una tabla de coeficientes asociados a cada covariable, que incluye el valor de su estadístico t y el valor p de su prueba de hipótesis.

	Estimación	Estadístico t	Valor p
EDAD	0.1043806	3.0687638	0.0029235
RINF	0.3352223	2.1544291	0.0341775
RRC	0.0287063	1.6150188	0.1101955
RRX	0.0209817	2.4662932	0.0157625
NCAMAS	-0.0106992	-2.4011396	0.0186350
PDP	0.0223642	4.5822850	0.0000165
NENFERM	-0.0060256	-2.0552029	0.0430819
FSD	0.0041605	0.2128384	0.8319878

De la tabla anterior se concluye que utilizando la prueba t los parámetros β_3 y β_8 no son significativos, es decir que la “Razón de rutina de cultivos” y “Facilidades y servicios disponibles” no ayudan a explicar la variabilidad de la “Longitud de permanencia” dado que las demás covariables no se encuentran en el modelo.

Prueba F para dos predictoras:

A partir de la tabla anterior se realiza una prueba F para las covariables “Censo promedio diario”, ####
Prueba F para test lineal general para el Censo promedio diario Se define el modelo completo:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \beta_8 X_{i8} + E_i, E_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2), i = 1, \dots, 90.$$

Se define el modelo reducido:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_7 X_{i7} + \beta_8 X_{i8} + E_i, E_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

A partir de los datos, se define:

$$F_{06} = \frac{SSR(X_6|X_1, X_2, X_3, X_4, X_5, X_7, X_8)}{MSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}, \quad (1)$$

$$F_{06} = \frac{SSE(X_1, X_2, X_3, X_4, X_5, X_7, X_8) - SSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}{MSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}, \quad (2)$$

$$F_{06} \sim f_{1,81} \quad (3)$$

Cálculo de valor p:

$$p\text{-value} = P(f_{1,81} > F_{06})$$

$$g.l.(SSE(X_1, X_2, X_3, X_4, X_5, X_7, X_8)) = 82$$

$$g.l.(SSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)) = 81$$

$$SSE(X_1, X_2, X_3, X_4, X_5, X_7, X_8) = 189.30$$

$$SSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) = 150.33$$

$$g.l.(SSR(X_6|X_1, X_2, X_3, X_4, X_5, X_7, X_8)) = 82 - 81 = 1$$

$$SSR(X_6|X_1, X_2, X_3, X_4, X_5, X_7, X_8) = 189.30 - 150.33 = 38.969$$

A partir de los valores previos se obtiene la siguiente tabla:

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
82	189.2993				
81	150.3299	1	38.9694680543685	20.9973358998904	1.6484475391443e-05

La columna ‘F’ presenta el valor de F_{06} y la columna $\text{Pr}(>F)$ su respectivo p-value.

Del resultado de la prueba anterior se concluye que el “Censo promedio diario” ayuda a explicar la “Longitud de permanencia” dado que el resto de las covariables se encuentran en el modelo.

Prueba F para test lineal general para la Edad Modelo completo:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \beta_8 X_{i8} + E_i, \quad E_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Modelo reducido:

$$Y_i = \beta_0 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \beta_8 X_{i8} + E_i, \quad E_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

A partir de los datos, se define:

$$F_{01} = \frac{SSR(X_1|X_2, X_3, X_4, X_5, X_6, X_7, X_8)}{MSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}, \quad (4)$$

$$F_{01} = \frac{SSE(X_2, X_3, X_4, X_5, X_6, X_7, X_8) - SSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}{MSE(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)}, \quad (5)$$

$$F_{01} \sim f_{1,81} \quad (6)$$

Se calcula el valor p:

$$p\text{-value} = P(f_{1,81} > F_{01})$$

```
lh2 <- linearHypothesis(modelo1, "EDAD = 0")
lh2[1, 3:6] <- c(" ", " ", " ", " ")
kable(lh2)
```

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
82	167.8077				
81	150.3299	1	17.4778182328601	9.41731151476297	0.00292351905114798

La columna ‘F’ presenta el valor de F_{06} y la columna $\text{Pr}(>F)$ su respectivo p-value.

Del resultado de la prueba anterior se concluye que la Edad ayuda a explicar la “Longitud de permanencia” dado que el resto de las covariables se encuentran en el modelo.

Sumas de cuadrados de tipo 1 y de tipo 2:

Sumas de cuadrados de tipo 1:

```
kable(anova(modelo1)[1:2])
```

	Df	Sum Sq
EDAD	1	13.2851764

	Df	Sum Sq
RINF	1	116.2244592
RRC	1	1.9362640
RRX	1	8.4433232
NCAMAS	1	31.8519622
PDP	1	37.8732896
NENFERM	1	7.8699352
FSD	1	0.0840737
Residuals	81	150.3298765

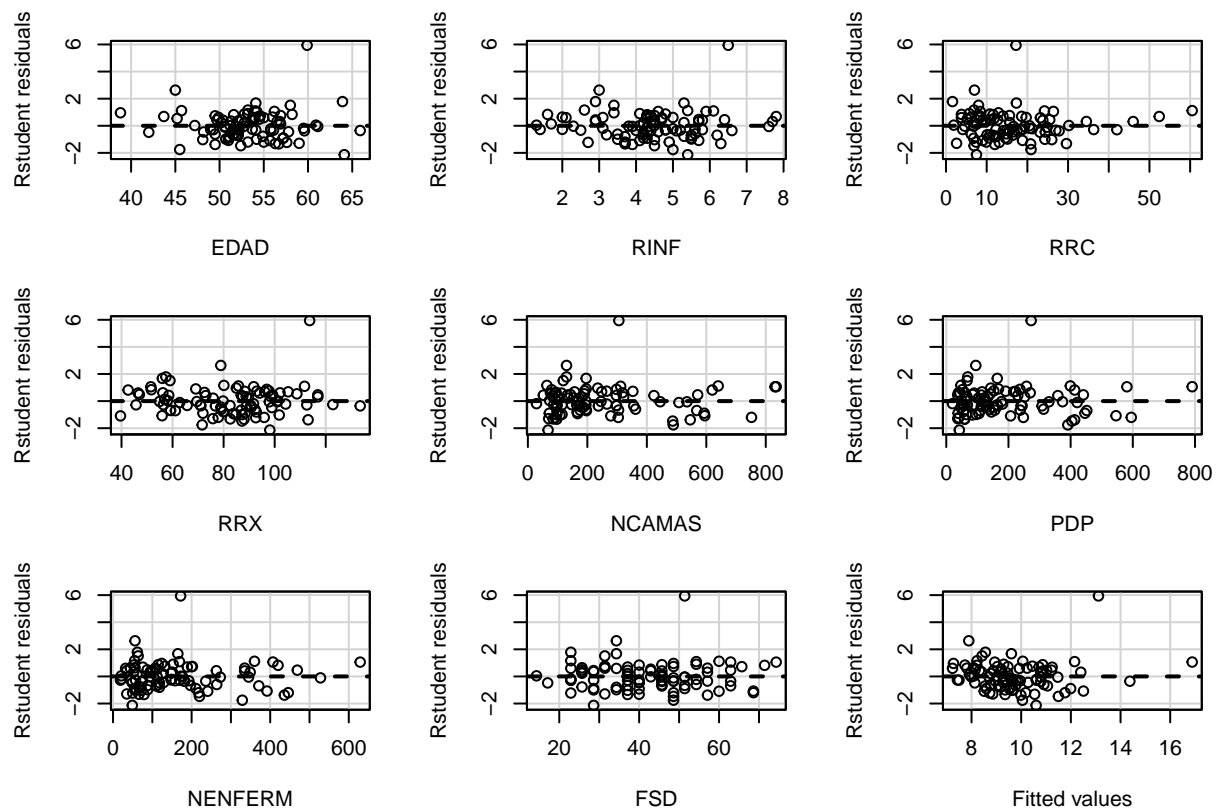
Sumas de cuadrados de tipo II:

```
kable(Anova(modelo1)[1:2])
```

	Sum Sq	Df
EDAD	17.4778182	1
RINF	8.6143936	1
RRC	4.8407814	1
RRX	11.2888497	1
NCAMAS	10.7002791	1
PDP	38.9694681	1
NENFERM	7.8391631	1
FSD	0.0840737	1
Residuals	150.3298765	81

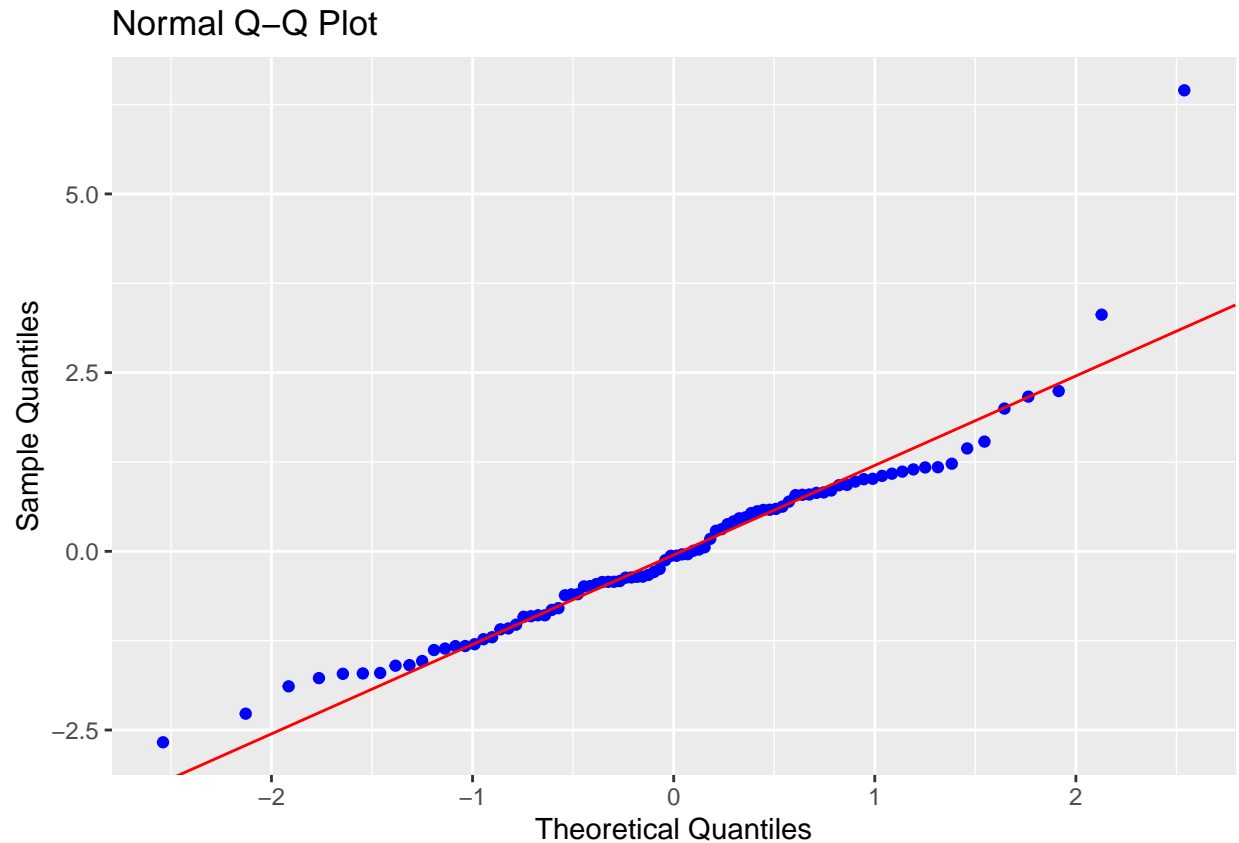
5. Gráfico de residuales estudentizados vs Valores ajustados

```
residualPlots(modelo1, type = "rstudent", tests = FALSE, quadratic = FALSE, fit = TRUE)
```



Gráfica de probabilidad normal para los residuales estudentizados.

```
ols_plot_resid_qq(model = modelo1)
```



```
normalidad <- shapiro.test(modelo1$residuals)
kable(t(normalidad[c("p.value" ,"statistic" ])))
```

p.value	statistic
1.83204082572401e-05	c(W = 0.914090225195161)

Según la gráfica anterior se observa un desajuste al inicio y al final de la gráfica que permite dudar sobre la hipótesis de normalidad sobre los errores del modelo.

7. Diagnóstico de la presencia de observaciones atípicas, de balanceo y/o influencias.

Se crea una tabla con las medidas de influencia:

```
influencias_modelo1 <- influence.measures(modelo1)
influencias_modelo1 <- (influencias_modelo1$infmat)
influencias_modelo1 <- cbind(influencias_modelo1, std.res = rstandard(modelo1), stud.res = rstudent(modelo1))
```

Observaciones atípicas mediante residuales estandarizados:

```
which(abs(influencias_modelo1[, "std.res"]) > 3 )
```

```
## 22
## 22
```

Según los residuales estandarizados la observación 22 es una observación atípica potencial.

Observaciones atípicas mediante residuales estudentizados:

```
which(abs(influencias_modelo1[, "stud.res"]) > 3 )
```

```
## 22
## 22
```

De nuevo se observa que según los residuales estudentizados la observación 22 es una observación potencialmente atípica.

Observaciones influyentes según las distancias de cook:

```
which(influencias_modelo1[, "cook.d"] > 1)
```

```
## named integer(0)
```

Según el criterio de las distancias de cook no se tienen observaciones influyentes.

Observaciones influyentes según los DFBetas:

Según el criterio de los DFBetas las siguientes observaciones son influyentes:

```
influencias_modelo1 <- as.data.frame(influencias_modelo1)
influenciales_dfbetas <- which(abs(influencias_modelo1[, 1:9]) > (2/sqrt(90)), arr.ind = TRUE)
kable(influencias_modelo1[unique(influenciales_dfbetas[, "row"]), 1:9])
```

	dfb.1__	dfb.EDAD	dfb.RINF	dfb.RRC	dfb.RRX	dfb.NCAM	dfb.PDP	dfb.NENF	dfb.FSD
22	-	0.7022585	0.5998931	-	0.6269032	-	0.7018199	-	0.3460527
	1.0554668			0.3475937		0.5409321		0.4175396	
40	0.3320045	-	-	-	0.0613822	0.1202810	-	-	0.1495173
		0.3901691	0.0255484	0.1062093			0.1114995	0.1478821	
55	0.6421624	-	-	-	0.0901574	0.0267936	0.0429387	-	0.0648500
		0.6371584	0.0369117	0.2969916				0.2630476	
60	-	0.2870884	-	0.0145697	0.1170199	-	-	-	0.1690624
	0.3328238		0.0651160			0.0072072	0.0690868	0.0210495	
74	-	0.4418508	-	-	-	0.1996062	-	0.0527205	-
	0.2495917		0.0371980	0.0448277	0.1013220		0.1651794		0.3295966

	dfb.1__	dfb.EDAD	dfb.RINF	dfb.RRC	dfb.RRX	dfb.NCAM	dfb.PDP	dfb.NENF	dfb.FSD
85	0.4375781	-	-	0.2715474	-	-	0.0825145	-	0.2150019
		0.4511554	0.3196276		0.1446658	0.0217175		0.1123090	
9	0.0902519	0.0177271	0.2902498	-	-	-	0.1217437	0.1375595	-
				0.0207575	0.3626614	0.1687336			0.1363087
25	-	0.0007032	-	0.7141728	-	0.6263768	-	-	-
	0.0078033		0.3519338		0.0278419		0.5035366	0.2313406	0.0308203
29	-	0.0610503	0.2457186	0.1203398	-	-	0.1201803	-	0.1256884
	0.0113679				0.4100964	0.0987417		0.2702677	
23	-	0.1479445	-	0.2292467	0.0331173	-	0.6778416	-	-
	0.0984735		0.1734543			0.3012805		0.3740053	0.1557559
36	-	0.0652693	-	-	0.2541075	0.3099330	-	0.0389461	-
	0.0849350		0.0762586	0.0385920			0.3642124		0.1382575
78	0.0207676	-	-	0.0000631	0.2943916	0.1226837	-	-	-
		0.0462453	0.0644256				0.0165751	0.0929246	0.2559914
32	0.0239483	-	-	0.0527249	0.0317064	-	0.0676942	0.3188200	0.0948138
		0.0381453	0.0773553			0.2471923			
83	-	0.0882718	-	-	-	0.2775600	-	0.2834297	-
	0.0500643		0.0370633	0.0714219	0.0857580		0.3307215		0.0971331
62	-	0.0983355	0.0042810	0.1604606	-	-	-	0.2856051	0.1011190
	0.0984058				0.1181551	0.0814065	0.1199827		

Observaciones influenciales según los DFFITS:

Según los dffits se tiene que las siguientes observaciones son influenciales:

```
influencias_dffits <- which(abs(influencias_modelo1[, "dffit"])>(2*(sqrt(9/90))))
kable(cbind(influencias_dffits,influencias_modelo1[ influencias_dffits,c("dffit")]))
```

influencias_dffits	
22	1.9386575
23	0.9924519
25	1.0390092
55	0.7983544
74	0.6795084
83	0.6484951
85	-0.7875551

Puntos influenciales según el COVRATIO:

Se verifica que $90 > 3(9)$, luego se puede concluir que una observación será cabdudata a ser inflencial sí $|COVRATIO_i - 1| > \frac{3(9)}{90}$, así, las siguientes observaciones son candidatas a ser influenciales:

```
influenciales_covratio <- which(abs(influencias_modelo1[, "cov.r"] -1)> (3*9/90))
kable(covratio(modelo1)[influenciales_covratio])
```

	x
3	1.3957971
14	1.5313864
21	1.4263831
22	0.0461483
23	1.8486647
25	1.8114627
34	1.3678717
48	1.3047817
52	1.3174769
55	0.5788661
58	1.5531299
76	1.3103089
83	1.3591291

Se ajusta un modelo para los datos sin incluir las observaciones cuyo ID no es 47 ni 112:

```
datos2 <- datos[-c(22,23),]
modelo2 <- lm(rlm_formula, data = datos2)
```

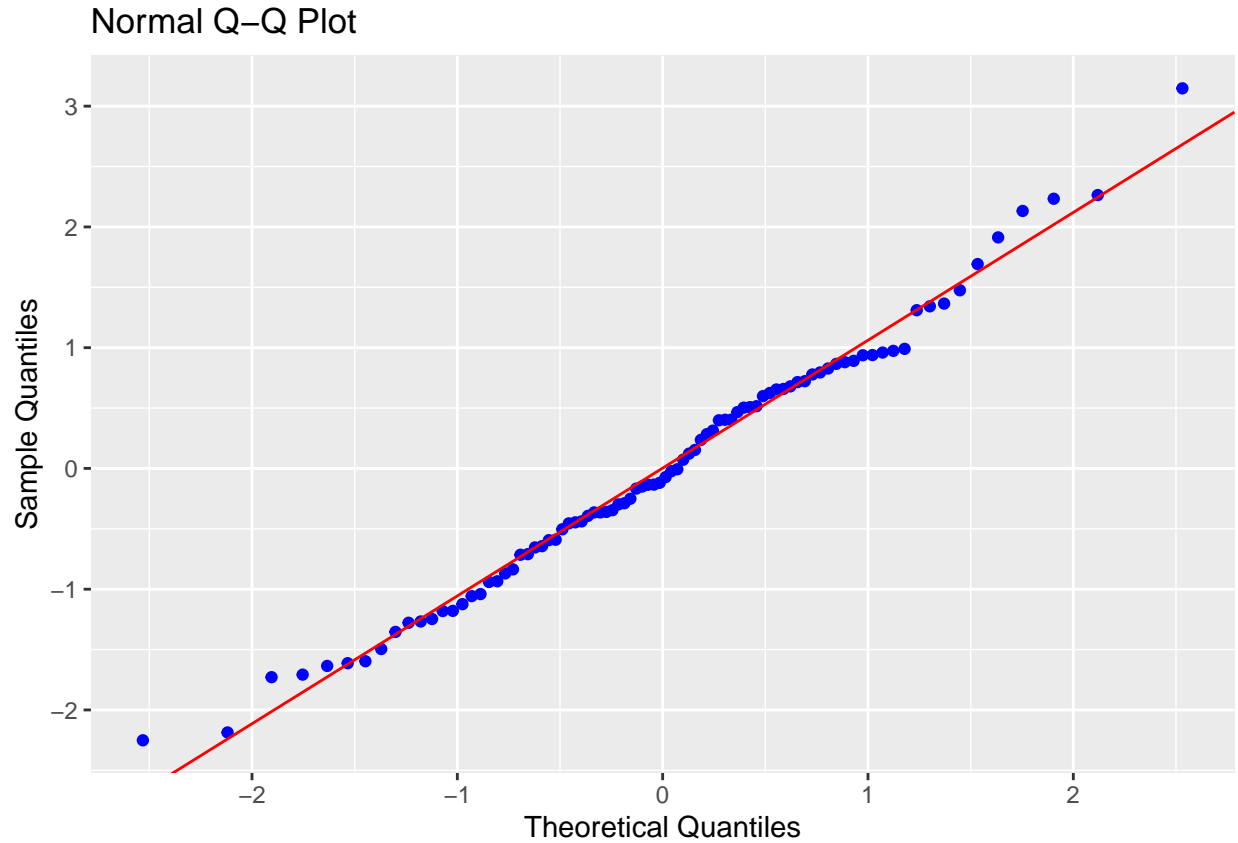
Se presenta la tabla de parámetros ajustados:

```
kable(summary(modelo2)$coefficients)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.9108818	1.6315588	1.1712001	0.2450388
EDAD	0.0750774	0.0284371	2.6401247	0.0099832
RINF	0.2979983	0.1297458	2.2967863	0.0242799
RRC	0.0273794	0.0149250	1.8344692	0.0703492
RRX	0.0158335	0.0070254	2.2537458	0.0269836
NCAMAS	-0.0063596	0.0038298	-1.6605654	0.1007650
PDP	0.0138772	0.0048219	2.8779749	0.0051460
NENFERM	-0.0031345	0.0025662	-1.2214791	0.2255373
FSD	0.0032664	0.0162206	0.2013754	0.8409227

Gráfico de normalidad para los residuales estudentizados:

```
ols_plot_resid_qq(model = modelo2)
```



Se observa una notoria mejoría en la gráfica de normalidad de los residuales estudentizados respecto a la gráfica anterior.

8. Diagnósticos de multicolinealidad:

Diagnósticos de multicolinealidad mediante la Matriz de correlación de las variables predictoras:

```
kable(cor(datos2[, numericas]))
```

	DPERM	EDAD	RINF	RRC	RRX	NCAMAS	PDP	NENFERM	FSD
DPERM	1.0000000	0.1096885	0.5889121	0.3658230	0.3924108	0.3894938	0.4253582	0.3584809	0.3764379
EDAD	0.1096885	1.0000000							
			0.0384364	0.2861650	0.0734373	0.1203036	0.1286958	0.1571305	0.0585730
RINF	0.5889121		1.0000000	0.5583204	0.4403791	0.4039787	0.4104520	0.4019360	0.4012267
			0.0384364						
RRC	0.3658230		0.5583204	1.0000000	0.3990202	0.1006151	0.0902737	0.1519585	0.1034185
			0.2861650						
RRX	0.3924108		0.4403791	0.3990202	1.0000000			0.0163626	0.0594719
			0.0734373			0.0210621	0.0063915		
NCAMAS	0.3894938		0.4039787	0.1006151		1.0000000	0.9825362	0.9228352	0.8227744
			0.1203036			0.0210621			
PDP	0.4253582		0.4104520	0.0902737		0.9825362	1.0000000	0.9231892	0.8152669
			0.1286958			0.0063915			

	DPERM	EDAD	RINF	RRC	RRX	NCAMAS	PDP	NENFERM	FSD
NENFERM	0.3764379	-	0.4019360	0.1519585	0.0163626	0.9228352	0.9231892	1.0000000	0.7958386
FSD	0.3764379	-	0.4012267	0.1034185	0.0594719	0.8227744	0.8152669	0.7958386	1.0000000

Diagnósticos de multicolinealidad mediante los VIF's:

```
kable(t(vif(modelo2)))
```

EDAD	RINF	RRC	RRX	NCAMAS	PDP	NENFERM	FSD
1.176584	2.151059	1.806731	1.376914	31.91805	31.61002	7.564419	3.290859

Diagnósticos de multicolinealidad mediante las proporciones de varianza:

```
ols_coll_diag(modelo2)
```

```
## Tolerance and Variance Inflation Factor
```

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

```
## Variables Tolerance VIF
## 1 EDAD 0.84991789 1.176584
## 2 RINF 0.46488734 2.151059
## 3 RRC 0.55348588 1.806731
## 4 RRX 0.72626193 1.376914
## 5 NCAMAS 0.03133024 31.918049
## 6 PDP 0.03163553 31.610025
## 7 NENFERM 0.13219787 7.564419
## 8 FSD 0.30387208 3.290858
```

```
##
```

```
##
```

```
## Eigenvalue and Condition Index
```

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

```
## Eigenvalue Condition Index intercept EDAD RINF
## 1 7.915777654 1.000000 7.849899e-05 8.950179e-05 0.0006106281
## 2 0.698769253 3.365735 5.212045e-04 6.414679e-04 0.0020301356
## 3 0.246418793 5.667739 1.944705e-03 3.233126e-03 0.0005206380
## 4 0.041315859 13.841667 1.135086e-02 1.093771e-02 0.4967510894
## 5 0.036450818 14.736459 2.048101e-03 4.885566e-03 0.0043693630
## 6 0.029206444 16.462942 2.720899e-04 3.346661e-03 0.4563946370
## 7 0.022970772 18.563463 1.914210e-02 2.951341e-02 0.0002436559
## 8 0.006124735 35.950352 4.892880e-03 3.106683e-04 0.0043104177
## 9 0.002965671 51.663673 9.597496e-01 9.470419e-01 0.0347694352
## RRC RRX NCAMAS PDP NENFERM FSD
## 1 0.002041220 0.0005843395 1.496514e-04 1.589298e-04 0.0006668985 4.338073e-04
## 2 0.026936931 0.0056664903 3.044047e-03 3.771330e-03 0.0144499980 2.734353e-05
## 3 0.495786196 0.0030218364 3.038221e-05 4.365977e-05 0.0026795661 3.783381e-03
## 4 0.239755798 0.0887123414 2.383517e-03 1.093003e-02 0.1025922493 3.603381e-02
```

```
## 5 0.080338776 0.3378944457 2.860204e-02 2.084011e-02 0.6135868447 2.196508e-02
## 6 0.044262645 0.5079756204 2.194668e-02 3.117629e-02 0.2428052382 4.573974e-03
## 7 0.003253317 0.0118498220 6.283190e-03 1.223564e-02 0.0205379916 9.195444e-01
## 8 0.014018709 0.0172199290 9.301177e-01 9.111510e-01 0.0006628867 1.078221e-02
## 9 0.093606407 0.0270751753 7.442804e-03 9.693012e-03 0.0020183268 2.856035e-03
```

9. Selección de modelo:

Tabla de todas las regresiones posibles:

```
regresiones <- ols_step_all_possible(modelo2)
```

```
head(regresiones)
```

##	Index	N	Predictors	R-Square	Adj. R-Square	Mallow's Cp
## 2	1	1	RINF	0.3468174	0.3392223	20.46379
## 6	2	1	PDP	0.1809296	0.1714055	46.99431
## 4	3	1	RRX	0.1539862	0.1441489	51.30337
## 5	4	1	NCAMAS	0.1517054	0.1418415	51.66814
## 8	5	1	FSD	0.1417055	0.1317253	53.26743
## 3	6	1	RRC	0.1338265	0.1237547	54.52753

Selección de modelo según el R_{adj}^2 :

Selección de modelo según el estadístico C_p

Selección de modelo según el metodo Stepwise:

Selección mediante el metodo forward:

Selección mediante el método backward:

10. ¿Cuál modelo sugiere para la variable respuesta?