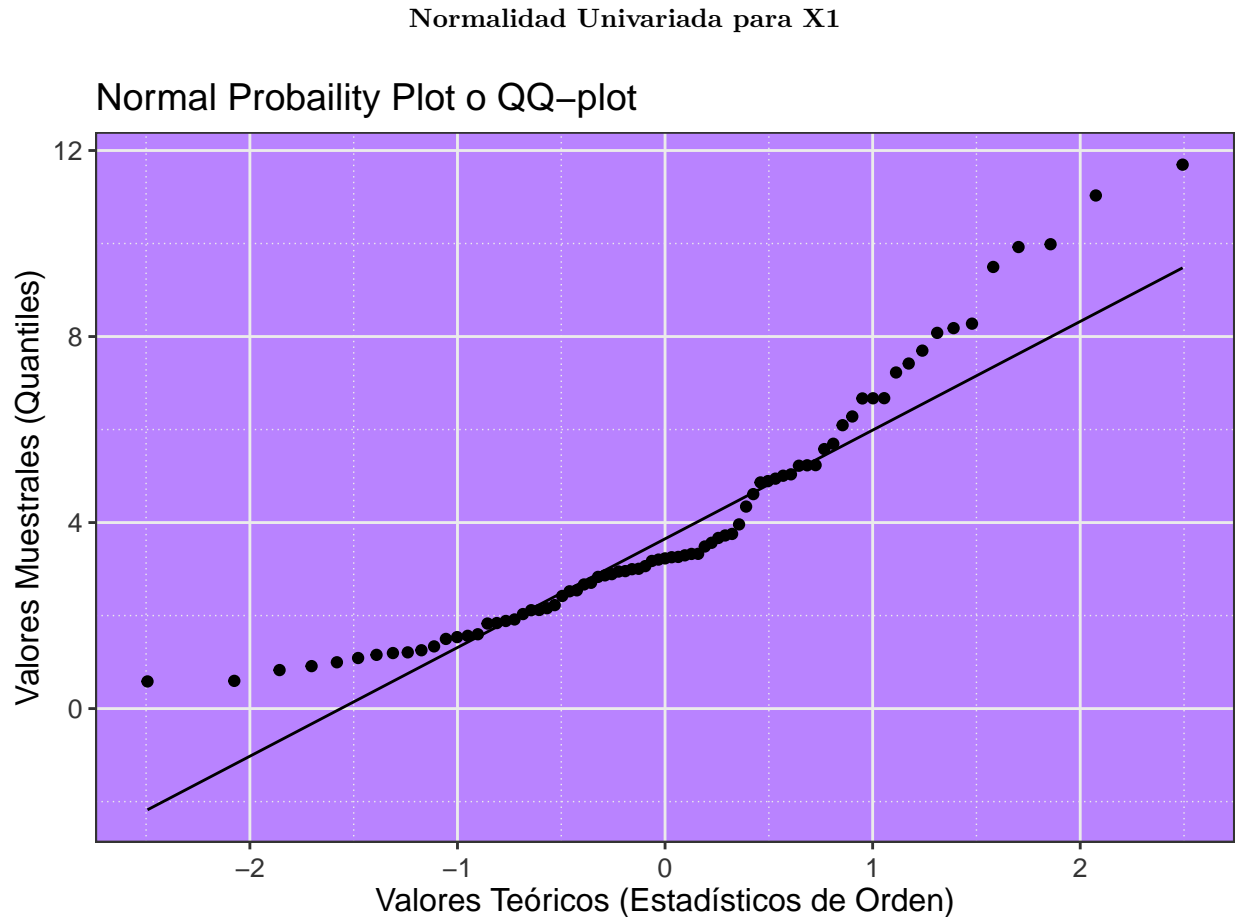


Daniela Pico Arredondo

Normalidad univariada para las variables 1, 2 y 3. En caso de no-Normalidad realizar una transformación de Box-Cox de ser posible



En la gráfica se observa un gran desajuste de los puntos con la recta, se observa una tendencia no lineal y las colas se alejan mucho de ella, por lo tanto se concluye que la variable X1 no es normal, Se utilizaran los siguientes test para Apoyar la respuesta obtenida en el gráfico de dispersión

$$H_0 = X_1 \sim Normal \quad vs \quad H_1 = X_1 \not\sim Normal$$

Resultados de la prueba de normalidad univariada vía la PH del coeficiente de Correlación:

Probability Plot Correlation Coefficient Test

data: bdparcialX1ppcc = 0.95191, $n = 79$, $p - value = 1e - 04$ alternative hypothesis : bdparcialX1 differs from a Normal distribution

Resultados de la prueba de normalidad univariada Test de Shapiro Wilk:

Shapiro-Wilk normality test

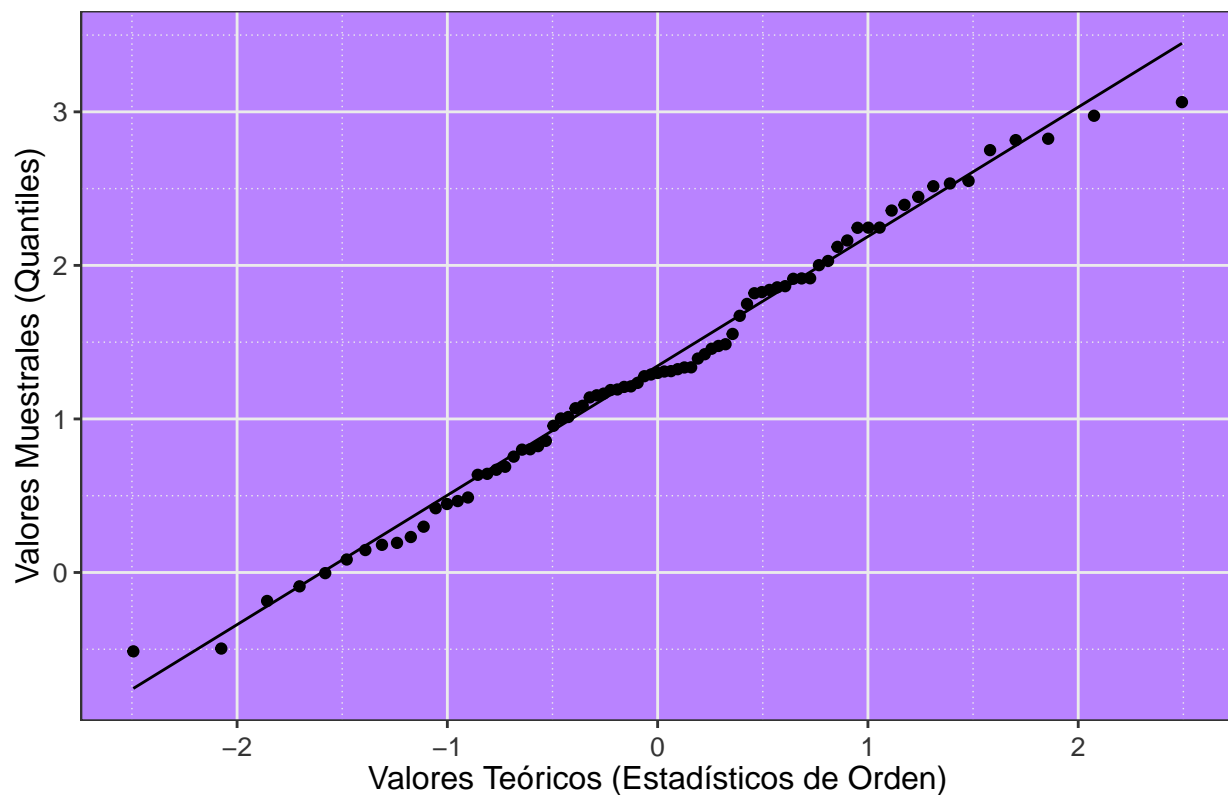
data: bdparcial\$X1 W = 0.90318, p-value = 1.858e-05

- La prueba de normalidad univariada del coeficiente de correlación se obtiene un valor $p\text{-value} = 2.2 \times 10^{-16} < 0.05$, Por tanto se rechaza la hipótesis nula y se concluye que X1 no es normal
- La prueba de normalidad univariada test de Shapiro Wilk se obtiene un $p\text{-value} = 1.859 \times 10^{-5} < 0.05$, Por tanto se rechaza la hipótesis nula y se concluye que X1 no es normal

Se decidió hacer una transformación Box Cox para normalizar las observaciones de la variable X1, se obtuvieron los siguientes resultados:

Datos transformados para X1

Normal Probability Plot o QQ-plot



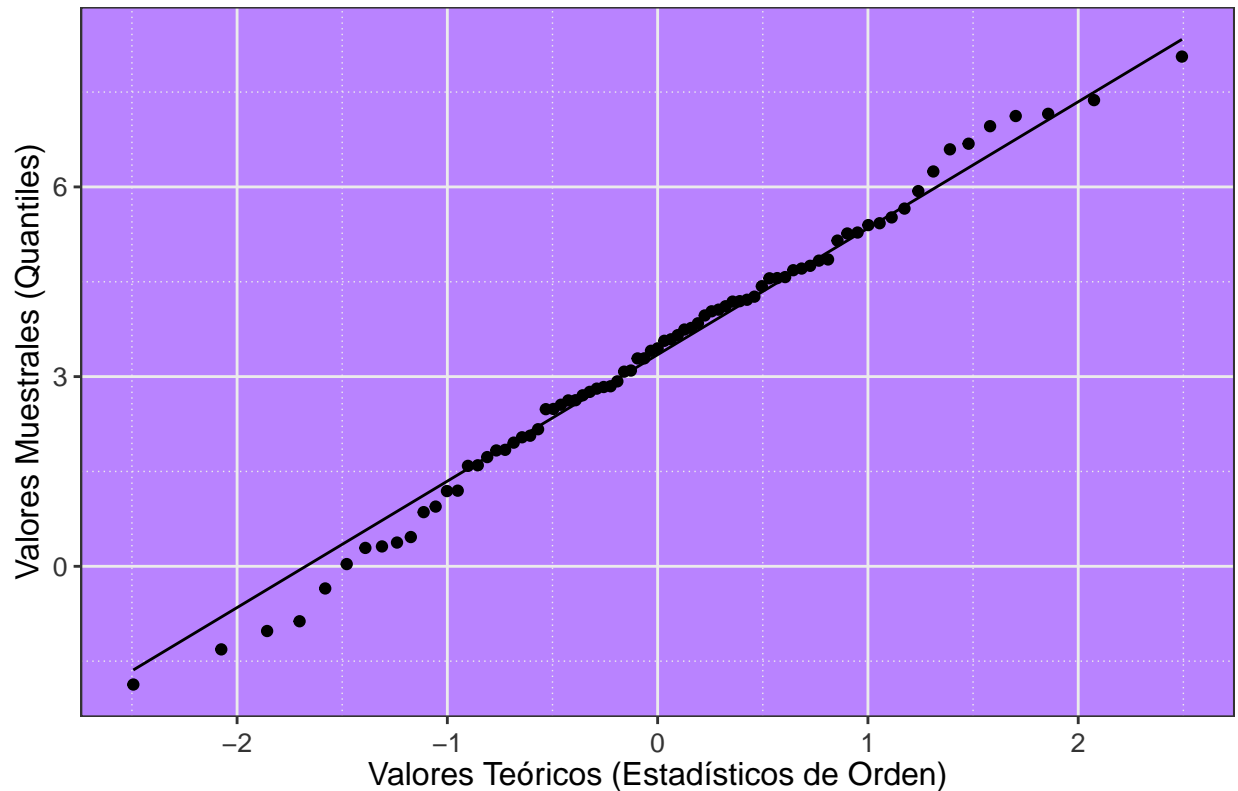
Shapiro-Wilk normality test

data: xt[, 1] W = 0.98757, p-value = 0.6445

De la gráfica de X1 transformada, considerando el 95% de los puntos del centro de la misma, observe que siguen la línea de referencia y no se alejan mucho de ella, por lo tanto se puede concluir que es normal.

Normalidad Univariada para X2

Normal Probability Plot o QQ-plot



De la gráfica, considerando el 95% de los puntos del centro de la misma, observe que siguen la línea de referencia y no se alejan mucho de ella, por lo tanto se puede concluir que la variable sigue una distribución normal.

$$H_0 = X_2 \sim Normal \quad vs \quad H_1 = X_2 \not\sim Normal$$

Resultados de la prueba de normalidad univariada vía la PH del coeficiente de Correlación:

Probability Plot Correlation Coefficient Test

data: bdparcialX2ppcc = 0.99668, $n = 79$, $p - value = 0.9284$ *alternative hypothesis : bdparcialX2 differs from a Normal distribution*

Resultados de la prueba de normalidad univariada Test de Shapiro Wilk:

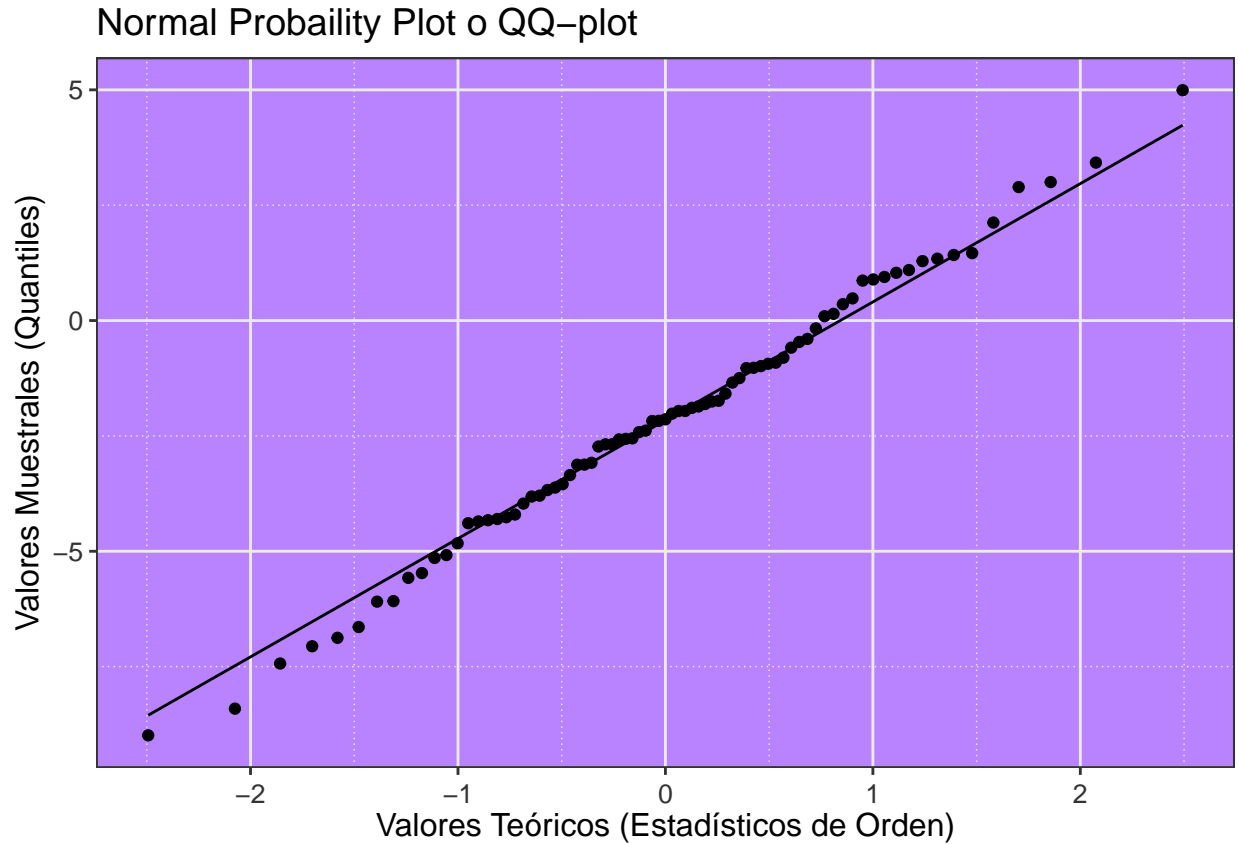
Shapiro-Wilk normality test

data: bdparcial\$X2 W = 0.99083, $p - value = 0.85$

- La prueba de normalidad univariada del coeficiente de correlación se obtiene un valor $p - value = 0.9275 > 0.05$, Por tanto no se rechaza la hipótesis nula y se concluye que X2 sigue una distribución normal

- La prueba de normalidad univariada test de Shapiro Wilk se obtiene un p-value=0.85>0.05, Por tanto no se rechaza la hipótesis nula y se concluye que X2 sigue una distribución normal

Normalidad univariada para X3



De la gráfica, considerando el 95% de los puntos del centro de la misma, observe que siguen la línea de referencia y no se alejan mucho de ella, por lo tanto se puede concluir que la variable sigue una distribución normal.

$$H_0 = X_3 \sim Normal \quad vs \quad H_1 = X_3 \not\sim Normal$$

Resultados de la prueba de normalidad univariada vía la PH del coeficiente de Correlación:

Probability Plot Correlation Coefficient Test

data: bdparcialX3 $p_{pcc} = 0.99763$, $n = 79$, $p - value = 0.9876$ *alternative hypothesis : bdparcialX3 differs from a Normal distribution*

Resultados de la prueba de normalidad univariada Test de Shapiro Wilk:

Shapiro-Wilk normality test

data: bdparcial\$X3 $W = 0.9951$, $p - value = 0.992$

- La prueba de normalidad univariada del coeficiente de correlación se obtiene un valor p-value= 0.9871>0.05, Por tanto no se rechaza la hipótesis nula y se concluye que X3 sigue una distribución normal
- La prueba de normalidad univariada test de Shapiro Wilk se obtiene un p-value=0.992>0.05, Por tanto no se rechaza la hipótesis nula y se concluye que X3 sigue una distribución normal

Normalidad bivariada para las variables 1 con 2, la 2 con 3 y la 1 con 4

X1 con X2

Table 1: Salidas Básicas PH NM- Royston

Test		H	p value	MVN	Test		Variable	Statistic	p value	Normality
Royston		17.80666	0.0001359	NO	Shapiro-Wilk		X1	0.9032	<0.001	NO
					Shapiro-Wilk		X2	0.9908	0.85	YES
n		Mean	Std.Dev	Median	Min	Max	25th	75th	Skew	Kurtosis
X1	79	3.951627	2.598114	3.22975	0.58402	11.69560	2.07403	5.22669	1.0476442	0.4621873
X2	79	3.332741	2.148609	3.44649	-1.86969	8.06127	1.99815	4.69584	-0.1764894	-0.3133897

$$H_0 = \underline{X_1, X_2} \sim Normal \quad vs \quad H_1 = \underline{X_1, X_2} \not\sim Normal$$

En la tabla 1, están las Salidas básicas de la función mvn usando la prueba multivariada de Royston que indican un valor-p=0.0001359<0.05 por tanto se rechaza la hipótesis nula y se concluye que no hay normalidad bivariada entre en X1 y X2, esto tiene sentido ya que X1 no es normal univariada. Además se visualiza el resumen de descriptivo de los datos.

Ahora veremos como se comporta la normalidad bivariada con X1 transformada

X1Transformada con X2

Table 2: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Royston	0.24871	0.88307	YES

Table 3: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	X1transform	0.9876	0.6445	YES
2	Shapiro-Wilk	X2	0.9908	0.8500	YES

$$H_0 = \underline{X_1}, \underline{X_2} \sim Normal \quad vs \quad H_1 = \underline{X_1}, \underline{X_2} \not\sim Normal$$

Según la prueba multivariada de Royston se observa un valor-p=0.88307>0.05, por lo tanto no se rechaza la hipótesis nula y se concluye que hay normalidad bivariada entre en X1 transformada y X2, esto tiene sentido ya que al transformar X1 se solucionó el problema de la normalidad univariada, aunque no siempre la normalidad univariada implica la normalidad multivariada.

X2 con X3

Table 4: Salidas Básicas PH NM- Royston

Test	H	p value	MVN	Test	Variable	Statistic	p value	Normality	
Royston	0.0359509	0.982328	YES	Shapiro-Wilk	X2	0.9908	0.850	YES	
				Shapiro-Wilk	X3	0.9951	0.992	YES	
n	Mean	Std.Dev	Median	Min	Max	25th	75th	Skew	Kurtosis
X2 79	3.332741	2.148609	3.44649	-1.86969	8.06127	1.998150	4.695840	-0.1764894	-0.3133897
X3 79	-2.156236	2.786198	-2.14004	-8.98955	4.99254	-3.890415	-0.431395	-0.0451146	-0.1330441

$$H_0 = \underline{X_2}, \underline{X_3} \sim Normal \quad vs \quad H_1 = \underline{X_2}, \underline{X_3} \not\sim Normal$$

Según la prueba multivariada de Royston se obtiene un valor-p=0.982328>0.05, por lo tanto no se rechaza la hipótesis nula y concluye que hay normalidad bivariada entre en X2 y X3.

X1 con X4

Table 5: Salidas Básicas PH NM- Royston

Test		H	p value	MVN	Test		Variable	Statistic	p value	Normality
Royston		18.04662	0.0001205	NO	Shapiro-Wilk		X1	0.9032	<0.001	NO
					Shapiro-Wilk		X4	0.9869	0.599	YES
n		Mean	Std.Dev	Median	Min	Max	25th	75th	Skew	Kurtosis
X1	79	3.951627	2.598114	3.22975	0.58402	11.6956	2.07403	5.22669	1.0476442	0.4621873
X4	79	3.173935	2.682228	3.18213	-3.16784	8.6487	1.59281	5.18977	-0.2110327	-0.4690189

$$H_0 = \underline{X_1}, \underline{X_4} \sim Normal \quad vs \quad H_1 = \underline{X_1}, \underline{X_4} \not\sim Normal$$

Usando la prueba multivariada de Royston que indican un valor-p=0.0001205<0.05, por lo tanto se rechaza la hipótesis nula y se concluye que no hay normalidad bivariada entre en X1 y X4, esto tiene sentido ya que X1 no es normal univariada. Además se visualiza el resumen de descriptivo de los datos.

Ahora veremos como se comporta la normalidad bivariada con X1 transformada

X1Transformada con X4

Table 6: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Royston	0.24871	0.88307	YES

Table 7: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	X1transform	0.9876	0.6445	YES
2	Shapiro-Wilk	X2	0.9908	0.8500	YES

$$H_0 = \underline{X_1}, \underline{X_4} \sim Normal \quad vs \quad H_1 = \underline{X_1}, \underline{X_4} \not\sim Normal$$

Según la prueba multivariada de Royston se obtiene un valor-p=0.88307>0.05, por tanto no se rechaza la hipótesis nula y se concluye que hay normalidad bivariada entre en X1transformada y X4, esto tiene sentido ya que al transformar X1 se solucionó el problema de la normalidad univariada, aunque no siempre la normalidad univariada implica la normalidad multivariada.

Normal Trivariada X2,X3 y X4

Table 8: Salidas Básicas PH NM- Royston

Test		H	p value	MVN	Test	Variable	Statistic	p value	Normality	
Royston		0.312766	0.9578078	YES	Shapiro-Wilk	X2	0.9908	0.850	YES	
					Shapiro-Wilk	X3	0.9951	0.992	YES	
					Shapiro-Wilk	X4	0.9869	0.599	YES	
n		Mean	Std.Dev	Median	Min	Max	25th	75th	Skew	Kurtosis
X2	79	3.332741	2.148609	3.44649	-1.86969	8.06127	1.998150	4.695840	-0.1764894	-0.3133897
X3	79	-2.156236	2.786198	-2.14004	-8.98955	4.99254	-3.890415	-0.431395	-0.0451146	-0.1330441
X4	79	3.173935	2.682228	3.18213	-3.16784	8.64870	1.592810	5.189770	-0.2110327	-0.4690189

$$H_0 = \underline{X_2, X_3, X_4} \sim Normal \quad vs \quad H_1 = \underline{X_2, X_3, X_4} \not\sim Normal$$

Según la prueba multivariada de Royston se obtiene un valor-p=0.9578078>0.05, por tanto no se rechaza la hipótesis nula y se concluye que hay normalidad bivariada entre en X2 y X3, Además se observa que X2, X3 y X4 son normales univariadas.

Normal Multivariada

$$H_0 = \underline{X_1, X_2, X_3, X_4, X_5} \sim Normal \quad vs \quad H_1 = \underline{X_1, X_2, X_3, X_4, X_5} \not\sim Normal$$

Table 9: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Mardia Skewness	66.8567259314768	0.000938908030120118	NO
2	Mardia Kurtosis	-0.704660353457317	0.481021629668812	YES
3	MVN	NA	NA	NO

- En la prueba de normalidad multivariada Mardia-skewness se obtiene un p-value=0.00093890<0.05, Por tanto se rechaza la hipótesis nula y se concluye que el conjunto de datos no sigue una distribución normal multivariada

En la tabla 10, se encuentran los resultados de las pruebas de normalidad univariadas para cada una de las variables del conjunto de datos p-variados.

Table 10: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	X1	0.9032	<0.001	NO
2	Shapiro-Wilk	X2	0.9908	0.85	YES
3	Shapiro-Wilk	X3	0.9951	0.992	YES
4	Shapiro-Wilk	X4	0.9869	0.599	YES
5	Shapiro-Wilk	X5	0.8178	<0.001	NO

Table 11: Salidas Básicas PH NM- Royston

Test	H	p value	MVN	Test	Variable	Statistic	p value	Normality
Royston	49.74698	0	NO	Shapiro-Wilk	X1	0.9032	<0.001	NO
				Shapiro-Wilk	X2	0.9908	0.85	YES
				Shapiro-Wilk	X3	0.9951	0.992	YES
				Shapiro-Wilk	X4	0.9869	0.599	YES
				Shapiro-Wilk	X5	0.8178	<0.001	NO

	n	Mean	Std.Dev	Median	Min	Max	25th	75th	Skew	Kurtosis
X1	79	3.951627	2.598114	3.22975	0.58402	11.69560	2.074030	5.226690	1.0476442	0.4621873
X2	79	3.332741	2.148609	3.44649	-1.86969	8.06127	1.998150	4.695840	-0.1764894	-0.3133897
X3	79	-2.156236	2.786198	-2.14004	-8.98955	4.99254	-3.890415	-0.431395	-0.0451146	-0.1330441
X4	79	3.173935	2.682228	3.18213	-3.16784	8.64870	1.592810	5.189770	-0.2110327	-0.4690189
X5	79	1.746180	1.799823	0.86253	0.07510	8.30886	0.426035	2.955365	1.2816303	1.0045934

Usando la prueba multivariada de Royston que indican un valor- $p=0<0.05$, por lo tanto se rechaza la hipótesis nula y se concluye que no hay normalidad Multivariada entre en X1,X2,X3 X4, esto tiene sentido ya que X1 y X5 no son normales univariadas.

Test de Shapiro Wilk Multivariada

FALSE Warning: package 'RVAideMemoire' was built under R version 4.0.5

Multivariate Shapiro-Wilk normality test

data: (X1,X2,X3,X4,X5) $W = 0.85838$, $p\text{-value} = 3.579e-07$

- La prueba de Shapiro Wilk indica que $p\text{-value}=3.57e07<0.05$ por lo tanto se rechaza la hipótesis nula y se concluye que no hay normalidad multivariada

Test de shapiro wilk multivariada con X1 transformada

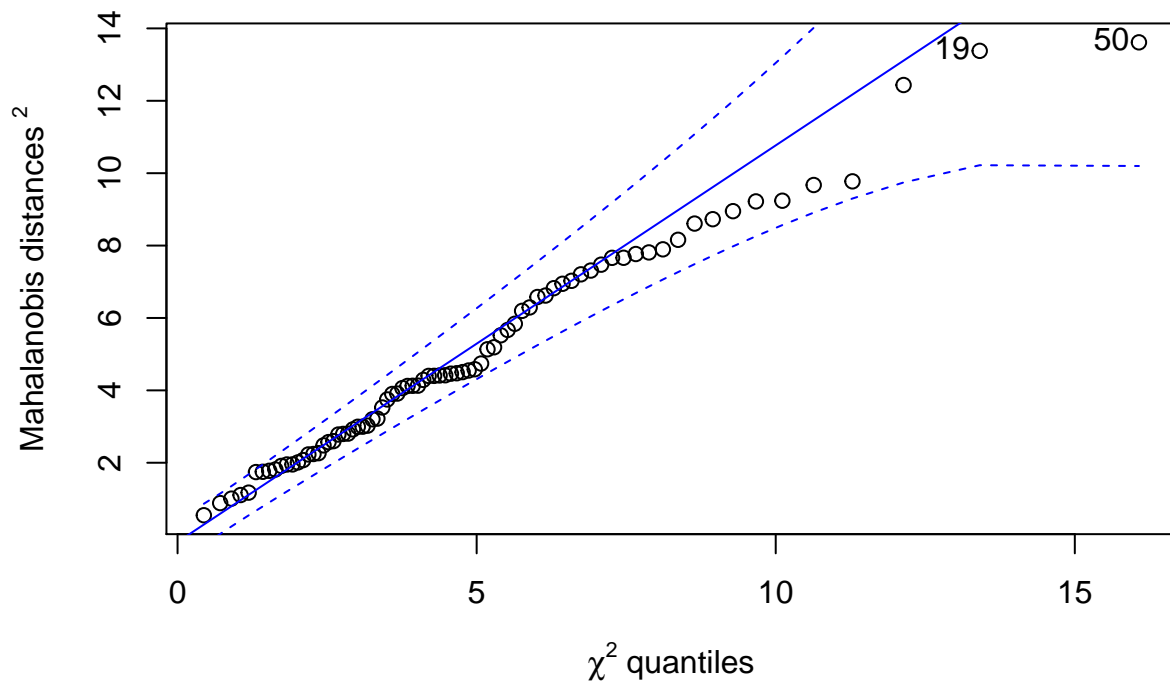
Se quiere ver el comportamiento de la normalidad multivariada si se usara la transformación en la variada X1

Multivariate Shapiro-Wilk normality test

data: (X1transform,X2,X3,X4,X5) $W = 0.96462$, $p\text{-value} = 0.02743$

- La prueba de Shapiro Wilk indica que $p\text{-value}=0.02743<0.05$ por lo tanto se rechaza la hipótesis nula y se concluye que no hay normalidad multivariada, claramente se ve que la transformación de X1 no es suficiente para asegurar la normalidad multivariada entre las variables, esto debido a que X5 no es normal univariada.

Gráfico QQ-Plot para Normalidad Multivariada



[1] 50 19

Se observa que los puntos no caen en una línea recta de pendiente 1. Por lo que no se apoya la normalidad Multivariada en estos datos