

Nombre: Salim, Nasim.

a) Para este problema el modelo a utilizar será el de ANOVA de 2 factores, que consiste en lo siguiente:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

con  $i=1,\dots,a$  (niveles de A);  $j=1,\dots,b$  (niveles de B);  $k=1,\dots,n$  (replicas por celda)

$y_{ijk}$ : Respuesta de la fila  $i$ , a la columna  $j$  y en la réplica  $k$

$\mu$ : Media global

$\tau_i$ : Efecto del factor en la fila  $i$

$\beta_j$ : Efecto del factor en la columna  $j$

$(\tau\beta)_{ij}$ : Efecto de la interacción entre la fila  $i$  y la columna  $j$

$\varepsilon_{ijk}$ : Error aleatorio del modelo

Supuestos: Homogeneidad de varianzas,  $\varepsilon_{ijk}$  independientes y  $\varepsilon_{ijk} \sim N(0, \sigma^2)$

Restricciones:

$$\sum_{i=1}^a \tau_i = 0 \quad \sum_{j=1}^b \beta_j = 0 \quad \sum_{i=1}^a (\tau\beta)_{ij} = 0 \quad \sum_{j=1}^b (\tau\beta)_{ij} = 0$$

Hipótesis:

Para el factor fila:

$$H_0) \tau_1 = \tau_2 = \dots = \tau_a = 0$$

$$H_1) \exists \tau_i \neq 0 \text{ para al menos un valor } i.$$

Para el factor columna:

$$H_0) \beta_1 = \beta_2 = \dots = \beta_b = 0$$

$$H_1) \exists \beta_j \neq 0 \text{ para al menos un valor } j.$$

Para la intersección:

$$H_0) (\tau\beta)_{ij} = 0 \quad \forall i, j$$

$$H_1) \exists (\tau\beta)_{ij} \neq 0$$

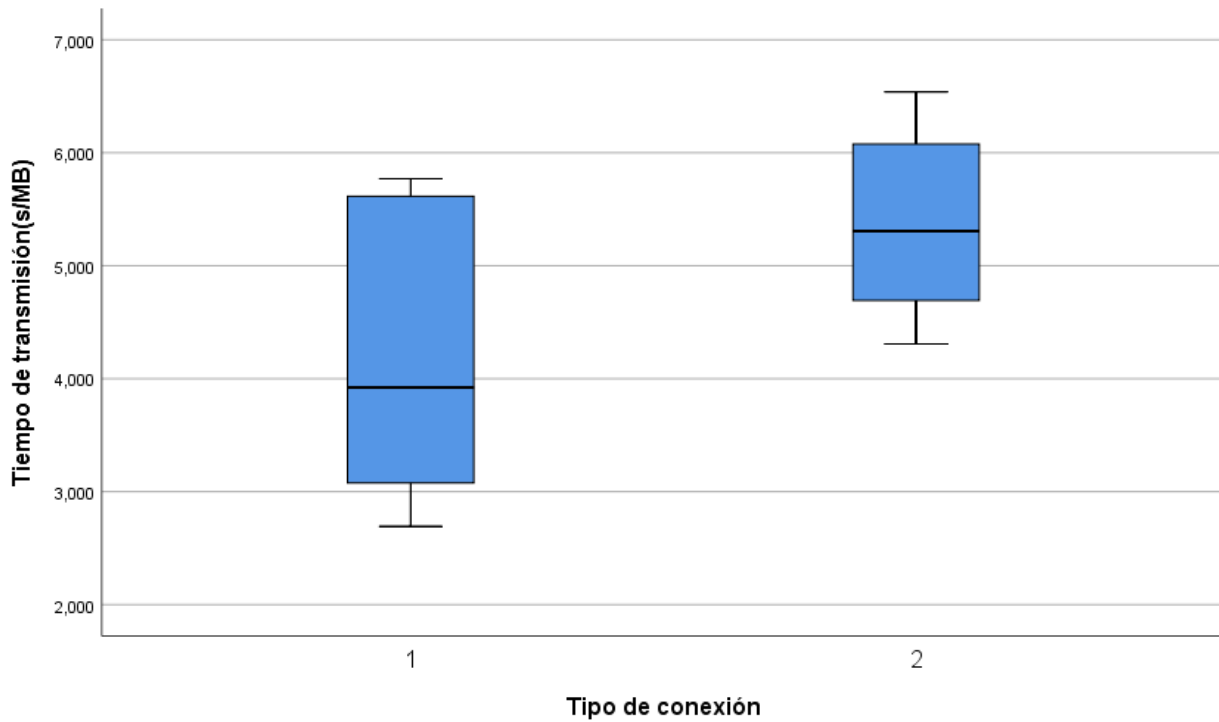
Para este caso:

Fila: Tipo de conexión

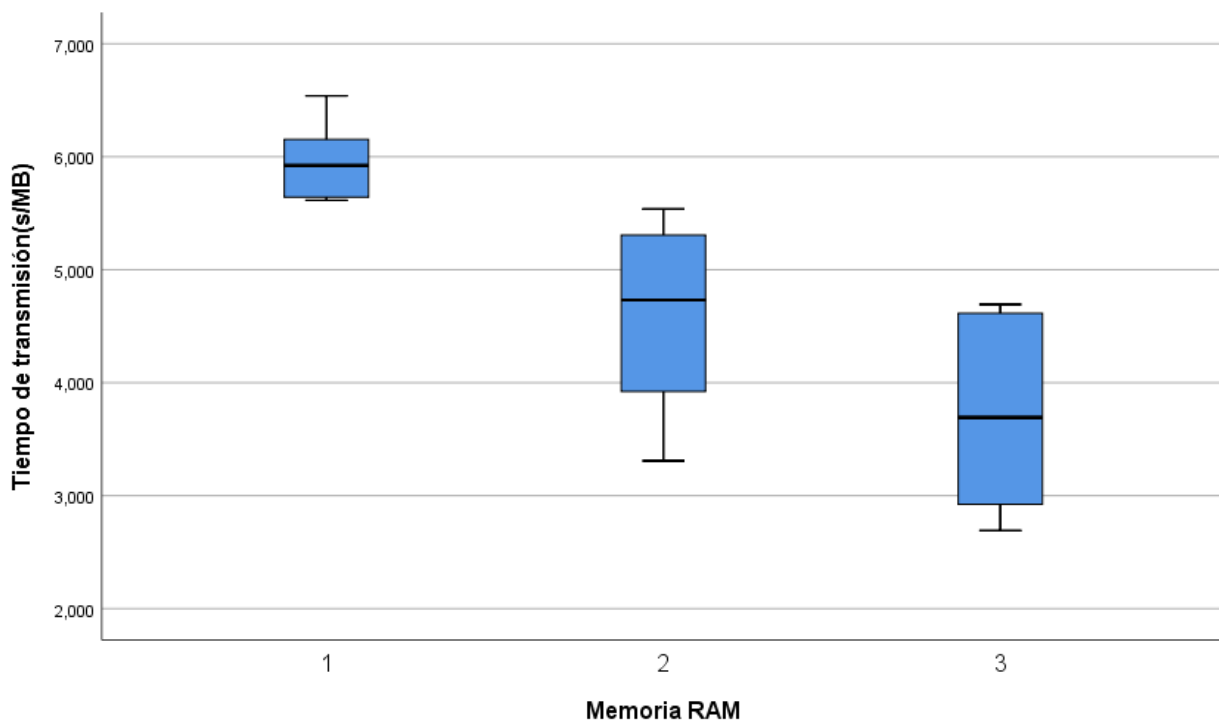
Columna: Cantidad de memoria RAM

Variable dependiente: Tiempo de transmisión (s/MB)

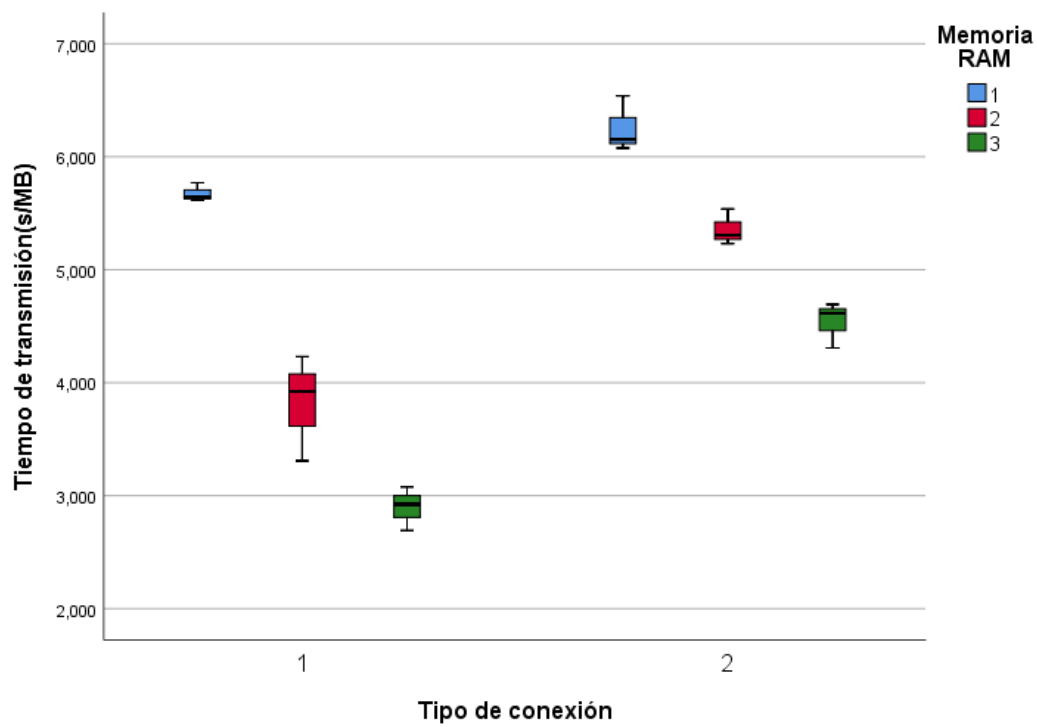
## Análisis exploratorio VarDep vs Factor fila:



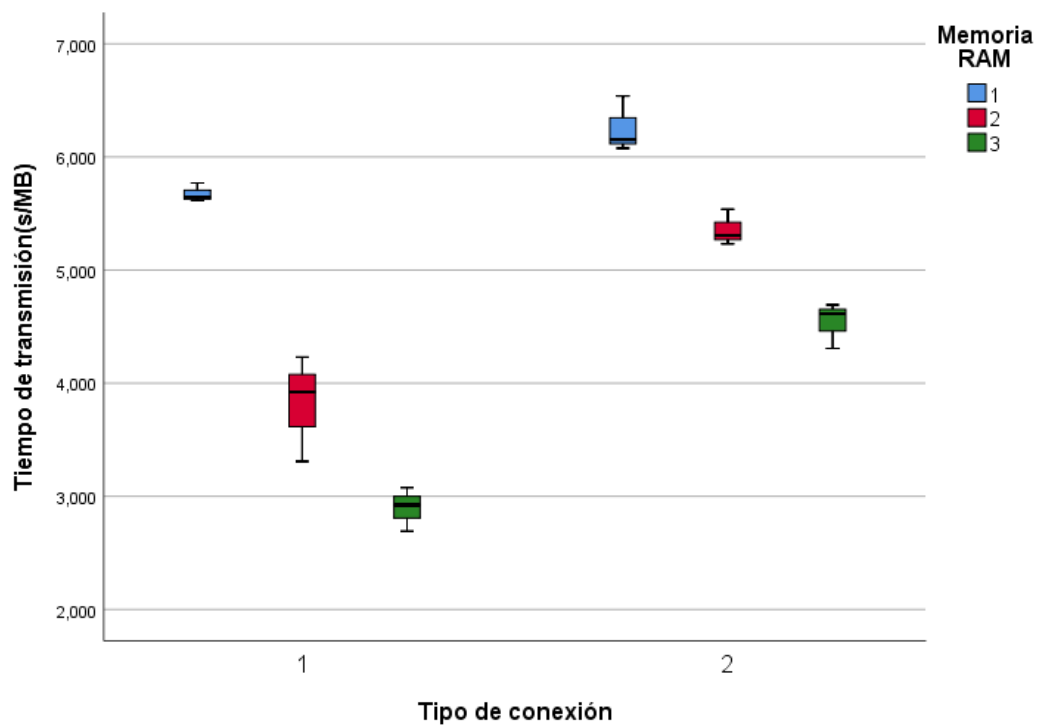
A simple vista parece que existe una diferencia significativa entre los tipos de conexión



Gráficamente se puede observar que existe diferencia significativa, principalmente en la memoria RAM 1, la de 64 Mb.



Nuevamente se puede observar gran diferencia entre las distintas cajas del gráfico.



Luego de procesar con el software, se construye la siguiente tabla ANOVA:

### Pruebas de efectos inter-sujetos

Variable dependiente: Tiempo de transmisión(s/MB)

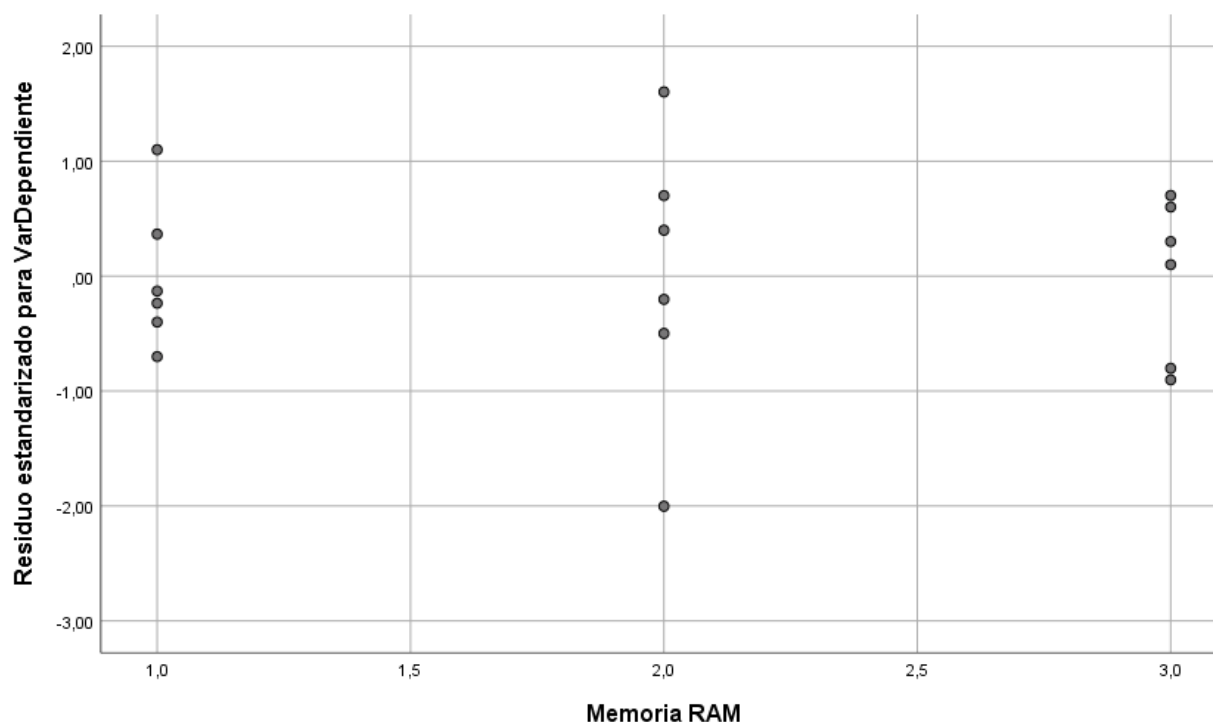
Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	23,509 <sup>a</sup>	5	4,702	71,747	,000
Intersección	407,447	1	407,447	6217,384	,000
Columna	15,417	2	7,708	117,627	,000
Fila	7,068	1	7,068	107,846	,000
Columna * Fila	1,025	2	,512	7,818	,007
Error	,786	12	,066		
Total	431,742	18			
Total corregido	24,296	17			

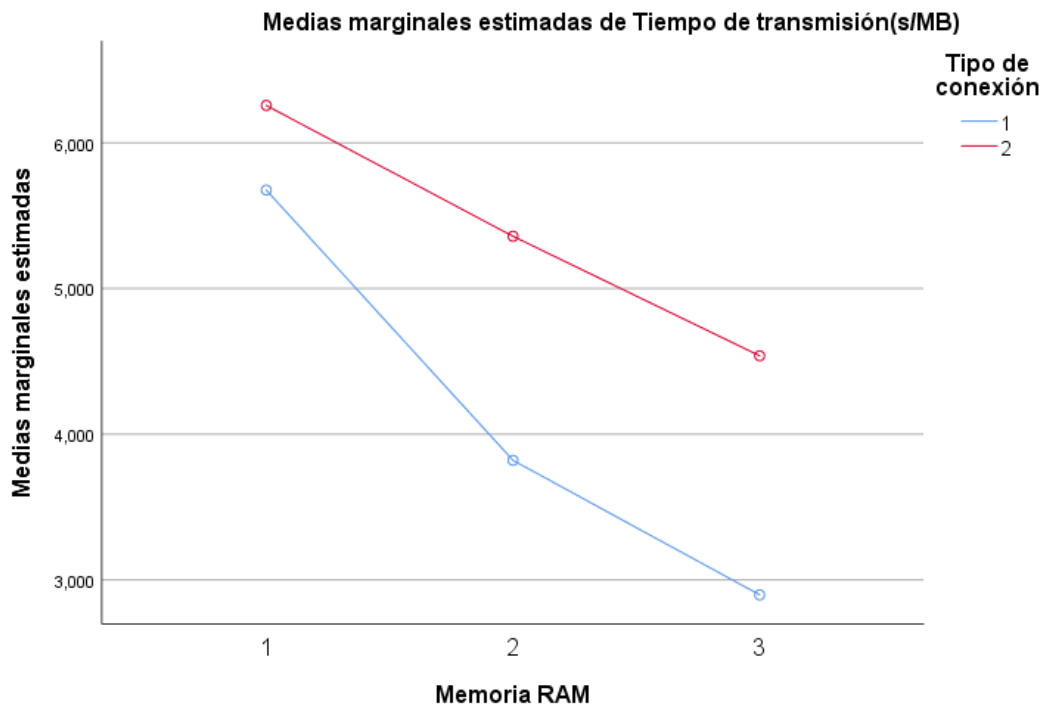
a. R al cuadrado = ,968 (R al cuadrado ajustada = ,954)

Primero, observando la celda roja, y siendo el  $p\text{-value} < 0,01$ , se puede concluir que existe un efecto por la intersección de los factores AB.

Y mirando las 2 celdas violetas se puede decir que hay un efecto significativo por los factores fila y columna respectivamente dados los  $p\text{-value}$  los cuales son muy pequeños.

Gráficos de perfiles:





Prueba de igualdad de varianzas:

**Prueba de igualdad de Levene de varianzas de error<sup>a,b</sup>**

		Estadístico de Levene	gl1	gl2	Sig.
Tiempo de transmisión(s/MB)	Se basa en la media	2,197	5	12	,123
	Se basa en la mediana	,700	5	12	,634
	Se basa en la mediana y con gl ajustado	,700	5	6,409	,642
	Se basa en la media recortada	2,052	5	12	,143

Prueba la hipótesis nula de que la varianza de error de la variable dependiente es igual entre grupos.

a. Variable dependiente: Tiempo de transmisión(s/MB)

b. Diseño : Intersección + Columna + Fila + Columna \* Fila

Normalidad de residuos y gráficos, observando Shapiro-Wilk:

**Pruebas de normalidad**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Residuo para VarDependiente	,091	18	,200*	,981	18	,956

\*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Gráfico Q-Q normal de Residuo para VarDependiente

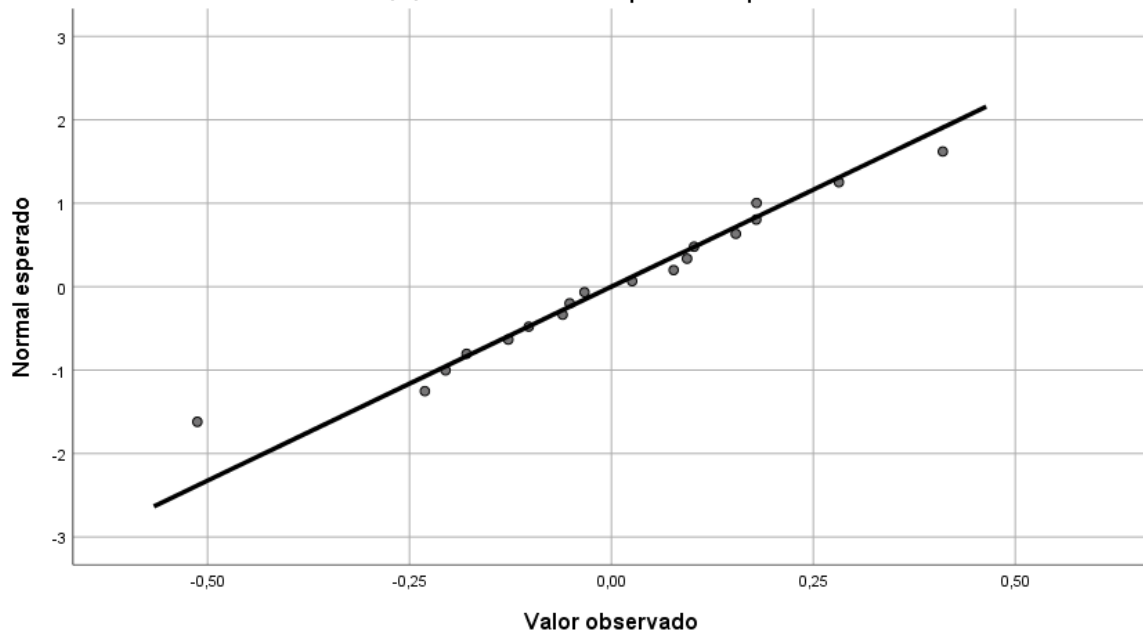
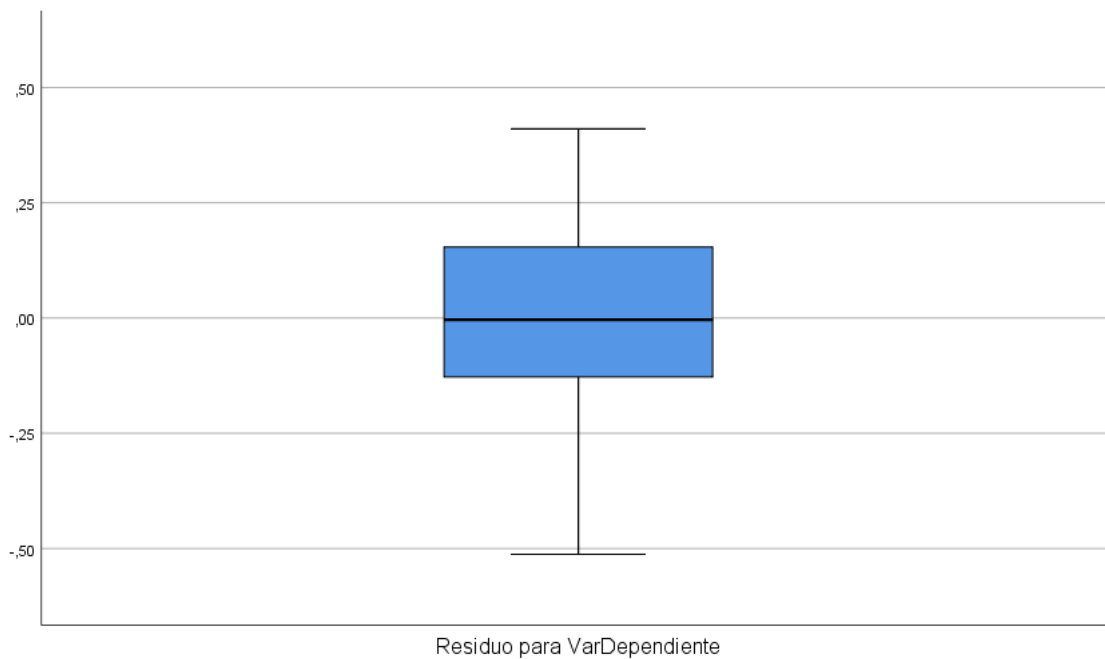
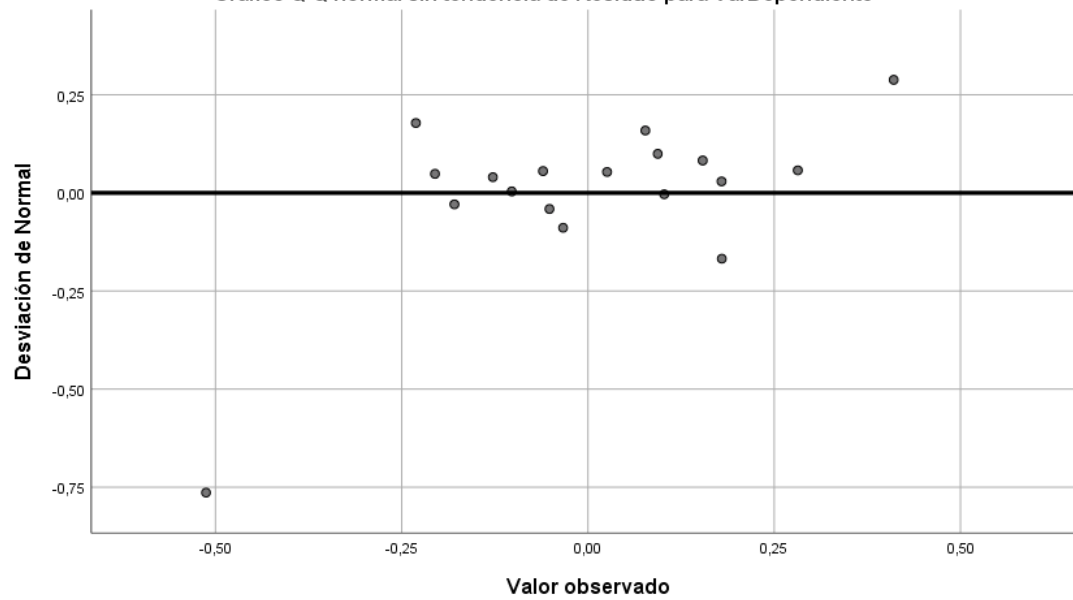
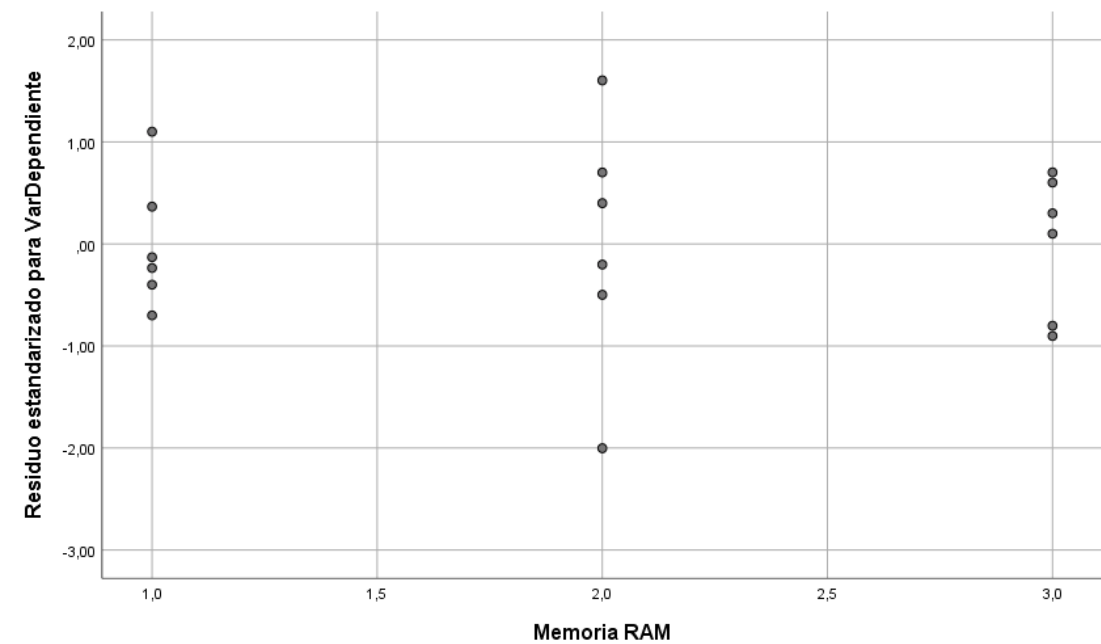
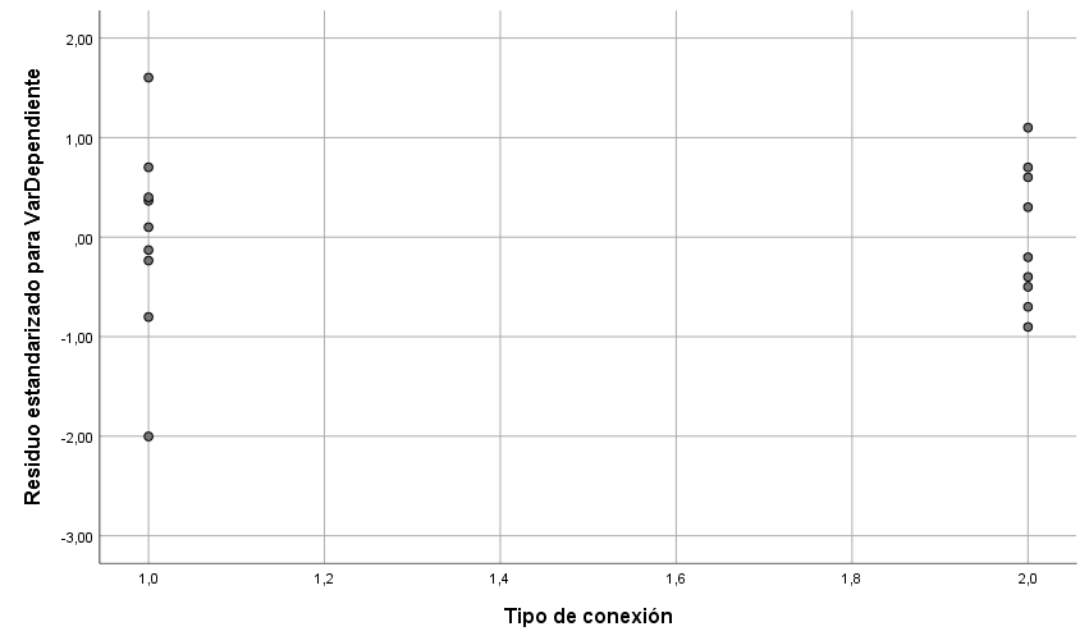
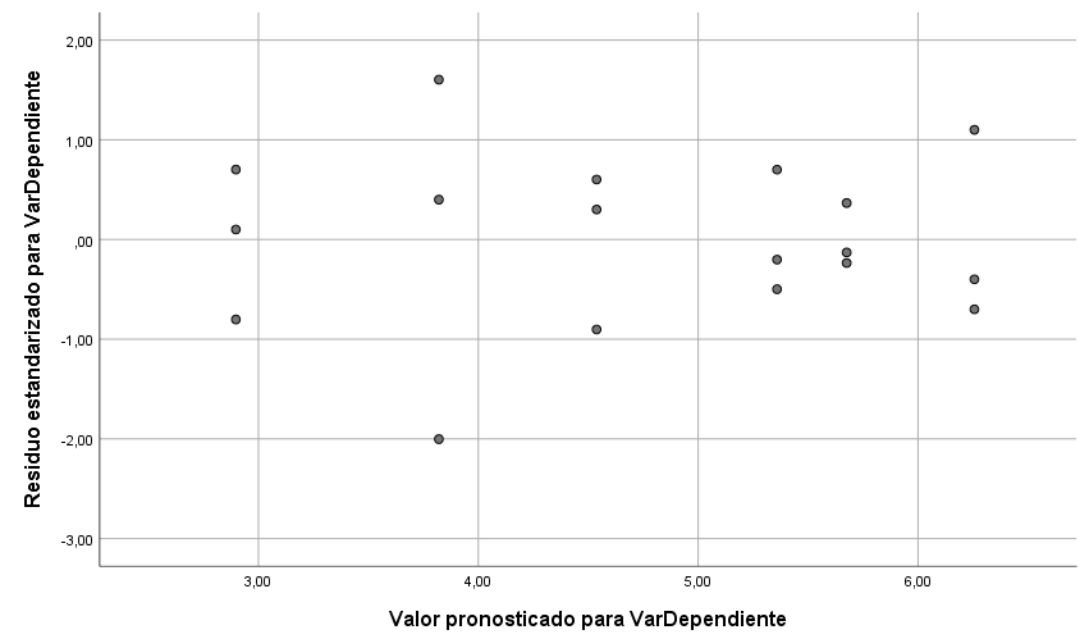


Gráfico Q-Q normal sin tendencia de Residuo para VarDependiente



Aleatoriedad e independencia de residuos:



Como todos los puntos se encuentran a menos de 3 desvíos del 0, se concluye que existe independencia en los residuos.

Se concluye entonces, que el modelo es correcto y valido.

b) Como existe efecto significativo por la intersección se deben realizar las comparaciones de media de manera manual mediante la tabla proveída por el software:

### Estadísticos descriptivos

Variable dependiente: Tiempo de transmisión(s/MB)

Memoria RAM	Tipo de conexión	Media	Desv. Desviación	N
1	1	5,67533	,082233	3
	2	6,25633	,246950	3
	Total	5,96583	,358283	6
2	1	3,82067	,469932	3
	2	5,35867	,159889	3
	Total	4,58967	,898996	6
3	1	2,89733	,193779	3
	2	4,53800	,203723	3
	Total	3,71767	,916055	6
Total	1	4,13111	1,251994	9
	2	5,38433	,765512	9
	Total	4,75772	1,195470	18

Primero se debe calcular el valor LSD para tener un umbral y definir si existe o no diferencia entre 2 medias:

$$LSD = t_{0,025;12} * \sqrt{\frac{MSE}{n}} = 2,179 * \sqrt{\frac{0,066}{12}} = 0,1616$$

No se compara fijando la RAM ya que solo existen 2 tipos de conexión y va a existir diferencia entre ambas por lo que ya se concluyó sobre ese factor.

Fijando el tipo de conexión 1:

3vs1:  $|2,897 - 5,675| = |-2,778| > 0,1616$  Existe diferencia

2vs1:  $|3,821 - 5,675| = |-1,854| > 0,1616$  Existe diferencia

3vs2:  $|2,897 - 3,821| = |-0,924| > 0,1616$  Existe diferencia

Fijando el tipo de conexión 2:

3vs1:  $|4,538 - 6,256| = |-1,718| > 0,1616$  Existe diferencia

2vs1:  $|5,359 - 6,256| = |-0,897| > 0,1616$  Existe diferencia

3vs2:  $|4,538 - 5,359| = |-0,821| > 0,1616$  Existe diferencia