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

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

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

y_{ijk}: Repuesta de la fila i, a la columna j y en la réplica k
μ: Media global
τ_i: Efecto del factor fila en i
β_j: Efecto del factor columna en j
$(\tau\beta)_{ij}$: Efecto del factor fila y columna en ij
ε_{ijk}: Error aleatorio del modelo

Supuestos: Homogeneidad de varianzas, ε_{ijk} independientes y ε_{ijk}~N(0, σ^2)

Restricciones:

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

Hipótesis:

Para el factor fila:
H_0) $\tau_1=\tau_2=\cdots=\tau_a=0$
H_1) $\exists \tau_i \neq 0$ para al menos un valor i.

Para el factor columna:

H_0) $\beta_1= \beta_2=\cdots= \beta_b=0$
H_1) $\exists\ \beta_j \neq 0$ para al menos un valor i.

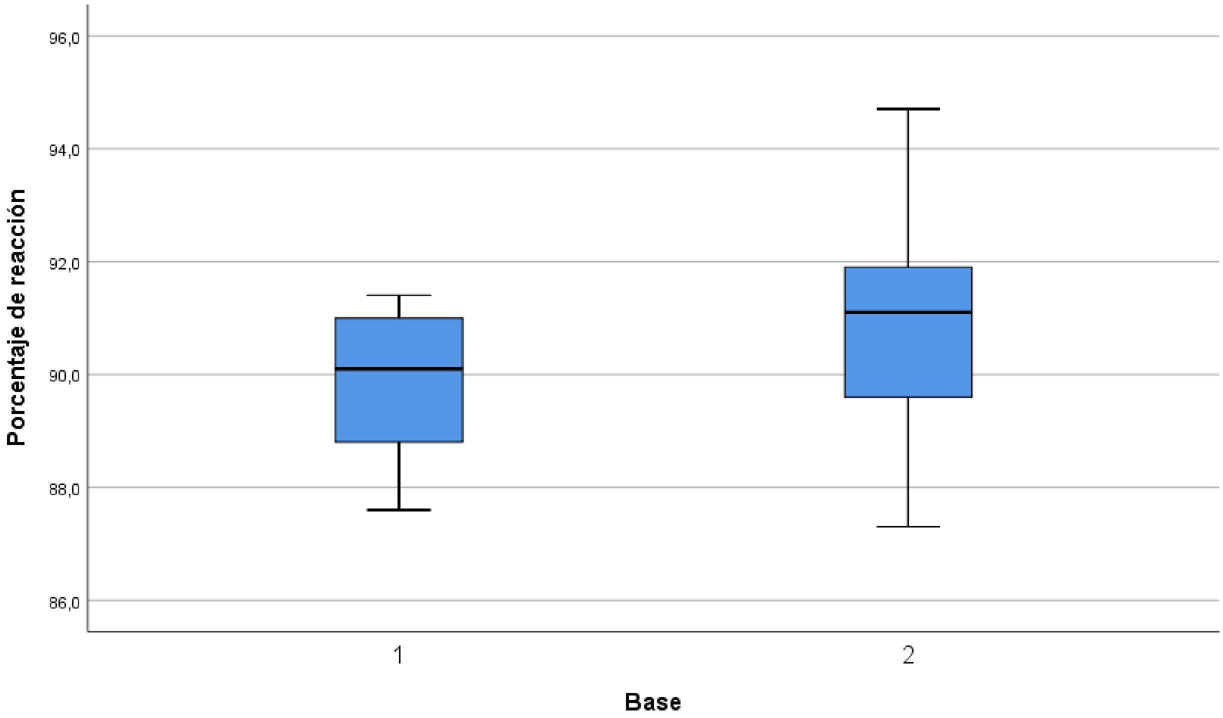
Para la intersección:

H_0) $(\tau\beta)_{ij}=0\ \forall i, j$
H_1) $\exists\ (\tau\beta)_{ij} \neq 0$

Para este caso:

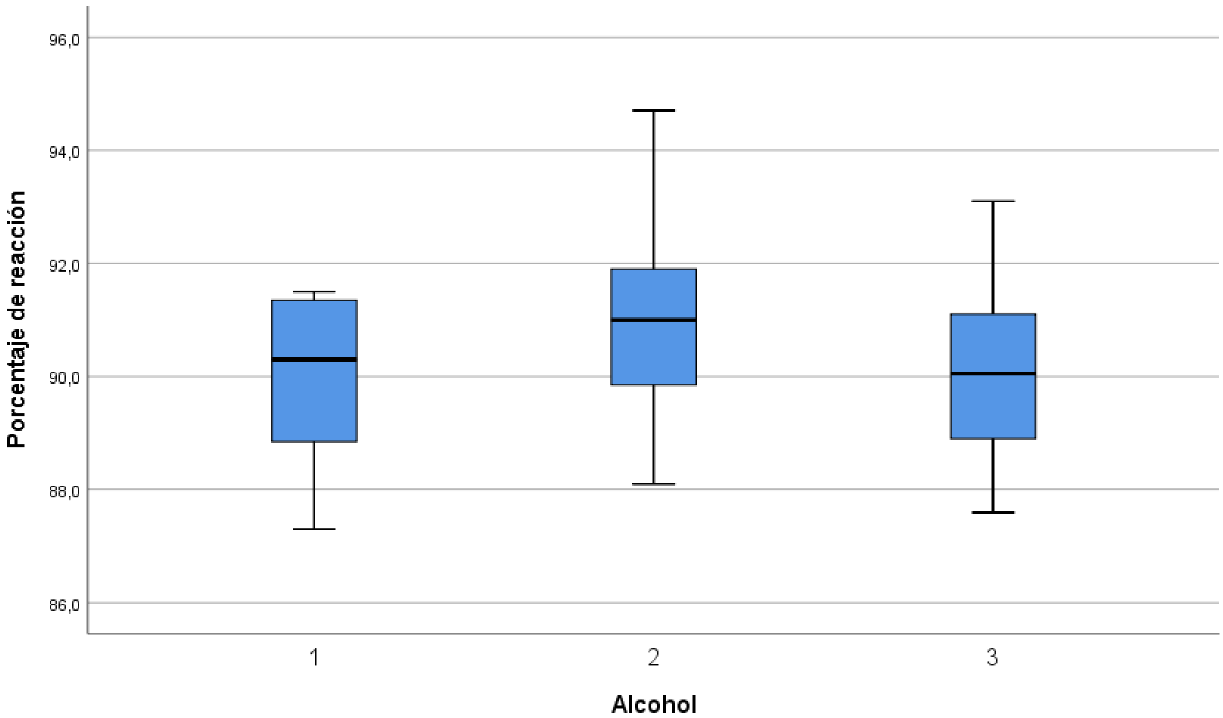
Fila: Base. Factor fijo. a=2
Columna: Alcohol. Factor fijo. b=3
Variable dependiente: Porcentaje de reacción.
Replicas por celda: n=4

Análisis exploratorio VarDep vs Factor fila:

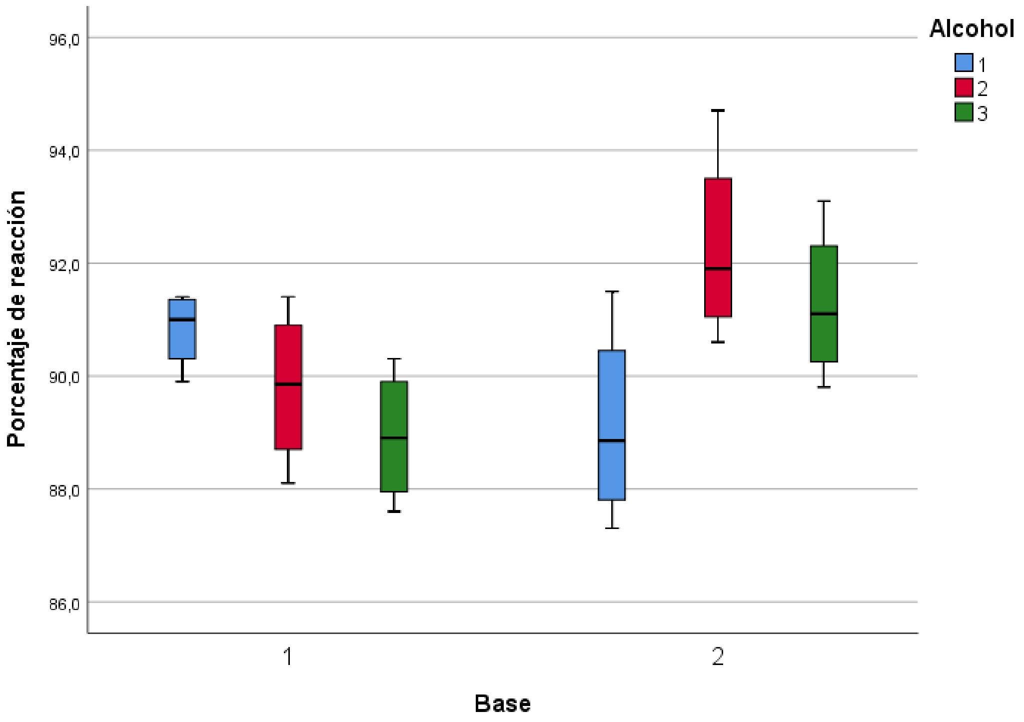


Resulta difícil determinar a simple vista si hubiese o no diferencia, se podría especular que si por la diferencia de altura entre las medianas y la distancia entre los primeros dos cuartiles.

Var dep vs factor columna:



Nuevamente es complicado determinar si existe diferencia, pareciera que si, al menos entre el primero y el segundo alcohol.



Ahora si parecería existir diferencia entre las cajas del gráfico.

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

Pruebas de efectos inter-sujetos

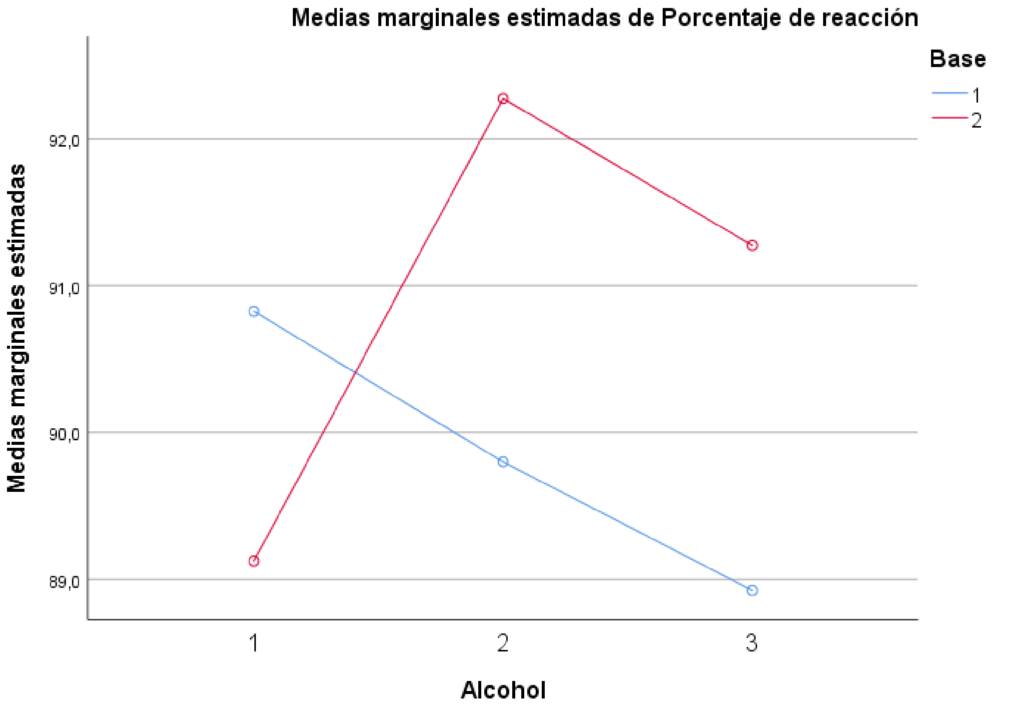
Variable dependiente: Porcentaje de reacción

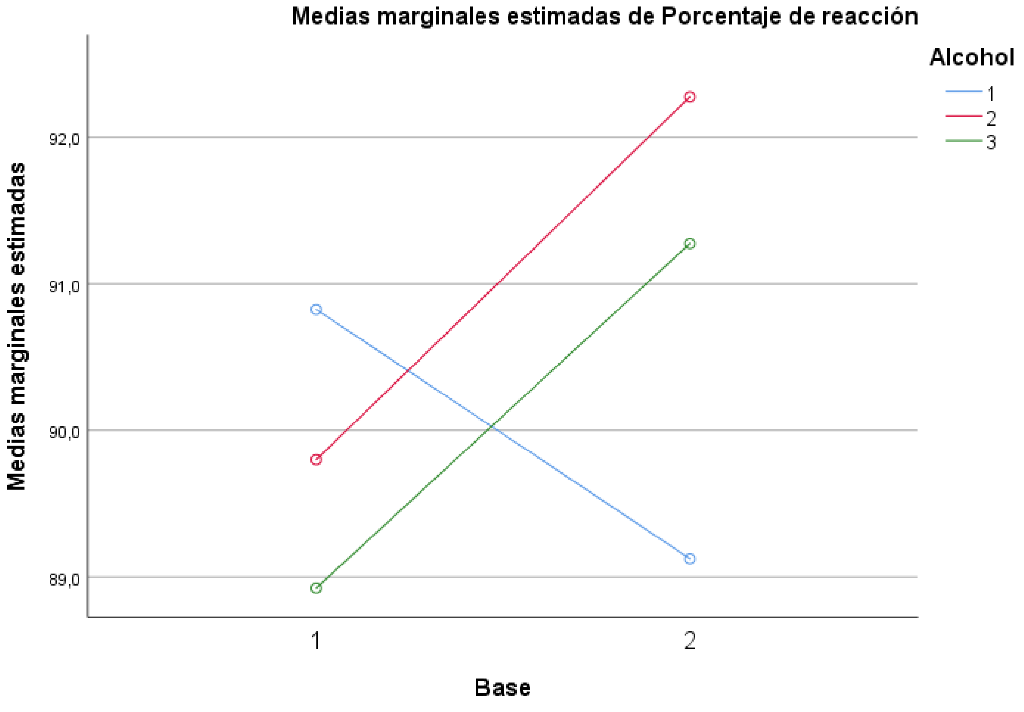
Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	34,472[a]	5	6,894	3,376	,025
Intersección	196005,300	1	196005,300	95983,008	,000
Columna	5,396	2	2,698	1,321	,291
Fila	6,510	1	6,510	3,188	,091
Columna * Fila	22,566	2	11,283	5,525	,013
Error	36,758	18	2,042		
Total	196076,530	24			
Total corregido	71,230	23			

a. R al cuadrado = ,484 (R al cuadrado ajustada = ,341)

Asumiendo una significancia del 10%.
Primero, observando la intersección entre el alcohol y la base, como el p-value=0,013<0,1 se puede concluir que existe un efecto significativo al usar un alcohol y una base distintos en el porcentaje de reacción.
Luego, observando la fila del factor base, como el p-value=0,091<0,1 se concluye que la base afecta significativamente al porcentaje de reacción.
Finalmente, el factor alcohol, no afecta significativamente al porcentaje de reacción por si solo, dado que el p-value=0,291>0,1.

Gráficos de perfiles:





d) Prueba de igualdad de varianzas:

$H_0)\ \sigma ij^2 = \sigma^2 \forall i, j$
$H_1)\ \exists \sigma_{ij}^2 \neq \sigma^2$

Prueba de igualdad de Levene de varianzas de error[a,b]

		Estadístico de Levene	gl1	gl2	Sig.
Porcentaje de reacción	Se basa en la media	,571	5	18	,721
	Se basa en la mediana	,512	5	18	,763
	Se basa en la mediana y con gl ajustado	,512	5	12,244	,762
	Se basa en la media recortada	,570	5	18	,722

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

a. Variable dependiente: Porcentaje de reacción

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

Dado el p-value elevado, se concluye no rechazar H0 y por lo tanto asumir igualdad de varianzas

Normalidad de residuos y gráficos, observando Shapiro-Wilk ya que N<50:

$H_0)\ e_{ijk} \sim N\ (0, \sigma^2)$
$H_1)\ No\ H_0$

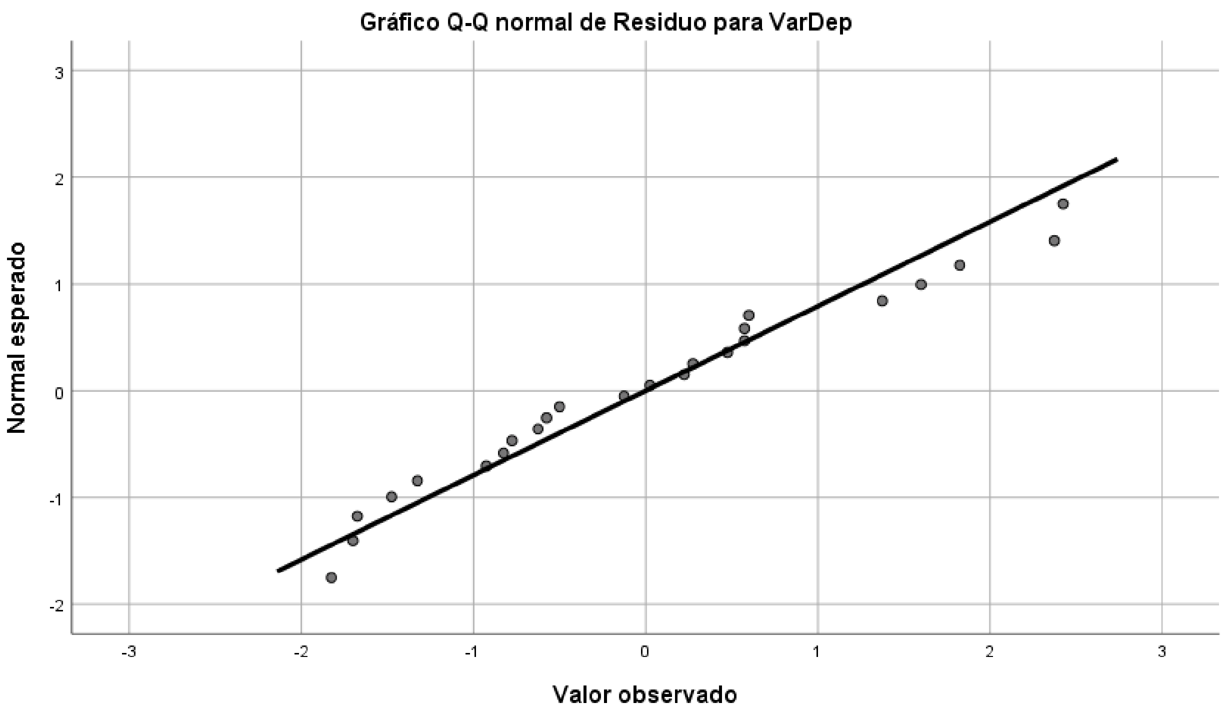
Pruebas de normalidad

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Residuo para VarDep	,112	24	,200*	,950	24	,278

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

a. Corrección de significación de Lilliefors

Como el p-value=0,278>0,2 no se puede rechazar H0 y se concluye la normalidad de los residuos



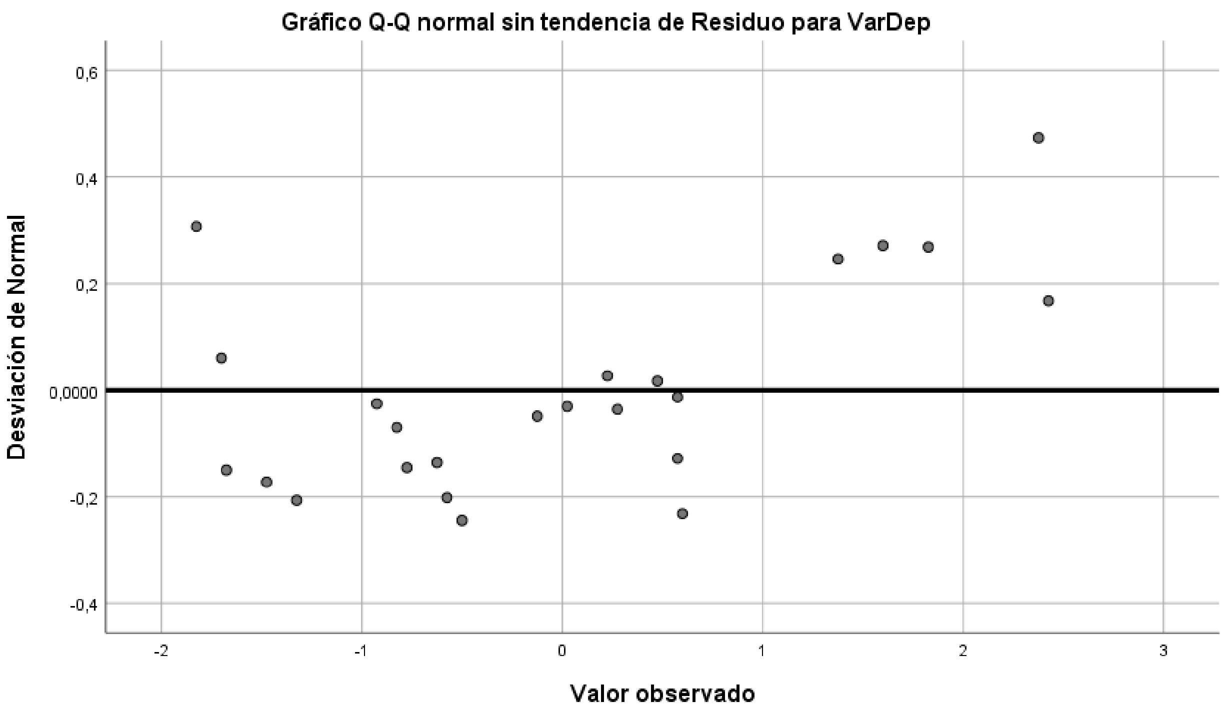
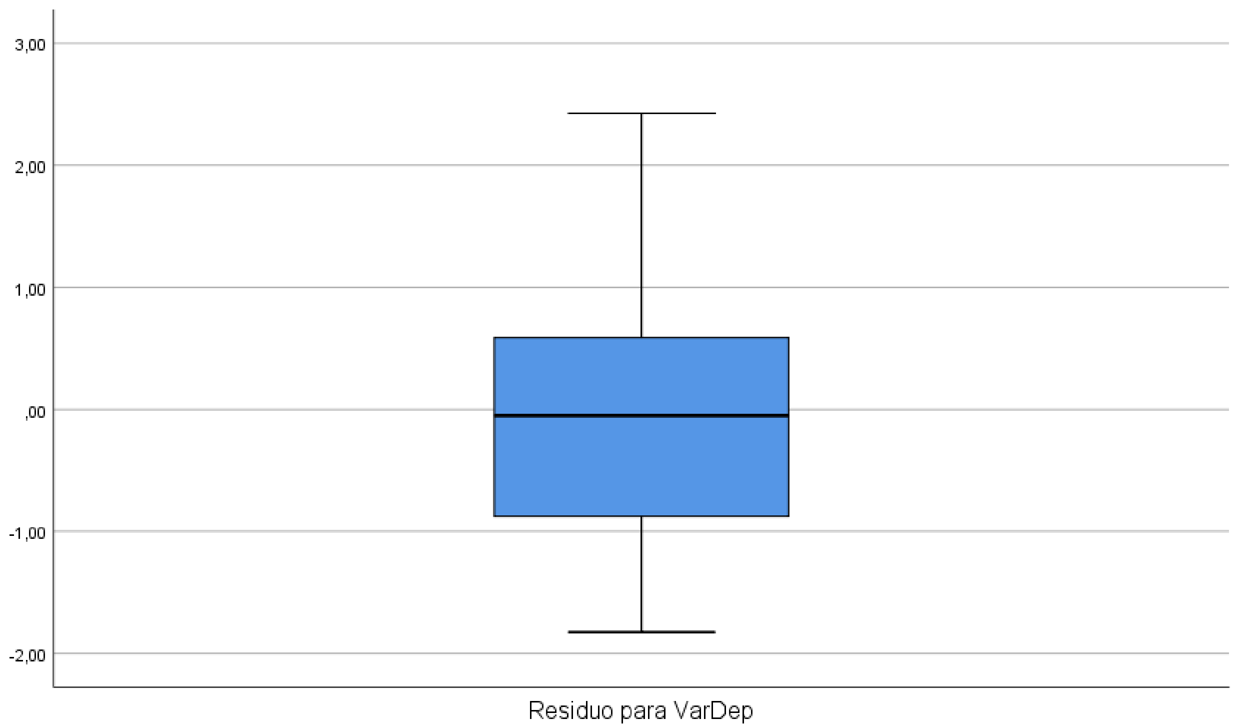
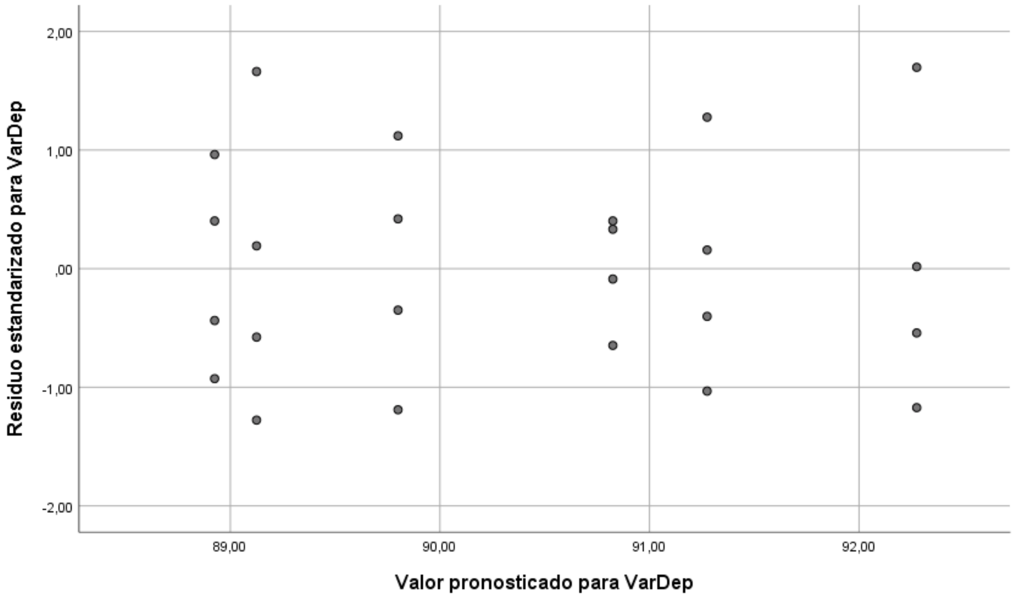


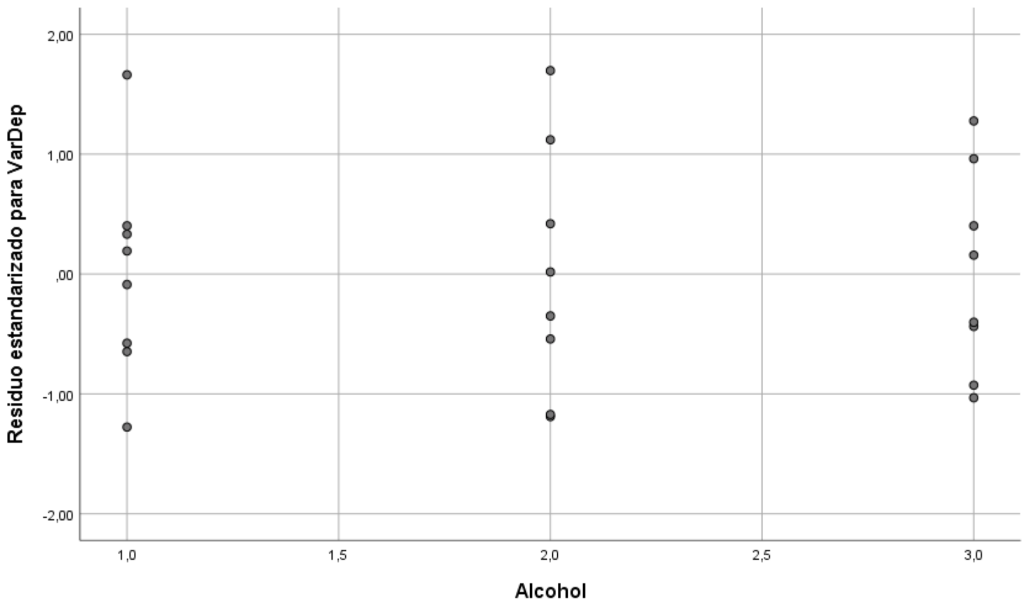
Gráfico Q-Q normal sin tendencia de Residuo para VarDep
Desviación de Normal
0,6
0,4
0,2
0,0000
-0,2
-0,4
-2
-1
0
1
2
3
Valor observado

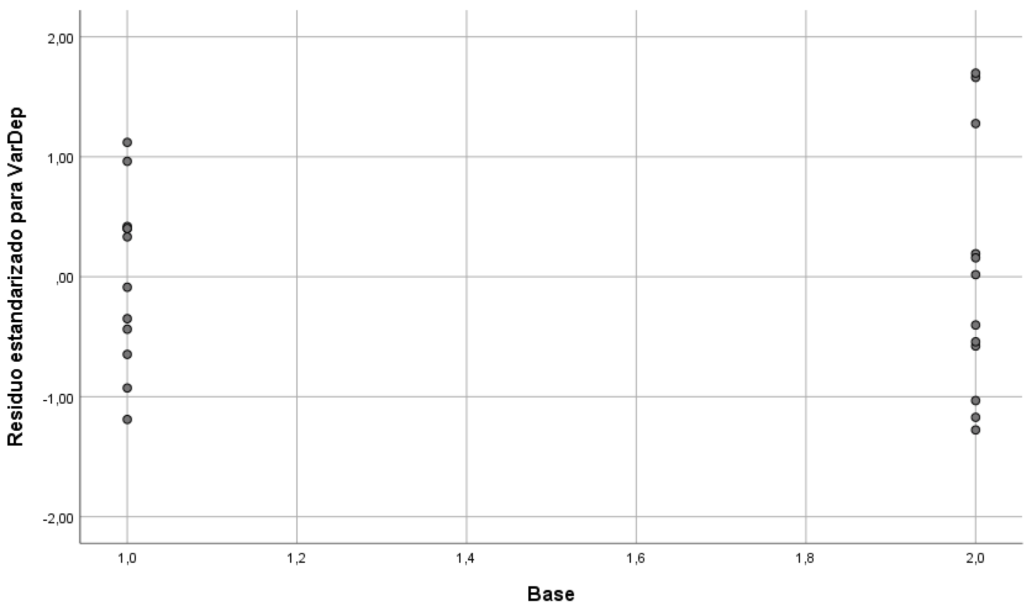


3,00
2,00
1,00
,00
-1,00
-2,00
Residuo para VarDep

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.

e) 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: Porcentaje de reacción

Alcohol	Base	Media	Desv. Desviación	N
1	1	90,825	,6898	4
	2	89,125	1,8007	4
	Total	89,975	1,5554	8
2	1	89,800	1,4213	4
	2	92,275	1,7595	4
	Total	91,038	1,9856	8
3	1	88,925	1,2066	4
	2	91,275	1,4009	4
	Total	90,100	1,7444	8
Total	1	89,850	1,3174	12
	2	90,892	2,0367	12
	Total	90,371	1,7598	24

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

$$LSD = t_{0,025;18} * \sqrt{\frac{MSE}{n}} = 2{,}101 * \sqrt{\frac{2{,}042}{4}} = 1{,}501$$

Solo se analizará diferencias entre medias fijando el alcohol, ya que las diferencias significativas están en la base utilizada.

Con el tipo de alcohol 1:

2vs1: |89,125-90,825|=|-1,7|>0,1616 Existe diferencia

Con el tipo de alcohol 2:

2vs1: |92,275-89,800|=|2,475|>0,1616 Existe diferencia

Con el tipo de alcohol 3:

2vs1: |90,892-89,850|=|1,042|<0,1616 No existe diferencia

f) La combinación para obtener una mayor reacción será con el alcohol 2 y la base 2.