

## Estadística: VERIFICACIÓN DE HIPOTESIS EN REGRESIÓN MÚLTIPLE

Viviana Barile

**MATERIALES:** Para este laboratorio necesita en su calculadora TI-89 Titanium el software de estadística

**OBJETIVO:** Verificar las hipótesis fundamentales sobre todo en áreas de la economía que si no se cumplen, puede conducir a la toma de decisiones en forma errónea. Se entregan herramientas que permiten utilizar la calculadora para comprobar estas hipótesis

**SITUACION:** Los datos a continuación corresponden a mediciones de la edad y nivel de poliamina en el plasma de 25 niños saludables (0 significa recién nacido):

Edad	0	1	2	3	4
Plasma	13.44	10.11	9.83	7.94	4.86
	12.84	11.38	9.00	6.01	5.10
	11.91	10.28	8.65	5.14	5.67
	20.09	8.96	7.85	6.90	5.75
	15.60	8.59	8.88	6.77	6.23

Ajuste un modelo de regresión lineal simple  $y = \beta_0 + \beta_1 x$  a estos datos, usando edad como predictor. Examine los residuos de este ajuste, y determine si los supuestos de normalidad, homogeneidad y homocedasticidad son apropiados

## DESARROLLO

1.- Coloque los datos en la calculadora de la forma que se muestra en la pantalla adjunta

F1+ Tools	F2+ Plots	F3+ List	F4+ Calc	F5+ Dist	F6+ Tests	F7+ Ints	
e		P					
0		11.91					
0		20.09					
0		15.6					
1		10.11					
1		11.38					
1		10.28					
e[8]=1							
MAIN		DEG AUTO		FUNC			

2 En modo **F4** seleccione regresión línea, complete los datos como se muestran en la pantalla adjunta.

LinRe3(a+bx)...	
X List:	e
Y List:	P
Store ResEqn to:	Y1(x) →
Freq:	1
Category List:	
Include Categories:	EQ
Enter=OK      ESC=CANCEL	
MAIN      DEG AUTO      FUNC      1/2	

Obtiene los siguientes datos

F1+ Tools	F2+ Plots	F3+ List	F4+ Calc	F5+ Dist	F6+ Tests	F7+ Ints
LinRe3(a+bx)...						
y=a+bx						
0		a				=13.4752
0		b				=-2.182
0		r <sup>2</sup>				=.753249
1		r				=-.867899
Enter=OK						
MAIN		DEG AUTO		FUNC		1/2

De aquí obtenemos los valores pedidos donde  $\beta_0 = 12,47$   $\beta_1 = -2,182$

$$y = \beta_0 + \beta_1 x$$

No olvidar que lo más importante en la aplicación del modelo es verificar si se cumplen la hipótesis que permitirán validar el modelo

## HIPOTESIS 1

**Los errores tienen una distribución normal  $u_i \sim N(0, \sigma^2)$ . Es decir, se distribuyen siguiendo una campana de Gauss**

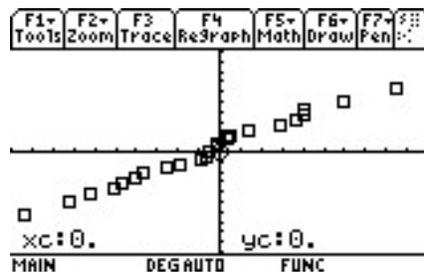
Para estudiar si se cumple esta hipótesis:

Al realizar la regresión lineal la **calculadora** le entrega una lista de residuos, debe copiar esta lista en una nueva con otro nombre, para esto proceda igual que en un computador marcando y en **F1** seleccione copiar y luego pegar

F1→ Tools	F2→ Plots	F3→ List	F4→ Calc	F5→ Distr	F6→ Tests	F7→ Ints	
e		P		resid	tt		
0		13.44		-.0352	<b>-.0352</b>		
0		12.84		-.6352	-.6352		
0		11.91		-1.565	-1.565		
0		20.09		6.6148	6.6148		
0		15.6		2.1248	2.1248		
1		10.11		-1.183	-1.183		
tt[1]=-.0352							
MAIN		DEG AUTO		FUNC		4/4	

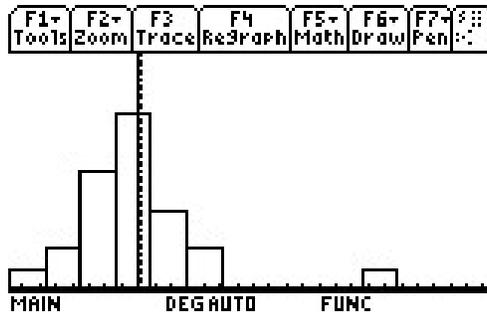
Ahora solicitamos en **F2** un histograma o un gráfico probabilístico normal de los residuos podemos verificar si éstos se distribuyen normalmente.

Obtenemos el grafico siguiente

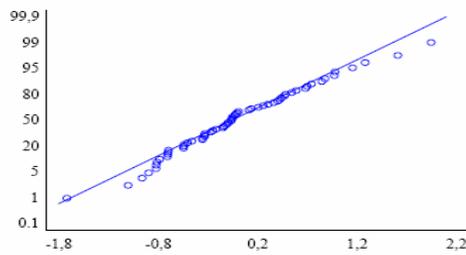
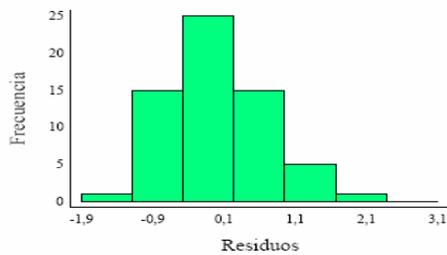


Estando los más próximo a una línea recta se cumple la hipótesis

También lo podemos analizar desde un punto de vista con Histograma



Los gráficos óptimos que nos indica la teoría son



## HIPOTESIS 2

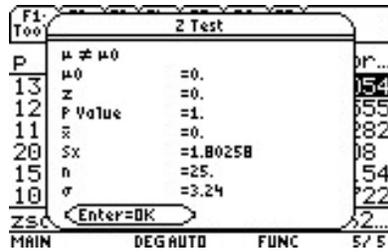
**Homogeneidad: Las perturbaciones tiene esperanza nula:  $E(u_i) = 0, i = 1, \dots, n$  Esto significa que el ajuste que se va a hacer está centrado respecto de los datos**

Para estudiar si se cumple esta hipótesis:

Para esto obtenemos las medidas descriptivas de la comuna de los residuos y realizamos una prueba considerando  $H_0: \mu = 0$  vs  $H_1: \mu \neq 0$

de hipótesis en **F6** tomamos 1





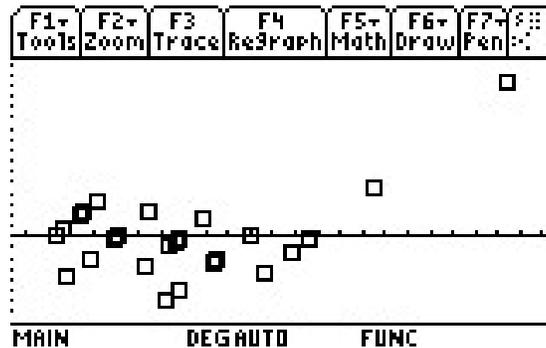
No hay evidencia para rechazar la hipótesis nula luego aceptamos que la medio es cero.

### HIPOTESIS 3

**Homocedasticidad:** La varianza de los errores es constante  $V(u_i) = \sigma^2 \quad i = 1, n$   
*Gráficamente, significa que la nube de puntos de los datos tiene una anchura más o menos constante a lo largo de la recta de regresión. En este caso, se dice que los datos son homocedásticos; en caso contrario, se dice que son heterocedásticos.*

Para estudiar si se cumple esta hipótesis:

Realizamos un grafico considerando en el eje X los valores predichos Y y en el eje Y los valores de los residuos.



Se trata que se encuentren cercanos a una recta aquí vemos que estamos relativamente bien pero tenemos un problema con un punto .

Lo idela seria el siguiente grafico

