

1	2	3	Nota

Apellido y Nombre:.....:

e-mail:.....:

LU N°

Estadística (Químicos) Segundo Parcial – Primer Cuatrimestre 2006

Puntaje y criterio de aprobación: El puntaje figura en cada ejercicio. Para aprobar es necesario obtener al menos 60 puntos.

Observación: En todos los ejercicios, defina claramente las variables aleatorias y los parámetros involucrados.

Ejercicio 1 (35 puntos)

Un mismo material, para el cual la temperatura de fusión es importante, puede ser producido por dos fabricantes (A y B). Cada fabricante genera una muestra de 21 observaciones de temperaturas de fusión del material y se obtienen, para cada fabricante, las siguientes temperaturas medias (en °C).

Fabricante A	420.1
Fabricante B	421.9

Por experiencias anteriores, se sabe que las temperaturas de fusión del material producidos por los dos fabricantes tienen distribución normal con igual desvío estándar de 4°C.

Se tiene interés en decidir si las temperaturas medias de fusión de los materiales ofrecidos por ambos fabricantes son distintas y se desea que la probabilidad de concluir que los materiales difieren en su temperatura de fusión media cuando en realidad no varía sea 0.01.

- Realice el test que considere adecuado para resolver este problema.
- Determine la probabilidad de que el test del ítem anterior detecte diferencia entre las temperaturas esperadas de fusión del material producido por los fabricantes A y B cuando la temperatura esperada del material producido por el fabricante B es 2 °C más que la temperatura media del fabricante A.
- ¿Cuántas determinaciones será necesario realizar si se decide realizar el mismo número de mediciones para cada material y se desea que la probabilidad del ítem b) sea por lo menos 0.90?

Ejercicio 2 (35 puntos)

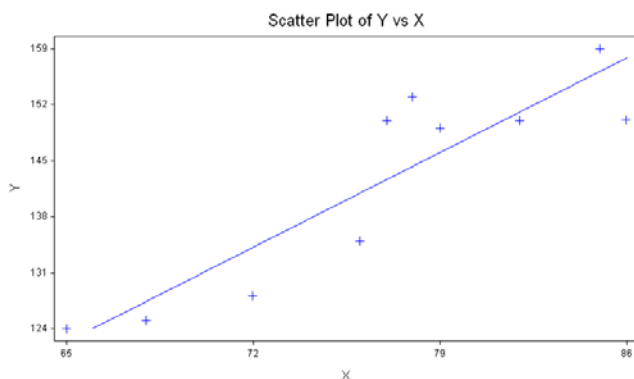
Un analista sospecha que la presión sanguínea de individuos de sexo masculino con edades entre 25 y 30 años está relacionada linealmente con su peso. Para estudiar esto, se seleccionan a 10 jóvenes del grupo de interés para los cuales se registra su peso (X, en kg) y su presión sanguínea (Y, en mm Hg) y se obtienen los siguientes resultados

x	78	76	86	68	82	85	65	77	72	79
y	153	135	150	125	150	159	124	150	128	149

(Para responder las siguientes preguntas puede utilizar las salidas del Statistix que se presentan al final del enunciado.)

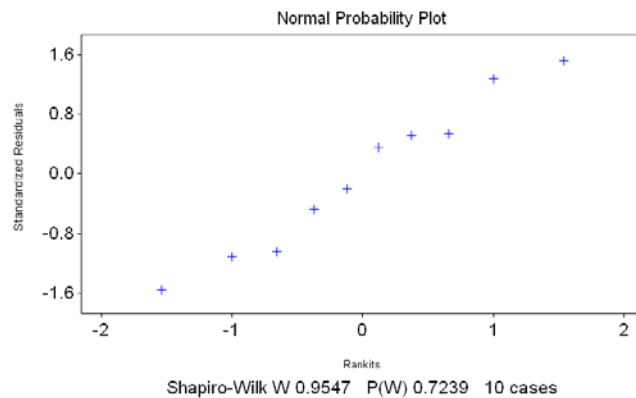
- Expresar el modelo propuesto indicando claramente los parámetros y variables involucradas.
- Obtenga la estimación de la recta de cuadrados mínimos. Interprete los valores obtenidos.
- Halle un intervalo de confianza para la pendiente y la ordenada al origen de la recta.
- ¿Cuál es el coeficiente de determinación R^2 ? Interprete el significado de esta medida en este ejemplo.
- ¿Es razonable suponer que existe una relación lineal entre X e Y? Explique en qué se basa su respuesta.
- Si es posible, halle una predicción para la presión sanguínea cuando el peso del joven es 70 kg y obtenga un intervalo de predicción al 99%. Si no es posible, explique por qué.
- Igual que el ítem anterior pero cuando el peso del joven es 80 kg.

Salidas Statistix



UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF Y

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	
-----	-----	-----	-----	-----	
CONSTANT	12.4474	22.5490	0.55	0.5960	
X	1.69079	0.29255	5.78	0.0004	
R-SQUARED	0.8068	RESID. MEAN SQUARE (MSE)		36.4260	
ADJUSTED R-SQUARED	0.7826	STANDARD DEVIATION		6.03539	
SOURCE	DF	SS	MS	F	P
-----	---	-----	-----	---	-----
REGRESSION	1	1216.69	1216.69	33.40	0.0004
RESIDUAL	8	291.408	36.4260		
TOTAL	9	1508.10			



PREDICTED/FITTED VALUES OF Y

LOWER PREDICTED BOUND	108.54	LOWER FITTED BOUND	121.55
PREDICTED VALUE	130.80	FITTED VALUE	130.80
UPPER PREDICTED BOUND	153.07	UPPER FITTED BOUND	140.05
SE (PREDICTED VALUE)	6.6352	SE (FITTED VALUE)	2.7568

UNUSUALNESS (LEVERAGE)	0.2086
PERCENT COVERAGE	99.0
CORRESPONDING T	3.36

PREDICTOR VALUES: X = 70.000

PREDICTED/FITTED VALUES OF Y

LOWER PREDICTED BOUND	126.24	LOWER FITTED BOUND	140.58
PREDICTED VALUE	147.71	FITTED VALUE	147.71
UPPER PREDICTED BOUND	169.18	UPPER FITTED BOUND	154.84
SE (PREDICTED VALUE)	6.3988	SE (FITTED VALUE)	2.1258

UNUSUALNESS (LEVERAGE)	0.1241
PERCENT COVERAGE	99.0
CORRESPONDING T	3.36

PREDICTOR VALUES: X = 80.000

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	VARIANCE
X	10	76.800	47.289
Y	10	142.30	167.57

Ejercicio 3 (30 puntos)

Los siguientes datos corresponden a 26 determinaciones repetidas de oxígeno disuelto en una muestra de agua, 8 de ellas fueron realizadas por un grupo de estudiantes (Grupo A),

otras 8 fueron realizadas por otro grupo de estudiantes (Grupo B) y las restantes por un tercer grupo (Grupo C).

Grupo A	5.9	6.1	6.3	6.0	6.2	5.8	6.1	6.4		
Grupo B	6.1	6.0	6.0	6.3	6.4	6.2	6.2	6.1		
Grupo C	5.9	5.8	5.7	6.0	6.1	5.6	6.1	6	5.9	5.8

Se decide aplicar un modelo de análisis de la varianza (ANOVA) para analizar si las diferencias observadas entre los grupos son sistemáticas.

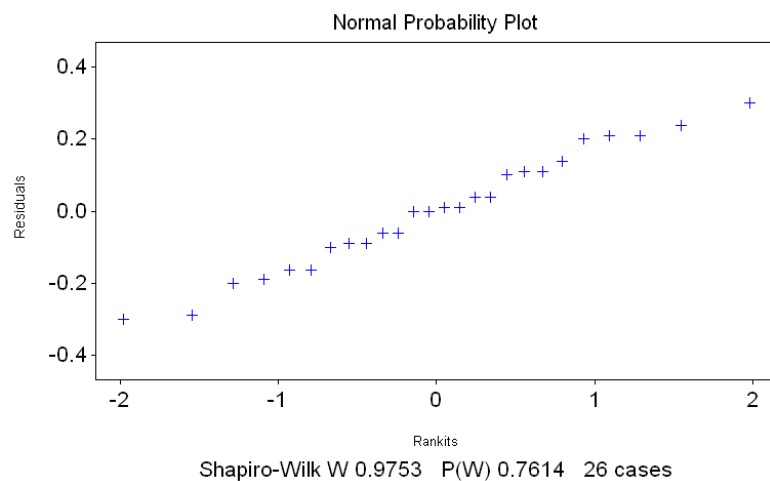
- Plantee la hipótesis que se desea testear definiendo claramente los parámetros involucrados en este problema.
- En base a los resultados obtenidos, es posible atribuir al azar la diferencia observada entre los grupos? ¿Con qué nivel de significación se basa su respuesta?

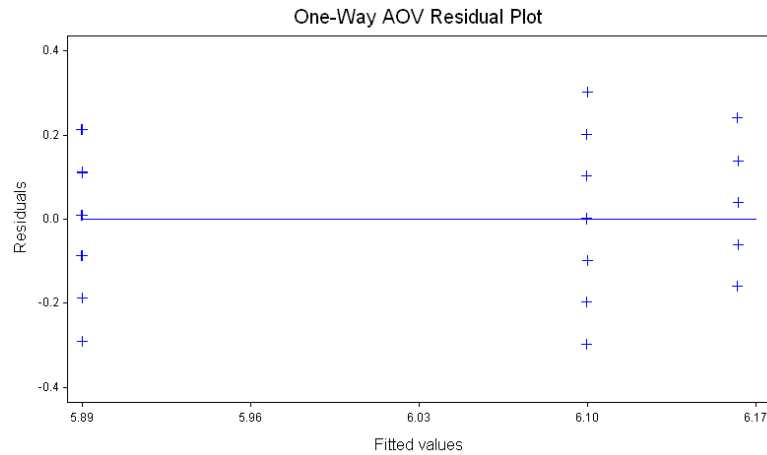
ONE-WAY AOV FOR: GRUPOA GRUPOB GRUPOC

SOURCE	DF	SS	MS	F	P
BETWEEN	2	0.37379	0.18689	6.44	0.0060
WITHIN	23	0.66775	0.02903		
TOTAL	25	1.04154			

- Indique los supuestos necesarios bajo los cuales el test F del ANOVA es válido.
- En base a las siguientes salidas del Statistix verifique si se cumplen los supuestos enunciados en el ítem anterior. Indique en cada caso en qué test y/o gráfico basa su conclusión.

BARTLETT'S TEST OF EQUAL VARIANCES	CHI-SQ	DF	P
	0.81	2	0.6656





- e) Construya los intervalos de confianza de nivel global 0.90 para las diferencias de medias. ¿Qué método propone para realizar esto? ¿En qué se basa su elección? ¿Es el nivel global exactamente de 0.90 en este caso? ¿Por qué?

TUKEY (HSD) COMPARISON OF MEANS

VARIABLE	MEAN
GRUPO A	6.1000
GRUPO B	6.1625
GRUPO C	5.8900
CRITICAL Q VALUE	3.054
REJECTION LEVEL	0.100

BONFERRONI COMPARISON OF MEANS

VARIABLE	MEAN
GRUPO A	6.1000
GRUPO B	6.1625
GRUPO C	5.8900
CRITICAL T VALUE	2.264
REJECTION LEVEL	0.100

- f) En base a los intervalos obtenidos en el ítem anterior, cuáles son las medias que resultan significativamente distintas?