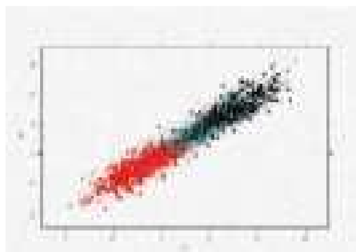
ESTADÍSTICA (SISTEMAS)

Profesores: Hilario Navarro. Jorge Martín



DEPARTAMENTO DE ESTADÍSTICA,
INVESTIGACIÓN OPERATIVA Y CÁLCULO
NUMÉRICO

Cuarta unidad didáctica.
Problemas propuestos de Otros Métodos Estadísticos

Curso 2004-2005

Problema 1. Con el fin de seleccionar el sistema más rápido de almacenamiento y recuperación de datos para un determinado tipo de procesos, se realizó un experimento consistente en hacer 4 pruebas con cada uno de los tres sistemas considerados: *CD, Disco* y *Cinta*. Los tiempos —en minutos— requeridos en cada ocasión se reflejan en la siguiente tabla

	CD	*Disco*	*Cinta*
	8.7	7.0	7.2
	9.3	6.4	9.1
	7.9	9.8	7.5
	8.0	8.2	7.7
Suma	33.9	31.4	31.5
Media	8.475	7.850	7.875
Varianza	0.429	2.250	0.709

Utilizando un nivel de significación $\alpha = 0.01$, contraste la hipótesis de igualdad de los tiempos medios.

Datos auxiliares: La variabilidad total es 11.167. $F_{2;9;0.01} = 8.\,0215$.

Problema 2. El número de trabajadores que diariamente integran una cadena de montaje varía a causa del nivel de absentismo. La tabla siguiente contiene los datos registrados en una muestra aleatoria de la producción diaria, siendo X el número de trabajadores ausentes e Y el número de productos defectuosos generados por dicha cadena.

X	1	3	5	0	2
Y	10	16	20	9	12

Si la recta de regresión estimada es $y = 8.26 + 2.34x$, calcule un intervalo de confianza (99 %) para la predicción de la cantidad de productos defectuosos que se obtendrán cuando el número de operarios ausentes sea 4.

Problema 3. Nos dicen que un programa de ordenador genera observaciones de una distribución *normal estándar*, es decir, $N(0;1)$. En una muestra aleatoria de 100 observaciones producidas mediante dicho programa se obtienen los siguientes resultados:

6 observaciones menores que -1.5,	20 entre -1.5 y -0.5,
30 entre -0.5 y 0,	25 entre 0 y 0.5,
15 entre 0.5 y 1.5,	4 mayores que 1.5

Sabiendo que una variable aleatoria Z con distribución $N(0;1)$ asigna las probabilidades: $P(Z < -1.5) = 0.07$, $P(-1.5 \leq Z \leq -0.5) = 0.24$, $P(-0.5 < Z < 0) = 0.19$, ¿qué regla utilizaría para decidir, al nivel $\alpha = 0.01$, si el programa funciona correctamente o no?

Problema 4. Se ha realizado un experimento orientado a detectar diferencias en el comportamiento de la respuesta Y en tres situaciones diferentes: A, B y C. Tras algunos cálculos con los datos recogidos, se han obtenido los siguientes resultados:

· Suma de cuadrados total $\boxed{172.93}$
· Suma de cuadrados "entre grupos" $\boxed{66.93}$
· Grados de libertad "dentro de los grupos" $\boxed{12}$
· Cuadrado medio "dentro de los grupos" $\boxed{8.83}$

Se pide:

(a) Construir la tabla de análisis de la varianza.

(b) ¿Cuántas unidades componen la muestra?

(c) ¿Qué conclusiones se obtienen a partir de la información contenida en dicha tabla?

Problema 5. En las poblaciones **M** y **H** se han recogido dos muestras aleatorias independientes de tamaño 50. La frecuencia de aparición de cada una de las dos clases de la variable cualitativa Y se refleja en la siguiente tabla (la frecuencia que figura entre paréntesis es la esperada):

	Y_1	Y_2
M	14 (12)	36 (38)
H	10 (12)	40 (38)
Total	24	76

Se pide:

(a) Calcular el valor del estadístico χ^2.

(b) ¿Qué se puede concluir sobre la distribución de la característica Y?

Problema 6. Se piensa que el tiempo de respuesta de un equipo informático, cuando se le solicita cierto tipo de información, sigue una distribución exponencial con parámetro $\lambda = 1$ seg. (por tanto, la densidad es $f(x) = e^{-x}$, para $x \geq 0$). Contraste dicha hipótesis utilizando los siguientes datos:

► Los puntos $0.22, 0.51, 0.92, 1.61$ determinan 5 clases equiprobables para la densidad citada.

► La frecuencia observada en cada una de estas clases, en un muestreo aleatorio de tamaño 40, es:

$$6, 8, 10, 7, 9$$

respectivamente.

Datos auxiliares: $\chi^2_{5;0.025} = 12.833$; $\chi^2_{4;0.05} = 9.488$; $\chi^2_{5;0.05} = 11.07$.

Problema 7. Para estudiar la posible relación entre dos variables, X e Y, se registraron los valores de dichas variables en una muestra aleatoria formada por 20 individuos. De la información recogida se obtuvieron los siguientes datos estadísticos:

$$\frac{1}{20}\sum x_i = 53.69 \qquad \frac{1}{20}\sum y_i = 58.81$$

$$\frac{1}{20}\sum x_i^2 = 2948.82 \qquad \frac{1}{20}\sum y_i^2 = 3505.15$$

$$\frac{1}{20}\sum x_i y_i = 3197.39$$

Suponiendo un modelo de regresión lineal, ¿se puede concluir, con un nivel de significación $\alpha = 0.05$, que la pendiente de la recta de Y sobre X es positiva?

Datos auxiliares: $t_{18;0.05} = 1.734$, $t_{20;0.025} = 2.086$; *la estimación insesgada de la varianza* σ^2 *vale* 31.36 ; cuando $\beta_1 = 0$,

$$\frac{\hat{\beta}_1}{S_R\sqrt{\frac{1}{nv_x}}} \sim t_{n-2} \ .$$

Problema 8. Se está realizando un estudio sobre los fallos de un dispositivo electrónico. Este elemento se puede montar en dos posiciones diferentes y hay cuatro tipos de fallos posibles. Un muestreo aleatorio proporciona la siguiente distribución de frecuencias:

	Tipo de fallo			
Posición de Montaje	*A*	*B*	*C*	*D*
1	14	18	8	20
2	6	12	12	10

¿Concluiría que el tipo de fallo es independiente de la posición de montaje?

Datos auxiliares: $t_{3;0.025} = 3.182$, $\chi_{8;0.05} = 15.507$, $\chi_{3;0.05} = 7.815$, $z_{0.05} = 1.645$