EXAMEN DE ESTADISTICA MATEMATICA. 29/6/01
2º DIPLOMATURA EN ESTADISTICA
FINAL

1.- Sea $X_1, X_2, \ldots, X_n$ m.a.s. de una población con ley de probabilidad uniforme en $(0, \theta)$. Se consideran los siguientes estimadores de θ:

$$T_1 = X_{(n)} \qquad T_2 = \frac{2(X_1 + X_2 + \ldots + X_n)}{n} \qquad T_3 = \frac{(n+1)X_{(n)}}{n}$$

Se pide:

a) Estudiar todas las propiedades que conozcas de estos estimadores y compararlos.
b) Construir un intervalo de confianza para θ del 95% utilizando la distribución exacta de algún estadístico.

(1.5 puntos)

2.- Sean $X_1, \ldots, X_n$ v.a.i.i.d. con media μ y varianza σ^2 y tales que existe el momento de orden cuatro $\mu_4 = E((X_1-\mu)^4)$, siendo $\mu_4 > \sigma^4$. Dar una distribución aproximada para la varianza muestral $S_n^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n}$ donde n se considera suficientemente grande. Para ello se aconseja comenzar utilizando la siguiente expresión $S_n^2 = \frac{\sum_{i=1}^{n} Y_i}{n} - (\overline{X}_n - \mu)^2$ donde $Y_i = (X_i - \mu)^2$

(1 punto)

3.- Diseñar un experimento y dar una regla de decisión que permita confirmar si una moneda es balanceada. Dicha regla de decisión debe cumplir las siguientes condiciones:

- La probabilidad de rechazar la hipótesis cuando en realidad es correcta es como máximo 0.05.
- La probabilidad de aceptar la hipótesis cuando en realidad la probabilidad de cara difiere de 0.5 por 0.1 o más es como máximo 0.05.

¿Podemos encontrar otra regla de decisión mejor que la dada?. Justificar la respuesta.
(1.5 puntos)

4.- Sea $X_1, X_2, \ldots, X_n$ m.a.s. de una población con ley de probabilidad Geométrica de parámetro p (que toma valores 0,1,2,…) se quieren contrastar las siguientes hipótesis sobre p en base a dicha muestra:

$$H_0: \; p = 0.5$$
$$H_1: \; p \neq 0.5$$

Se pide:

a) Calcular la razón de verosimilitudes.
b) Construir a partir de ella dos reglas de decisión de nivel 0.05. Las reglas deben estar indicadas explícitamente, la respuesta no se considerará correcta si sólo se dejan indicadas de forma implícita.

(1.5 puntos)

5.- Once niños pasaron un test de aritmética, después de tres semanas de enseñanza especial pasaron un nuevo test de igual dificultad que el primero. La puntuación de cada prueba se da a continuación:

Primer test:	45	61	33	29	21	47	53	32	37	25	81
Segundo test:	53	67	47	34	31	49	62	51	48	29	86

¿Indican los datos que la mejora media en la puntuación debida a la enseñanza especial es de diez puntos?. Contestar a esta pregunta bajo los dos supuestos siguientes:

a) Suponiendo normalidad.
b) Sin suponer normalidad.

(1.5 puntos)

6.- Dos robots A y B envasan pipas en bolsas de 200 gramos. Se ha tomado una muestra de la producción de cada robot obteniéndose los siguientes datos en gramos:

A 202.2 204.3 197.3 200.2 204.9 204.1 195.0 196.1 199.1 198.0 196.2 196.7 200.3 196.8 204.8 201.2 201.9 203.8 201.1 201.7

B 204.9 195.3 200.1 199.0 202.1 203.1 195.0 198.3 204.3 201.9 197.5 196.1 196.7 196.9 201.7 196.8 201.4 200.9

Para cada robot agrupamos la muestra en tres clases en función de lo que difieren las bolsas de 200 grs.

Clase 1: Faltan dos o más gramos.
Clase 2: La diferencia está entre –2 y 2 gramos.
Clase 3: Sobrepasa en dos o más gramos.

a) Contrastar con los datos agrupados si los errores en el peso para ambos robots son similares. Proponer un segundo procedimiento alternativo al que se haya utilizado (no es necesario hacer los cálculos de este segundo procedimiento).
b) Se sospecha que los errores de B siguen una ley uniforme (θ, 5). Contrastar esta hipótesis utilizando la anterior división en clases. Para estimar θ utilizar dos procedimientos distintos indicando cuál es preferible y porqué.
c) ¿Qué contraste es más apropiado para el ajuste del apartado b) si se utilizan los datos sin agrupar?. Detallar el método sin hacer los cálculos (R.C. y cómo se obtendría la distribución del estadístico test).
(1.5 puntos)

7.- De entre los individuos tratados con analgésicos no esteroideos, se tomaron 85 que habían padecido una hemorragia digestiva y 10295 que no la habían padecido. De entre los primeros 42 tenían menos de 60 años y de los segundos 7496. ¿Es la edad un factor de riesgo?.
Comenta la salida de SPSS para estos datos y explica cada uno de los resultados que en ella aparecen, destacando e interpretando los más adecuados para contestar a la pregunta planteada. No es necesario escribir las expresiones de los estadísticos.

Tabla de contingencia EDAD * ENFERMED

Recuento

		ENFERMED		
		NO	SI	Total
EDAD	< 60	7496	42	7538
	> 60	2799	43	2842
Total		10295	85	10380

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	23.217[b]	1	.000		
Corrección por continuidad[a]	22.055	1	.000		
Razón de verosimilitud	20.643	1	.000		
Estadístico exacto de Fisher				.000	.000
N de casos válidos	10380				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 23.27.

Estimación de riesgo

	Valor	Intervalo de confianza al 95%	
		Inferior	Superior
Razón de las ventajas para EDAD (< 60 / > 60)	2.742	1.788	4.204
Para la cohorte ENFERMED = NO	1.010	1.005	1.015
Para la cohorte ENFERMED = SI	.368	.241	.562
N de casos válidos	10380		

Estimación de riesgo

	Valor	Intervalo de confianza al 95%	
		Inferior	Superior
Razón de las ventajas para ENFERMED (NO / SI)	2.742	1.788	4.204
Para la cohorte EDAD = < 60	1.474	1.188	1.828
Para la cohorte EDAD = > 60	.537	.435	.665
N de casos válidos	10380		

(1.5 puntos)

Notas:

$$\mathrm{Var}_{H_0} W_s^* = \frac{mn(N+1)}{12} - \frac{mn\sum_{i=1}^{e}(d_i^3 - d_i)}{12N(N-1)}$$

$$\mathrm{Var}_{H_0} V_s^* = \frac{N'(N'+1)(2N'+1)}{24} - \frac{\sum_{i=1}^{e}(d_i^3 - d_i)}{48}$$