Métodos Estadísticos en Economía y Empresa
2º LADE. Curso Académico 2005-06
Prácticas Tema 1

Problema 1.- Durante cien semanas se ha observado el número de veces a la semana que se ha averiado una máquina, presentándose los resultados en la siguiente tabla.

Número de averías	0	1	2	3	4	5 o más
Número de semanas	10	24	32	23	6	5

a) Suponiendo que el número de averías sigue una distribución de Poisson, hallar el estimador de máxima verosimilitud de lambda. (tomar como representante de la última clase el 6).
b) Ajustar a los datos una distribución de Poisson, obtener las frecuencia esperadas y compararlas con las frecuencias observadas. Obtener el estadístico chi-cuadrado.

Problema 2- Los siguientes datos representan el número de horas entre fallos sucesivos del sistema de aire acondicionado de determinado tipo de aviones.

23	261	87	7	120	14	62
47	225	71	246	21	42	20
5	12	120	11	3	14	71
11	14	11	16	90	1	16
52	95					

a) Dibujar el histograma de los datos. Tomar como extremos de clase: 40, 80, 120, 160, 200, 240, 280.
b) Se supone que los datos proceden una distribución exponencial con función de densidad:

$$f(x;\theta) = \theta e^{-\theta x}, \quad x > 0$$

Hallar el estimador de momentos y de máxima verosimilitud del parámetro θ.
c) Hallar el EMV de la probabilidad de que el número de horas entre fallos sea inferior a 40.

Problema 3.- Los siguientes datos representan el número de camiones pesados que pasan por un determinado punto de una autovía en intervalos de un minuto.

0	3	3	0	0	5	2
3	3	0	2	1	0	0
2	2	0	1	0	0	3
4	2	1	4	1	4	3

a) Disponer los datos en forma de distribución de frecuencias

b) Ajustar los datos a una distribución de Poisson con parámetro igual a la media muestral (EMV). Obtener las probabilidades teóricas y latos frecuencias esperadas de acuerdo a la ley de Poisson. Representar gráficamente los resultados. A la vista de los resultados ¿se trata de un ajuste satisfactorio?

c) Si X_0 es el número de ceros, se considera el estimador de λ:

$$\hat{\lambda} = -\log\left(\frac{X_0}{n}\right)$$

Ajustar una distribución de Poisson usando este nuevo estimador.

Problema 4.- Los siguientes datos (Greenwood y Yule) representan las frecuencias de los accidentes de trabajo ocurridos a 647 mujeres en 5 meses

Número de Accidentes	Frecuencia Observada
0	447
1	132
2	42
3	21
4	3
5 o más	2

a) Media y varianza del número de accidentes (tomar como representante de la última clase el 5)

b) Ajustar los datos a una distribución de Poisson, tomando como estimador de la media muestral. Obtener las probabilidades teóricas y las frecuencias esperadas de acuerdo a la ley de Poisson.

c) Se propone como modelo alternativo una distribución binomial negativa cuya función de densidad es:

$$P(X = k) = \binom{k + r - 1}{k} p^r q^k, \quad k = 0,1,2,...$$

donde $q=1-p$. Sabiendo que la media y la varianza son:

$$E(X) = \frac{r(1-p)}{p}; \quad \mathrm{var}(X) = \frac{r(1-p)}{p^2}$$

probar que los estimadores de momentos de los parámetros r y p son:

$$\hat{p} = \frac{\overline{X}}{S^2}; \ \hat{r} = \frac{\overline{X}^2}{S^2 - \overline{X}}$$

d) Ajustar los datos a la distribución binomial negativa usando los estimadores anteriores. Para el cálculo de las probabilidades teóricas usar la recurrencia:

$$P(X = 0) = p^r$$
$$P(X = k) = \frac{(1-p)(k+r-1)}{k} P(X = k-1); \quad k = 1,2,...$$

e) Estudiar la bondad de ajuste de los dos modelos. ¿Cuál de los dos modelos es preferible?

Problema 5.- En un conjunto de familias alemanas con 8 hijos, se registraron el número de varones. Los datos se recogen a continuación (datos tomados de un estudio de Geissler)

Número X de varones en Familias Alemanas de N=8 hijos	*Frecuencias Observadas*
0	215
1	1.485
2	5.331
3	10.649
4	14.959
5	11.929
6	6.678
7	2.092
8	342
	53.680

a) Justificar que se trata de una distribución binomial con parámetros N=8 y p=P(varón).
b) Hallar un estimador de p basado en el método de los momentos. Ajustar una distribución binomial con parámetros $N=8$ y p. Representar gráficamente las frecuencias observadas y las frecuencias teóricas (es decir, $e_x=nP(X=x)$)
c) Ajustar la distribución anterior usando un estimador de p basado en la frecuencia de ceros, es decir,

$$\hat{p} = 1 - \left(\frac{X_0}{n}\right)^{1/N}$$

Justificar la elección del estimador. (X_0 es la frecuencia de ceros en la muestra)
d) Hallar el EMV de p.
e) Hallar un estimador suficiente de p.