Modelos no paramétricos y de regresión | Estadística II

Tarea 2

Fecha de entrega: 23 de febrero

1. Mostrar que el estadístico X^2 de la prueba χ^2 de Pearson se puede escribir como

$$X^2 = \sum_{k=1}^{c} \frac{O_k^2}{E_k} - n,$$

donde n es el tamaño de muestra, c es el número de categorías, O_k y E_k son las frecuencias observada y esperada en la $k-ésima$ categoría, respectivamente.

2. Simular 74 observaciones de la distribución $Exp(1)$. Agrupar las observaciones en las categorías dadas por los cuartiles de la distribución $Exp(1)$. Contrastar si la verdaera distribución de los datos es $Exp(1)$. Contrastar si la verdadera distribución es $Ga(2, 1.5)$. Utilizar $\alpha = 0.1$

3. Para comprobar que un dado está balanceado se propuso lanzarlo 125 veces y contrastar las fecuencias obtenidas con las esperadas si el dado en verdad está balanceado. Los resultados se presentan a continuación.

   a. Contrastar ls hipótesis de que el dato está balanceado. Usar $\alpha = 0.07$.
   b. Calcular el *p-value* de la prueba anterior.

1	**2**	**3**	**4**	**5**	**6**
12	18	23	13	20	39

   Aproximar la distribución del estadístico X^2 con simulación (utilizar $m = 5,000$ repeticiones) para calcular el valor crítico y el *p-value* de la prueba.

4. Suponer que se tiene la siguiente realización de una muestra aleatoria de una distribución F

2.38	0.49	3.43	1.02	0.71
0.55	2.01	1.31	0.67	0.60

Contrastar las siguientes hipótesis con una prueba de Kolmogorov-Smirnov de tamaño $\alpha = 0.01$.

$$H_0 : F(x) = \text{LogNormal}(x \,|\, 0, 1) \quad \text{vs.} \quad H_1 : F(x) \neq \text{LogNormal}(x \,|\, 0, 1)$$

5. Soponer que los siguientes datos son una realización de una muestra aleatoria de una distribución desconocida. Utilizar la prueba de Lilliefors para contrastar si la verdadera distribución es normal. Considerar un tamaño $\alpha = 0.02$.

3.36	2.76	2.73	4.33	3.71
2.44	2.06	2.50	3.47	3.98
-2.05	2.75	2.48	1.74	4.98

Aproximar la distribución del estadístico de prueba con simulación (utilizar $m = 5,000$ repeticiones) para calcular el valor crítico y el *p-value* de la prueba.

6. Suponer que los siguientes datos son una realización de una muestra aleatoria de una distribución desconocida F. Utilizar la prueba exponencial de Lilliefors para contrastar si la verdadera distribución es exponencial. Considerar un tamaño $\alpha = 0.08$.

Aproximar la distribución del estadístico de prueba con simulación (utilizar $m = 5,000$ repeticiones) para calcular el valor crítico y el *p-value* de la prueba.

2.07	1.49	0.84	0.56	0.76	0.65	1.87
0.83	0.40	0.84	1.25	0.66	0.95	0.52

7. Mostrar que el estadístico de Cramér-von Mises es

$$W_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(\frac{2i-1}{2n} - F_0(x_{(i)}) \right)^2 .$$

8. Utilizar la prueba de Cramér-von Mises para contrastar las hipótesis

$$H_0 : F(x) = \text{Weibull}(x \,|\, 5, 2) \quad \text{vs.} \quad H_1 : F(x) \neq \text{Weibull}(x \,|\, 5, 2)$$

a partir de los siguientes datos, que se puede asumir son una realización de una muestra aleatoria de F. Considerar un tamaño $\alpha = 0.03$.

1.85	1.59	1.55	1.98	1.73
1.75	2.27	1.39	1.68	1.53
1.47	1.07	1.66	2.05	2.12

Aproximar el valor crítico de la distrbución de W_n^2 y el *p-value* de la prueba con simulación (usar $m = 5,000$ repeticiones).

9. Mostrar que el estadístico de Anderson-Darling es

$$A_n^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left[\log F_0(x_{(i)}) + \log(1 - F_0(x_{(i)}) \right]$$

10. Utilizar la prueba de Anderson-Darling para contrastar las hipótesis

$$H_0 : F(x) = \text{Gumbel}(x \,|\, 3, 4) \quad \text{vs.} \quad H_1 : F(x) \neq \text{Gumbel}(x \,|\, 3, 4)$$

a partir de los siguientes datos, que se puede asumir son una realización de una muestra aleatoria de F. Considerar un tamaño $\alpha = 0.06$.

11.05	3.65	6.40	-2.56	-4.90	-1.35
-2.30	-0.11	-0.92	7.14	-0.20	-2.53

Aproximar el valor crítico de la distrbución de A_n^2 y el *p-value* de la prueba con simulación (usar $m = 5,000$ repeticiones).