Se tiene la muestra de 25 datos (ordenados de menor a mayor por filas) de la Tabla 1.

| | | | | |
|---|---|---|---|---|
| 0.01 | 0.06 | 0.08 | 0.09 | 0.11 |
| 0.16 | 0.22 | 0.23 | 0.29 | 0.30 |
| 0.35 | 0.38 | 0.40 | 0.41 | 0.42 |
| 0.48 | 0.57 | 0.66 | 0.71 | 0.75 |
| 0.78 | 0.79 | 0.82 | 0.88 | 0.90 |

Tabla 1

**a)** Utiliza el estadístico de Kolmogorov-Smirnov para contrastar, a nivel $\alpha = 0.05$, si la población que generó los datos es uniforme en el intervalo [0,1].

**b)** A nivel $\alpha = 0.05$, determina si la población que generó los datos es uniforme en el intervalo [0,1] utilizando un contraste $\chi^2$. (Sugerencia: Utilizar la partición de [0,1] dada por [0,1/3), [1/3,2/3), [2/3,1].

**c)** Compara los resultados de (**a**) y (**b**).

---

**Solución:** En los apartados (**a**) y (**b**) queremos realizar el contraste de bondad de ajuste de la muestra de 25 observaciones de una v.a. $X$ a la distribución uniforme en el intervalo [0,1]:

$$\begin{array}{ll} H_0: & X \sim U[0,1] \\ H_0: & X \nsim U[0,1]. \end{array}$$

**a)** El estadístico de Kolmogorov-Smirnov es $D_n = \|\hat{F} - F\|_\infty$, la distancia del supremo entre la función de distribución empírica $\hat{F}$ y la de la distribución poblacional $F$ cuya bondad de ajuste a la muestra queremos contrastar. La función de distribución uniforme en el intervalo [0,1] es

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0,1) \\ 1 & \text{si } x \geq 1. \end{cases}$$

La función de distribución empírica $\hat{F}$ asociada a la muestra del enunciado es la que vale 0 a la izquierda de $x_{(1)} = 0.01$, 1 a la derecha de $x_{(25)} = 0.90$, es constante a trozos, no decreciente, continua por la derecha y va incrementando su valor en 1/25 en cada punto de la muestra. Los valores de $\hat{F}$ en las observaciones de la muestra aparecen en la Tabla 2 y los valores de $\hat{F} - F$ en la Tabla 3. Por tanto, $D_n = 0.18$. A nivel $\alpha = 0.05$, la región de rechazo es $R = \{D_n > c_{0.05}\}$,

| | | | | |
|---|---|---|---|---|
| 0.04 | 0.08 | 0.12 | 0.16 | 0.20 |
| 0.24 | 0.28 | 0.32 | 0.36 | 0.40 |
| 0.44 | 0.48 | 0.52 | 0.56 | 0.60 |
| 0.64 | 0.68 | 0.72 | 0.76 | 0.80 |
| 0.84 | 0.88 | 0.92 | 0.96 | 1 |

Tabla 2

donde $c_{0.05} = 0.26404$ (ver tabla de la distribución del estadístico $D_n$). Por tanto, no hay suficiente evidencia para rechazar la hipótesis nula.

| | | | | |
|---|---|---|---|---|
| 0.03 | 0.02 | 0.04 | 0.07 | 0.09 |
| 0.08 | 0.06 | 0.09 | 0.07 | 0.10 |
| 0.09 | 0.10 | 0.08 | 0.15 | 0.18 |
| 0.16 | 0.11 | 0.06 | 0.05 | 0.05 |
| 0.06 | 0.09 | 0.10 | 0.08 | 0.10 |

Tabla 3

**b)** Para la partición dada en el enunciado las frecuencias absolutas observadas son: $o_1 = 10$, $o_2 = 8$, $o_3 = 7$ y las esperadas son $e_1 = e_2 = e_3 = 25/3 = 8.33$. El estadístico del contraste $\chi^2$ es

$$\chi^2 = \sum_{i=1}^{3} \frac{o_i^2}{e_i} - 25 = 0.56.$$

La región de rechazo es $R = \{\chi^2 > \chi^2_{2;0.05}\}$, con $\chi^2_{2;0.05} = 5.99$, luego no rechazamos $H_0$.

**c)** Las decisiones de los contrastes de K-S y $\chi^2$ son coherentes, pero la decisión del segundo contraste depende de la elección de la partición que discretiza la distribución uniforme continua. El contraste K-S, sin embargo, está diseñado para hacer contrastes de bondad de ajuste a distribuciones continuas. Por esa razón podría suceder que el test $\chi^2$ fuera menos “sensible” a las diferencias entre la muestra y la distribución uniforme.