Exam 1

1. ## Pregunta

Se ha medido el peso a 8 recien nacidos que participaron en un estudio. Los valores se muestran en la siguiente tabla:

id	peso
1	2,29
2	2,70
3	2,90
4	2,35
5	2,35
6	2,40
7	2,30
8	3,30

La $\overline{\mathrm{X}}$ de la variable peso es 2,57375 y la DE es 0,36559.

NOTA: se utilizan muchos decimales para la $\overline{\mathrm{X}}$ y la **DE**, para evitar dudas debido al redondeo en los cálculos de esta pregunta.

En la siguiente tabla se muestra los valores que son necesarios calcular para realizar la prueba de ``Kolmogorov-Smirnov'' (K-S) que permita decidir si se puede asumir que los valores de **peso** provienen de una población que sigue una distribución normal.

peso	zpeso	observ	expec	dif1	dif2
2,29	-0,78	0,125	0,218	-0,093	-0,218
2,30	-0,75	0,250	0,227	0,023	-0,102
2,35	Valor 1	0,500	0,271	0,229	-0,021
2,35	-0,61	Valor 2	0,271	0,229	0,229
2,40	-0,48	0,625	Valor 3	0,309	0,184
2,70	0,35	0,750	0,637	0,113	-0,012
2,90	0,89	0,875	0,813	0,062	-0,063
3,30	1,99	1,000	0,977	0,023	Valor 4

Observe que faltan algunos valores ¿Cuál de las siguientes respuestas es **FALSA**?

a. Valor 2 = 0,500
b. Valor 3 = 0,316
c. Para realizar la prueba de K-S se ha de calcular "el máximo de las diferencias, en valor absoluto" (max ($|dif|$)). En este ejercicio "el máximo de las diferencias, en valor absoluto" es = 0,206.
d. Valor 4 = -0,063
e. Valor 1 = -0,61

Solución

La primera columna (**peso**) son los valores de peso ordenados de mayor a menor.

La segunda columna (**zpeso**) son los valores de peso *estandarizados*. Se le pide calcular (zx_3). Éste se obtiene:

$$\mathrm{zx}_3 = \frac{\mathrm{x}_3 - \overline{\mathrm{X}}}{\mathrm{DE}} = \frac{2,35 - 2,57375}{0,36559} = -0,61$$

La tercera columna (**observ**) es la proporción de individuos que tiene valores de peso ≤ que el peso de la fila. El valor de $peso_4$ es 2,35 años. Hay 4 individuos que tienen peso ≤ 2,35. Por lo tanto la proporción = $\frac{4}{8} = 0,5$

La cuarta columna (**expec**) se obtiene, utilizando la tabla de la distribución normal, respondiendo a: Si una variable estandarizada siguiera una distribución normal ¿que proporción de individuos tendría valores ≤ al valor que hay en la columna (**zpeso**)?

La figura al final del documento ilustra lo que se acaba de explicar. En esta figura, note que la proporción de individuos que tendría valores ≤ -0,48 es 0,316.

La quinta columna (**dif1**) se obtiene calculando la diferencia entre la proporción acumulada *observada* (**observ**) y la proporción acumulada *esperada* (**expec**) si siguiera una distribución normal. Por ejemplo, para el peso 2,3:

dif1 =0,25 - 0,227 = 0,023

La sexta columna (**dif2**) se obtiene de forma muy similar, pero en lugar de utilizar la proporción acumulada *observada* de la misma fila, la operación se hace con la de la fila anterior. Por ejemplo, para el $peso_8$:

dif2 = 0,875 - 0,977 = -0,102

Una vez realizados todos los cálculos de la tabla, se debe calcular un valor, llámese “*z*”, que es el resultado de aplicar la fórmula:

$$\mathrm{z} = \sqrt{\mathrm{n}} * \text{máx}(|\ dif\ |)$$

Dónde **máx ($|dif|$)** es el máximo, en valor *absoluto*, de las columnas **dif1** y **dif2**. Concretamente, en este ejemplo, máx ($|dif|$) = 0,309.

Para calcular “z”:

z = $\sqrt{8} * 0,309 = 0,874$

Una vez obtenido este valor “z”, se compara en la siguiente tabla:

Nivel de significación	·····	0,2	0,1	0,05	0,02	0,01	0,002	·····
Valor de “z”	·····	1,0727	1,2238	1,3581	1,5174	1,6276	1,8585	·····

Los valores de “z” ≥ 1,3581 conducen a la conclusión que asumir que la muestra de individuos se ha obtenido de una población que sigue una distribución normal es poco probable.

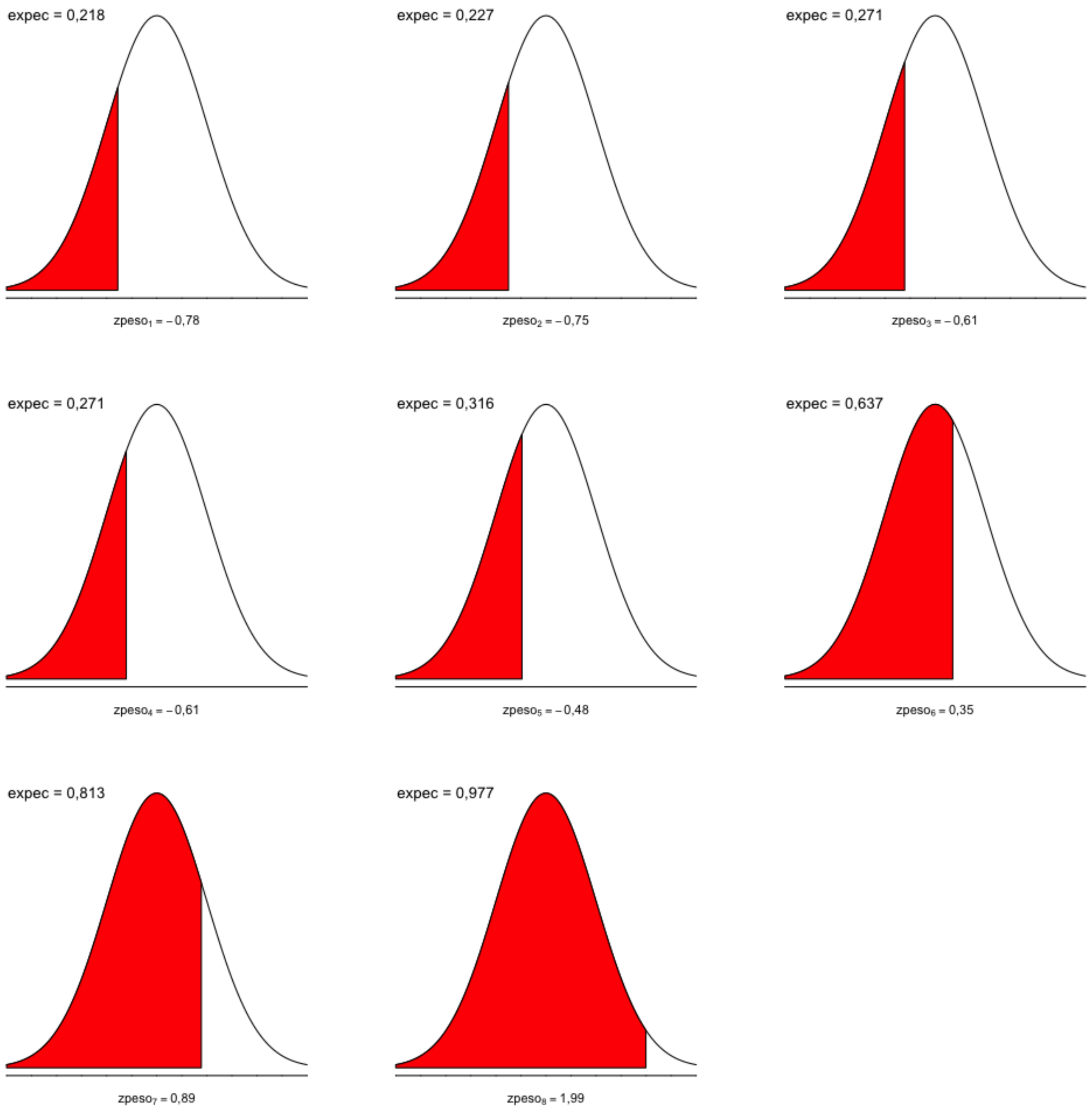
Esta *poca probabilidad* la indica la fila **Nivel de significación**. Obsérvese que los valores ≥ 1,3581 tienen una probabilidad ≤ 0,05, los valores ≥ 1,5174 tienen una probabilidad ≤ 0,02, los valores ≥ 1,6276 tienen una probabilidad ≤ 0,01, etc.

Tradicionalmente el punto de corte para decidir si una situación es *poco* probable se ha situado en probabilidades ≤ 0,05. Por lo tanto los valores de *z* ≥ 1,3581 nos llevaran a la conclusión que lo que se está midiendo (en el ejemplo: peso) es poco probable que se haya obtenido de una una población que sigue una distribuación normal.

Los valores de $z < 1{,}3581$ nos llevaran a la conclusión que nada se opone a que lo que se está midiendo (en el ejemplo: peso) se haya obtenido de una una población que sigue una distribución normal.

El valor de *z* obtenido en este ejercicio ha sido de 0,874, por lo tanto la conclusión es: ***no hay evidencias en contra de que los valores de peso se han obtenido de una población que sigue una distribución normal***.

Explicación gráfica de los valores **expec** que se muestran en la tabla de los cálculos de la prueba de KS:



a. Respuesta Correcta
b. Respuesta Correcta
c. Respuesta FALSA
d. Respuesta Correcta
e. Respuesta Correcta