ESTADÍSTICA PARA TOD@S: ADAPTACIÓN A LA DIVERSIDAD FUNCIONAL

ESTADÍSTICOS DE DISPERSIÓN

UNIVERSIDAD PABLO DE OLAVIDE

ÁREA DE ESTADÍSTICA E IO

Autoras:

Nieves Aquino Llinares

Mª del Pilar Moreno Navarro

UNIVERSIDAD PABLO DE OLAVIDE

Octubre de 2023

ISBN: 978-84-09-55248-1

Introducción

Esta publicación forma parte de un proyecto más ambicioso que se denomina ***Estadística para tod@s: adaptación a la diversidad funcional*** y a través del cual las autoras, conscientes de la falta de material específico en estadística adaptado a diversas discapacidades, quieren poner su granito de arena para alcanzar una universidad más inclusiva, con la realización de píldoras formativas y publicaciones de apoyo al estudio accesibles a diversas discapacidades. En concreto este anexo es un complemento al vídeo formativo sobre los estadísticos de dispersión que las autoras han realizado adaptado a personas con discapacidad auditiva. Además, para personas con discapacidad visual, esta píldora docente cuenta con material de apoyo en soporte informatizado Edico disponible en el Servicio bibliográfico de ONCE.

Cápsula formativa: Estadísticos de dispersión

1ª DIAPOSITIVA **Minuto: 00:10**

Esta píldora formativa pertenece a la serie “Estadística para tod@s: adaptación a la diversidad funcional”, cuyas autoras son Nieves Aquino Llinares y Mª del Pilar Moreno Navarro, profesoras del área de Estadística e Investigación Operativa de la Universidad Pablo de Olavide. Su objetivo es hacer una Universidad más inclusiva.

2ª DIAPOSITIVA **Minuto: 00:30**

Concretamente, en este vídeo hablaremos sobre los estadísticos de dispersión, que nos informan sobre la homogeneidad del grupo y cuya utilidad es complementar la información aportada por otros estadísticos descriptivos como pueden ser los de centralización o forma.

La estadística descriptiva trata de organizar, sintetizar y describir un conjunto de datos a través de diferentes herramientas como las tablas de frecuencias, los gráficos y los estadísticos descriptivos. En concreto, los estadísticos de dispersión que estudiamos a continuación nos informan sobre el grado en el que una distribución de datos se comprime o se estira.

3ª DIAPOSITIVA **Minuto: 01:07**

Por tanto, son medidas que nos informan sobre la variabilidad u homogeneidad de un grupo de valores. Cuando una muestra de datos tiene mayor dispersión que otra indica que sus valores están más alejados de su media aritmética. En el ejemplo mostrado se indican las notas de dos clases.

En la primera clase las notas son 1, 3, 5, 7, 2, 0 y 9 mientras que en la segunda clase las notas son 6, 5, 5, 4, 5, 4 y 5,

Si observamos las notas de la primera clase son muy heterogéneas pues hay estudiantes que obtienen un 0 y otros que obtienen un 9. En este grupo hay mucha variabilidad de las notas.

Sin embargo, en las notas de la segunda clase se aprecia mucha más homogeneidad pues sus valores tienen poca variación pues la nota mínima es 4 y la máxima 6.

4ª DIAPOSITIVA **Minuto: 02:04**

Las características principales de los estadísticos de dispersión es que ofrecen un valor numérico que nos informa sobre el grado de variabilidad de una variable, tomando valores entre 0 e infinito.

Cuando toman el valor 0 todos los valores de la distribución son iguales. En este caso no hay dispersión ninguna y se da la máxima homogeneidad.

Cuanto mayor es el valor del estadístico de dispersión, mayor dispersión hay en los datos.

Por ejemplo, si en una clase todos los estudiantes obtienen un 5 en el examen el grupo será completamente homogéneo. La nota media de la clase será 5 puntos y los estadísticos de dispersión tomarán el valor 0.

5ª DIAPOSITIVA **Minuto: 02:45**

Por otro lado, los estadísticos de dispersión complementan la información dada por las medidas de tendencia central y ayudan a entender hasta qué punto son realmente representativas.

Por ejemplo:

En la clase 1 las notas obtenidas por los estudiantes son 2, 3, 10, 7, 0, 4 y 9 por lo que la nota media es de 5 puntos.

En la clase 2 las notas obtenidas por los estudiantes son 5, 5, 5, 5, 5, 5 y 5 por lo que la nota media de esta clase también es de 5 puntos.

¿Son igual de representativas ambas medias? La respuesta es no. En el grupo 2 hay mucha más homogeneidad en las notas por lo que la media de 5 puntos es mucho más representativa que en el grupo 1.

6ª DIAPOSITIVA **Minuto: 03:36**

Otra característica de los estadísticos de dispersión es que su aplicación es fácil y rápida, permitiendo comparar la dispersión de varias variables con diferentes unidades de medida.

Por ejemplo, si tomamos por un lado la nota del examen de matemáticas de 7 personas obteniéndose los valores 3, 2, 10, 7, 0, 4, y 9 y por otro el valor del nivel de colesterol en sangre de 5 pacientes, obteniéndose los valores 134’6, 89’7, 213’7, 156’3 y 167’9, ¿qué variable tiene los datos más dispersos?

El coeficiente de variación de Pearson permite saber si la variable “Nota en el examen de matemáticas” es más homogénea, o no, que la variable “Nivel de colesterol”.

7ª DIAPOSITIVA **Minuto: 04:32**

Los estadísticos de dispersión más comunes son el Rango, el Rango intercuartílico, la Varianza, la Desviación típica, y el Coeficiente de variación de Pearson. Cada uno de ellos tiene una formulación concreta y sus resultados tienen diferentes unidades de medida. A continuación estudiaremos cada uno de ellos en profundidad.

8ª DIAPOSITIVA **Minuto: 04:57**

El Rango o Recorrido nos informa sobre la diferencia de los valores extremos de una distribución permitiendo conocer entre cuántas unidades de medida se encuentran todos los datos analizados de una variable **X.**

Por ejemplo, si medimos la altura de un grupo de personas el Rango, recorrido o amplitud nos informará sobre entre cuántos centímetros varían todos ellos.

9ª DIAPOSITIVA **Minuto: 05:21**

Se calcula restando al valor máximo el valor mínimo y su unidad de medida es la misma que la de la variable.

Por ejemplo, si en el grupo de estudiantes a la que le hemos medido la altura se observa que la persona más bajita mide 1 metro y 60 centímetros, 1,60 m. y la persona más alta mide 1 metro y 92 centímetros, 1,92 m., el Rango se calcularía como:

$$R = 1{,}92 - 1{,}60 = 0{,}32 \text{ metros}$$

El Rango nos indica que en 32 centímetros oscilan las alturas de todos los estudiantes investigados.

10ª DIAPOSITIVA **Minuto: 06:05**

Otro estadístico de dispersión que nos informa sobre la variación absoluta que sufren las unidades centrales de la distribución, es el Rango Intercuartílico.

Es un estadístico parecido al Rango, pero mientras que el Rango tenía en cuenta a todos los valores observados, el Rango intercuartílico solo tiene en cuenta el 50% de datos centrales, eliminando el efecto de los valores extremos.

Para entender el concepto en la diapositiva se presenta una imagen con 8 bebés ordenados según el peso de menor a mayor, de forma que el primer bebé es el que pesa menos y el último es el que pesa más.

A modo de ejemplo, si el bebé de menor peso alcanza los 2 kilos 260 gramos y el que más pesa alcanza el valor de 5 kilos y 100 gramos, el Rango se calculaba como la diferencia entre el mayor valor y el menor, es decir, 5'100 kilos menos 2'260 kilos obteniéndose un valor de 2 kilos con 840 gramos. Esto quiere decir que el peso de los 8 bebés varía en 2'840 kilogramos.

11ª DIAPOSITIVA Minuto: 07:17

Pero, ¿entre cuántos kilogramos oscilan los bebés que, ni pesan poco ni pesan mucho, es decir, los que se encuentran en el 50% central?

Pues este dato nos lo da el Rango Intercuartílico.

12ª DIAPOSITIVA Minuto: 07:33

Para su cálculo necesitamos conocer el primer y tercer cuartil pues el Rango Intercuartílico es el Percentil 75 menos el Percentil 25.

Para entender el concepto en la diapositiva se presenta de nuevo la imagen con 8 bebés ordenados según el peso de menor a mayor, de forma que el primer bebé es el que pesa menos y el último es el que pesa más. Ahora nos interesa el grupo central de bebés, que ni pesa mucho ni pesa poco, es decir, eliminamos del cálculo el 25% de bebés de cada extremo.

A modo de ejemplo si sabemos que el 75% de los 8 bebés pesa como mucho 3 kilos y 920 gramos y el 25% de ellos solo llega a pesar 3 kilos y medio, el Rango Intercuartílico será 3'920 menos 3'500 que es igual a 420 gramos.

¿Qué significa entonces este valor?

13ª DIAPOSITIVA **Minuto: 08:30**

Que el Rango Intercuartílico sea 0’420 kg nos indica que el peso de los bebés que ni pesan poco ni pesan mucho al nacer, es decir, que ocupan las posiciones centrales en la distribución, tiene una variación de solo 420 gramos de peso.

14ª DIAPOSITIVA **Minuto: 08:47**

Como solo toma para su cálculo los valores centrales, una de sus principales ventajas es que no se ve afectado por valores extremos.

Como ejemplo podemos estudiar dos distribuciones de bebés con el mismo Rango intercuartílico, 0’420 pero que sin embargo difieren mucho en el valor del Rango, 2’840 kilogramos frente a 3’350 kilogramos.

Esto es debido a que la segunda distribución tiene un dato extremo mucho mayor que en la primera distribución pero que no afecta a la variación del 50% central afectando al cálculo del Rango pero no al valor del Rango Intercuartílico.

En esta ocasión se muestran dos grupos de 8 bebés ordenados de menor a mayor peso con la única diferencia de que el último bebé del segundo grupo tiene un mayor peso que el del primer grupo. Esta diferencia en el peso del último bebé hace que las dos distribuciones tengan Rangos diferentes pero el mismo Rango Intercuartílico.

15ª DIAPOSITIVA **Minuto: 09:46**

Otra característica del Rango Intercuartílico es que si se tienen los valores del primer y tercer cuartil su cálculo es muy sencillo, teniendo la misma unidad de medida que la variable. El estudio de la diferencia entre el valor del Rango y el Rango Intercuartílico nos informa de la dispersión y concentración de la distribución. Si es pequeña la diferencia entre estos valores nos indica que la distribución de datos está muy dispersa. Un valor del Rango Intercuartílico mucho menor que el valor del Rango indica que la distribución está muy concentrada.

16ª DIAPOSITIVA **Minuto: 10:22**

Otro estadístico de dispersión muy importante es la Varianza que representa la variabilidad de un conjunto de “n” datos con respecto a su media aritmética.

Para su cálculo se hallan **las distancias de cada valor (x_i) a la media $\bar{x}$** y se realiza la media aritmética de dichas distancias al cuadrado.

Cuando un grupo de datos está muy disperso las distancias de cada valor a la media serán muy grandes y por tanto nos dará un valor de la varianza grande.

Cuando el grupo de datos está muy concentrado las distancias de cada valor a la media serán muy pequeñas y ello permitirá que la varianza nos dé un valor pequeño.

Cuando todos los valores son iguales la varianza será 0 pues las distancias de cada valor a la media darán todas 0.

Se representa por S^2 ó σ^2

y su fórmula es S^2 igual al cociente entre el sumatorio de las distancias de cada valor a la media al cuadrado dividido entre el total de individuos o unidades investigadas, n.

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

17ª DIAPOSITIVA **Minuto: 11:25**

La unidad de medida de la varianza es la de la variable al cuadrado, es decir, en el caso del cálculo de la variabilidad en las alturas, medidas en metros, la varianza tendrá como unidad de medida metros al cuadrado, m^2.

Nunca podrá dar un valor negativo pues es una media de distancias al cuadrado. Por tanto, la varianza toma valores entre 0 e infinito de forma que cuanto más cercano a cero sea su valor más homogéneo será el grupo de datos estudiado.

No obstante, la interpretación aislada de la varianza es compleja siendo más fácil su interpretación cuando comparamos su valor en dos o más grupos de datos.

18ª DIAPOSITIVA **Minuto: 12:06**

Es necesaria para otros cálculos estadísticos.

Si los datos están agrupados, para su cálculo se multiplica cada distancia del valor a la media, todo ello al cuadrado, por la frecuencia absoluta del valor.

Es decir, S^2 es igual al cociente entre el sumatorio de las distancias de cada valor a la media, todo ello al cuadrado, multiplicado por la frecuencia absoluta de cada valor, dividido entre el total de individuos o unidades investigadas, n.

$$S^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{n}$$

Si operamos con las fórmulas vistas hasta ahora obtenemos dos formas más sencillas de calcular la varianza:

Si los datos no están agrupados S^2 es igual al cociente entre la suma de cada valor de la variable al cuadrado dividido entre n y a este cociente hay que restarle la media al cuadrado.

$$S^2 = \frac{\sum {x_i}^2}{n} - \bar{x}^2$$

Si los datos están agrupados en frecuencias la fórmula de la varianza será igual al cociente entre la suma de cada valor de la variable al cuadrado, multiplicado por la frecuencia absoluta, dividido entre n y a este cociente hay que restarle la media al cuadrado.

$$S^2 = \frac{\sum {x_i}^2 n_i}{n} - \bar{x}^2$$

19ª DIAPOSITIVA **Minuto: 13:14**

Repasemos los conceptos de Rango y Varianza con un pequeño ejemplo.

Se recogen las notas de 8 estudiantes de la clase obteniéndose los valores 6, 3, 8, 4, 4, 9, 5, y 9.

Para calcular el Rango nos fijamos en la nota más baja, 3 puntos, y en la nota más alta, 9 puntos. El Rango será 9 menos 3 igual a 6 puntos. Esto quiere decir que la nota del 100% de los estudiantes de la clase varía en 6 puntos.

Para calcular la varianza necesitamos el valor de la media aritmética por lo que para su cálculo sumamos todas las notas (3+4+4+5+6+8+9+9) y las dividimos entre el número de estudiantes analizados, 8, dando un resultado de 6 puntos.

20ª DIAPOSITIVA **Minuto: 14:05**

Para calcular la varianza sumamos las diferencias al cuadrado entre cada nota y la media y todo ello lo dividimos ente 8, es decir,

La varianza sería el cociente entre $(3-6)^2+(4-6)^2+(4-6)^2+(5-6)^2+(6-6)^2+(8-6)^2+(9-6)^2+(9-6)^2$ todo ello dividido entre 8 al ser ocho los estudiantes a los que se les ha analizado la nota. Este cociente da un resultado de 5 puntos^2

Mientras que el Rango nos dice que las notas de la clase oscilan en 6 puntos, la varianza es un valor que mide la dispersión de un grupo y cuanto más cercano a 0 sea este valor más homogéneo será el conjunto de notas.

21ª DIAPOSITIVA **Minuto: 15:03**

Otro estadístico de dispersión muy utilizado en estadística es la desviación típica o desviación estándar. Su cálculo es muy sencillo si se dispone de la varianza pues no es más que la raíz cuadrada positiva de la varianza. Se representa por S, sigma o SD del inglés Standard Deviation.

22ª DIAPOSITIVA **Minuto: 15:2**

La desviación típica tiene la misma unidad de medida que la variable y al ser una raíz cuadrada de un número positivo nunca puede tomar un valor negativo. De igual forma que la varianza cuando más cercano a cero sea ese valor más homogeneidad tiene el conjunto de datos pues su cálculo se fundamenta en la media aritmética de las distancias al cuadrado de cada valor a la media, es decir, en la varianza.

23ª DIAPOSITIVA **Minuto: 15:47**

Para entender cómo se calcula y su interpretación estudiamos las notas de los estudiantes de dos clases. En la clase 1 las calificaciones obtenidas han sido 8, 9, 7 y 5 y en la clase 2 los estudiantes han obtenido un 9, 8, 10 y un 2. Para saber qué grupo de notas es más homogéneo calculamos el valor de la desviación típica en cada uno de los grupos.

24ª DIAPOSITIVA **Minuto: 16:13**

La clase 1 tiene una nota media de 7’25 puntos obtenido como la suma de 5+7+8+9 y dividido entre 4.

La varianza se calcula como el cociente determinado cuyo numerador es $((5-7'25)^2+(7-7'25)^2+(8-7'25)^2+(9-7'25)^2$ y cuyo denominador es 4. Este cociente toma el valor 2'19 puntos^2.

$$s^2=\frac{(5-7'25)^2+(7-7'25)^2+(8-7'25)^2+(9-7'25)^2}{4}=2'19\ puntos^2$$

Para esta clase la desviación típica se calculará como la raíz cuadrada de 2'19 que da como resultado 1'48 puntos.

Si realizamos los mismos cálculos para la clase 2 obtenemos que la media aritmética de las notas obtenidos por sus estudiantes asciende a 7'25 puntos también pues su cálculo sería 2+8+9+10 dividido entre 4.

Para esta clase la varianza se calcularía como el cociente cuyo numerador es $(2-7'25)^2+(8-7'25)^2+(9-7'25)^2+(10-7'25)^2$ todo ello dividido entre 4. Este cálculo nos da un valor para la varianza de la clase 2 de 9'69 puntos^2. La raíz de este valor alcanza un valor de 3'11 puntos y este sería el valor de la desviación típica para la clase 2.

$$s^2=\frac{(2-7'25)^2+(8-7'25)^2+(9-7'25)^2+(10-7'25)^2}{4}=9'69\ puntos^2$$

25ª DIAPOSITIVA **Minuto: 17:55**

Estudiados los estadísticos de centralización y dispersión de ambas clases, ¿qué clase presenta más variabilidad en las notas?

La respuesta a esta pregunta nos la da la varianza y la desviación típica. Dado que la desviación típica en la clase 1 es 1'48 puntos y que en la clase 2 ha sido de 3'11 puntos se puede afirmar que la clase 2 tiene más variabilidad en las notas y por tanto presenta mayor heterogeneidad.

26ª DIAPOSITIVA **Minuto: 18:26**

Pero, dado que ambas clases tienen la misma calificación media obtenida por sus estudiantes, 7'25 puntos, ¿son las notas de ambas clases similares? Es decir, las dos medias representan de igual forma a las calificaciones del grupo?

La menor variabilidad obtenida en el grupo 1, al tener un valor de la desviación típica de 1’48 puntos, frente a los 3’11 puntos de la clase 2, nos indica que su media es más representativa al ser un grupo con las calificaciones más homogéneas pues el valor del estadístico de dispersión, S, se acerca más a cero.

27ª DIAPOSITIVA Minuto: 19:00

Hasta ahora todos los estadísticos de dispersión estudiados tenían la misma unidad de medida que la variable salvo la varianza que era esta misma pero al cuadrado. Ello nos permitía comparar la dispersión entre grupos de variables con la misma unidad de media, pero ¿qué pasa si tengo dos variables con diferente unidad de medida? ¿Cómo sabemos cuál es más homogéneo que otro?

La respuesta a estas preguntas nos la da el Coeficiente de Variación de Pearson pues es una medida de dispersión relativa que mide la relación que hay entre la media y la desviación típica de una variable. Es decir:

El coeficiente de variación es el cociente entre la desviación típica y la media en valor absoluto de forma que si la media toma un valor negativo utilizaremos su valor positivo. Para facilitar su interpretación podemos multiplicar dicho cociente por 100 para convertirlo en un porcentaje.

28ª DIAPOSITIVA Minuto: 19:53

Entre las características del Coeficiente de variación destaca que es adimensional y cuanto más cercano a cero sea su valor más homogeneidad hay en el grupo. Frecuentemente obtenemos valores entre 0 y 100 aunque puede obtenerse valores superiores a 100 cuando haya mucha heterogeneidad en el grupo.

29ª DIAPOSITIVA Minuto: 20:13

Hay que resaltar que cuando la media es cero o cercana a cero, al estar en el denominador, el cociente no se puede calcular o no representa la variabilidad relativa. Un valor menor que 0’3 o menor que el 30% indica que la media es muy representativa pues el grupo presenta unos valores muy homogéneos.

30ª DIAPOSITIVA Minuto: 20:33

Lo importante de este coeficiente es que no tiene dimensiones, es adimensional, y al ser un porcentaje y no depender de las unidades de medida de las variables, nos permite comparar entre diferentes tipos de medición.

Por ejemplo, ¿qué variable es más homogénea, la altura o el consumo de frutas?

Si utilizáramos estadísticos de dispersión como la varianza o el Rango no podríamos responder la pregunta pues la altura está medida en metros y las porciones de fruta en piezas de fruta.

31ª DIAPOSITIVA Minuto: 21:01

A continuación estudiaremos un ejemplo donde para estudiar si un grupo hay más dispersión que en otro debemos hacer uso del Coeficiente de Variación.

Se dedea estudiar la producción en dos granjas, una primera granja apícola, donde se produce miel y otra segunda granja avícola, donde se producen huevos. ¿**Qué producción es más homogénea**, la de la granja 1 o la de la granja 2?

32ª DIAPOSITIVA Minuto: 21:25

La variable medida en la granja 1 indica los kilogramos de miel al mes y se han tomado 6 mediciones que son: 2, 1’2, 3, 3’8, 4’1 y 2’9 kilos.

En la segunda granja la variable medida es el número de huevos producidos al día y se han obtenido los siguientes datos: 253, 246, 310, 225 y 336 huevos al día.

¿Qué granja tiene una producción más homogénea? La respuesta nos la dará el Coeficiente de Variación pues se comparan kilogramos de miel con número de huevos al día.

33ª DIAPOSITIVA Minuto: 22:11

Calculamos la media, varianza y desviación típica para cada granja para poder calcular el coeficiente de variación de Pearson.

En la granja 1, tenemos una media de 2’833 kilogramos, una varianza de 0’988 kilogramos al cuadrado y una desviación típica e 0’994 kilogramos. Con estos datos obtenemos que el coeficiente de variación es el cociente entre 0’994 y 2’833 todo ello multiplicado por 100 dando un resultado de 35,10%.

En la granja 2 la media es 274 huevos/día, la varianza es 1753’2 huevos/día al cuadrado y la desviación típica toma el valor 41’87 huevos/día. Con estos datos el Coeficiente de Variación para la granja 2 toma el valor 15’28%.

34ª DIAPOSITIVA **Minuto: 23:15**

Ya estamos en disposición de responder a la pregunta de qué producción es más homogénea, ¿la de miel o la de huevos? Para ello utilizamos el valor del coeficiente de variación calculado. Como en la granja 1 toma el valor 35’1% y en la granja 2 toma del valor 15’28% se puede decir que la producción de huevos es más homogénea al tener menor coeficiente de variación y ello implica que la media de huevos producidos en la granja 2 es más representativa que la media de kilogramos de miel producido por la granja 1.

35ª DIAPOSITIVA **Minuto: 23:48**

Para finalizar el estudio de los estadísticos descriptivos de dispersión analicemos el siguiente ejemplo. Sofia está investigando las variaciones del precio de un litro de aceite de oliva, de una determinada marca, en 30 supermercados de su municipio. Los datos recogidos los tiene agrupados en intervalos de forma que entre 2’5 euros y menos de 3’5 hay 4 supermercados. Encuentra 5 establecimientos con precios menores de 5 euros pero superiores o iguales a 3’5 euros. Sofía descubre que en 9 supermercados la marca de aceite investigada costaba 5 o más euros pero menos de 8’5. Por último hay 12 centros con un precio entre 8’5 y 10 euros.

Con estos datos se pide calcular la media y la varianza de la distribución de datos.

36ª DIAPOSITIVA **Minuto: 24:39**

Como los precios están agrupados en intervalos es necesario calcular la marca de clase para cada intervalo como medida representativa del precio del litro de aceite y de esta forma obtener los valores de los estadísticos descriptivos solicitados.

La marca de clase del intervalo 2’5 y 3’5 es de 3 euros y se ha calculado sumando los dos extremos y dividiendo entre 2: (2’5+3’5) dividido entre 2 es igual a 3.

De igual forma la marca de clase del intervalo 3’5 y 5 es 4’25 euros. Si nos fijamos en el intervalo de precios entre 5 y 8’5 euros la marca de clase es 6’75 euros y por último la marca de clase o valor representativo del intervalo 8’5 y 10 toma el valor 9’25 euros.

37ª DIAPOSITIVA **Minuto: 25:31**

Para calcular la media aritmética calculamos la suma de cada valor o precio por el número de veces que se repite o frecuencia absoluta y lo dividimos entre el número total de mediciones. Por ello, la media de los datos analizados por Sofía se calcula como $3 * 4 + 4'25 * 5 + 6'75 * 9 + 9'25 * 12$ lo que es igual a 205 y ello se divide entre 30, lo que nos da un precio medio por litro de aceite de 6’83 euros.

Marca de Clase	Precio / litro	Frecuencia
3	[2’5 , 3’5)	4
4’25	[3’5 , 5)	5
6’75	[5 , 8’5)	9
9’25	[8’5 , 10]	12
		Total 30

38ª DIAPOSITIVA **Minuto: 26:07**

La varianza se calcula como

$$\sigma^2 = S^2 = \frac{3^2 * 4 + 4'25^2 * 5 + 6'75^2 * 9 + 9'25^2 * 12}{30} - 6'83^2 = 5'4\ €^2$$

Tomando el valor de 5’4 euros^2.

39ª DIAPOSITIVA **Minuto: 26:41**

Si Marcos ha realizado la misma investigación en otro municipio obteniendo una media de 7’35 euros y una varianza de 2’97 euros^2, ¿Qué media es más representativa? la que obtuvo Sofía en su municipio que recordemos fue 6’83 euros o la que ha obtenido Marcos?

Pues como la varianza obtenida en los datos de Sofía fue de 5’4 euros^2 en el municipio 1 y esta varianza es superior a la obtenida en el municipio 2 por Marcos, nos indica que los precios del litro de aceite en el municipio de Sofía tienen mucha más variabilidad que los precios obtenidos por Marcos. Esto determina que hay más homogeneidad en el precio del aceite en el municipio de Marcos y aunque el precio por litro es mayor en media que en el municipio de Sofía, pues alcanza 7’35 euros frente a 6’83 euros, también sabemos que es más representativa.

40ª DIAPOSITIVA **Minuto: 27:42**

Muchas gracias por la atención.

Esperamos que la serie “Estadística para tod@s” sea un granito más para alcanzar una Universidad más inclusiva.