

ESTADÍSTICA PARA TOD@S: ADAPTACIÓN A LA DIVERSIDAD FUNCIONAL

ESTADÍSTICOS DESCRIPTIVOS: LA MODA



UNIVERSIDAD PABLO DE OLAVIDE

ÁREA DE ESTADÍSTICA E IO

Autoras:

Nieves Aquino Llinares

M^a del Pilar Moreno Navarro

UNIVERSIDAD PABLO DE OLAVIDE

Año 2021

ISBN: 978-84-09-34305-8

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 la moda que las autoras han realizado adaptado a personas con discapacidad auditiva. Además, dicho anexo se encuentra traducido a Braille, estando el documento EDICO, disponible para su consulta en la ONCE. La cápsula docente puede visualizarse en el siguiente enlace:

Cápsula formativa: [La Moda](#)

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 M^a 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: 00:32

Concretamente, esta píldora formativa trata sobre **la moda**, que es un estadístico descriptivo de posición central muy utilizado para conocer los valores o categorías más frecuentes en la variable estudiada.

3ª DIAPOSITIVA

Minuto: 00:46

Definimos “**La moda**” como el valor de la variable con mayor frecuencia absoluta en un conjunto de datos. Es decir, la moda es el valor o categoría de la variable que más se repite.

En el ejemplo se observan varios valores de la variable de forma gráfica siendo la moda la categoría que está rodeada ya que es el valor o categoría que más se repite de todos, en este caso 3 veces.

4ª DIAPOSITIVA

Minuto: 01:11

La moda se calcula de forma diferente dependiendo de si los datos son cualitativos o cuantitativos y si dichos datos están agrupados o no agrupados.

Comencemos viendo cómo se calcula la moda cuando los datos son cualitativos o cuantitativos no agrupados, pero expresados según su tabla de frecuencias. En estos casos la moda es la modalidad o valor de la variable cuya frecuencia absoluta es el máximo de todas las frecuencias absolutas.

x_i	n_i
x_1	n_1
\vdots	\vdots
x_j	n_j
\vdots	\vdots
x_k	n_k
	n

En los gráficos de frecuencias absolutas, relativas y en los histogramas, la moda coincide con el valor donde se alcanza la máxima altura.

5ª DIAPOSITIVA

Minuto: 01:54

Veamos un ejemplo para el cálculo de la moda cuando los datos son cualitativos.

Deseamos conocer el color preferido de 9 personas y los datos son VERDE ROJO GRIS AZUL ROSA ROJO AZUL VERDE y ROJO, observamos que la modalidad que más se repite es ROJO. Por tanto, la moda de la variable “color preferido” es ROJO.

Para su cálculo lo primero que hacemos es construir la tabla de frecuencias, donde en la 1ª columna indicamos las distintas modalidades de la variable, es decir, los colores, y en la 2ª las veces que se repite cada una de las modalidades, la frecuencia absoluta.

x_i	n_i
AZUL	2
GRIS	1
ROJO	3
ROSA	1
VERDE	2
	$n=9$

Una vez que tenemos la tabla de frecuencias, la moda se calcula buscando la modalidad de la variable a la que le corresponde la máxima frecuencia. En este

ejemplo, la máxima n_i es 3 que corresponde a la modalidad ROJO, que es el valor que toma la moda.

6ª DIAPOSITIVA

Minuto: 03:00

En el segundo ejemplo observamos cómo se calcula la moda cuando los datos son cuantitativos no agrupados.

En este caso se recoge el número de veces que una familia asiste al cine en un año y los datos obtenidos son: 6 5 2 4 3 4 5 6 y 5, observamos que la modalidad que más se repite es 5. Por tanto, la moda es 5.

Cuando construimos la tabla de frecuencias, donde en la 1ª columna indicamos las distintas modalidades o valores que toma la variable, es decir, el número de veces que la familia va al cine en un año, y en la 2ª las veces que se repite cada una de las modalidades, la moda se calcula buscando la modalidad de la variable a la que le corresponde la máxima frecuencia.

x_i	n_i
2	1
3	1
4	2
5	3
6	2
	n=9

En este ejemplo, la máxima n_i es 3 que corresponde a el valor de la variable 5 veces al año, por lo que la moda de la distribución toma el valor 5.

7ª DIAPOSITIVA

Minuto: 04:08

Cuando los datos son cuantitativos pero están agrupados en intervalos el cálculo de la moda es más complejo. Para su cálculo es preciso añadir dos columnas a la tabla de frecuencias absolutas:

- Una columna que indica la amplitud del intervalo a_i , la cual se calcula restando el extremo superior del intervalo y el inferior, es decir, el cálculo de la amplitud es igual a:

$$a_i = L_i - L_{i-1}.$$

- Otra columna necesaria es la altura asociada a los intervalos o la densidad y se calcula dividiendo la frecuencia absoluta del intervalo entre la amplitud, es decir, calculamos h_i como n_i dividido entre a_i .

$(L_{i-1}, L_i]$	n_i	$a_i = L_i - L_{i-1}$	$h_i = n_i / a_i$
$[L_0, L_1]$	n_1	a_1	h_1
\vdots	\vdots	\vdots	\vdots
$(L_{j-2}, L_{j-1}]$	n_{j-1}	a_{j-1}	h_{j-1}
$(L_{j-1}, L_j]$	n_j	a_j	h_j
$(L_j, L_{j+1}]$	n_{j+1}	a_{j+1}	h_{j+1}
\vdots	\vdots	\vdots	\vdots
$(L_{k-1}, L_k]$	n_k	a_k	h_k
	n		

En la columna h_i , buscamos el valor máximo, que corresponde al intervalo donde pertenece la moda al que denominamos INTERVALO MODAL. A continuación, la moda se calcula con esta fórmula:

La moda es igual al extremo inferior del intervalo modal L_{j-1} más el cociente entre la diferencia de la altura máxima calculada h_j y la altura del intervalo anterior al modal h_{j-1} , dividido entre la suma de las dos diferencias siguientes, $(h_j$ menos $h_{j-1})$ más $(h_j$ menos $h_{j+1})$. Este cociente hay que multiplicarlo por la amplitud del intervalo modal, a_j .

$$Mo = L_{j-1} + \frac{h_j - h_{j-1}}{h_j - h_{j-1} + h_j - h_{j+1}} a_j$$

8ª DIAPOSITIVA

Minuto: 05:44

Entre las ventajas que tiene la moda destaca que puede calcularse para variables cualitativas y cuantitativas, su cálculo es sencillo y es estable con un n.º grande de observaciones.

Resumiendo:

- Puede calcularse para variables cualitativas y cuantitativas.
- Es sencilla de calcular.
- Es estable con un n.º grande de observaciones.

9ª DIAPOSITIVA

Minuto: 05:59

Pero la moda también tiene inconvenientes:

El 1^{er} inconveniente es que puede haber más de una moda si dos o más valores de la variable presentan la misma frecuencia máxima, es decir, puede que no sea única.

El 2º inconveniente es que la moda no existe si todos los valores de la variable tienen la misma frecuencia.

Y como 3^{er} inconveniente y más destacable es que la moda no siempre se sitúa en el centro de la distribución.

10ª DIAPOSITIVA

Minuto: 06:28

A continuación, vemos tres propiedades de la moda.

La 1ª de ellas es que, si en los datos existen 2 valores que presentan la misma frecuencia y es máxima, la distribución es bimodal.

11ª DIAPOSITIVA

Minuto: 06:44

La 2ª propiedad de la moda es que, si en los datos existen 3 o más valores que presentan la misma frecuencia y es máxima, la distribución es multimodal.

12ª DIAPOSITIVA

Minuto: 06:57

La 3ª propiedad de la moda es que, si en los datos todos los valores presentan la misma frecuencia, la distribución se denominará “amodal”.

13ª DIAPOSITIVA

Minuto: 07:08

Veamos un ejemplo de una distribución bimodal analizando las notas de un estudiante que realiza 12 pruebas y en ellas obtiene las siguientes calificaciones:

x_i : nota obtenida	n_i : frecuencia absoluta
6	5
7,5	2
10	5

- Tiene un 6 en 5 de las pruebas.
- Alcanza un 7,5 en 2 de las pruebas.
- Y obtiene un 10 en 5 de las pruebas realizadas.

¿Qué valor tienen la moda estadística?

En 1^{er} lugar, identificamos la o las frecuencias absolutas máximas. En el ejemplo es 5.

En 2^o lugar, identificamos los valores de la variable asociados a esa frecuencia absoluta máxima 5, que son 6 puntos y 10 puntos.

Por tanto, en el ejemplo estudiado la moda alcanza dos valores diferentes, 6 y 10 puntos. Es decir, se obtienen dos modas por lo que la distribución es bimodal. Los valores que más se repiten son 6 puntos y 10 puntos en los exámenes de dicho estudiante.

14ª DIAPOSITIVA**Minuto: 08:17**

A continuación se proponen tres ejemplos según el tipo de datos y si están agrupados o no. Clica sobre cada uno de los ejemplos para acceder a ellos.

Vídeo 1: Datos discretos (no agrupados)**Minuto: 08:32**

Un analista de riesgos desea estudiar el número de partes de accidente que presentan al año sus clientes con el objeto de hacer una comparación con las aseguradoras de la competencia. Los datos de sus clientes analizados son los siguientes:

$x_i = n^{\circ}$ de partes de accidente	$n_i = n^{\circ}$ de clientes	$f_i = n_i / n$	N_i	$F_i = N_i / n$	$x_i * n_i$
0	132	0,55	132	0,55	0
1	60	0,25	192	0,80	60
2	36	0,15	228	0,95	72
3	12	0,05	240	1	36
	n = 240	1			168

En este ejemplo observamos cómo un analista de riesgos desea estudiar el número de partes de accidente que presentan al año sus 240 clientes, disponiendo de los siguientes datos: 132 clientes han presentado cero partes de accidente en el período estudiado, 60 clientes presentan un parte, 36 clientes han llegado a presentar a la compañía 2 partes en un año y por último 12 clientes presentan 3 partes. ¿Cómo calculamos la moda de la distribución?

Para calcular la moda buscamos la modalidad donde se alcanza el valor máximo de la frecuencia absoluta, n_i . En este caso, el valor máximo de n_i es 132, que

corresponde a la modalidad 1, cero partes. Por tanto, la modalidad más frecuente es no presentar ningún parte de accidentes y esa sería la moda estadística.

Vídeo 2: Datos cuantitativos agrupados en intervalos Minuto: 09:52

A través del siguiente ejemplo vamos a comprobar cómo se calcula la moda en el caso más complejo y es cuando se tienen datos cuantitativos agrupados en intervalos. Para ello vamos a trabajar con los datos recogidos a 35 nadadores que realizan una prueba de 50 metros espalda en una piscina de 25 metros.

Tiempo	Nº de nadadores (n_i)	a_i	$h_i = n_i / a_i$
[22 , 24]	7	2	3,5
(24 , 28]	12	4	3,0
(28 , 30]	10	2	5,0
(30 , 32]	6	2	3,0
	n = 35		

Como podemos observar en la tabla hay 7 nadadores que realizan la prueba en un tiempo que va desde los 22 a los 24 segundos. 12 nadadores tardan entre 24 y 28 segundos, 10 nadadores tardan entre 28 y 30 segundos y por último hay 6 nadadores que tardan entre 30 y 32 segundos. En la tabla presentada se encuentran tanto los valores de la variable agrupados en intervalo como la frecuencia absoluta obtenida, n_i . ¿Qué valor toma la moda?

En primer lugar debemos identificar el INTERVALO MODAL que es aquel que tiene una mayor altura, densidad o h_i . Calculamos por tanto dos columnas nuevas, una con las amplitudes de los intervalos, a_i y otra con las densidades h_i que se calculan como el cociente entre la frecuencia absoluta y la amplitud, es decir, n_i dividido entre a_i .

a_1 será 24 menos 22 por lo que la amplitud del primer intervalo es 2. La amplitud del segundo intervalo es 4 que resulta de restar a 28 el extremo inferior 24. La

amplitud del tercer intervalo, a_3 , será 2 y a_4 será 2 también siguiendo el mismo procedimiento.

Para calcular los valores de h_i se divide la frecuencia entre la amplitud. 7 en el primer intervalo dividido entre su amplitud, 2, lo que resulta un valor de 3.5. Para calcular h_2 realizamos el cociente entre 12 y 4 obteniéndose un valor de 3. De la misma forma se realizan los cálculos para h_3 que vale 5 y para h_4 que toma el valor de 3, resultado de dividir 6 entre 2.

Con estos resultados podemos afirmar que el intervalo modal de la distribución es (28, 30] pues el que alcanza un valor más alto, cinco, en la columna de las h_i .

Para calcular la moda hay que utilizar la siguiente fórmula pues el tipo de datos con el que trabajamos así lo requiere.

La moda se calcula como el Intervalo inferior del intervalo modal más el siguiente cociente: en el numerador se calcula la densidad del intervalo modal, h_i , menos la densidad del intervalo anterior y en el denominador se suma h_i del intervalo modal menos la h del intervalo anterior más h_i del intervalo modal menos la h del intervalo posterior, y ese cociente se multiplica por la amplitud del intervalo modal, a_i .

Siguiendo la fórmula obtenemos que la moda saldrá como resultado de calcular:

$$Mo = 28 + \frac{5 - 3}{5 - 3 + 5 - 3} 2$$

28 más 5 menos 3 dividido entre 5 menos 3 más 5 menos 3 y todo el cociente multiplicado por la amplitud del intervalo modal que es 2.

De esta forma obtenemos un valor de la moda de 29 segundos por lo que podemos concluir que lo más frecuente es realizar la prueba de natación en 29 segundos.

Vídeo 3: Datos cualitativos**Minuto: 15:18**

La siguiente tabla recoge el nº total de medallas del medallero olímpico oficial de los diez países que han obtenido mejor posición en las Olimpiadas de Tokio 2020:

País	Total de medallas (n_i)	$f_i = n_i / n$
Alemania	37	0,06
Australia	46	0,08
Comité Olímpico Ruso	71	0,12
Estados Unidos de América	113	0,19
Francia	33	0,06
Gran Bretaña	65	0,11
Italia	40	0,07
Japón	58	0,10
Países Bajos	36	0,06
República Popular de China	88	0,15
	$n = 587$	1,00

En la siguiente tabla de datos podemos observar el número de medallas que han obtenido en las últimas olimpiadas de Tokio 2020 algunos de los participantes. ¿Qué valor toma la moda estadística? ¿Es única?

Para responder a estas preguntas y calcular la moda en los datos analizados buscamos la modalidad donde se alcanza el valor máximo de la frecuencia absoluta n_i . En este caso, el valor máximo de n_i es 113, que corresponde a la modalidad Estados Unidos de América. Por tanto, la moda es Estados Unidos de América. Solo existe una moda por lo que la distribución es unimodal.

15ª DIAPOSITIVA

Minuto: 16:28

Muchas gracias por la atención.

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