



Tutorial
MT-m6

Matemática 2006



Tutorial Nivel Medio

Estadística Descriptiva



CEPECH
ESPECIALISTAS DE LA PSU

Grupo Educacional Cepech

Estadística descriptiva

Marco Teórico

1. *Estadística descriptiva*

Es la ciencia que trata de la recolección, presentación, análisis y uso de datos para tomar decisiones. Incluye las técnicas que se relacionan con el resumen y la descripción de datos numéricos, gráficos, tablas y diagramas que muestran los datos y facilitan su interpretación.

a) Población:

Es una colección de todos los elementos que estamos estudiando y acerca de los cuales intentamos establecer conclusiones. Poseen alguna característica común que permite agruparlos.

b) Muestra:

Es una colección de algunos de los elementos que componen una población y que son representativos de dicha población.

c) Muestra aleatoria:

Es una muestra al azar. Para que se considere propia y representativa de la población, deberá ser al azar.

d) Frecuencia absoluta: Número de veces que se repite un dato.

e) Distribución de frecuencia es la representación estructurada, en forma de tabla, de toda la información que se ha recogido sobre la variable que se estudia.

1.1 Media aritmética o promedio ($\bar{X}$): La media se calcula al sumar los valores de un conjunto y al dividir el valor de su suma entre el número de valores del mismo.

1.2 Mediana (M_d): Es el valor central de una serie cuando los valores se ordenan según su magnitud, y es aquel que divide a una serie de tal forma que 50% de los valores son menores o iguales que él, y el 50% de los valores son mayores o iguales que él.

1.3 Moda (Mo): Es el valor que con más frecuencia se presenta en un conjunto de datos. Es muy fácil de determinarlo: basta con observar detenidamente el conjunto de datos y ver cuál es el que más se repite. Además puede ocurrir que una distribución tenga dos o más valores que se repitan con la misma frecuencia, en tal caso se tiene dos o más modas. También puede ocurrir que no exista ningún valor que se repita y entonces no habrá moda.

1.4 Marca de clase: Algunos datos se encuentran agrupados en forma de intervalos. En estos casos para el cálculo de media, mediana y moda, no utilizamos ninguno de los extremos del intervalo, sino que el promedio de ellos, al que llamamos marca de clase. A veces los datos nos los entregan en forma gráfica en lo que conocemos como histograma, que son una especie de gráfico de barras en el cuál también se encuentran presentes los intervalos y, por lo tanto, la marca de clase.

Ejercicios:

1. Dado los datos 2, 2, 3, 3, 5, 1, 5, 1, 2, 4, ¿cuál es la frecuencia de la moda?

A) 2
B) 3
C) 4
D) 5
E) La muestra no posee moda
2. En una balanza, se pesan, al mismo tiempo, 20 personas. Si la balanza registra 1200 kilogramos, ¿cuál es el peso promedio de cada persona?

A) 60
B) 50
C) 40
D) 30
E) 20

Resuelva los ejercicios 3, 4, 5 y 6 utilizando la siguiente tabla de frecuencias, que representa las notas de una prueba

Nota	fi
3	1
4	5
5	2
6	1
7	1

3. ¿Cuántas personas rindieron la prueba? (número de datos de la muestra)

- A) 5
- B) 6
- C) 8
- D) 10
- E) 11

4. ¿Cuál es la moda?

- A) 7
- B) 6
- C) 5
- D) 4
- E) 3

5. ¿Cuál es la mediana?

- A) 7
- B) 6
- C) 5
- D) 4
- E) 3

6. ¿Cuál es la media (promedio)?

- A) 4,6
- B) 4,5
- C) 4,4
- D) 4,3
- E) 4,2

7. Una serie de camisas de iguales características valen 5.000, 8.000, 10.000, 10.000 y 15.000. Dados estos datos, ¿cuál (es) de las siguientes características es (son) verdaderas?

- I. La moda es 10.000
- II. La mediana es 10.000
- III. La media (promedio) es 9600

- A) Sólo I
- B) Sólo II
- C) Sólo III
- D) Sólo I y II
- E) I, II y III

8. Dada la siguiente tabla:

Nota	$f(i)$
N1	f1
N2	f2
N3	f3
N4	f4

La media (promedio) es:

- A) $\frac{N1+N2+N3+N4}{4}$
- B) $\frac{N1 \cdot f1 + N2 \cdot f2 + N3 \cdot f3 + N4 \cdot f4}{4}$
- C) $\frac{N1 \cdot f1 + N2 \cdot f2 + N3 \cdot f3 + N4 \cdot f4}{f1 + f2 + f3 + f4}$
- D) $\frac{N1 \cdot f1 + N2 \cdot f2 + N3 \cdot f3 + N4 \cdot f4}{4}$
- E) $\frac{N1+N2+N3+N4}{f1+f2+f3+f4}$
9. Dado los datos 1, 1, 2, 3, 5, 1, 5, 2, calcule la media aritmética, la moda y mediana, respectivamente
- A) 2; 1; 5
- B) 2; 1; 2
- C) 2,5 ; 1 ; 2
- D) 2,5 ; 1 ; 4
- E) 2,5 ; 5 ; 4

Utilice la siguiente tabla de frecuencia para resolver los ejercicios 10 y 11:

Datos	f i
7,0	1
6,0	5
5,0	3
3,0	1

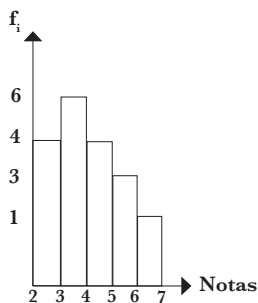
10. Calcule la media aritmética de los datos de la tabla de frecuencia

- A) 5
- B) 5,5
- C) 6,0
- D) 6,5
- E) 6,7

11. Calcule la mediana de los datos de la tabla de frecuencia

- A) 4,5
- B) 5
- C) 5,5
- D) 6,0
- E) 6,6

Para resolver los ejercicios 12,13 y 14, utilice el siguiente histograma, que representa las notas obtenidas por un curso en una prueba de historia:



12. ¿Cuál es la moda de los datos del histograma?

- A) 3
- B) 3,5
- C) 4
- D) 4,5
- E) 5

13. ¿Cuál es la mediana de los datos del histograma?

- A) 2
- B) 3
- C) 3,5
- D) 4
- E) 4,5

14. ¿Cuál es la media aritmética de los datos del histograma?

- A) 3
- B) 3,5
- C) 3,77
- D) 4
- E) 4,5

15. Con los siguientes datos A, A, A, B, B, B, C, C, D, D, D, E, E es posible afirmar que:

- A) No existe moda.
- B) Existen dos modas.
- C) Existen tres modas.
- D) Todos los datos son moda.
- E) El promedio es 1

Alternativas:

Pregunta	Respuesta
1	B
2	A
3	D
4	D
5	D
6	A
7	E
8	C
9	C
10	B
11	D
12	B
13	C
14	D
15	C

Solucionario:

1. Alternativa correcta letra B.

Dado que la moda es el dato que posee la mayor frecuencia (el que más se repite) en este caso la moda es el número 2 y al repetirse tres veces, su frecuencia es precisamente tres.

2. Alternativa correcta letra A.

Como la media aritmética (promedio) se calcula dividiendo la suma de todos los datos de la muestra por el número de datos de la misma, tenemos que este caso los datos suman 1200 kilogramos (peso de todas las personas). Dado que hay 20 personas el número de datos es 20

$$\text{Peso promedio} = \frac{1200}{20} = \quad (\text{Dividiendo})$$

60 kilogramos

3. Alternativa correcta letra D.

Para este ejercicio, basta con sumar la frecuencia de cada nota:

$$1+5+2+1+1 = \quad (\text{Sumando})$$

10 datos

4. Alternativa correcta letra D.

Dado que la moda es el dato que posee la mayor frecuencia (el que más se repite), en este caso la moda es la nota 4 (que se repite 5 veces).

5. Alternativa correcta letra D.

Como la mediana es el valor central de una serie cuando los valores se ordenan según su magnitud, al existir 10 datos, los valores centrales son la nota ubicada en el lugar 5 y la nota ubicada en el lugar 6, en ambos casos una nota 4. Si promediamos los dos valores centrales (los dos cuatros), tenemos que la mediana es también cuatro.

6. Alternativa correcta letra A.

Como la media aritmética (promedio) se calcula dividiendo la suma de todos los datos de la muestra por el número de datos de la misma, tenemos que este caso la suma de todos los datos se obtiene al resolver la suma de cada dato multiplicada por su frecuencia:

$$(3 \cdot 1 + 4 \cdot 5 + 5 \cdot 2 + 6 \cdot 1 + 7 \cdot 1) = \quad (\text{Multiplicando})$$

$$3 + 20 + 10 + 6 + 7 = \quad (\text{Sumando})$$

$$46$$

Y como el número de datos es 10 el promedio es:

$$\frac{46}{10} = \quad (\text{Dividiendo})$$

$$4,6$$

7. Alternativa correcta letra E.

Como los datos ya se encuentran ordenados, para encontrar la mediana basta con observar el valor central, en este caso 10.000, con lo cual la afirmación II es verdadera.

Como el dato más repetido es 10.000 (repetido dos veces), la moda es 10.000, con lo cual la afirmación I es verdadera.

Al sumar los datos el resultado es 48.000 ($5.000 + 8.000 + 10.000 + 10.000 + 15.000$). Como hay cinco datos el promedio es:

$$\frac{48.000}{5} = \quad (\text{Dividiendo})$$

9.600, con lo cual la afirmación III es verdadera

8. Alternativa correcta letra C.

Nota	$f(i)$
N1	f1
N2	f2
N3	f3
N4	f4

Como la media aritmética (promedio) se calcula dividiendo la suma de todos los datos de la muestra por el número de datos de la misma, tenemos que este caso la suma de todos los datos se obtiene al resolver la suma de cada dato, multiplicada por su frecuencia:

$$(N1 \cdot f1 + N2 \cdot f2 + N3 \cdot f3 + N4 \cdot f4)$$

Como el número de datos se obtiene al sumar las frecuencias, tenemos que el número de datos es:

$$N1 + N2 + N3 + N4$$

Entonces el promedio es:

$$\left(\frac{N1 \cdot f1 + N2 \cdot f2 + N3 \cdot f3 + N4 \cdot f4}{N1 + N2 + N3 + N4} \right)$$

9. Alternativa correcta letra C

Como la media se calcula al sumar los valores de un conjunto y al dividir el valor de su suma entre el número de valores del mismo, y como tenemos 8 datos la media es:

$$\frac{1+1+2+3+5+1+5+2}{8} = \quad \text{(Sumando el numerador)}$$

$$\frac{20}{8} = \quad \text{(Simplificando)}$$

$$\frac{5}{2} = \quad \text{(Dividiendo)}$$

2,5

Ya que la mediana es el valor central de una serie cuando los valores se ordenan según su magnitud, lo primero será ordenar los datos, de donde resulta: 1, 1, 1, 2, 2, 3, 5, 5

De donde los valores centrales son los 2. Luego, si los sumamos y promediamos, resulta

$$\frac{2+2}{2} + \frac{4}{2} = 2$$

Finalmente, como la moda es el valor que con más frecuencia se presenta en un conjunto de datos, la moda en este caso es el 1.

10. Alternativa correcta letra B.

Datos	f i
7,0	1
6,0	5
5,0	3
3,0	1

La frecuencia indica el número de veces que se repite un dato, así entonces existe un 7,0, cinco 6,0 ,tres 5,0 y un 3,0 ,o sea, un total de 10 datos. Luego calculando la media

$$\frac{7 \cdot 1 + 6 \cdot 5 + 5 \cdot 3 + 3 \cdot 1}{10} =$$

(Multiplicando el numerador)

$$\frac{7 + 30 + 15 + 3}{10} =$$

(Sumando el numerador)

$$\frac{55}{10} =$$

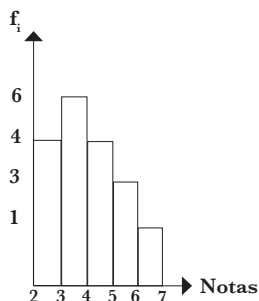
(Dividiendo)

$$5,5$$

11. Alternativa correcta letra D.

La mediana es el valor central, una vez ordenados los datos. Dado que los valores en cuestión son 7, 6, 6, 6, 6, 6, 5, 5, 3, los valores centrales son 6 y 6. Si los promediamos nos da 6 que es el valor de la mediana.

12. Alternativa correcta letra B.



En este caso, los datos que más se repiten son los del intervalo entre 3 y 4. Si calculamos el promedio de ese intervalo, obtendremos su marca de clase, la que equivale a 3,5. Por lo tanto la moda es 3,5

13. Alternativa correcta letra C.

Observemos que hay 4 datos entre 2 y 3, 6 datos entre 3 y 4, 4 datos entre 4 y 5, 3 datos entre 5 y 6, y un dato entre 6 y 7, lo que da un total de 18 datos. Luego, la mediana es el promedio de los datos 9 y 10 que corresponden a datos del intervalo entre 3 y 4, cuya marca de clase es 3,5. Luego la mediana es 3,5

14. Alternativa correcta letra D.

Para calcular la media, es recomendable ordenar los datos del histograma en forma de una tabla de frecuencia, de donde tenemos que:

Notas	f _i
2-3	4
3-4	6
4-5	4
5-6	3
6-7	1

Luego, calculando la marca de clase de cada intervalo

Notas	fi	Marca de clase
2-3	4	2,5
3-4	6	3,5
4-5	4	4,5
5-6	3	5,5
6-7	1	6,5

luego calculando la media

$$\frac{4 \cdot 2,5 + 6 \cdot 3,5 + 4 \cdot 4,5 + 3 \cdot 5,5 + 1 \cdot 6,5}{18} = \quad \text{(Multiplicando el numerador)}$$

$$\frac{10 + 21 + 18 + 16,5 + 6,5}{18} = \quad \text{(Sumando el numerador)}$$

$$\frac{72}{18} = \quad \text{(Dividiendo)}$$

4

15. Alternativa correcta letra C

Los datos que más se repiten son A, B, D con una frecuencia de tres cada uno. Luego estos datos tienen tres modas o, si se prefiere, la muestra es “trimodal”.