Dejamos por ahora el resto de geometría (parte 13) (Vectores y recta) que nunca ha salido y nos centraremos en el bloque siguiente. Al final, si sobra tiempo, retomaremos la parte 13.

Estadística Descriptiva

1 Introducción Estadística Descriptiva

La estadística es una parte de la matemática aplicada y tiene como función ordenar, analizar y decidir sobre la estructura matemática asociada a masas de datos numéricos, obtenidos de la observación de fenómenos (físicos, económicos. psicológicos, sociológicos, etc.)

La tarea de describir y procesar de modo adecuado la masa de datos numéricos, proveniente de observaciones y experimentos, constituye el objetivo de la estadística descriptiva, que es la rama de la estadística que vamos a ver en este tema.

La estadística descriptiva trata de determinar ciertos parámetros estadísticos, parámetros que nos dan información sobre los rasgos de un elemento tipo de colectivo (o población) estudiado, así como otros parámetros que miden la desviación respecto de este elemento tipo.

Resumen:

Es decir disponemos de muchos datos y tenemos que inventar un método para poder entender un conjunto de datos, el método es resumir esa información en unos parámetros, parámetros que llamamos parámetros estadísticos.

2 Parámetros estadísticos.

Una población estadística es un conjunto de individuos, objetos, etc.; sobre los que recae observaciones de un número finito de características.

Veamos cual sería la población y cada uno de los individuos de los siguientes ejemplos.

Ejemplo 1

El peso de los alumnos de una clase
Los individuos serían: cada uno de los alumnos
La población sería: los alumnos que hay en una clase.
Cada fenómeno (o lo que llamamos variable estadística) es el peso del alumno.

Ejemplo 2

Si le hacemos varios análisis de colesterol a un solo alumno de forma seguida
Individuos:
Población:
Fenómeno que estamos midiendo:

Ejemplo 3

Si hacemos unos análisis en un huerto de limoneros, para saber la cantidad de potasio en hoja, para hacer el experimento cogemos 10 hojas de cada árbol.
Individuo:
Población:
Fenómeno que estamos midiendo:

Llamaremos variable estadística al conjunto de valores que adopta una cualidad o propiedad de los elementos de la población estudiada.

2.1 Media de la población

Supongamos que tenemos diez alumnos en clase, los pesamos y obtenemos los siguientes datos

Alumno	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Peso	70	54	57	63	56	78	86	54	69	71

La media de esta población es:

$$media = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{70+54+57+63+56+78+86+54+69+71}{10} = \frac{658}{10} = 6'58$$

¿Pero cómo calculamos el peso medio de todos los alumnos de la universidad?

Tenemos la posibilidad de pesar a todos los alumnos y calcular la media del peso de los alumnos, pero eso es muy costoso. Pero podríamos coger solamente unos cuantos, no podremos calcular la media real, pero tendríamos una media aproximada.

2.2 Concepto de muestra

Se entenderá por muestra una colección finita de elementos de la población estudiada

Entonces podemos obtener la media muestral. (la media de 10 alumnos elegidos entre los miles de alumnos de la universidad)

Alumno	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Peso	70	54	57	63	56	78	86	54	69	71

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{70+54+57+63+56+78+86+54+69+71}{10} = \frac{658}{10} = 6'58$$

Esta media muestral no es la media de la población, pero el parámetro estadístico que es la media muestral se acerca al valor de la media poblacional a medida que el número de alumnos en la muestra que cojamos sea más representativo (en principio más grande).

Supongamos que hacemos lo mismo para las notas de los alumnos, tomamos las notas de 10 alumnos.

Alumno	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Nota	7	5	5	6	7	7	4	8	8	8

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{7+5+5+6+7+7+4+8+8+8}{10} = \frac{65}{10} = 6'5$$

Se llama frecuencia absoluta n_i de x_i al número de veces que aparece repetido dicho valor en los elementos de la muestra. En este caso tendremos los valores discretos de las notas, y los valores de las frecuencias absolutos.

x_i	4	5	6	7	8
n_i	1	2	1	3	3

Podemos entonces calcular la media multiplicando los valores por sus frecuencias absolutas y dividiendo por el número de alumnos.

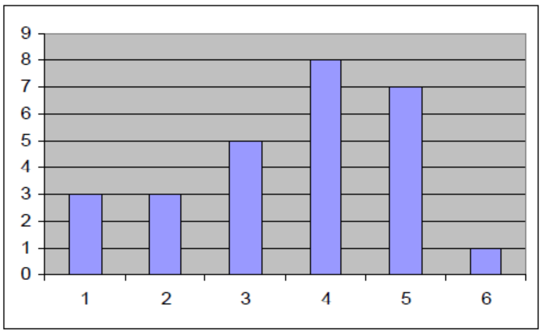
$$\bar{x} == \frac{\sum x_i n_i}{10} = \frac{4*1+5*2+6*1+7*3+8*3}{10} = \frac{65}{10} = 6'5$$

2.3 Varianza de la muestra

<u>1º Ejemplo</u>

Puede haber dos poblaciones que tengan la misma media, pero que sean muy diferentes.

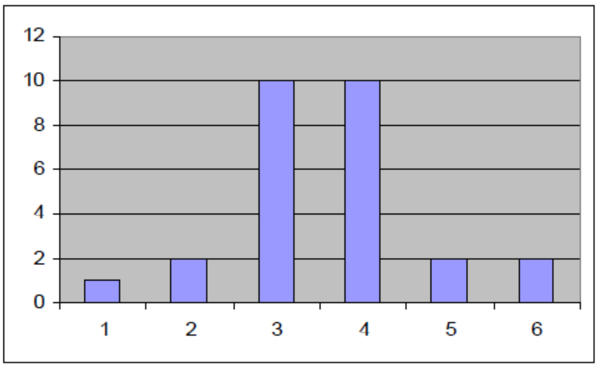
Valores	1	2	3	4	5	6
Frecuencias	3	3	5	8	7	1



$$\bar{x}=\frac{\sum x_i n_i}{27}=\frac{97}{27}=3'59$$

<u>2º Ejemplo</u>

Valores	1	2	3	4	5	6
Frecuencias	1	2	10	10	2	2



$$\bar{x}=\frac{\sum x_i n_i}{27}=\frac{97}{27}=3'59$$

Los dos muestreos tienen la misma media pero en cambio tienen formas muy diferentes en el segundo se concentran todos en los valores 3 y 4. La concentración se mide con la varianza. Y se calcula de la siguiente forma:

$$\sigma^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{m}$$ en el caso de disponer de frecuencias absolutas $$\sigma^2=\frac{\sum_{i=1}(x_i-\bar{x})^2 n_i}{m}$$

Calculémoslo para los dos ejemplos anteriores

1º Ejemplo

Valores	1	2	3	4	5	6		
Frecuencias	3	3	5	8	7	1	27	
$n_i x_i$	3	6	15	32	35	6	97	
							3,59	**media**
$(x_i-\bar{x})$	-2,59	-1,59	-0,59	0,41	1,41	2,41		
$(x_i-\bar{x})^2$	6,72	2,54	0,35	0,17	1,98	5,80		
$(x_i-\bar{x})^2 n_i$	20,16	7,61	1,76	1,33	13,87	5,80	50,52	
							1,87	**Varianza**

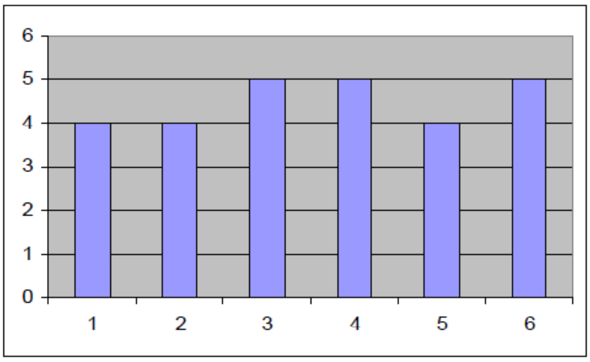
2º Ejemplo

Ejemplo 2

Valores	1	2	3	4	5	6	
Frecuencias	1	2	10	10	2	2	27
$n_i x_i$	1	4	30	40	10	12	97
							3,59 **media**
$(x_i - \bar{x})$	-2,59	-1,59	-0,59	0,41	1,41	2,41	
$(x_i - \bar{x})^2$	6,72	2,54	0,35	0,17	1,98	5,80	
$(x_i - \bar{x})^2 n_i$	6,72	5,07	3,51	1,66	3,96	11,59	32,52
							1,20 **Varianza**

Tercer ejemplo

Valores	1	2	3	4	5	6
Frecuencias	4	4	5	5	4	5



Ejemplo 3

Valores	1	2	3	4	5	6	
Frecuencias	4	4	5	5	4	5	27
$n_i x_i$	4	8	15	20	20	30	97
							3,59 **media**
$(x_i - \bar{x})$	-2,59	-1,59	-0,59	0,41	1,41	2,41	
$(x_i - \bar{x})^2$	6,72	2,54	0,35	0,17	1,98	5,80	
$(x_i - \bar{x})^2 n_i$	26,89	10,15	1,76	0,83	7,92	28,98	76,52
							2,83 **Varianza**

Podemos observar como a medida que se concentran los datos la varianza es menor. La varianza nos da como de puntiaguda es la representación gráfica.

EJERCICIOS DE APLICACIÓN

1 Primer ejercicio sin frecuencias.

Tenemos los siguientes pesos de una muestra de alumnos de una clase

65	67	89	56	45	67	56	57	66	45

Calcula la media muestral y la varianza.

2 Segundo ejercicio sin frecuencias.

Tenemos los siguientes pesos de una muestra de alumnos de una clase

42	93	98	40	51	66	100	98	65	45

Calcula la media muestral y la varianza.

3 Ejercicio

De las muestras de los ejercicios 1 y 2, ¿Cuál de las dos tiene mayor varianza?.

4 Primer ejercicio con frecuencias.

Tenemos la valoración de un líder político H, hemos preguntado a 100 personas y la valoración obtenida del 1 al 4 es la siguiente.

Valoración	1	2	3	4
Frecuencia	25	30	40	5

Calcula la media muestral y la varianza.

5 Segundo ejercicio con frecuencias

Tenemos la valoración de otro líder político J, hemos preguntado a 100 personas y la valoración obtenida del 1 al 4 es la siguiente.

Valoración	1	2	3	4
Frecuencia	55	30	10	5

Calcula la media muestral y la varianza.

6 Con los resultados obtenidos en los ejercicios 4 y 5, ¿cuál es el político más valorado?.