

# Ejemplo aplicación de estadística descriptiva en variables cuantitativas discretas

Se tiene la siguiente muestra representativa de la población:

Numero de computadores por casa en concepción:

4 5 3 2 1 4 5 4 4 3 1 0 4 0 2

# Determinar:

- 1.- variable en estudio y su clasificación
- 2.- unidad o elemento de observación
- 3.- población o muestra
- 4.- ordenamiento de los datos Tabla o distribución de frecuencia respectiva
- 5.- estadísticos de tendencia central, de dispersión y de posición
- 6.- frecuencias: absoluta, absoluta %, acumulada, absoluta acumulada relativa %
- 7.- Simetria, homogeneidad, normalidad.

# Graficas

- Graficar adecuadamente para representar los datos muestrales: grafico de líneas, de barras separadas, circular.
- Diagramas para análisis exploratorio de datos: tallo hoja, de caja.
- Suponiendo homogeneidad : aplicar regla empírica de shebyshev

# Propiedades de la media y la varianza

# Desarrollo:

- 1.-Sea  $X$  la variable en estudio, que representa el NUMERO de computadores por casa , variable cuantitativa discreta.
- 2.- unidad o elemento de observación: casa
- 3.- muestra
- 4.- Tabla de frecuencias respectiva:

Xi N   de comput	Fi frec abs	Frec rel %	Frec abs acum		Xi*fi	(xi-media) <sup>2</sup>	(xi -media) <sup>2</sup> * fi =
0	2		2		0	$(0-2,8)^2=7,84$	$7,84*2=15,68$
1	2		4		2	$(1-2,8)^2=3,24$	$3,24*2=6,48$
2	2		6		4	$(2-2,8)^2=0,64$	$0,64*2=1,28$
3	2		8		6	$(3-2,8)^2=0,04$	$0,04*2=0,08$
4	5		13		20	$(4-2,8)^2=1,44$	$1,44*5=7,2$
5	2		15		10	$(5-2,8)^2=4,84$	$4,84*2=9,68$

# Estadísticos de tendencia central

- Media = (sumatoria de  $x_i$  por  $f_i$ ) / sumatoria de las  $f_i$  =  $42/15 = 2,8$
- Sumatoria de las  $f_i$  =  $n$  = tamaño de la muestra
- El promedio de comp por casa es de 3 (v. cuantitativa discreta)
  
- Mediana = es el dato central con los datos ordenados = 3
- 0 0 1 1 2 2 3 **3** 4 4 4 4 4 5 5
- El 50% de las casas que poseen más computadores, como mínimo tienen 3 computadores
- Moda = dato que mas se repite = 4 , el numero de comp. Que mas casa tienen es 4 aparatos.

# Estadísticos de dispersión

- Varianza= (sumatoria de  $x_i$  menos la media ) al cuadrado por  $f_i$ )/n  
 $= 40,4 / 15 = 2,69$
- Desviación estándar o desviación típica= Raíz cuadrada de la varianza= $\sqrt{2,69} = 1,6$
- Coeficiente de variación o de variabilidad = CV (desviación estándar)/media =  $1,6 / 2,8 = 0,57$  CV%= coef de var. porcentual = 57%



# Estadísticos de posición

- Mediana = 3
- Cuartiles : los datos muestrales agrupados en cuatro o divididos en cuatro grupos y los datos divisorios son los cuartiles: Q1, Q2 y Q3
- $Q1 = 1$
- $Q2 = 3$
- $Q3 = 4$
- Percentiles : los datos divididos en 100 partes o grupos , P1,P2, .....P99

# Simetría, homogeneidad y normalidad

- Simetría: si media = mediana = moda
- En nuestro ejemplo media 2,8 menor a la mediana=3 menor a la moda =4
- Luego no hay simetría, hay asimetría positiva
- Homogeneidad : CV% entre 35% y 45 %
- En nuestro caso es muy superior a 45 % luego nuestros datos son muy heterogéneos
- Si el CV% es inferior a 35 % los datos son muy homogéneos