Probabilidad y Estadística- Unidades 1, 2 y 3

Etapas del Método Científico:

1) Definición del Problema: Las notas obtenidas en PyE- ISI- UTN-FRC- Ultimo turno de examen final
2) Población o muestra:
Población: Todos los elementos o individuos que poseen la característica bajo estudio.
Ejemplo: Todos los estudiantes de PyE- ISI- UTN- FRC- Ultimo turno de ex. Final
Muestra: Parte de la población seleccionada para obtener la información.
3) Recolección de los datos.
Unidad estadística: Cada elemento de la población. Cada estudiante
Unidad de relevamiento: Entidad o individuo a quien recurrimos para que nos brinde la información. Ej: UTN-FRC- Depto. de Exámenes
4) Organización de los datos, presentación (tablas o gráficos), resumen, descripción
5) Si la info. proviene de la población. Análisis y conclusiones.
Si la info. proviene de la muestra- Inferencia Estadística Ej: Dócima de Hipótesis (U10)

Datos estadísticos

Datos cuantitativos: Surgen de respuestas expresada numéricamente.

Datos cuantitativos discretos (Variables cuantitativas discretas) Conteo
Condiciones: Admiten valores enteros y no están expresadas en cierta unidad de medida
Cantidad de hijos

Datos cuantitativos continuos (Variables cuantitativas continuas) Medición
Estatura de un niño: 1 metro 1,20 mts Pueden admitir valores con cifras decimales
Y admiten unidad de medida (Debe cumplirse al menos una de estas dos condiciones.
Edad: 20 años 20 meses 20 días (variable continua)
Vida útil de un artículo (variable continua) 30 meses

Cantidad de materias aprobadas de un estudiante: Variable discreta

Cantidad de materias aprobadas:

5 estudiantes: 2 2 4 3 1
Promedio:
Cantidad de materias aprobadas Promedio: 12/5= 2,4

Si el promedio se obtiene a partir de una variable discreta (cantidad de materias aprobadas) entonces no lleva unidad de medida, a pesar de ello es una variable continua porque al calcularlo podemos obtener un resultado con cifras decimales.

Otro ejemplo:
5 estudiantes: 2 2 2 3 1
Promedio: 2

Si una variable puede admitir tanto valores enteros como valores con cifras decimales entonces se considera que es una variable continua

Datos cualitativos o categóricos: Están expresados en forma literal

Datos cualitativos nominales: No podemos especificar determinado orden o jerarquía en las categorías que comprende.

Estado civil: Casado soltero viudo
Género: Masculino- Femenino

Datos cualitativos Ordinales: Es posible establecer un orden en las categorías
Nivel de satisfacción con el empleo: MS-S-PS-I I-PS-S-MS

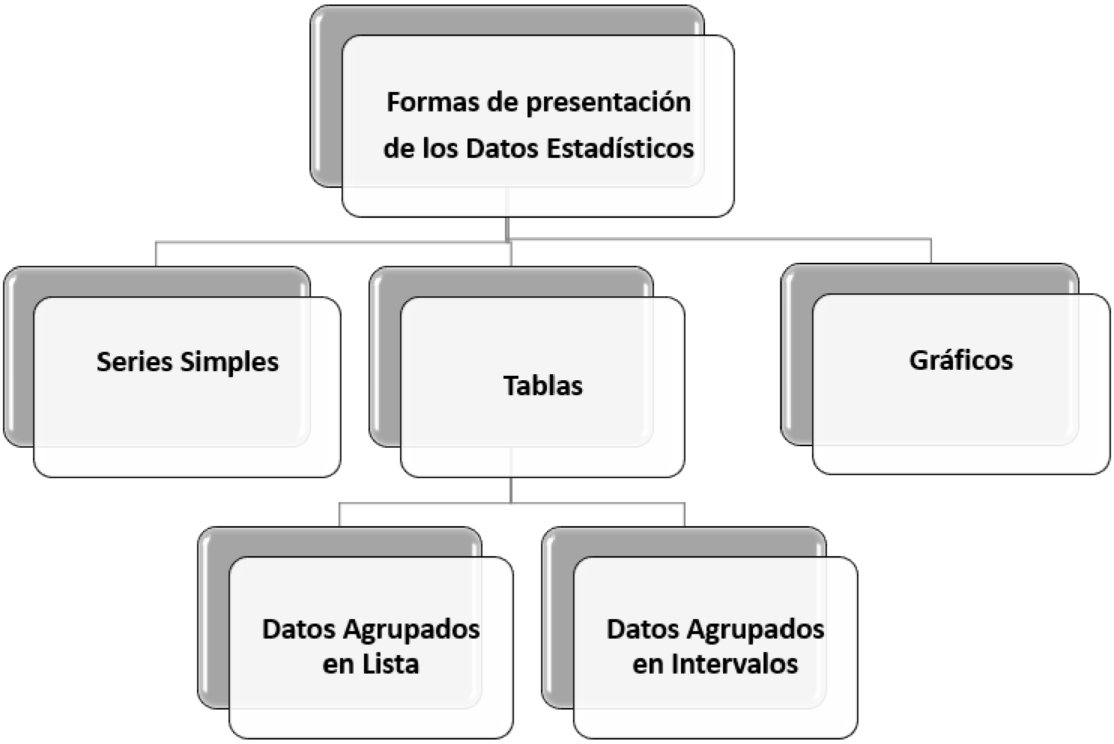
Se desea realizar un estudio referido a los empleados de una fábrica textil, para ello fueron encuestados los 50 primeros empleados del registro (fueron seleccionados
50 empleados)

Si el enunciado hubiera expresado:
Se encuestaron **los** empleados de la fábrica, entonces no habría muestra, se trabajaría con toda la población.

1) Haciendo uso del vocabulario técnico, defina cada concepto, **siempre en relación al caso propuesto,** y detalle el tipo de dato estadístico que se presenta.
   a. Población: Todos los empleados de la fábrica textil.
   b. Muestra: 50 primeros empleados
   c. Unidad estadística: Cada empleado
   d. Unidad de relevamiento: Cada empleado
   e. Características de las variables bajo estudio.

   Características de las variables bajo estudio:

   Salario: variable continua
   Cantidad de hijos: variable discreta
   Área: dato cualitativo nominal



Serie
Datos agrupados en lista:
Este tipo de tabla resulta adecuado cuando se presentaron pocos valores distintos de la variable bajo estudio y se repiten.

Valores observados: Aquellos valores que se presentaron al recolectar los datos. **De entre todos los posibles valores que puede asumir una variable, los valores observados son <u>solo</u> aquellos valores obtenidos al recolectar los datos.**
Valores observados: Pueden incluir sólo a algunos de los posibles valores de la variable En nuestro caso, los valores observados están comprendidos entre 1 y 5

Valores posibles: Todos los valores que puede asumir la variable

Cantidad de hijos: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 o más hijos

Notas en un cierto examen de 5 estudiantes encuestados:
6 7 8 9 9 Los valores observados comprenden solo a algunos de los posibles valores de la variable nota

Posibles valores de la variable nota: Entre 1 y 10

Los valores observados siempre están comprendidos dentro de los posibles valores.

2) A partir de la **Base de Datos** suministrada y mediante el empleo del InfoStat construya la/s tabla/s de Distribución de Frecuencias con agrupación en lista o por intervalos, según corresponda. Calcule todas las frecuencias que conozca de las siguientes variables:

   a. Cantidad de hijos
   b. Salario (tabla con 5 intervalos de amplitud 8)
   c. Área

Base de Datos

ID	Salarios en $	Hijos	Área
1	153	3	Administración
2	134	3	Administración
3	138	2	Administración
4	137	4	Administración
5	128	5	Administración
6	123	5	Administración
7	148	2	Administración
8	138	2	Diseño
9	146	1	Diseño
10	127	1	Diseño
11	129	1	Diseño
12	125	3	Diseño
13	122	2	Diseño
14	138	3	Comunicación
15	146	3	Comunicación
16	132	3	Comunicación
17	139	3	Comunicación
18	146	1	Comunicación
19	146	1	Ingeniería
20	144	1	Ingeniería
21	147	1	Ingeniería
22	146	5	Ingeniería
23	137	2	Ingeniería
24	140	3	Ingeniería
25	137	3	Ingeniería
26	138	2	Ingeniería
27	145	3	Ingeniería
28	151	3	Administración
29	137	4	Administración
30	128	2	Administración
31	137	4	Administración

32	148	4	RRHH
33	145	1	Administración
34	129	4	Administración
35	142	2	Administración
36	134	4	Administración
37	135	4	Administración
38	124	4	Administración
39	126	2	Administración
40	141	4	Administración
41	131	4	Administración
42	152	5	Administración
43	132	4	Diseño
44	117	3	Diseño
45	136	4	Diseño
46	147	3	Diseño
47	128	4	Diseño
48	138	3	Diseño
49	136	3	Diseño
50	138	2	RRHH

a. Datos agrupados en lista: Se utiliza cuando al recolectar los datos se obtuvieron pocos valores diferentes (Pocos valores diferentes y muchas repeticiones)

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Fuente: BD- clase PyE

MC: Clases: Diferentes valores que se presentaron al recolectar los datos.

b. Datos agrupados en intervalos: Se utiliza cuando al recolectar los datos se obtuvieron muchos valores diferentes

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

Fuente: BD- clase práctica PyE

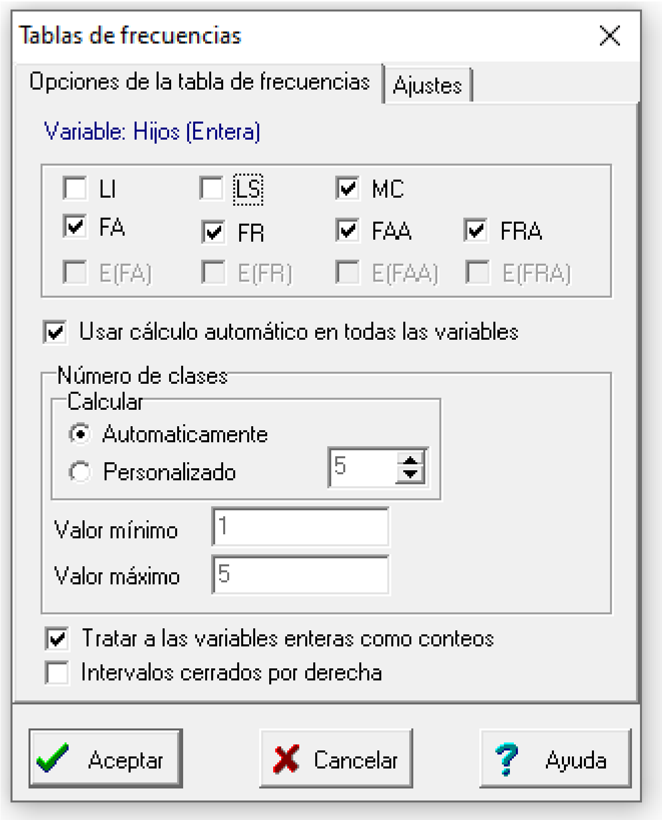
c. Para datos cualitativos nominales no deben obtenerse las frecuencias acumuladas.

Variable	Clase	Categorías	FA	FR
Área	1	Administración	21	0.42
Área	2	Comunicación	5	0.10
Área	3	Diseño	13	0.26
Área	4	Ingeniería	9	0.18
Área	5	RRHH	2	0.04

Fuente: BD- clase práctica PyE

FA= ni Frecuencia absoluta simple
FR= hi Frecuencia relativa simple
FAA= Ni Frecuencia absoluta acumulada
FRA= Hi Frecuencia relativa acumulada

a. Datos agrupados en lista para la variable cantidad de hijos.



Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Fuente: BD- clase PyE

Infostat efectúa automáticamente los siguientes cálculos:

i (orden de la clase)	yi (valor de la clase)	n_i (FA) cantidad	N_i (FAA) cantidad	hi= **ni**/**n** (FR) proporción	H_i= **N_i**/**n** (FRA) proporción
1	1	**8**	**8**	**8**/50=0,16	**8**/50=0,16
2	2	**n_2=10**	**N_2**= 8+10= **18**	**10**/50= 0,20	0,16+0,2 o **18**/50=0,36
3	3	**15**	**N_3** = 8+10+15= **33**	**15**/50= 0,30	**33**/50=0,66
4	4	**n_4=13**	8+10+15+13= **46**	**13**/50= 0,26	**H_4** =**46**/50= 0,92
5	5	**4**	**n= 50**	**h_5=4**/50= 0,08	**50**/50

yi: cantidad de hijos

Clases (yi): Cada uno de los diferentes valores observados para la variable.
n= 50 (cantidad total de valores observados)

Frecuencia Absoluta Simple o Frecuencia Absoluta: ni (FA)
Cantidad de veces que se presenta cada valor observado

¿Cuántos empleados tienen 2 hijos? **n_2=10**
Hay 10 empleados que tienen 2 hijos

¿Cuántos empleados tienen 3 hijos? **n_3=15**
Hay 15 empleados que tienen 3 hijos

Frecuencia Absoluta Acumulada: Ni (FAA)
Ni= n_1+ … + ni
Cantidad de observaciones que hay desde la primera clase hasta una clase i

¿Cuántos empleados tienen 3 o menos hijos? **N_3** = 8+10+15= **33**
Hay **33** empleados que tienen 3 o menos hijos

¿Cuántos empleados tienen menos de 4 hijos? **N_3** = 8+10+15= **33**
Hay **33** empleados que tienen 3 o menos hijos

¿Cuántos empleados tienen 2 o menos hijos? **N_2**= 8+10= **18**
Hay 18 empleados que tienen 2 o menos hijos.

¿Cuántos empleados tienen menos de 3 hijos? **N_2**= 8+10= **18**
Hay 18 empleados que tienen menos de 3 hijos.

Frecuencias absolutas: Cantidades

Frecuencias relativas: Proporciones= cantidad parcial / cantidad total(n)
Frecuencia relativa= frecuencia absoluta/**n**

Frecuencia Relativa simple: h_i= n_i/n (FR)
Proporción de observaciones que hay en una clase i

¿Qué proporción de empleados tienen 3 hijos? **15**/50= 0,30 **h_3= 0,30**
0,30 es la proporción de empleados que tienen 3 hijos

¿Qué porcentaje de empleados tiene 3 hijos? **h_3*100= 0,30*100**
= 30%
El 30% de los empleados tienen 3 hijos.

¿Qué proporción de empleados tienen 4 hijos? **13**/50= 0,26
0,26 es la proporción de empleados que tienen 4 hijos
No es correcto 0,26%

¿Qué porcentaje de empleados tienen 4 hijos? 0,26*100= 26%
El 26% de los empleados tienen 4 hijos

Frecuencia relativa acumulada: Hi= **Ni**/**n** o Hi= h_1 + ….. + hi **(FRA)**

Proporción de observaciones comprendidas hasta una clase i

Se obtiene sumando sumando las Frecuencias Relativas (simples) hasta una clase i
H_i= h_1 + ….. + hi **(FRA)**

También puede obtenerse mediante el cociente entre la Frecuencia Absoluta Acumulada sobre el **total de observaciones (n):** H_i= **Ni**/**n**

¿Qué proporción de empleados tienen menos de 3 hijos?
¿Qué proporción de empleados tienen 2 o menos hijos?
0,20 + 0,16= 0,36 o bien:
18/50= 0,36
H_2 = 0,36
0,36 es la proporción de empleados que tienen 2 o menos hijos

¿Qué proporción de empleados tienen 3 o menos hijos? Son los empleados comprendidos hasta la clase 3
¿Qué proporción de empleados tienen menos de 4 hijos? Son los empleados comprendidos hasta la clase 3

0,16 + 0,20 + 0,30= 0,66 o bien:
33/50= 0,66
$\mathbf{H_3}$ = 0,66

0,66 es la proporción de empleados que tienen 3 o menos hijos.
0,66 es la proporción de empleados que tienen menos de 4 hijos.

En el caso planteado:

yi: cantidad de hijos
ni: cantidad de empleados que tienen cierta cantidad de hijos.
Ejemplos:
n_2=10 Es la cantidad de empleados que tienen 2 hijos.
Hay 10 empleados que tienen 2 hijos
¿Cuántos empleados tienen 2 o menos hijos? 8+10= 18 FAA

Otro ejemplo:
y_1: primera clase y_2: segunda clase

i (orden de la clase)	yi (valor de la clase
1	y_1= 2
2	y_2= 4
3	y_3= 7
4	11
5	13

y_1= 2 El valor de la primera clase es 2
y_2= 4 El valor de la segunda clase es 4
y_3= 7 El valor de la tercera clase es 7 y así sucesivamente

b. Salario (tabla con 5 intervalos de amplitud 8)

Tabla de datos agrupados en intervalos:

Resulta adecuado agrupar los datos en intervalos cuando existen muchos valores observados diferentes y dichos valores se repiten.

Pasos para obtener la tabla con intervalos:

Valor mínimo: 117

Valor máximo: 153

Recorrido: Valor máximo - Valor mínimo = 153 – 117

R= 36

Amplitud: Es la diferencia entre el límite superior y el límite inferior de cada intervalo (Definición)

Amplitud= límite superior – límite inferior (Definición)

Fórmula de cálculo de la Amplitud = Recorrido/cantidad de intervalos

= R/m

= 36/5

Amplitud = 7,2

Si obtenemos un resultado con cifras decimales, **siempre** se redondea al entero inmediato superior.

Amplitud: ci = 8

Recorrido ampliado: R'= ci * m **m: cantidad de intervalos**

R' = 8 * 5

R' = 40

Diferencia entre Recorrido Ampliado y Recorrido original:

R' – R= 40 – 36
R' – R= 4

La diferencia (R´- R= 4) se reparte equitativamente entre el primero y el último intervalos de la siguiente manera:

Valor mínimo – 2 = 117 – 2
= 115
Límite inferior del primer intervalo = 115

Valor máximo + 2 = 153 + 2
= 155
Límite superior del último intervalo = 155

Si la diferencia R´- R fuese por ejemplo igual a 5, se repartiría de la siguiente manera:

Valor mínimo – 2
Valor máximo + 3

O bien:

Valor mínimo – 3
Valor máximo + 2

i	yi-1	yi
1	**115**	
2		
3		
4		
5		**155**

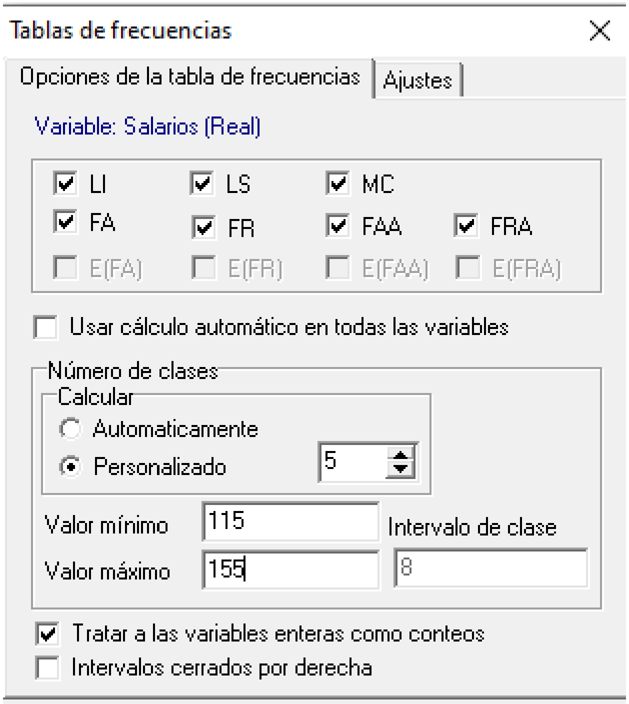
Límite superior de cada intervalo= Límite inferior + amplitud

i	yi-1	yi
1	**115**	**115 + 8 = 123**
2	**123**	**123 + 8 = 131**
3	**131**	**131 + 8 = 139**
4	**139**	**119 + 8 = 147**
5	**147**	**147 + 8 = 155**

Cada intervalo incluye los valores mayores o iguales que el límite inferior

y menores que el límite superior.

La tabla con intervalos puede generarse haciendo uso de Infostat:



Para **obtener intervalos abiertos a derecha**, la opción “Intervalos cerrados por derecha” debe dejarse **destildada.**

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

Infostat efectúa automáticamente los siguientes cálculoS

i	yi-1 yi	FA (n_i) Frecuencia Absoluta simple	FR (h_i)=ni/n Frecuencia Relativa simple	FAA (N_i) Frecuencia Absoluta Acumulada	FRA (Hi)= Ni / n Frecuencia Relativa Acumulada
1	[115 123)	2	2/50=0,04	2	2/50=0,04
2	[123 131)	10	10/50=0,20	N_2= 2+10=12	0,04+0,20=0,24 o 12/50=0,24
3	[131 139)	19	19/50=0,38	N_3=2+10+19= 31	31/50=0,62
4	[139 147)	12	12/50=0,24	2+10+19+12= 43	43/50=0,86
5	[147 155]	7	7/50=0,14	50	1

Amplitud: LS – LI = 8

Frecuencia Absoluta simple:

¿Cuántos empleados tienen salarios mayores o iguales que 131 y menores que 139? **n_3= 19 Frecuencia Absoluta simple**

Hay 19 empleados que tienen salarios mayores o iguales que 131 y menores que 139 (entre 131 y 139)

¿Cuántos empleados tienen salarios mayores o iguales que 139 y menores que 147? **n_4= 12 Frecuencia Absoluta simple**

Frecuencia Absoluta Acumulada

¿Cuántos empleados tienen salarios menores que 147? Es decir:

¿Cuántos empleados tienen salarios menores mayores o iguales que 115 y menores que 147?

Son los empleados comprendidos hasta el cuarto intervalo:

2+10+19+12= 43

Hay 43 empleados que tienen salarios mayores o iguales que 115 y menores que 147

Hay 43 empleados que tienen salarios menores que 147

N_4= 43

¿Cuántos empleados tienen salarios menores que 139? **N_3=2+10+19= 31**

¿Cuántos empleados tienen salarios mayores o iguales que 115 y menores que 139? **N_3=2+10+19= 31**

Hay 31 empleados que tienen salarios menores que 139

Frecuencia Relativa simple:

¿Qué proporción de empleados tienen salarios mayores o iguales que 139 y menores que 147? Es decir:

¿Qué proporción de empleados tienen salarios comprendidos entre 139 y 147? **Frecuencia Relativa simple**

h_4= 12/50 h_4=0,24

0,24 es la proporción de empleados tienen salarios mayores o iguales que 139 y menores que 147

0,24 es la proporción de empleados tienen salarios comprendidos entre 139 y 147

¿Qué proporción de empleados tienen salarios comprendidos entre 123 y 131? Es decir:

¿Qué proporción de empleados tienen salarios mayores o iguales que 123 y menores que 131? **h_2 = 10/50=0,20 Frecuencia Relativa simple**

0,20 es la proporción de empleados tienen salarios mayores o iguales que 123 y menores que 131

¿Qué proporción de empleados tienen salarios menores que 139? Es decir:

¿Qué proporción de empleados tienen salarios mayores o iguales que 115 menores que 139? **Frecuencia Relativa Acumulada**

H_3= 31/ 50 H_3=0,62

0,62 es la proporción de empleados tienen salarios mayores o iguales que 115 y menores que 139

0,62 es la proporción de empleados tienen salarios menores que 139

i	yi-1	yi	FA (ni) Frecuencia absoluta simple	FAD (N´ i) frecuencia absoluta desacumulada	FR (hi)=ni/n frecuencia relativa simple	FRD (N´ i) = N´ i /n frecuencia relativa desacumulada
1	115	123	2	50	2/50=0,04	50/50= 1
2	123	131	10	50 - 2= 48	10/50=0,20	48/50= 0,96
3	131	139	19	48 - 10= 38	19/50=0,38	38/50= 0,76
4	139	147	12	38 – 19= 19	12/50=0,24	19/50= 0,38
5	147	155	7	19- 12= 7	7/50=0,14	7/50= 0,14

Cada intervalo está comprendido entre un límite inferior (LI o yi-1) y un límite superior (LS o yi).

La diferencia entre el límite superior y el límite superior de cada intervalo se denomina Amplitud.

Cada intervalo está comprendido entre un límite inferior y un límite superior.

Los límites son los valores dentro de los cuales se encuentra comprendido cada intervalo.

Cada intervalo es cerrado a izquierda y abierto a derecha.

Cada intervalo incluye los valores mayores o iguales que el límite inferior y menores que el límite superior.

Cada intervalo incluye al límite inferior pero no incluye al límite superior.

Ejemplo:

El segundo intervalo incluye los salarios mayores o iguales que 123 y menores que 131.

Si un valor observado coincide con un límite superior, entonces queda incluido en el siguiente intervalo.

FA (ni): Frecuencia Absoluta o Frecuencia Absoluta Simple

Definición: Cantidad de elementos que hay en cada intervalo i. Cantidad de elementos mayores o iguales que el límite inferior y menores que el límite superior.

En el caso planteado:

Cantidad de salarios que hay en cada intervalo i. Cantidad de salarios mayores o iguales que el límite inferior y menores que el límite superior.

Ejemplo: Frecuencia absoluta del segundo intervalo: n_2= 10

Interpretación: Cantidad de empleados que tienen salarios mayores o iguales que 123 y menores que 131

FR (hi): Frecuencia Relativa

Definición: Proporción de elementos que hay en cada intervalo i.

En el caso planteado:

Proporción de salarios que hay en cada intervalo i. Proporción de salarios mayores o iguales que el límite inferior y menores que el límite superior.

Cálculo: Cociente entre la Frecuencia Absoluta y la cantidad total de observaciones (n):

FR= FA / n

Ejemplo: Frecuencia relativa del segundo intervalo: h_2= 10 / 50

Interpretación: Proporción de empleados que tienen salarios mayores o iguales que 123 y menores que 131

FAA (Ni): Frecuencia Absoluta Acumulada

Definición: Cantidad de elementos menores que el límite inferior del intervalo i

En el caso planteado:

Cantidad de salarios menores que el límite inferior del intervalo i

Cantidad de salarios mayores o iguales que el límite inferior y menores que el límite superior.

Cálculo: La Frecuencia Absoluta Acumulada se obtiene sumando las frecuencias absolutas desde el primer intervalo hasta un intervalo i.

Ejemplo: Frecuencia Absoluta Acumulada del tercer intervalo:

$N_3 = 2 + 10 + 19$ $N_3 = 31$

FRA (Hi): Frecuencia Relativa Acumulada

Definición: Proporción de elementos que están comprendidos hasta un intervalo i.

En el caso planteado:

Proporción de salarios que están comprendidos hasta un intervalo i. Proporción de salarios menores que el límite superior de un intervalo i.

Cálculo: Cociente entre la Frecuencia Absoluta Acumulada y la cantidad total de observaciones (n):

FRA= FAA / n

Ejemplo: Frecuencia relativa del segundo intervalo: $H_2= 12 / 50$

Interpretación: Proporción de empleados que tienen salarios menores que 131

Indicar tipo, simbología y valor de las siguientes frecuencias:

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

¿Cuántos obreros perciben semanalmente menos de 131 pesos?

Tipo: Frecuencia absoluta acumulada.

Simbología: N_2 (Salarios comprendidos hasta el intervalo 2)

Valor: 2+10= 12

Interpretación:

Hay 12 obreros que perciben semanalmente menos de 131 pesos.

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

¿Cuántos obreros perciben semanalmente entre 139 y 147 pesos?

Tipo: Frecuencia absoluta simple.

Simbología: n_4 (Salarios comprendidos en el intervalo 4)

Valor: 12

Interpretación:

Hay 12 obreros que perciben semanalmente 139 pesos o más y menos de 147 pesos.

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

¿Qué proporción de obreros perciben semanalmente menos de 139 pesos?

Tipo: Frecuencia relativa acumulada.

Simbología: H_3

Valor: 31/50= 0,62 (Salarios comprendidos hasta el intervalo 3)

Interpretación:

Es la proporción de obreros que perciben semanalmente menos de 139 pesos

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

¿Qué proporción de obreros perciben semanalmente entre 147 y 155 pesos?

Tipo: Frecuencia relativa simple.

Simbología: h_5

Valor: 7/50= 0,14 (Salarios comprendidos en el intervalo 5)

Interpretación: Es la proporción de obreros que perciben semanalmente 147 pesos o más y 155 pesos o menos **(este intervalo incluye al límite superior por ser el último intervalo)**

Seleccione un renglón de la tabla de distribución de frecuencia e interprete en términos del problema planteado, para cada variable del punto 2), las frecuencias obtenidas (para un valor o un intervalo de la variable).

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

Interpretación correspondiente a las frecuencias del segundo intervalo:

n_2= 10 Frecuencia absoluta simple del segundo intervalo. Es la cantidad de empleados que tienen salarios mayores o iguales que 123 y menores que 131 pesos.

h_2= 0,20 Frecuencia relativa simple del segundo intervalo. Es la proporción de empleados que tienen salarios mayores o iguales que 123 y menores que 131 pesos.

N_2= 12 Frecuencia absoluta acumulada del segundo intervalo. Es la cantidad de empleados que tienen salarios menores que 131 pesos.

H_2= 0,24 Frecuencia relativa acumulada del segundo intervalo. Es la proporción de empleados que tienen salarios menores que 131

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Interpretación correspondiente a las frecuencias de la tercera clase:

n_3= 15 Frecuencia absoluta simple de la tercera clase. Es la cantidad de empleados que tienen 3 hijos.

h_3= 0,30 Frecuencia relativa simple de la tercera clase. Es la proporción de empleados que tienen 3 hijos.

N_3= 33 Frecuencia absoluta acumulada de la tercera clase. Es la cantidad de empleados que tienen 3 o menos hijos.

H_3= 0,66 Frecuencia relativa acumulada de la de la tercera clase. Es la proporción de empleados que tienen 3 o menos hijos.

Salario (tabla con 5 intervalos de amplitud 8)

Tabla de datos agrupados en lista:

Resulta adecuado agrupar los datos de esta manera cuando existen pocos valores observados diferentes y dichos valores se repiten.

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Fuente: BD-clase PyE

Para los datos cualitativos nominales solo deben calcularse las frecuencias simples:

Variable	Clase	Categorías	FA	FR
Área	1	Administración	21	0.42
Área	2	Comunicación	5	0.10
Área	3	Diseño	13	0.26
Área	4	Ingeniería	9	0.18
Área	5	RRHH	2	0.04

Interpretación correspondiente a las frecuencias de la cuarta clase

n_4= 9 Frecuencia absoluta simple de la cuarta clase. Es la cantidad de empleados que pertenecen al área Ingeniería.

h_4= 0,18 Frecuencia relativa simple de la cuarta clase. Es la proporción de empleados que pertenecen al área Ingeniería.

Para los datos cualitativos ordinales deben calcularse las frecuencias simples y las frecuencias acumuladas.

Con los datos tabulados o mediante el uso de funciones en InfoStat, responda en forma completa y clara, utilizando vocabulario técnico adecuado en la explicación.

a. ¿Qué cantidad de empleados tienen 3 o menos hijos?

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Hay **33** empleados que tienen 3 o menos hijos.

b. ¿En qué proporción de empleados tiene un salario comprendido entre 123 y menos de 131 pesos? Equivale a decir 123 o más y menos de 131.

d. ¿Qué proporción de empleados tienen más de 2 hijos?

Para generar esta Tabla de Datos agrupados en lista, debemos verificar en Infostat que Cantidad de Hijos sea Tipo de Dato: Entero

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

(15+13+4)/50 = 32/50

= **0,64**

También lo podemos resolver de la siguiente manera:

0,30 + 0,26 + 0,08 = 0,64 Es la proporción de empleados que tienen más de 2 hijos.

d. ¿Cuántos empleados tienen un salario menor que 139 pesos?

Variable	Clase		LI	LS		MC	FA	FR	FAA	FRA
Salarios	1	[	115.00	123.00	)	119.00	2	0.04	2	0.04
Salarios	2	[	123.00	131.00	)	127.00	10	0.20	12	0.24
Salarios	3	[	131.00	139.00	)	135.00	19	0.38	31	0.62
Salarios	4	[	139.00	147.00	)	143.00	12	0.24	43	0.86
Salarios	5	[	147.00	155.00	]	151.00	7	0.14	50	1.00

31 es la cantidad de empleados que cobran menos de 139

e. ¿Cuántos obreros perciben semanalmente menos de 131 pesos?

Variable	Clase		LI	LS		MC	FA	FR	FAA	FRA
Salarios	1	[	115.00	123.00	)	119.00	2	0.04	2	0.04
Salarios	2	[	123.00	131.00	)	127.00	10	0.20	12	0.24
Salarios	3	[	131.00	139.00	)	135.00	19	0.38	31	0.62
Salarios	4	[	139.00	147.00	)	143.00	12	0.24	43	0.86
Salarios	5	[	147.00	155.00	]	151.00	7	0.14	50	1.00

Hay **12** obreros que perciben semanalmente menos de 131 pesos.

f. ¿Qué proporción de empleados pertenecen al área Ingeniería?

Variable	Clase	Categorías	FA	FR
Área	1	Administración	21	0.42
Área	2	Comunicación	5	0.10
Área	3	Diseño	13	0.26
Área	4	Ingeniería	9	0.18
Área	5	RRHH	2	0.04

9 empleados pertenecen al área Ingeniería.

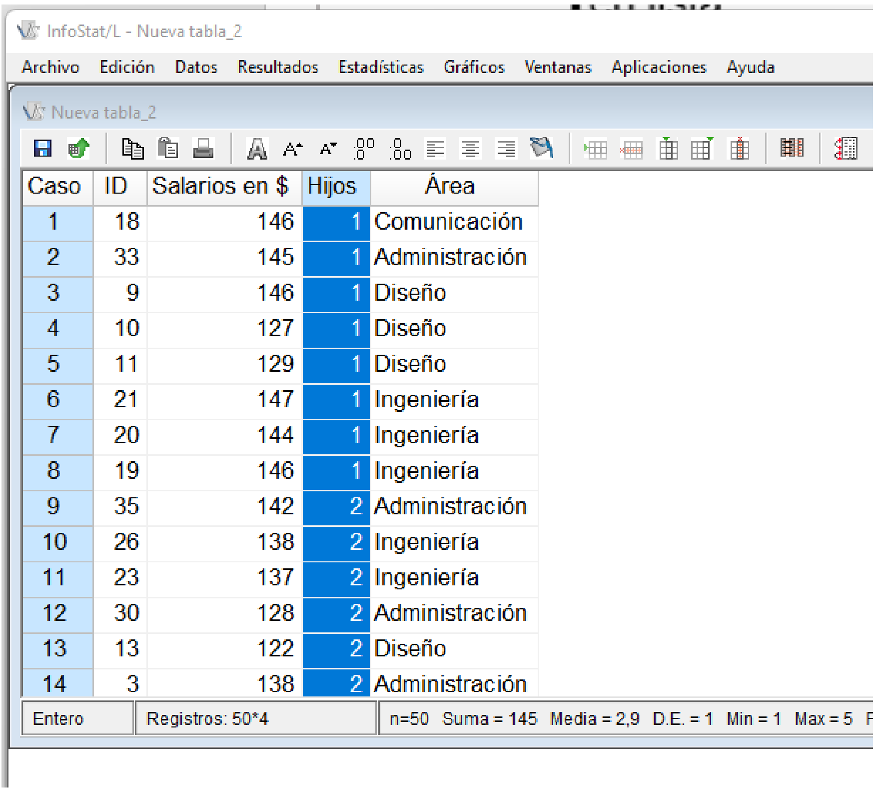
9/50= 0,18 Es la proporción de empleados que pertenecen al área Ingeniería.

Gráficos

Gráficos para Datos cualitativos	- Círculo Radiado (Gráfico circular) - Barra Porcentual - Gráfico de barras	
Gráficos para	**frecuencias absolutas simples (ni)** frecuencias relativas simples (hi= **ni**/n)	**frecuencias absolutas acumuladas (Ni)** frecuencias relativas acumuladas (Hi= **Ni**/n)
Datos agrupados en lista	❖ Diagrama de bastones	❖ Gráfico acumulativo de frecuencias
Datos agrupados en intervalos	❖ Histograma ❖ Polígono de frecuencias	❖ Diagrama escalonado ❖ Ojiva

El Histograma y la Ojiva solo se utilizan para representar datos agrupados en intervalos. No se utilizan para representar datos agrupados en lista.

Antes de construir utilizando Infostat el Diagrama de Bastones, debemos ordenar de menor a mayor los valores de la variable.



InfoStat/L - Nueva tabla_2

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

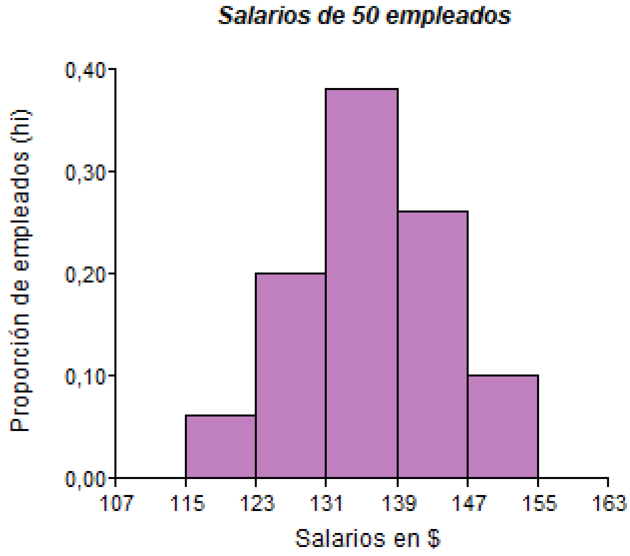
Nueva tabla_2

Caso	ID	Salarios en $	Hijos	Área
1	18	146	1	Comunicación
2	33	145	1	Administración
3	9	146	1	Diseño
4	10	127	1	Diseño
5	11	129	1	Diseño
6	21	147	1	Ingeniería
7	20	144	1	Ingeniería
8	19	146	1	Ingeniería
9	35	142	2	Administración
10	26	138	2	Ingeniería
11	23	137	2	Ingeniería
12	30	128	2	Administración
13	13	122	2	Diseño
14	3	138	2	Administración

Entero | Registros: 50*4 | n=50 Suma = 145 Media = 2,9 D.E. = 1 Min = 1 Max = 5

Nombre: Histograma
Fuente: BD- clase práctica PyE

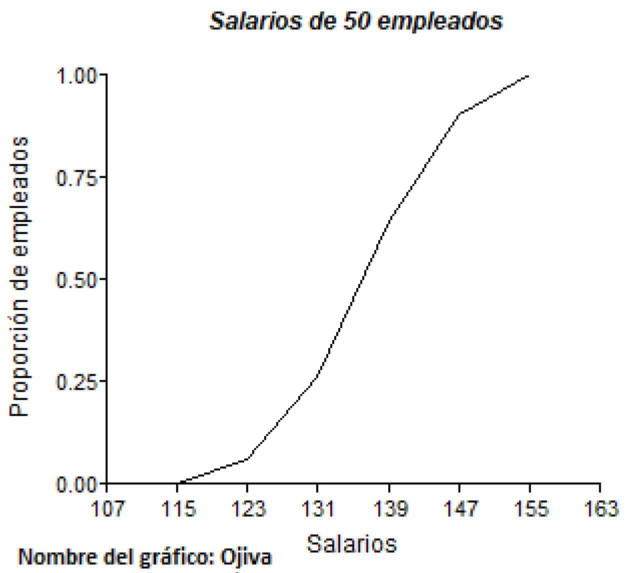
Interpretación: Hasta el tercer intervalo a medida que se incrementan los salarios aumenta la cantidad de empleados. A partir de dicho intervalo a media que se incrementan los salarios disminuye la cantidad de empleados. La mayor cantidad de empleados tiene salarios comprendidos entre 131 y 139 pesos, mientras que la menor cantidad de empleados se registra para el primero y el último intervalos.



Nombre: Histograma
Fuente: BD- clase práctica PyE

Interpretación de un gráfico que representa frecuencias simples: Hasta el tercer intervalo a medida que se incrementan los salarios aumenta la proporción de empleados. A partir de dicho intervalo a media que se incrementan los salarios disminuye la cantidad de empleados. La mayor proporción de empleados tiene salarios comprendidos entre 131 y 139 pesos, mientras que la menor proporción de empleados se registra para el primero y el último intervalos.

Nombre: Diagrama Escalonado
Fuente: BD- clase práctica PyE

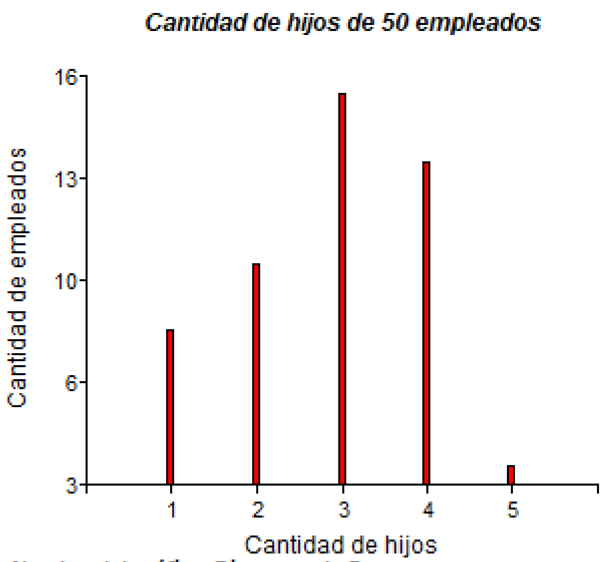


Nombre del gráfico: Ojiva
Fuente: BD- clase práctica PyE

Nombre: Ojiva
Fuente: BD- clase práctica PyE

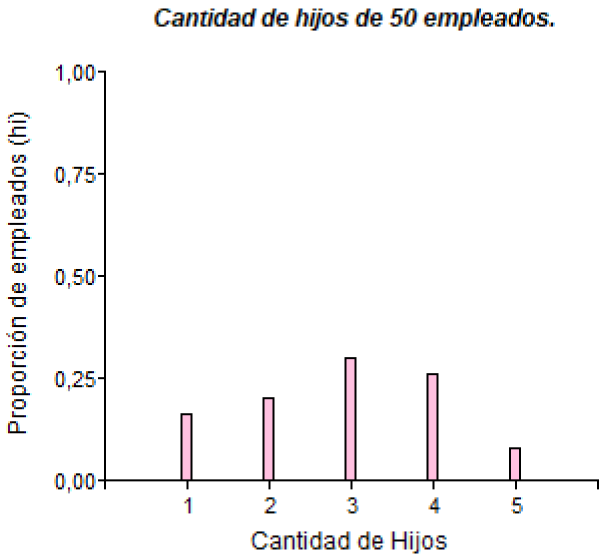
Interpretación en un gráfico que representa frecuencia acumulada: En el segundo intervalo el incremento en la cantidad de empleados se vuelve más acentuado, es más abrupto aún en los tres siguientes intervalos y se vuelve más suave en el último intervalo.

El Diagrama de Bastones y el Gráfico acumulativo de frecuencias se utilizan para representar datos agrupados en lista.



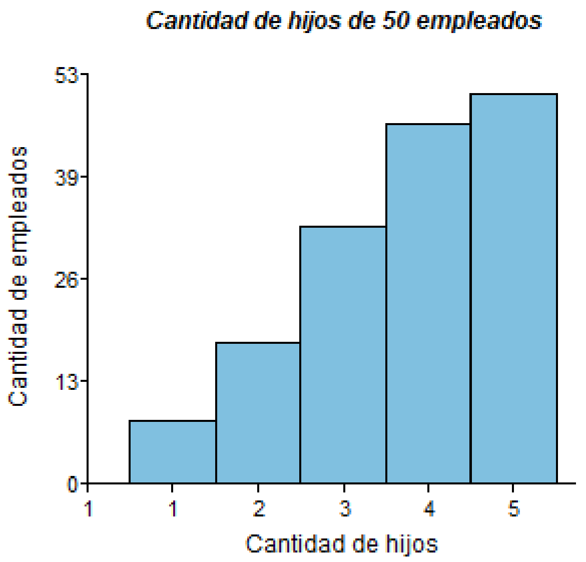
Nombre del gráfico: Diagrama de Bastones
Fuente: BD- clase práctica PyE

Interpretación: Hasta la tercera clase, a medida que se incrementa la cantidad de hijos. A partir de dicha clase a medida que aumenta la cantidad de hijos, disminuye la cantidad de empleados. La mayor cantidad de empleados tiene 3 hijos. La menor cantidad de empleados tiene 5 hijos.



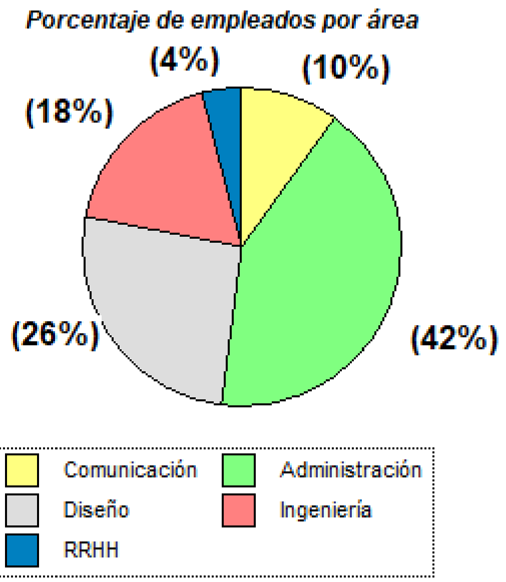
Nombre: Diagrama de Bastones
Fuente: BD- clase práctica- PyE

Interpretación: Hasta la tercera clase, a medida que se incrementa la cantidad de hijos. A partir de dicha clase a medida que aumenta la cantidad de hijos, disminuye la proporción de empleados. La mayor proporción de empleados tiene 3 hijos. La menor proporción de empleados tiene 5 hijos.



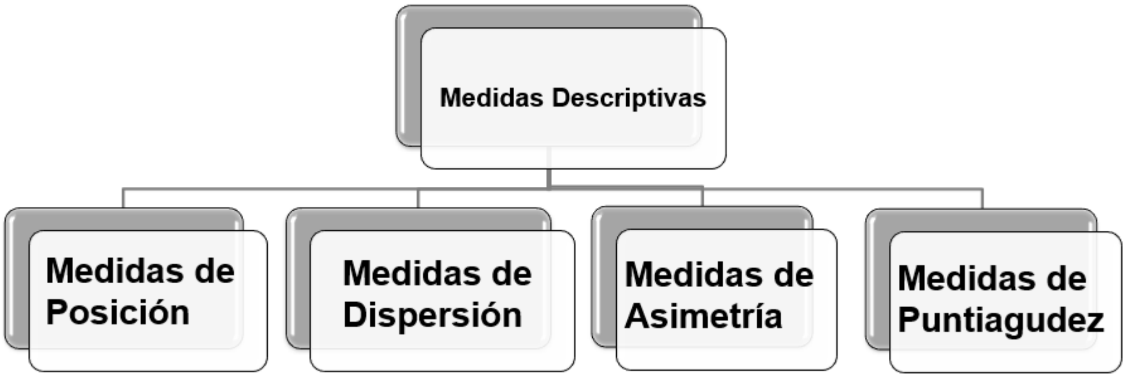
Nombre del gráfico: Gráfico acumulativo de frecuencias
Fuente: BD- clase práctica PyE

Interpretación: El incremento en la cantidad de empleados se produce de manera muy pareja hasta la cuarta clase, y se vuelve más leve desde la cuarta clase a la siguiente.

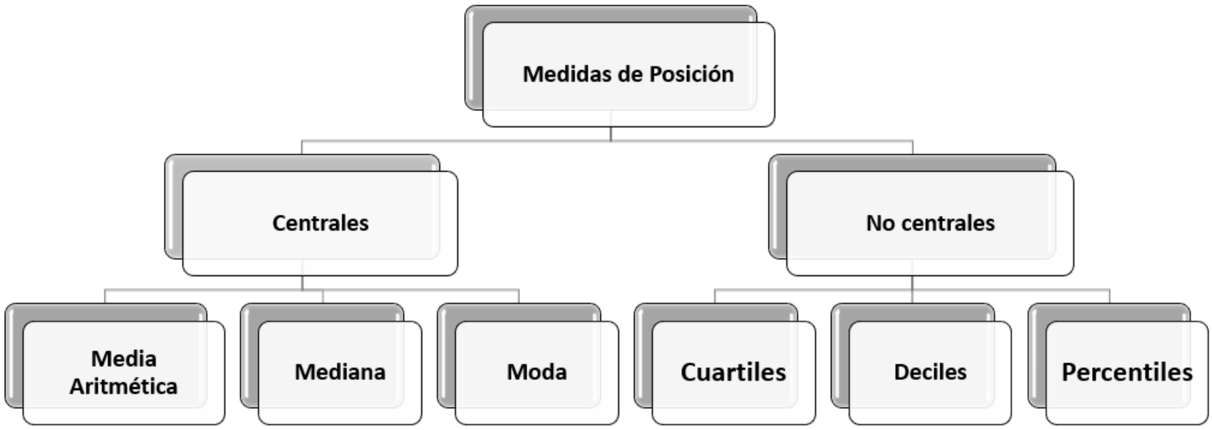


El mayor porcentaje de empleados corresponde al área Administración y disminuyen para las áreas de Ingeniería y Diseño, en tanto que los menores porcentajes se registran para las áreas de Comunicación y RRHH.

Medidas Descriptivas



Medidas de Posición



Medidas de Posición Centrales

Medidas de Posición Centrales para Datos no agrupados

Media Aritmética (Promedio)

Suma de todos los valores observados dividido la cantidad total de observaciones.

Media poblacional: $\mu = \frac{\sum X}{N}$

N: tamaño de la población (cantidad de elementos o individuos que forman la población)

Media muestral: $\bar{x} = \frac{\sum x}{n}$

n: tamaño de la muestra (cantidad de elementos o individuos que forman la muestra)

Ejemplo:

Cantidad de hijos de 7 familias seleccionadas para ser encuestadas:

4 2 1 0 6 5 3 n=7

$\bar{x} = \frac{\sum x}{n}$ $\bar{x} = \frac{21}{7}$ $\bar{x} = 3$

Interpretación: En promedio cada familia tiene 3 hijos (Cantidad de hijos promedio)

En un planilla Excel no deben completarse con ceros las celdas vacías.

Mediana

Es un valor tal que el 50% de las observaciones son menores que la Mediana y el 50% de las observaciones son mayores que la Mediana.

Ejemplo:

Edad de 7 personas encuestadas:

23 43 19 61 18 22 75 n= 7

Se ordenan los valores observados de acuerdo a su magnitud en forma creciente o en forma decreciente:

1°	2°	3°	4°	5°	6°	7°
18	19	22	**23**	43	61	75

Se determina la ubicación o posición de la Mediana (no es el valor de la Mediana):

$\left(\frac{n+1}{2}\right)^{0} = \left(\frac{7+1}{2}\right)^{0}$ = 4° La Mediana es el cuarto valor

El valor de la Mediana es 23

Mediana= 23

Interpretación:

La mitad de las edades son menores que 23 años y la otra mitad de las edades son mayores que 23 años.

La mitad de las personas tienen una edad menor que 23 años y la otra mitad de las personas tienen una edad mayor que 23 años.

Cantidad de hijos de 8 familias seleccionadas para ser encuestadas:

4 2 1 0 6 5 3 1 n= 8

Ordenamos los valores:

1°	2°	3°	4°	5°	6°	7°	8°
0	1	1	2	3	4	5	6

Determinamos la posición de la Mediana:

$$\left(\frac{n+1}{2}\right)^0 = \left(\frac{8+1}{2}\right)^0 = \mathbf{4,5}$$

por lo tanto el valor de la Mediana es el promedio entre el **cuarto** y el **quinto** valor de la serie:

1°	2°	3°	**4°**	**4,5**	**5°**	6°	7°	8°
0	1	1	**2**	**2,5**	**3**	4	5	6

Mediana: (**2** **+** **3**) / 2= **2,5 (Valor de la Mediana)**

La mitad de las familias tienen menos de 2,5 hijos y otra la mitad de las familias tienen más de 2,5 hijos.

Moda:

Valor más frecuente:

Cantidad de hijos de 8 familias encuestadas en una Provincia "C":

4 2 1 0 6 5 3 1 Moda= 1

La cantidad de hijos más frecuente es igual a 1

Cantidad de hijos de 9 familias encuestadas en una Provincia “D”:

1 1 2 **3 3** 4 5 6 7 Dos Modas: 1 y 3

Cantidad de hijos de 7 familias encuestadas en una Provincia “E”:

1 2 3 4 5 6 7

No hay Moda

Medidas de Posición no centrales (Fractiles)

Son medidas que dividen al conjunto de datos observados en partes porcentualmente iguales.

Cuartiles:

Dividen al conjunto de datos en cuatro partes porcentualmente iguales, cada parte contiene el 25% de las observaciones.

Deciles:

Dividen al conjunto de datos en diez partes porcentualmente iguales, cada parte contiene el 10% de las observaciones.

Percentiles:

Dividen al conjunto de datos en cien partes porcentualmente iguales, cada parte contiene el 1% de las observaciones.

Fractiles para datos no agrupados (Series Simples)

Cuartiles:

Dividen al conjunto de datos en cuatro partes porcentualmente iguales, cada parte contiene el 25% de las observaciones.

25% 25% 25% 25%

|_____|_____|_____|_____|

Q1 Q2 Q3

Q1: Primer Cuartil

Es un valor tal que 25% de los valores observados son menores que el Primer Cuartil o bien podemos decir que:

Es un valor tal que 75% de los valores observados son mayores que el Primer Cuartil.

25% 25% 25% 25%

|_____|_____|_____|_____|

Q1 Q2 Q3

La Mediana es el Segundo Cuartil (Q2)

Primer paso: Ordenamos los valores (de menor a mayor, o de mayor a menor)

Segundo paso: Se determina la ubicación del cuartil en la serie.

Fórmula para determinar la ubicación o posición de los cuartiles:

$[k(n + 1) / 4]^{\circ}$

k es el número de cuartil.

Ejemplos:

Para el **Primer** cuartil **k** vale **1**.

Para el **Segundo** cuartil **k** vale **2**.

Para el **Tercer** cuartil **k** vale **3**

Tercer paso: Determinamos el valor del cuartil.

Primer Cuartil o Cuartil 1(Q1)

Definición:

Es un valor tal que el 25% de las observaciones son menores que el primer cuartil y el 75% de las observaciones son mayores que el primer cuartil.

El primer cuartil supera al 25% de las observaciones y es superado por el 75% de las observaciones.

25% 25% 25% 25%

|_____|_____|_____|_____|

Q_1

Ejemplo:

Dada la cantidad de materias aprobadas de los estudiantes de cierta carrera:

9 7 5 1 6 2 6 9 4 n=9

Calcular el Primer Cuartil.

Se determinará el **Primer** Cuartil (Cuartil **uno**)

Paso 1: Se ordenan los valores

1 2 4 5 6 6 7 9 9

n: cantidad total de observaciones n= 9

Paso 2: Se determina la **posición** o **ubicación** del **Primer Cuartil**

[**k**(n + 1) / **4**]**°** **Fórmula para determinar la posición de un Cuartil**

k= **1**

Fórmula para determinar la **posición** del **Primer Cuartil**

[**1**(n + 1) / **4**]**°** = [**1**(9 + 1) / **4**]**°**

= **2,5** No es valor del Primer Cuartil, este resultado indica que el Primer Cuartil estará comprendido entre el segundo y el tercer valor.

Paso 3: Se Obtiene el valor del Primer Cuartil

El primer cuartil estará comprendido entre el segundo y el tercer valor.

1° 2° **2,5** 3° 4°

1 2 4 5 6 6 7 9 9

El Primer Cuartil será el promedio entre el segundo y el tercer valor:

Q1= (2 + 4)/ 2

Q1= 3 Valor del Primer Cuartil

Q1=3

Interpretación:

Para realizar la interpretación debe tenerse en cuenta la definición del cuartil y la característica bajo estudio.

Si el fractil no coincide con algún valor de la serie entonces se expresa la interpretación de la siguiente manera:

Interpretación en el caso planteado:

Aproximadamente el 25% de los estudiantes tienen menos de 3 materias aprobadas.

Aproximadamente el 75% de los estudiantes tienen más de 3 materias aprobadas.

El 25% de los estudiantes tienen menos de 3 materias aprobadas.

El 75% de los estudiantes tienen más de 3 materias aprobadas.

25% 25% 25% 25%

|_____|_____|_____|_____|

Q_1

Tercer Cuartil (Q3):

25% 25% 25% 25%

|____|____|____|____|

Q_3

Es un valor tal que el 75% de las observaciones son menores que el tercer cuartil y el 25% de las observaciones son mayores que el tercer cuartil.

El Tercer Cuartil supera al 75% de las observaciones y es superado por el 25% de las observaciones.

Ejemplo:

Dada la cantidad de materias aprobadas de 9 estudiantes:

1 2 4 5 6 6 7 9 9

Calcular el Tercer Cuartil

Primer paso:

Se ordenan los valores (En este caso ya se encuentran ordenados)

1 2 4 5 6 6 7 9 9

Segundo paso:

Se determina la posición o ubicación del **Tercer Cuartil** a través de la siguiente fórmula:

El **Tercer Cuartil** es el cuartil **3** por lo tanto **k**=**3**

[k(n + 1) / 4]° = [3(9 + 1) / 4]°

= [30 / 4]°

[30 / 4]° = 7,5 No es el valor del tercer cuartil

7,5 indica que el tercer cuartil será un valor comprendido entre el séptimo y el octavo valor.

1°	2°	3°	4°	5°	6°	7°	7,5	8°	9°
1	2	4	5	6	6	**7**		**9**	9

Tercer paso:

Se determina el valor del **Tercer Cuartil**

Si al calcular **[k(n + 1) / 4]°** se obtiene un resultado con cifras decimales, entonces se determina el valor del Cuartil mediante un promedio.

Tercer cuartil: Q3= (7 + 9)/2

Q3= 8 Valor del Tercer Cuartil

Interpretación:

Aproximadamente el 75% de los estudiantes tienen menos de 8 materias aprobadas, y aproximadamente el 25% de los estudiantes tienen más de 8 materias aprobadas.

A lo sumo el 75% de los estudiantes tienen menos de 8 materias aprobadas, y a lo sumo el 25% de los estudiantes tienen más de 8 materias aprobadas.

El 75% de los estudiantes tienen menos de 8 materias aprobadas.

El 25% de los estudiantes tienen más de 8 materias aprobadas.

25% 25% 25% 25%

|____|_____|_____|_____|

Q3=8

Otros ejemplos:

Dada la cantidad de cuentas por cobrar en un conjunto de empresas muestreadas:

1 8 8 7 5 6 8 5 4 3 4

Calcular e interpretar Primer Cuartil y Tercer Cuartil

Primer paso: Se ordenan los valores observados

1 3 4 4 5 5 6 7 8 8 8

n= 11

Segundo paso: Se identifica la **ubicación** del **Primer Cuartil** (**Cuartil 1**):

[k(n + 1) / 4]°

k=1 n=11

[k(n + 1) / 4]°= [1(11 + 1) / 4]°

= (12 / 4)°

= **3°**

3 no es el valor del primer cuartil, este resultado indica que el primer cuartil es el tercer valor de la serie (previamente ordenada)

1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°
1	3	4	4	5	5	6	7	8	8	8

El tercer valor de la serie es 4, por lo tanto el primer cuartil es igual a 4

Q1= **4**

25% 25% 25% 25%

|_____|_____|_____|_____|

Q1= 4

Interpretación del Primer Cuartil:

Para realizar la interpretación se tiene en cuenta cuál es la variable bajo estudio, en este caso es la cantidad de cuentas por cobrar de 11 empresas.

Si el cuartil coincide con algún valor de la serie entonces se expresa la interpretación de la siguiente manera:

Aproximadamente el 25% de las observaciones son menores o iguales que el valor del cuartil y el 75% de las observaciones son mayores o iguales que el valor del fractil.

Interpretación en el caso planteado:

Aproximadamente el 25% de las empresas tienen 4 o menos cuentas por cobrar, y aproximadamente el 75% de las empresas tienen 4 o más cuentas por cobrar.

Tercer Cuartil:

1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°
1	3	4	4	5	5	6	7	8	8	8

Tercer Cuartil (Cuartil 3)

[k(n + 1) / 4]°= **[3(11 + 1) / 4]°**

= (36 /4)°

= 9°

9 no es el valor del tercer cuartil, este resultado indica que el tercer cuartil es el noveno valor de la serie

Si al calcular **[k(n + 1) / 4]°** se obtiene un resultado exacto, entonces se busca en la serie el valor del Cuartil al cual le corresponda la posición calculada.

1°	2°	3°	4°	5°	6°	7°	8°	**9°**	10°	11°
1	3	4	4	5	5	6	7	**8**	8	8

El noveno valor de la serie es 8, por lo tanto el tercer cuartil es igual a 8

25% 25% 25% 25%

I_____I______I______I_____I

Q1 Q2 Q3=8 cuentas por cobrar

Característica bajo estudio: Cantidad de cuentas por cobrar de 11 empresas.

Interpretación:

Aproximadamente el 75% de las empresas tienen 8 o menos cuentas por cobrar y aproximadamente el 25% restante de las empresas tienen 8 o más cuentas por cobrar.

Medidas Descriptivas con Infostat

Estadísticas ___ Medidas de resumen

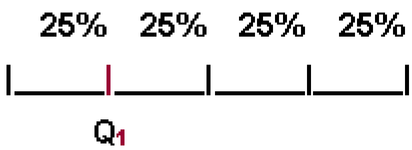
Medidas resumen

Variable	n	Media	Mediana	Q1	Q3
Salarios en $	50	137,22	137,50	131,00	145,00

Media Aritmética: 137,22 En promedio cada empleado percibe por hora un salario de $137,22 (Salario promedio)

Mediana: El 50% de los empleados tienen un salario menor que 137,50 y el otro 50% de los empleados tienen un salario mayor que 137,50

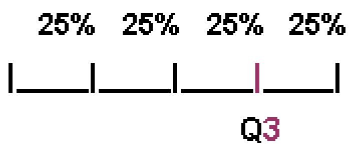
Q1 (Primer Cuartil):



El 25% de los empleados tienen un salario menor que 131

El 75% de los empleados tienen un salario mayor que 131

Q3 (Tercer Cuartil):



El 75% de los empleados tienen un salario menor que 145

El 25% de los empleados tienen un salario mayor que 145

En Infostat no podemos identificar el **Salario más frecuente (Moda)**.

En un gráfico de frecuencias simples (ni o hi) podemos identificar el intervalo en el cual está comprendida la Moda

Nombre del gráfico: Histograma

La **Moda** está ubicaba en el intervalo al cual le corresponde la mayor frecuencia absoluta simple (ni)

La **Moda** está comprendida entre 131 y 139

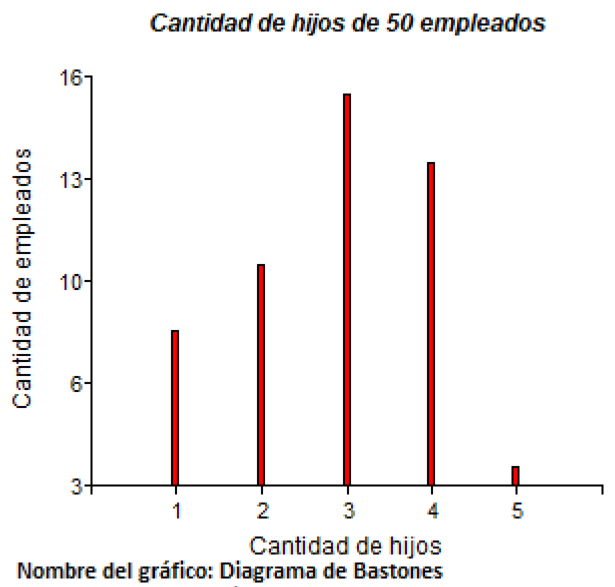
El **Salario más frecuente** está comprendido entre 131 y 139

Tablas de frecuencias

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

En la tabla con intervalos podemos identificar el intervalo al cual le corresponde la mayor frecuencia absoluta simple, dicho intervalo es el que contiene la **Moda**

El **Salario más frecuente** está comprendido entre 131 y 139



Nombre del gráfico: Diagrama de Bastones
Fuente: BD- clase práctica PyE

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Fuente: BD-clase PyE

Moda= 3 Es el valor de la variable al cual le corresponde la mayor frecuencia absoluta simple (FA)

La Moda es igual a 3 porque a este valor le corresponde la mayor frecuencia que es 15

Miden grado de concentración de los valores de la variable con respecto a alguna medida de tendencia central.

Permiten comparar dos o más grupos de datos para determinar en cuál de ellos los valores son más homogéneos, es decir en cuál grupo los valores presentan menor variabilidad. En dicho grupo la media es más representativa de los valores.

Recorrido o Rango:

Valor Máximo – Valor Mínimo

Es la máxima separación entre el máximo y el mínimo valor observado de la variable.

Temperaturas registradas en el sur del país en Julio de 2021

Se muestrearon 5 días: -5 0 -1 2 3

Valor máximo= 3

Valor mínimo= -5

Recorrido o Rango: R= 3 – (-5)

R= 8

Interpretación: Es la máxima separación entre los valores observados de la variable *temperatura*.

Rango Intercuartílico:

Diferencia entre el Tercer Cuartil y el Primer Cuartil.

El 50% de los valores observados están comprendidos entre el Primer Cuartil y el Tercer Cuartil.

Ejemplo:

Cantidad de cuentas por cobrar en un conjunto de empresas muestreadas:

1 8 8 7 5 6 8 5 4 3 4

Q1= 4

Q3= 8

Rango Intercuartílico: RI= Q3 – Q1

= 8 – 4

= 4

El 50% de los valores correspondientes a la variable *cantidad de cuentas por cobrar* está comprendido entre 4 y 8.

Varianza Poblacional: σ^2 En Infostat: Var(n)

$$\sigma^2 = \frac{\sum(\mathbf{x}-\mu)^2}{N}$$

μ: Media Poblacional

Media poblacional: μ: **Media o promedio calculado a partir de una población.**

$\mathbf{x}-\mu$: Desvío con respecto a la Media: Diferencia entre cada valor observado y la Media

N: cantidad de elementos de la población (tamaño de la población)

Varianza: Promedio del cuadrado de los desvíos con respecto a la Media Aritmética.

Desviación Poblacional: σ

$$\sigma = \sqrt{\frac{\sum(\mathbf{x}-\mu)^2}{N}}$$

Varianza muestral corregida. En Infostat: Var(n-1)

$$\hat{s}^2 = \frac{\sum(\mathbf{x}-\bar{x})^2}{n-1}$$

$$\hat{s}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$$ Fórmula de cálculo

Ejemplo:

Variable x: nota

Ejemplo:

$\sum x^2$: Sumatoria del cuadrado de las notas (Esta sumatoria puede ser un dato)

Ejemplo: $\sum x$: Sumatoria de las notas

$\overline{x} = \frac{\sum x}{n}$ **Media muestral: Media o promedio calculado a partir de una muestra.**

Desviación Estandar:

Raíz cuadrada positiva de la Varianza.
Indica la distancia promedio que hay entre los valores observados de la variable y la Media Aritmética.

Comparemos las notas de dos cursos:

x: notas en puntos

Curso "A" (N=5)	Curso "B" (N=5)
4 4 7 10 10	6 6 7 8 8
Media(A)= 35/5 = 7 puntos	**Media(B)= 35/5 = 7 puntos**

En el curso "B" las **notas son más homogéneas** (más similares entre sí)
En el curso "B" las notas presentan menor variabilidad.
En el curso "B" las notas se encuentran más concentradas alrededor de la Media
En el curso "B" las notas se encuentran a **menor distancia** con respecto a la **Media**, por lo tanto en dicho curso las notas son más homogéneas.
En el curso "B" la **media es más representativa de los valores observados.**
En el curso "B" el rendimiento fue más parejo.

En el curso "A" hay mayor dispersión de las notas con respecto a la media.
En el curso "A" las notas presentan mayor variabilidad, son más desparejas.

Aquel conjunto de datos en el cual los valores se encuentran a menor distancia con respecto a la Media, es el más homogéneo.

Aquel conjunto de datos en el cual los valores se encuentran más concentrados, más cercanos con respecto a la Media, es el más homogéneo.

Curso "A" (N=5) Curso "B" (N=5)

4 4 7 10 10 6 6 7 8 8

Media(A)= 35/5 = 7 puntos

Media(B)= 35/5 = 7 puntos

$$\sigma^2 = \frac{\sum(\mathbf{x} - \boldsymbol{\mu})^2}{N}$$

A través de la siguiente diferencia: (x- **Media**) calculamos la distancia que hay entre cada valor y la Media. A esta diferencia le llamamos Desvío con respecto a la Media.

$$\frac{(4-7)+(4-7)+(7-7)+(10-7)+(10-7)}{5} = 0$$

$\sum$(x- Media)= 0 La sumatoria de los desvíos con respecto a la Media siempre es igual a cero. Este resultado no resulta útil para obtener alguna conclusión referida a la dispersión de los datos.

Para que todos los términos sean positivos, elevamos al cuadrado cada desvío y así surge la varianza:

$$Varianza = \frac{\sum(\mathbf{x} - \mathbf{Media})^2}{N}$$

$$\sigma^2 = \frac{(\mathbf{4}-\mathbf{7})^2 + (\mathbf{4}-\mathbf{7})^2 + (\mathbf{7}-\mathbf{7})^2 + (\mathbf{10}-\mathbf{7})^2 + (\mathbf{10}-\mathbf{7})^2}{5}$$

$\sigma^2(A) = 7{,}2$ (puntos)^2

Definición de la Varianza:

Es la sumatoria del cuadrado de los desvíos con respecto a la media aritmética sobre la cantidad total de observaciones.

También podemos definir la Varianza como el promedio de los desvíos al cuadrado con respecto a la Media Aritmética.

En el caso planteado es el promedio del cuadrado de los desvíos de las notas con respecto a la nota promedio.

Varianza en curso “B”:

$$\sigma^2(B) = \frac{(\mathbf{6}-\mathbf{7})^2 + (\mathbf{6}-\mathbf{7})^2 + (\mathbf{7}-\mathbf{7})^2 + (\mathbf{8}-\mathbf{7})^2 + (\mathbf{8}-\mathbf{7})^2}{5}$$

$\sigma^2(B) =$ 0,8 (puntos)^2

Interpretación:
Promedio del cuadrado de los desvíos de las notas con respecto a la nota promedio.

Varianza en el curso “A”:

$$\sigma^2(A) = \frac{(\mathbf{4}-\mathbf{7})^2 + (\mathbf{4}-\mathbf{7})^2 + (\mathbf{7}-\mathbf{7})^2 + (\mathbf{10}-\mathbf{7})^2 + (\mathbf{10}-\mathbf{7})^2}{5}$$

$\sigma^2(A) =$ 7,2 (puntos)^2

La Varianza no refleja correctamente la distancia que hay entre los valores observados y la Media, debido a que la varianza no está expresada en la misma unidad de medida que posee la variable bajo estudio (notas en puntos y varianza expresada en puntos al cuadrado).

Ejemplos:

Salario en \$ La varianza en \$^{2}
Estatura en mts La varianza en mts^2

La Varianza nunca puede ser negativa.

Si calculamos la raíz cuadrada de la Varianza entonces obtendremos una medida expresada en la misma unidad que posee la variable bajo estudio, y así llegamos a la Desviación Estándar.

Desviación Estándar en el curso “A”:

Curso “A” (N=5)

4 4 7 10 10
Media= 7 (Nota Promedio)

DE(A)= $\sqrt{\text{Varianza}}$

DE(A) = $\sqrt{7{,}2\ (\text{puntos})^2}$

DE(A)= 2,68 puntos

La Desviación Estándar indica la distancia promedio que hay entre los valores observados y la media aritmética.

En promedio cada nota del curso “A” se encuentran a 2,68 puntos con respecto a la nota promedio.

En el curso "A" hay notas que se encuentran a más de 2,68 puntos con respecto a la nota promedio.
En el curso "A" hay notas que se encuentran a menos de 2,68 puntos con respecto a la nota promedio.
En promedio las notas del curso "A" se encuentran a 2,68 puntos con respecto a la nota promedio.

En el curso "A" hay notas que coinciden con la nota promedio.

En promedio cada nota del curso "B" se encuentran a 0,89 puntos con respecto a la nota promedio.
En el curso "B" hay notas que se encuentran a más de 0,89 puntos con respecto a la nota promedio.
En el curso "B" hay notas que se encuentran a menos de 0,89 puntos con respecto a la nota promedio.
En el curso "B" hay notas que coinciden con la nota promedio.

Curso "A" (N=5)

4 4 7 10 10

Media(A)= 7 puntos

DE(A)= 2,68

Curso "B" (N=5)

6 6 7 8 8

Media (B)= 7 puntos

DE(B)= 0,89
En "B" las notas son más homogéneas
En "B" las notas presentan menor variabilidad (notas más similares entre si)
En "B" la media es más representativa

Para realizar esta comparación de dos o más grupos de datos utilizando la Desviación Estándar los grupos a comparar deben tener:

- **Igual Media Aritmética**
- **La variable bajo estudio expresada en la misma de unidad de medida.**

CV(A)= 38,29%

CV(B)= 6,23%

Coeficiente de Variación: Aquel grupo con menor CV será el más homogéneo.

CV(A)= **Desviación Estandar / Media**

CV(A)=(**Desviación Estandar / Media**)***100** **CV expresado en porcentaje**

CV(A)= (2,68/7)*100
= 0,3829*100

CV(A)= 38,29% es mayor al 26% por lo tanto en el curso “A” las notas son muy heterogéneas.

En el curso “A” la Desviación Estándar que es de 2,68 equivale al 38,29% del valor de la Media que es 7-

CV(B)= (**0,89**/7)*100
= 0,0623 *100

CV(B)= 6,23% está comprendido entre 0% < CV < 11% por lo tanto en el curso “B” las notas son muy homogéneas.

0% < CV < 11% Datos son muy homogéneos (son muy similares entre si) 11% ≤ CV < 16% Datos homogéneos 16% ≤ CV < 26% Datos heterogéneos CV ≥ 26% Datos muy heterogéneos (presentan mucha variabilidad, son muy desparejos, muy diferentes entre si)

En el curso “A” la **Desviación Estándar** que es de **2,68** equivale al 38,29% (Valor del CV) del valor de la Media que es 7

En el curso “B” la **Desviación Estándar** que es de **0,89** equivale al 6,23% (Valor del CV) del valor de la Media que es 7

En el curso “B” el rendimiento es más parejo (las notas son más homogéneas, es decir que presentan menor variabilidad)

Aquel grupo con **menor Desviación Estándar**, es el grupo en el cual los valores se encuentran a menor distancia con respecto a la media, y en consecuencia es el que presenta valores más homogéneos, lo que equivale a decir que los valores presentan menor variabilidad, por lo tanto en dicho grupo la media es más representativa.

Deben cumplirse dos condiciones para comparar dos o más conjuntos de observaciones utilizando la Desviación Estándar:

Los grupos a comparar deben tener:

-Igual Media y

- La variable bajo estudio expresada en la misma unidad de medida.

Si no se cumplen **ambas** condiciones entonces debemos utilizar el Coeficiente de Variación para efectuar la comparación.

> La **Desviación Estándar** permite comparar dos o más grupos de datos si poseen la **misma Media Aritmética y** tienen la **variable bajo estudio expresada en la misma unidad de medida**. Si no se cumplen ambas condiciones, entonces debemos utilizar el Coeficiente de Variación para comparar los grupos de datos. Aquel grupo en el cual se obtiene el menor Coeficiente de Variación, es el que posee valores más homogéneos.

Coeficiente de Variación: Cociente entre **Desviación Estándar (DE)** y Media Aritmética (Promedio)

El valor del Coeficiente de Variación no está expresado en la unidad de medida correspondiente a la variable bajo estudio. No lleva unidad de medida.

Aquel grupo de datos con menor Coeficiente de Variación es el más homogéneo.

A su vez, en cada conjunto de datos puede identificarse el nivel de homogeneidad:

> 0% < CV < 11% Datos son muy homogéneos (son muy similares entre si)
>
> 11% ≤ CV < 16% Datos homogéneos
>
> 16% ≤ CV < 26% Datos heterogéneos
>
> CV ≥ 26% Datos muy heterogéneos (presentan mucha variabilidad, son muy desparejos, muy diferentes entre si)

A medida que se incrementa el Coeficiente de Variación los valores incrementan su variabilidad se vuelven más desparejos, son menos homogéneos, es decir son más heterogéneos, y la media será menos representativa de los valores observados.

Mientras menor sea el Coeficiente de Variación, los valores observados de la variable estarán más concentrados alrededor del valor de la Media correspondiente a dicha variable, la Media será más representativa de los valores observados.

CV= (DE / Media)*100 %

Por ejemplo:

CV= 13%

La valor de la **Desviación Estándar** equivale al 13% del valor de la Media

Ejercicio:

Fueron registrados los precios de 6 artículos seleccionados en un comercio "A"

(La palabra "seleccionados" indica que los valores provienen de una muestra).

Calcular medidas de dispersión:

3 6 5 9 2 9

(Precios en una muestra de 6 artículos de un comercio "A") n=6

$\overline{x}$ = **34 /** 6

$$\overline{x} = 5,667$$

$$\sum x^2 = 3^2 + 6^2 + 5^2 + 9^2 + 2^2 + 9^2$$

$$\sum x^2 = 236$$

Varianza muestral:

$$\hat{s}^2 = \frac{\sum(\mathbf{x} - \overline{x})^2}{n-1}$$

$$\hat{s}^2 = \frac{\sum x^2 - n\overline{x}^2}{n-1}$$ Fórmula de cálculo

$$\hat{s}^2 = \frac{\mathbf{236 - 6 * 5{,}667^2}}{6-1}$$

$$\hat{s}^2 = \frac{\mathbf{236 - 6 * 5{,}667^2}}{6-1}$$

$\hat{s}^2$= 8,66 Varianza a partir de una muestra.

Desviación Estándar: Raíz cuadrada de la varianza:

$$\hat{s} = \sqrt{8{,}66}$$

$\hat{s}$=2,94

$\hat{s}$ simboliza Desviación Estándar Muestral

$$CV = \frac{\hat{s}}{\bar{x}} * 100$$

$$CV = \frac{2{,}94}{5{,}667} * 100$$

CV=51,87%

Los precios en la muestra de 6 artículos son muy heterogéneos es decir que presentan mucha variabilidad, son muy diferentes entre sí.

El valor de la Desviación Estándar equivale al 51,87% del valor de la Media

Si en otro comercio “B” la Desviación estándar es de $4,75 y la Media es de $12,54. ¿Cuál grupo de precios es más homogéneo?

¿Utilizamos Desviación Estándar o Coeficiente de Variación?

Los precios tienen distinta Media por lo tanto utilizaremos el Coeficiente de Variación para efectuar la comparación. Aquel grupo con menor Coeficiente de Variación será el más homogéneo.

CV(B)= DE/Media*100

= $4,75 / $12,54*100

CV(B)= 37,88%

En el comercio “B” se obtuvo menor Coeficiente de Variación, por lo tanto en dicho comercio los precios se encuentran a menor distancia con respecto al precio promedio, por lo tanto en el comercio “B” los precios son más homogéneos, es decir presentan menor variabilidad que en el comercio “A”.

Medidas de Asimetría

Coeficiente de Asimetría o Coeficiente de Pearson

$$CP = 3 . \frac{\textbf{Media} - \textbf{Mediana}}{\textbf{Desviación Standar}}$$

El tipo de Distribución también puede identificarse mediante la observación de un Histograma o de un Polígono de Frecuencias (gráficos que permiten representar frecuencias simples) (ni: frecuencia absoluta simple) o hi (frecuencia relativa simple)

El tipo de distribución no pueden identificarse a partir de gráficos que representen frecuencias acumuladas.

Distribución Asimétrica Derecha

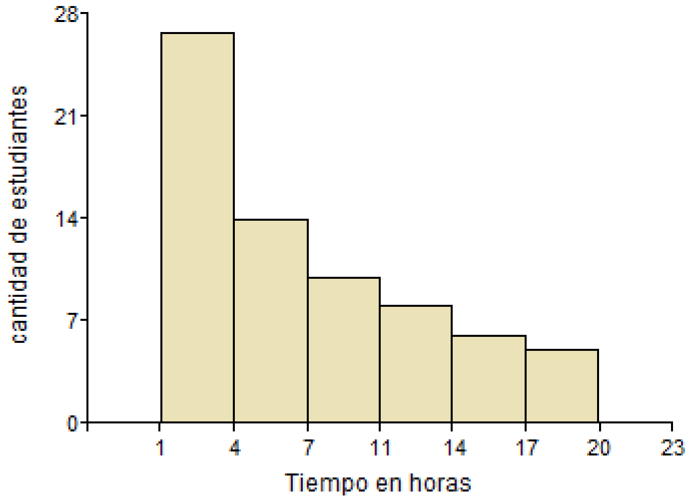
(Media > Mediana) $\Rightarrow$ CP > 0 $\Rightarrow$ Distribución Asimétrica Derecha

Las mayores frecuencias corresponden a los valores de la variable más bajos.

Las menores frecuencias corresponden a los valores de la variable más altos.

Predominan los valores de la variable que son más bajos.

Tiempo dedicado al estudio semanalmente por 70 estudiantes.



A medida que se incrementan los tiempos, disminuye la cantidad de estudiantes.
Predominan los estudiantes que dedican menor tiempo al estudio.
La mayor cantidad de estudiantes dedican menor tiempo al estudio.
La menor cantidad de estudiantes dedican mayor tiempo al estudio.
La Media será mayor que la Mediana
El **Coeficiente de Asimetría** o **Coeficiente de Pearson** será **mayor que cero**.

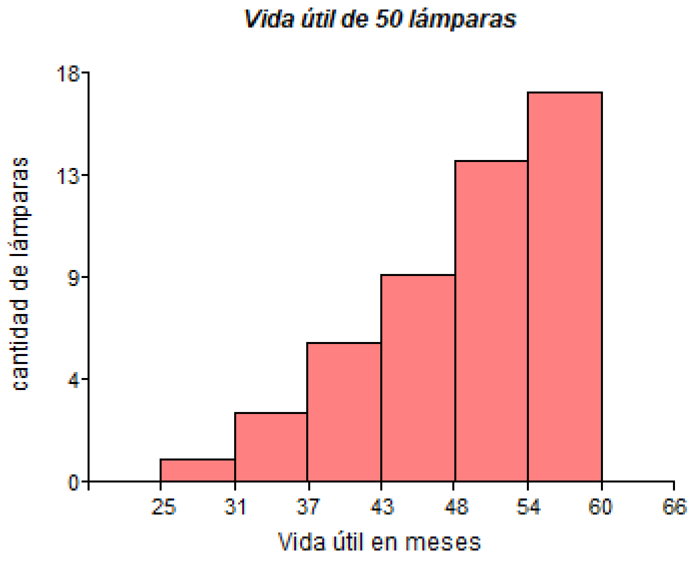
Distribución Asimétrica Izquierda

(Media < Mediana) ⇒ CP < 0 ⇒ Distribución Asimétrica Izquierda

Las mayores frecuencias corresponden a los valores de la variable más altos.

Las menores frecuencias corresponden a los valores de la variable más bajos.

Predominan los valores de la variable que son más altos.



A medida que se incrementa la vida útil, aumenta la cantidad de lámparas.
Predominan las lámparas que tienen menor vida útil.
La mayor cantidad de lámparas tienen mayor vida útil.
La menor cantidad de lámparas tienen menor vida útil.
La Media será menor que la Mediana.
El **Coeficiente de Asimetría** o **Coeficiente de Pearson** será menor que cero.

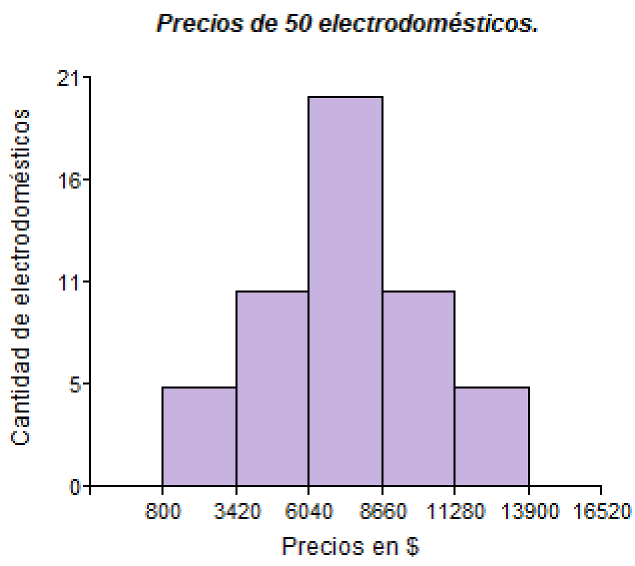
Distribución Simétrica

(Media = Mediana) ⇒ CP = 0 ⇒ Distribución Simétrica

Las mayores frecuencias corresponden a los valores de la variable intermedios.

Las menores frecuencias corresponden a los valores de la variable más bajos y más altos.

Predominan los valores de la variable que son intermedios.



En el primer tramo de la distribución comprendido hasta el tercer intervalo, a medida que aumentan los precios, se incrementa la cantidad de electrodomésticos.

En el primer tramo de la distribución comprendido desde el tercer intervalo, a medida que aumentan los precios, se disminuye la cantidad de electrodomésticos.

Predominan los electrodomésticos que tienen precios intermedios.

Si el valor del coeficiente de asimetría es muy cercano a 0 ello indica que la distribución posee una leve asimetría, tiende a la simetría.

Si es positivo y muy cercano a 0 entonces la distribución presenta una leve asimetría derecha. Ejemplo: 0,02 0,09

Si es negativo y muy cercano a 0 entonces la distribución presenta una leve asimetría izquierda. Ejemplo: -0,02 -0,09

Coeficiente de curtosis (Kurtosis en Infostat)

Menor que 0 $\Rightarrow$ Distribución Platicúrtica.

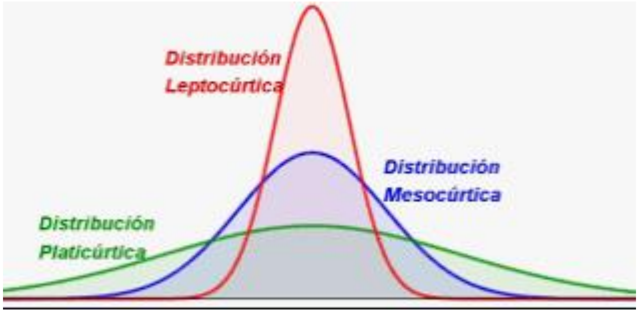
La distribución presenta un reducido grado de concentración con respecto a los valores centrales de la variable.

Igual que 0 $\Rightarrow$ Distribución Mesocúrtica

La distribución presenta un grado medio de concentración con respecto a los valores centrales de la variable.

Mayor que 0 $\Rightarrow$ Distribución Leptocúrtica

La distribución presenta un alto grado de concentración con respecto a los valores centrales de la variable.



Interpretación de las Medidas Descriptivas para la variable *Cantidad de hijos.*

Medidas resumen

Resumen	Hijos
n	50.00
Media	2.90
D.E.	1.20
Var(n-1)	1.44
CV	41.36
Mín	1.00
Máx	5.00
Mediana	3.00
Q1	2.00
Q3	4.00
Asimetría	-0.10
Kurtosis	-0.91

Media= 2,9 En promedio cada empleado tiene 2,9 hijos

Mediana= 3 La mitad de los empleados tienen 3 o menos hijos y la otra mitad de los empleados tienen 3 o más hijos.

También son correctas las siguientes interpretaciones:

Aproximadamente el 50% de los empleados tienen 3 o menos hijos y aproximadamente el otro 50% tienen 3 o más hijos.

A lo sumo el 50% de los empleados tienen 3 o menos hijos y a lo sumo el 50% de los empleados tienen 3 o más hijos.

No más del 50% de los empleados tienen 3 o menos hijos

Moda= **3 La cantidad de hijos más frecuente**

La identificamos en la tabla de Datos Agrupados en Lista. La clase (**yi**) con mayor frecuencia absoluta simple (FA= ni) es el valor de la Moda.

Variable	Clase	MC	FA	FR	FAA	FRA
Hijos	1	1	8	0.16	8	0.16
Hijos	2	2	10	0.20	18	0.36
Hijos	3	3	15	0.30	33	0.66
Hijos	4	4	13	0.26	46	0.92
Hijos	5	5	4	0.08	50	1.00

Fuente: BD- TPED- PYE- 2020- ISI- FRC- UTN

La mayor frecuencia absoluta simple es igual a **15** en consecuencia la Moda es igual a **3**

Recorrido: Máx – Min = 5 – 3

R =2

Es la máxima separación entre los valores de la variable cantidad de hijos.

Primer Cuartil: Q1= 2

El 25% de los empleados tienen 2 o menos hijos.

El 75% de los empleados tienen 2 o más hijos.

Primer Cuartil: Q3= 4

El 75% de los empleados tienen 4 o menos hijos.

El 25% de los empleados tienen 4 o más hijos.

Rango Intercuartílico: Diferencia entre el Tercer Cuartil y el Primer Cuartil.

El 50% de los valores observados de la variable están comprendidos entre el Primer Cuartil y el Tercer Cuartil

25% **25%** **25%** 25%

|_____|_____|_____|_____|

Q1 **Q3**

Q3 – Q1= 4 – 2

=2

25% **25%** **25%** 25%

|_____|_____|_____|_____|

2 **4**

El 50% de los valores de la variable cantidad de hijos están comprendidos entre 2 y 4.

Varianza muestral: Var (n-1)= 1,44

Promedio del cuadrado de los desvíos de la cantidad de hijos con respecto a la cantidad de hijos promedio.

Desviación Estándar: D.E= **1,20**

Distancia promedio que hay entre la cantidad de hijos y la cantidad de hijos promedio. Existen valores de la cantidad de hijos que se diferencian de la cantidad de hijos promedio en **más de** 1,20

Existen valores de la cantidad de hijos que se diferencian de la cantidad de hijos promedio en **menos de** 1,20

Coeficiente de variación: CV= 41,36%

La desviación estándar que es de **1,20** equivale al **41,36%** del valor de la media que es de **2,90**

Coeficiente de Variación: CV= 41,36 Es mayor que el 26%

0% < CV < 11% Datos son muy homogéneos

11% ≤ CV < 16% Datos homogéneos

16% ≤ CV < 26% Datos heterogéneos

CV ≥ 26% Los valores correspondientes a la variable *cantidad de hijos* son muy heterogéneos (presentan mucha variabilidad, son muy desparejos, muy diferentes entre si)

Var(n-1) Varianza Muestral corregida

D.E: Raiz cuadrada de Var(n-1)

Medidas resumen

Variable	n	D.E.	Var(n-1)	Mín	Máx	Q1	Q3
Salarios en $	50	8,56	73,36	117,00	153,00	131,00	145,00

Asimetría= -0,10

El Coeficiente de Asimetría es menor que cero por lo tanto la distribución correspondiente a la variable cantidad de hijos es Asimétrica Izquierda.

También puede identificarse de la siguiente manera:

Media < Mediana ⟹ Distribución Asimétrica Izquierda

2,9 < 3

Kurtosis= -0,91

El Coeficiente de Curtosis es menor que cero por lo tanto la distribución correspondiente a la variable *cantidad de hijos* es Platicúrtica es por ello que presenta un reducido grado de concentración con respecto a los valores centrales de dicha variable.

Interpretación de las Medidas Descriptivas para la variable *Salario.*

Resumen	Salarios en $
n	50,00
Media	137,22
D.E.	8,56
Var(n-1)	73,36
CV	6,24
Mín	117,00
Máx	153,00
Mediana	137,50
Q1	131,00
Q3	145,00
Suma	6861,00
Asimetría	-0,22
Kurtosis	-0,65

Media= 137,22 En promedio cada empleado percibe un salario de 137,22
Mediana= 137,50 La mitad de los empleados percibe un salario menor que 137,50 y la otra mitad de los empleados percibe un salario mayor que 137,50

Moda: Valor más frecuente.
En una tabla con intervalos no es posible determinar el valor. Sólo puede identificarse el intervalo en el cual está comprendida la Moda.
El valor de la Moda se encuentra comprendido en el intervalo al cual le corresponde la mayor frecuencia absoluta simple.

Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Salarios	1	[115.00	123.00)	119.00	2	0.04	2	0.04
Salarios	2	[123.00	131.00)	127.00	10	0.20	12	0.24
Salarios	3	[131.00	139.00)	135.00	19	0.38	31	0.62
Salarios	4	[139.00	147.00)	143.00	12	0.24	43	0.86
Salarios	5	[147.00	155.00]	151.00	7	0.14	50	1.00

Moda: El salario más frecuente está comprendido entre 131 y 139.

Recorrido: Máx – Min = 153 – 117

R =36

Es la máxima separación entre los valores de la variable salario.

Primer Cuartil: Q1= 131

El 25% de los empleados tienen un salario menor que 131.

El 75% de los empleados tienen un salario mayor que 131.

Primer Cuartil: Q3= 145

El 75% de los empleados tienen un salario menor que 145

El 25% de los empleados tienen un salario mayor que 145

Rango Intercuartílico: Diferencia entre el Tercer Cuartil y el Primer Cuartil.

El 50% de los valores de la variable están comprendidos entre el Primer Cuartil y el Tercer Cuartil

```
  25%   25%   25%   25%
|_____|_____|_____|_____|
     Q1          Q3
```

Q3 – Q1= 145 - 131

= 14

```
  25%   25%   25%   25%
|_____|_____|_____|_____|
     145         131
```

El 50% de los valores de los salarios están comprendidos entre 131 y 145

Varianza muestral: Var (n-1)= 73,36

Promedio del cuadrado de los desvíos de los salarios con respecto al salario promedio.

Desviación Estándar = **8,56**

Distancia promedio que hay entre los valores de la variable salario y la media (salario promedio).

En promedio, el valor de un salario se diferencia del salario promedio (**137,50**) en aproximadamente **8,56.**

Algunos valores de salarios observados se separan en más de **8,56** con respecto al valor de la media. Algunos valores de salarios observados se separan en menos de **8,56** con respecto al valor de la media**.**

Coeficiente de variación: CV= 6,24%

La desviación estándar que es de **8,56** equivale al **6,24%** del valor de la media que es de **137,50**.

0% < CV < 11% Datos son muy homogéneos (son muy similares entre si)

11% ≤ CV < 16% Datos homogéneos

16% ≤ CV < 26% Datos heterogéneos

0% < CV < 11% Datos muy homogéneos

Asimetría= -0,22

El Coeficiente de Asimetría es menor que cero por lo tanto la distribución correspondiente a la variable *salario* es Asimétrica Izquierda.

También puede identificarse de la siguiente manera:

Media < Mediana $\Rightarrow$ Distribución Asimétrica Izquierda

137,22 < 137,50

Kurtosis= -0,65

El Coeficiente de Curtosis es menor que cero por lo tanto la distribución correspondiente a la variable *salario* es es Platicúrtica es por ello que presenta un reducido grado de concentración con respecto a los valores centrales de dicha variable.

Interpretación de las Medidas Descriptivas para Datos Cualitativos

Para Datos Cualitativos Nominales **solo** puede identificarse la categoría más frecuente (Moda).

Ejemplo:

Variable	Clase	Categorías	FA	FR
Área	1	Administración	21	0.42
Área	2	Comunicación	5	0.10
Área	3	Diseño	13	0.26
Área	4	Ingeniería	9	0.18
Área	5	RRHH	2	0.04

Administración **es el área a la cual le corresponde la mayor frecuencia absoluta simple (simbolizada como FA o ni)**

Para Datos Cualitativos Ordinales pueden identificarse la Moda y la Mediana.

Ejemplo:

Nivel máximo de estudios que poseen los habitantes de cierta región

Moda

Nivel de estudios más frecuente.

Mediana

La mitad de los empleados tienen estudios secundarios completos o un nivel inferior, y la otra mitad de los empleados tienen estudios secundarios completos o un nivel superior.

Propiedades de la Media y de la Varianza

Las propiedades de la Media y de la Varianza se utilizan para calcular la nueva media y la nueva varianza luego de corregir cierto error detectado en los valores de la variable

Propiedades de la Media Aritmética:

1) **M(c) = c**

   La media de una constante es igual a la constante misma.

2) **M(cx) = c M(x) Esta propiedad se aplica cuando el error está expresado en %**

   La media del producto de una constante por una variable es igual al producto de la constante por la media de la variable.

3) $M(x \pm c) = M(x) \pm c$

La media de la suma de una variable y una constante es igual suma de la media de la variable más constante. Esta propiedad se aplica de manera similar en la media de la diferencia.

Propiedades de la Varianza:

1) $V(c) = 0$

La varianza de una constante es igual a cero.

2) $V(cx) = c^2 V(x)$ **Esta propiedad se aplica cuando el error está expresado en %**

La varianza del producto de una constante por una variable es igual al producto del cuadrado de la constante por la varianza de la variable.

3) $V(x \pm c) = V(x)$

La varianza de la suma de una variable más una constante es igual a la varianza de la variable. Esta propiedad se aplica de manera similar en la varianza de la diferencia.

¿Cuál sería el valor del promedio y el de la desviación estándar correspondiente a la variable **Salario,** si se ha comprobado que existe un error en más (por exceso del 3,21% en el registro de los datos?

Error por exceso → error en más

Tendremos en cuenta la siguiente simbología:

w = salario correcto

x = precio con el error

Un error en más significa que está sobrando el 3,21% por lo tanto, al valor "x" le debemos restar el 3,21%.

En primer lugar debemos calcular el 3,21% de x, para ello planteamos una regla de 3 simple:

100% ___ x

3,21% ____ (3,21%*x)/100% = 0,0321x

Si al valor "x" le restamos "0,0321x" le estaremos restando el 3,21% y de esa manera obtendremos el precio correcto "w":

w = x- 0,0321x

Sacamos factor común a la variable "x" obteniendo:

w = (1- 0,0321)x

w = 0,9679x

Si a "x" lo multiplicamos por 0,9679 le estaremos restando el 3,21% que está sobrando.

Cálculo del nuevo Promedio:

Aplicamos media en ambos miembros de la igualdad

M(w)= M(0,9679x)

En el segundo miembro de la igualdad nos queda planteada la media de una constante por una variable por lo tanto corresponde aplicar la propiedad 2: "La media de una constante por una variable es igual a la constante por la media de la variable": M(kx)= kM(x)

M(w)=M(0,9679x)

=0,9679 M(x)

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana	Q1	Q3	Asimetría	Kurtosis
Salarios	50	137.22	8.56	73.36	6.24	117.00	153.00	137.50	131.00	145.00	-0.22	-0.65

M(x) = 137,22 Media calculada con Infostat a partir de la columna Salarios en la B.D.

M(w) = 0,9679 * 137,22

M(w) = 132,82 Nuevo salario promedio luego de corregir el error del 3,21% en más.

Cálculo de la nueva Desviación Standar:

Antes de calcular la nueva Desviación Estándar debemos obtener la nueva Varianza.

Aplicamos varianza en ambos miembros de la igualdad

V(w) = V(0,9679x)

En el segundo miembro de la igualdad nos queda planteada la varianza de una constante por una variable por lo tanto corresponde aplicar la propiedad 2: “La varianza de una constante por una variable es igual al cuadrado de la constante por la varianza de la variable”

V(w) = V(0,9679x)

= $(0,9679)^2$ V(x)

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana	Q1	Q3
Salarios	50	137.22	8.56	$\boxed{73.36}$	6.24	117.00	153.00	137.50	131.00	145.00

V(x) = 73,36 Esta es la varianza que debe calcularse a través del Infostat a partir de la columna Salarios en la B.D.

V(w) = $(0,9679)^2$ * 73,36

V(w) = 68,73 Nueva varianza luego de corregir el error del 3,21% en más.

Se pide la Desviación Standar, debe obtenerse calculando la raíz cuadrada de la varianza:

DS(w) = $\sqrt{68,73}$

DS(w) = 8,29

Nueva Desviación Estándar luego de corregir el error del 3,21% en más.

¿Cuáles sería el valor del promedio y el de la desviación estándar correspondiente a la variable **cantidad de hijos** si se ha comprobado que existe un error por defecto de 1 hijo en el registro de los datos?

Error por defecto = error en menos

Un error en más significa que está faltando 1 por lo tanto, al valor “x” le debemos sumar los dos pesos que están faltando.

Cabe aclarar que en este caso no corresponde el planteo de la regla de 3 porque el error no está expresado en % (es un error de 1 en menos y no de un 1% en menos) por lo tanto directamente sumamos 1:

w= x + 1

A continuación aplicamos media en ambos miembros de la igualdad

M(w) = M(x + 1)

En el segundo miembro de la igualdad nos queda planteada la media de la suma entre una variable y una constante por lo tanto corresponde aplicar la propiedad 3: “La media de la suma entre una variable y una constante es igual a la media de la variable más la constante”

M(w)= M (x + 1)

=M(x) + 1

M(x) = 2,90 Esta es la media que debe calcularse a partir de la columna Hijos en la B.D. utilizando Infostat:

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana	Q1	Q3
Hijos	50	$\boxed{2.90}$	1.20	1.44	41.36	1.00	5.00	3.00	2.00	4.00

M(w) = 2,90 + 1

M(w) = 3,90 Nueva cantidad de hijos promedio luego de corregir el error de 1 hijo por defecto.

Cálculo de la nueva Desviación Standar:

Antes de calcular la nueva Desviación Estándar debemos obtener la nueva Varianza.

Aplicamos varianza en ambos miembros de la igualdad

V(w) = V(x - 1)

En el segundo miembro de la igualdad nos queda planteada la media de la diferencia entre una variable y una constante por lo tanto corresponde aplicar la propiedad 3: “La varianza de la diferencia entre una variable y una constante es igual a la varianza de la variable” por lo tanto, la varianza no se modifica

V(w) = V(x - 1)

= V(x) La varianza no se modifica.

Medidas resumen

Variable	n	Media	D.E.	Var(n-1)	CV	Mín	Máx	Mediana	Q1	Q3
Hijos	50	2.90	1.20	1.44	41.36	1.00	5.00	3.00	2.00	4.00

V(x) = 1,44 Esta es la varianza que debe calcularse a partir de la columna Hijos, utilizando el Infostat

V(w) = 1,44 (la varianza no se modifica si incrementamos ó restamos una suma fija, solo cambia su valor cuando la modificación se expresa en porcentaje)

En consecuencia la DS tampoco se modifica:

DS(x)= √1,44

= 1,20 La desviación stándar no se modifica luego de corregir el error de1 hijo por defecto.

¿Cuál es el valor de la nueva media y el valor de la nueva varianza luego de corregir un error del 7,21% en menos y un error de medio punto en más en los valores originales de las notas?

x: nota antes de la corrección (notas originales en las cuales se descubrió el error)

w: nota después de la corrección

Error del 7,21% en menos (error por defecto) →

Debemos sumar el 7,21% que está faltando

X + (7,21% de x)

~~**X + 7,21**~~

Debemos obtener la expresión que representa al 7,21%:

100% ___ x

7,21% ___ (7,21 . X) / 100= 0,0721x

0,0721x equivale al 7,21% de x

w= x + (7,21% de x)

w= x 0,0721x

w = 1x + 0,0721x (de esta manera a x le sumamos el 7,21%)

Sacamos factor común x:

w= (1+ 0,0721)x

w= 1,0721x (de esta manera a x le sumamos el 7,21%)

Si a un valor lo multiplicamos por 1,0721 entonces le estamos sumando el 7,21%

Además debemos corregir un error de 0,5 punto en más (error por exceso) para ello debemos restar el medio punto que está sobrando:

w = 1,0721x – 0,5

Propiedades de la Media

c: constante
x: variable

1) **M(c)= c**
2) **M(cx)= cM(x)** **se aplica cuando el error está expresado en porcentaje**
3) **M(x+c)= M(x)+ M(c)**
   = M(x)+ c

La propiedad 3 se aplica de manera similar para la diferencia:

M(x- c)= M(x)- M(c)

Aplicamos Media en cada término:

M(w) = M(1,0721x) – M(0,5)

En el primer término aplicamos la propiedad 2 y en el segundo término

aplicamos la propiedad 1:

= 1,0721 M(x) – 0,5

M(w) = 1,0721* $4,5$ **– 0,5**

M(w) = 4,3244 Nueva nota promedio luego de corregir un error de 7,21% en menos y un error de medio punto en más.

Propiedades de la Varianza:

1) V(c)= 0 Ejemplo: 7 7 7 7 Media= 7 Varianza = 0

2) V(cx)= c^2 V(x) se aplica cuando el error está expresado en porcentaje

3) V(x+c)= V(x)+ V(c)
= V(x)+ 0
V(x+c)= V(x)

V(x- c)= V(x)- V(c)
= V(x)

Nueva varianza:

w= 1,0721x – 0,5
V(w) = V(1,0721x) – V(0,5)
= V(1,0721x) – V(0,5)

En el primer término aplicamos la propiedad 2.

En el segundo término aplicamos la propiedad 1: La varianza de una constante es igual a cero.

V(w) = $(1,0721)^2$ V(x) – 0
= $(1,0721)^2$ (1,67)

V(w) = 1,919 Nueva varianza luego de corregir un error de 7,21% en menos y un error de medio punto en más.

d. ¿Cuál es el valor de la media y el valor de la nueva varianza luego de corregir un error de medio punto en más en los valores originales de las notas?

Media:

M(w)= M(x – 0,5)
= M(x) – 0,5
= 4,5 – 0,5
M(w)= 4

Varianza:

V(w)= V(x – 0,5)
= V(x) –V(0,5)

En el segundo término aplicamos la propiedad 1: La varianza de una constante es igual a cero.

= 1,67 – 0
V(w)= 1,67   La Varianza no varía.

Ejercitación

Ejercicio N° 1:

Espesor en mm. de 50 piezas

0,26	0,23	0,79	0,28	0,22	0,26	0,37	0,28	0,41	0,35
0,36	0,24	0,73	0,28	0,30	0,48	0,38	0,28	0,38	0,23
0,33	0,27	0,66	0,27	0,32	0,25	0,36	0,31	0,59	0,33
0,25	0,23	0,28	0,29	0,36	0,23	0,29	0,65	0,58	0,23
0,27	1,17	0,29	0,28	0,33	0,23	0,30	0,54	0,59	0,30

Calcular el Recorrido o Rango.

Recorrido o Rango: Valor máximo – Valor mínimo

Diferencia entre el valor máximo y el valor mínimo. Es la mayor separación entre todos los valores observados

Respuesta: 1,17- 0,22= 0,95

La mayor separación entre **todos** los valores de la variable espesor es de 0,95 mm.

Ejercicio N° 2:

En la empresa "A" la antigüedad promedio de un grupo de empleados seleccionados fue de 10 años y la desviación estándar de 1,2 años.

En la empresa “B” la antigüedad promedio de un grupo de empleados seleccionados fue de 24 años y la desviación estándar de 2,2 años.

¿En cuál empresa los valores de las antigüedades son más homogéneos?

¿En cuál empresa los valores de las antigüedades presentan menor variabilidad?

¿En cuál empresa la media es más representativa?

La Desviación Estándar (DS) solo puede utilizarse cuando las distribuciones a comparar tienen igual media y la variable se encuentra expresada en la misma unidad de medida.

Si las distribuciones tienen distinta media entonces para efectuar la comparación debemos utilizar el Coeficiente de Variación

La distribución con menor Coeficiente de Variación es la más homogénea, por lo tanto los valores son más similares entre si, es por ello que presentan menor variabilidad y están más concentrados con respecto a la media, en consecuencia en dicha distribución la media es más representativa.

CV= DS/Media

El valor del CV también puede expresarse en porcentaje, para ello al cociente lo multiplicamos por 100:

CV= (DS/Media)100

Infostat genera el resultado del CV expresado en porcentaje

	Empresa "A"	Empresa "B"
Media	10 años	24 años
Desviación estándar	1,2 años	2,2 años

Empresa "A" $C.V. = \frac{1,2}{10}$ Empresa "B" $C.V. = \frac{2,2}{24}$

$C.V. = 0,12$ $C.V. = 0,09$

En la **empresa “B”** se obtuvo el menor Coeficiente de Variación por lo tanto los valores de la variable antigüedad son **más homogéneos** que en la empresa “A”.

En la empresa “B” se obtuvo el menor Coeficiente de Variación, por lo tanto en dicha empresa, los valores de la variable antigüedad presentan menor variabilidad (son más similares entre si) que en la empresa “A”.

En la empresa “B” los valores de la variable antigüedad se encuentran a menor distancia con respecto a la Media, en consecuencia la media (antigüedad promedio) es más representativa en esta empresa.

En la empresa “B” se obtuvo el menor Coeficiente de Variación, por lo tanto la antigüedad promedio es más representativa que en la empresa “A”.

En la empresa “A”:

CV= 0,12*100= 12%

Del 11% a menos del 16%	Homogéneo

CV= 12% Por lo tanto en la empresa “A” la distribución de las antigüedades es homogénea.

CV= 12% El valor de la **desviación estándar** que es de **1,2** equivale a un **12%** del valor de la media aritmética que es de 10

En la empresa “A” la **Desviación estándar** es de **1,2** es la distancia promedio que hay entre los valores observados de antigüedad y la media (antigüedad promedio).

En promedio, el valor de una antigüedad se diferencia de la antigüedad promedio (10) en aproximadamente **1,20** años.

La distancia promedio que hay entre los valores de la antigüedad y la antigüedad promedio es de aproximadamente 1,20 años

Algunos valores se separan en más de 1,20 años con respecto a la antigüedad promedio y algunos valores se separan en menos de 1,20 años con respecto a la antigüedad promedio.

Distancia promedio entre los valores de x y la media

1,2 equivale al **12%** de 10

Algunos valores observados de antigüedad se separan en más de **1,20** con respecto al valor de la media. Algunos valores de antigüedad se separan en se separan en menos de **1,20** con respecto al valor de la media (antigüedad promedio)

En la empresa “B”:

CV= 0,09*100= 9%

0% a menos del 11%	Muy Homogéneo

CV= 9% Por lo tanto en la empresa “B” la distribución de las antigüedades es muy homogénea.

CV= 9% El valor de la **desviación estándar** que es de **2,2** equivale a un **9%** del valor de la media que es de 24

En la empresa “B” la **Desviación estándar** es de **2,2** es la distancia promedio que hay entre los valores observados de antigüedad y la media (antigüedad promedio= 24).

En promedio, el valor de una antigüedad se diferencia de la antigüedad promedio (24) en aproximadamente **2,2** años.

Algunos valores observados de antigüedad se encuentran a más **2,2** años con respecto a la antigüedad promedio y otros valores se encuentran a menos de **2,2** años con respecto a la antigüedad promedio.

Ejercicio N° 3:

Fueron seleccionados 4 estudiantes y se obtuvo que la sumatoria de las notas es 18 y la sumatoria del cuadrado de las notas es 86

Fueron seleccionados por lo tanto los 4 estudiantes forman una muestra

a. **Calcular Media y Varianza.**
b. **¿Cuál es el valor de la media y el valor de la nueva varianza luego de corregir un error del 7,21% en menos y un error de medio punto en más en los valores originales de las notas?**
c. **¿Cuál es el valor de la media y el valor de la nueva varianza luego de corregir un error de medio punto en más en los valores originales de las notas?**

Datos:

X: notas n= 4

$\sum x = 18$

$\sum x^2 = 86 = (x1)^2 + (x2)^2 + (x3)^2 + (x4)^2$

a.

Media muestral:

$$\bar{x} = \frac{\sum x}{n}$$

= **18 /** 4

$\bar{x} = 4{,}5$ Nota promedio. En promedio cada estudiante obtuvo una nota igual a 4,5.

Varianza muestral:

$$\hat{S}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$$

$$= \frac{86 - 4 * 4{,}5^2}{4-1}$$

$\hat{S}^2 = 1{,}67$

Definición:

La varianza es el promedio del cuadrado de los desvíos de los valores con respecto a la media

Interpretación:

La varianza es el promedio del cuadrado de los desvíos de los valores de las notas con respecto a la nota promedio

Propiedades de la Media y la Varianza

Las propiedades de la Media y de la Varianza se utilizan para calcular la nueva media y la nueva varianza luego de corregir cierto error detectado en los valores de la variable

b. ¿Cuál es el valor de la nueva media y el valor de la nueva varianza luego de corregir un error del 7,21% en menos y un error de medio punto en más en los valores originales de las notas?

x: nota antes de la corrección (notas originales en las cuales se descubrió el error)

w: nota después de la corrección

Error del 7,21% en menos (error por defecto) →

Debemos sumar el 7,21% que está faltando

x + (7,21% de x)

~~x + 7,21~~

Debemos obtener la expresión que representa al 7,21%:

100% ___ x
7,21% ___ (7,21~~% .~~ X) / 100~~%~~= 0,0721x
0,0721x equivale al 7,21% de x

w= x + (7,21% de x)
w= x + 0,0721x
w = 1x + 0,0721x (de esta manera a x le sumamos el 7,21%)

Sacamos factor común x:
w= (1+ 0,0721)x
w= 1,0721x (de esta manera a x le sumamos el 7,21%)

Si a un valor lo multiplicamos por 1,0721 entonces le estamos sumando el 7,21%

Error de 0,5 punto en más (error por exceso) → Debemos restar el medio punto que está sobrando:

w = 1,0721x – 0,5

Propiedades de la Media

c: constante
x: variable

1) **M(c)= c**
2) **M(cx)= cM(x) se aplica cuando el error está expresado en porcentaje**
3) **M(x+c)= M(x)+ M(c)**
   = M(x)+ c

La propiedad 3 se aplica de manera similar para la diferencia:

$M(x- c)= M(x)- M(c)$

$= M(x)-c$

Aplicamos Media en cada término:

$w= 1,0721x$

$w=M(1,0721x)$

$w = 1,0721x – 0,5$

En el primer término aplicamos la propiedad 2 y en el segundo término aplicamos la propiedad 1:

$M(w) = M(1,0721x) – M(0,5)$

$= 1,0721\ M(x) – 0,5$

$M(w) = 1,0721* 4,5 – 0,5$

$M(w) = 4,3244$ Nueva nota promedio luego de corregir un error de 7,21% en menos y un error de medio punto en más.

Propiedades de la Varianza:

1) $V(c)= 0$ Ejemplo: 7 7 7 7 Media= 7 Varianza = 0

2) $V(cx)= c^2\ V(x)$ se aplica cuando el error está expresado en porcentaje

3) $V(x+c)= V(x)+ V(c)$
$= V(x)+ 0$
$V(x+c)= V(x)$

$V(x- c)= V(x)- V(c)$
$= V(x)-0$
$= V(x)$

Nueva varianza:

w= 1,0721x – 0,5
V(w) = V(1,0721x) – V(0,5)
= V(1,0721x) – V(0,5)

En el primer término aplicamos la propiedad 2.

En el segundo término aplicamos la propiedad 1: La varianza de una constante es igual a cero.

V(w) = $(1{,}0721)^2$ V(x) – 0
= $(1{,}0721)^2$ (1,67)
V(w) = 1,919 Nueva varianza luego de corregir un error de 7,21% en menos y un error de medio punto en más.

c. ¿Cuál es el valor de la media y el valor de la nueva varianza luego de corregir un error de medio punto en más en los valores originales de las notas?

En el punto c estamos supiendo que solo se detectó el error de medio punto en más.

Error de medio en más ___ error de medio punto por exceso

M(x- c)= M(x)- M(c)

= M(x)-c

Debemos restar el medio punto que está sobrando

w= x- 0,5

Media:

M(w) = M(x – 0,5)
= M(x) – 0,5
= 4,5 – 0,5
M(w) = 4

Varianza:

V(w)= V(x – 0,5)

= V(x) –V(0,5)

En el segundo término aplicamos la propiedad 1: La varianza de una constante es igual a cero.

= 1,67 – 0

V(w)= 1,67   La Varianza no varía.

Ejercicio N° 4:

Dados: 3 6 5 9 2 9

Precios en una muestra de 6 artículos de un comercio "A"   n=6

(Precios de 6 artículos seleccionados de un comercio "A")

Calcular Varianza, Desviación Estándar y Coeficiente de Variación:

$\bar{x}$ = **34** / 6

$$\bar{x} = 5,667$$

$$\sum x^2 = 3^2 + 6^2 + 5^2 + 9^2 + 2^2 + 9^2$$

$$\sum x^2 = 236$$

Varianza muestral:

$$\hat{s}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$$

$$\hat{s}^2 = \frac{236 - 6 * 5,667^2}{6-1}$$

$\hat{s}^2$= 8,66 Varianza a partir de una muestra.

Desviación Estándar: Raíz cuadrada de la varianza:

$\hat{s}= \sqrt{8,66}$

$\hat{s}$=2,94

$\hat{s}$ simboliza Desviación Estándar Muestral: Raíz cuadrada de la varianza

Coeficiente de Variación: Cociente entre Desviación Estándar y Media Aritmética (Promedio):

$CV= \frac{\hat{s}}{\bar{x}} *100$

$CV= \frac{2,94}{5,667} *100$

CV=51,87%

Los precios en la muestra de 6 artículos son muy heterogéneos es decir que presentan mucha variabilidad, son muy diferentes entre sí.

El valor de la Desviación Estándar equivale al 51,87% del valor de la Media

Si en otro comercio “B” la Desviación estándar es de $4,75 y la Media es de $12,54. ¿Cuál grupo de precios es más homogéneo?

¿Utilizamos Desviación Estándar o Coeficiente de Variación?

Los precios tienen distinta Media por lo tanto utilizaremos el Coeficiente de Variación para efectuar la comparación. Aquel grupo con menor Coeficiente de Variación será el más homogéneo.

CV(B)= DE/Media*100

= $4,75 / $12,54*100

CV(B)= 37,88%

En el comercio “B” se obtuvo menor Coeficiente de Variación, por lo tanto en dicho comercio los precios se encuentran a menor distancia con respecto al precio promedio, por lo tanto en el comercio “B” los precios son más homogéneos, es decir presentan menor variabilidad que en el comercio “A”.

Distancias semanales en kilómetros recorridas por los distribuidores de tres empresas:

	Empresa A	Empresa B	Empresa C
Media	**321,7**	**427,9**	**375,3**
Varianza	**10465,29**	**12122,01**	**254016**
Desviación Standar	**102,3**	**110,1**	**5,04**
Coeficiente de variación	**(102,3/321,7)*100= 31,8%**	**25,7%**	**1,34%**

Coeficiente de variación= (Desviación Estándar / Media)*100

Realizar un análisis comparativo entre las tres distribuciones.

Resolución:

Variable bajo estudio: Distancia recorrida

0 < CV < 11%   Distribución muy homogénea, presenta poca variabilidad
11% ≤ CV < 16%   Distribución homogénea
16% ≤ CV < 26%   Distribución heterogénea
CV ≥ 26% Distribución muy heterogénea, presentan mucha variabilidad

CV= 31,8% En la Empresa "A" la distribución de variable *distancia recorrida* es muy heterogénea, por lo tanto los valores de la variable *distancia recorrida* presentan mucha variabilidad, es decir que los valores observados son muy diferentes entre sí.

CV= 25,7% En la Empresa "B" la distribución es heterogénea

CV= 1,34% En la Empresa "C" la distribución es muy homogénea, por lo tanto los valores de la variable *distancia recorrida* presentan poca variabilidad. **Esta**

distribución es la más homogénea de las tres. En esta distribución la media es más representativa de los valores observados.

En la Empresa "C "se obtuvo el menor Coeficiente de Variación por lo tanto en dicha empresa las distancias recorridas son más homogéneas, es decir que presenta menor variabilidad, y en consecuencia la media es más representativa.

Ejercicio N° 5:

Completar la siguiente tabla:

Ingresos	Ni	hi	Ni	Hi
110 120	**1**	1/50= 0,02	1	0,02
120 130	0,22*50= 11	**0,22**		0,24
130 140	**20**			
140 150			0,94*50=47	**0,94**
150 **160**			**n=50**	1

Para completar la tabla debemos tener en cuenta que:

En el primer intervalo, ni y Ni son iguales

En el primer intervalo, hi y Hi son iguales

Ni del último intervalo es igual a n (total de valores observados)

Hi del último intervalo es igual a 1

ni: es la frecuencia absoluta simple

Ni: es la frecuencia absoluta acumulada

Fórmulas que utilizaremos para completar la tabla

hi: frecuencia relativa simple hi= ni/n

Hi: frecuencia relativa acumulada Hi= Ni/n

Fórmula para calcular una frecuencia relativa simple a partir de una frecuencia absoluta simple:

$hi= ni/n \rightarrow ni= hi * n$

Fórmula para calcular una frecuencia relativa acumulada a partir de una frecuencia absoluta acumulada:

$Hi= Ni/n \rightarrow Ni= Hi * n$

Fórmula para calcular una frecuencia absoluta simple a partir de frecuencias absolutas acumuladas:

$ni= Ni – N_{i-1}$ Frecuencia absoluta acumulada del mismo intervalo de la frecuencia absoluta simple que debemos calcular menos Frecuencia absoluta acumulada del intervalo anterior

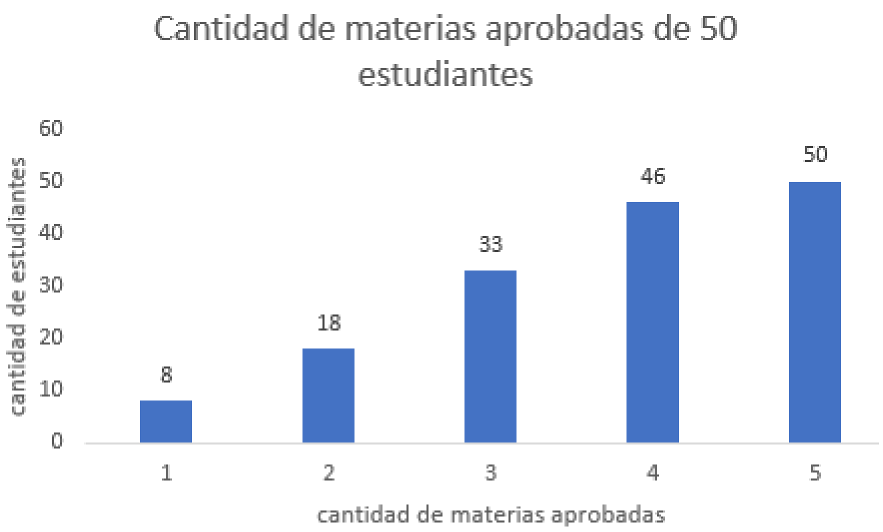
$hi= Hi – H_{i-1}$ Frecuencia relativa acumulada del mismo intervalo de la frecuencia relativa simple que debemos calcular menos Frecuencia relativa acumulada del intervalo anterior

Ingresos	ni	hi	Ni	Hi
110 120	**1**	1/50=0,02	1	1/50=0,02
120 130	**0,22*50=11**	**0,22**	1+11=12	0,02+0,22 o bien 12/50= 0,24
130 140	**20**	20/50= 0,40	1+11+20= 32	32/50= 0,64
140 150	0,30*50= 15	0,94-0,64= 0,30	0,94*50= 47	**0,94**
150 **160**	50-47= 3	1-0,94= 0,06	**n=50**	1

Ingresos	ni	hi	Ni	Hi
110 120	**1**	1/50= 0,02	1	0,02
120 130	0,22*50= 11	**0,22**	1+11=12	12/50= 0,24

130	140	**20**	0,40	1+11+20=32	32/50=0,64
140	150	47-32=15	15/50=0,30	0,94*50=47	**0,94**
150	160	50-47=3	3/50=0,06	**n=50**	1

Ejercicio N° 6:



El gráfico acumulativo de frecuencias viene dado como dato.

Se pide:

a. ¿Cuántos estudiantes tienen 4 o menos materias aprobadas?
b. ¿Cuántos estudiantes tienen al menos 3 materias aprobadas?
c. ¿Cuántos estudiantes tienen al más de 2 materias aprobadas?
d. ¿Cuántos estudiantes tienen 5 materias aprobadas?
e. ¿Cuántos estudiantes tienen entre 3 y 4 (se incluyen ambos valores) materias aprobadas?

Respuesta:

Dadas Ni calcular ni

Antes de responder las preguntas es conveniente calcular las **frecuencias absolutas simples**. En el gráfico están representadas las **frecuencias absolutas acumuladas**, por lo tanto tenemos como dato dichas frecuencias

¿Cómo calcular frecuencias absolutas simples a partir de frecuencias absolutas acumuladas?

La frecuencia absoluta simple de un intervalo i (n_i) se obtiene mediante la diferencia entre la frecuencia absoluta acumulada del intervalo i menos la frecuencia absoluta acumulada del intervalo anterior: $n_i = N_i – N_{i-1}$

yi	ni	Ni (dato)
1	8	8
2	**18- 8=** 10	18
3	**33- 18=** 15	33
4	**46- 33=** 13	46
5	**50- 46=** 4	50

a. ¿Cuántos estudiantes tienen 4 o menos materias aprobadas?
   8+ 10+ 15+ 13= 46

b. ¿Cuántos estudiantes tienen al menos 3 materias aprobadas?
   Al menos 3 equivale a 3 ó más: 15 + 13 + 4= 32

c. ¿Cuántos estudiantes tienen más de 2 materias aprobadas?
   Más de 2 equivale a 3 ó más: 32

d. ¿Cuántos estudiantes tienen 5 materias aprobadas?
   4

e. ¿Cuántos estudiantes tienen entre 3 y 5 (se incluyen ambos valores) materias aprobadas?
   3, 4 o 5 materias aprobadas: 15+ 13+ 4= 32

Ejercicio N° 7:

Gastos en cierto rubro de un grupo de familias encuestadas:

351.8	266.6	344.8	381.2	258.4	393.8	344.8
395.6	311.2	381.4	345	285.2	256.8	358.8
258.6	314.4	215.6	257	335	262.2	384.6
262	344.6	260.2	390	392.4	271.8	389.8
265.8	211	331.2	233.4	393.4	311.6	393.6

Completar la tabla:

ni= FA hi= FR Ni= FAA Hi= FRA

y´i-1 y' i	yi	ni	Ni	hi	Hi
[200 - 240)	220	3	3	0,09	0,09
[240 - 280)	260	10	13	0,29	0,37
[280 - 320)	300	4	17	0,11	0,49
[320 - 360)	340	8	25	0,23	0,71
[360 - 400]	380	10	**35**	0,29	1

yi: punto medio (y'i-1 + y'i)/2 (En Infostat está simbolizado como MC: marca de clase)

Copiamos los datos y los pegamos en Excel:

351.8	266.6	344.8	381.2	258.4	393.8	344.8
395.6	311.2	381.4	345	285.2	256.8	358.8
258.6	314.4	215.6	257	335	262.2	384.6
262	344.6	260.2	390	392.4	271.8	389.8
265.8	211	331.2	233.4	393.4	311.6	393.6

Los ubicamos en una **única** columna:

A
351.8
395.6
258.6
262
265.8
266.6
311.2
314.4
344.6
211
344.8
381.4
215.6
260.2
331.2

Posteriormente copiamos la columna en Infostat.

Caso	Gastos	Colum
1	351,80	
2	395,60	
3	258,60	
4	262,00	
5	265,80	
6	266,60	
7	311,20	
8	314,40	
9	344,60	
10	211,00	
11	344,80	
12	381,40	
13	215,60	
14	260,20	
15	331,20	
16	381,20	
17	345,00	
18	257,00	
19	390,00	
20	[illegible]	

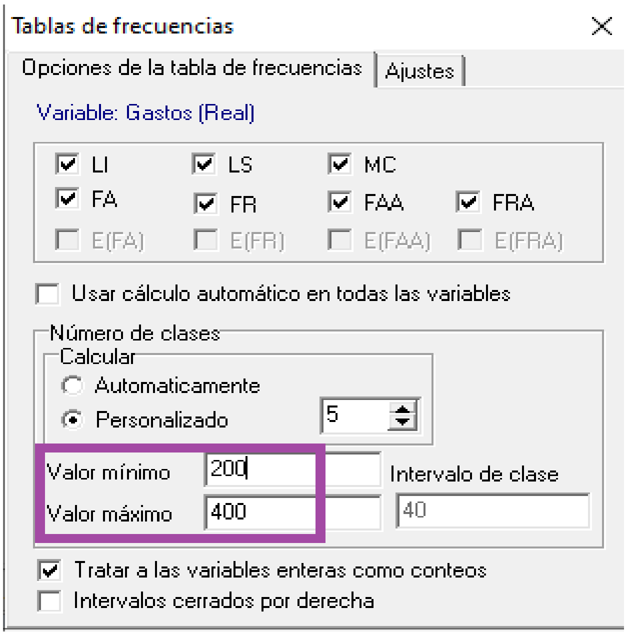
Pasos para generar la tabla con intervalos:

Estadísticas → Tablas de frecuencias

Luego:

Ingresamos como valor mínimo al límite inferior del primer intervalo, e ingresamos como valor máximo al límite superior del último intervalo (son los **límites** que vienen dados como datos):

y´i-1 y’ i	yi	Ni	Ni	Hi	Hi
200 - 240					
240 - 280					
280 - 320					
320 - 360					
360 - **400**					



Variable	Clase	LI	LS	MC	FA	FR	FAA	FRA
Gastos	1	[200.00	240.00)	220.00	3	0.09	3	0.09
Gastos	2	[240.00	280.00)	260.00	10	0.29	13	0.37
Gastos	3	[280.00	320.00)	300.00	4	0.11	17	0.49
Gastos	4	[320.00	360.00)	340.00	8	0.23	25	0.71
Gastos	5	[360.00	400.00]	380.00	10	0.29	35	1.00

Completamos la tabla con los valores obtenidos en Infostat:

y´i-1 - y' i	yi	Ni	Ni	Hi	Hi
200 - 240	220	3	3	0,09	0,09
240 - 280	260	10	13	0,29	0,37
280 - 320	300	4	17	0,11	0,49
320 - 360	340	8	25	0,23	0,71
360 - **400**	380	10	35	0,29	1,00

Ejercicio N° 8: Completar la tabla:

y´i-1 - y'i	yi	ni	Ni	hi	Hi
0,45 - 1,61					**0,80**
1,61 - 2,77					**0,84**

2,77 - 3,93					**0,92**
3,93 - 5,09					**0,96**
5,09 - 6,25			**50**		**1**

Para completar la tabla de datos agrupados, debemos tener en cuenta las siguientes fórmulas:

hi= ni/n → ni= hi * n

Hi= Ni/n → Ni= Hi * n

ni= Ni – Ni-1

hi= Hi – Hi-1

y´i-1 - y’i	yi	ni	Ni	hi	Hi
0,45 - 1,61	1,03	0,80*50= 40	40	0,80	**0,80**
1,61 - 2,77		0,04*50= 2	40+2= 42	0,84-0,80= 0,04	**0,84**
2,77 - 3,93		0,08*50= 4	40+2+4= 46	0,92-0,84= 0,08	**0,92**
3,93 - 5,09		0,04*50= 2	0,96*50= 48	0,96-0,92= 0,04	**0,96**
5,09 - 6,25		0,04*50= 2 o bien: 50-48= 2	**50**	1-0,96= 0,04	**1**

0,80+0,04+0,08+0,04+0,04= 1

y_1: punto medio del primer intervalo = (0,45 + 1,61) /2

=1,03

Respuesta:

y´i-1	y'i	yi	ni	Ni	hi	Hi
0,45 -	1,61	1,03	0,80*50=40	40	0,80	**0,80**
1,61 -	2,77	2,19	42- 40= 2	0,84*50= 42	2/50= 0,04	**0,84**
2,77 -	3,93	3,35	46- 42= 4	0,92*50= 46	4/50= 0,08	**0,92**
3,93 -	5,09	4,51	48- 46= 2	0,96*50= 48	2/50= 0,04	**0,96**
5,09 -	6,25	5,67	50- 48= 2	**50**	2/50= 0,04	**1**

Ejercicio N° 9: Salarios diarios de 100 empleados de cierta empresa:

Completar:

y´i-1	y' i	Ni	Hi	Ni	Hi
1008	1170	**19**			
1170	1332		**0,16**		
1332	1494			**47**	
1494	1656				**0,60**
1656	1818			**82**	
1818	1980	**18**			

Para completar la tabla de datos agrupados, debemos tener en cuenta las siguientes fórmulas:

hi= ni/n → ni= hi * n

Hi= Ni/n → Ni= Hi * n

ni= Ni – Ni-1

hi= Hi – Hi-1

En el primer intervalo, ni y Ni son iguales

En el primer intervalo, hi y Hi son iguales

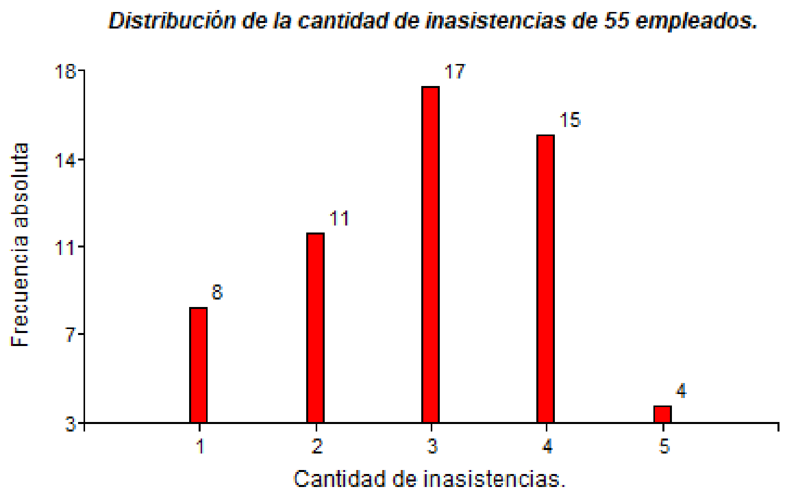
Ni del último intervalo es igual a n (total de valores observados)

Hi del último intervalo es igual a 1

Respuesta:

y´i-1	y' i	Ni	Hi	Ni	Hi
1008	1170	**19**	19/100= 0,19	19	19/100= 0,19
1170	1332	0,16*100= 16	**0,16**	19 + 16= 35	35/100= 0,35
1332	1494	47- 35= 12	12/100= 0,12	**47**	47/100= 0,47
1494	1656	60- 47= 13	13/100= 0,13	0,60*100= 60	**0,60**
1656	1818	82- 60= 22	22/100= 0,22	**82**	82/100= 0,82
1818	1980	**18**	18/100= 0,18	100	1

Ejercicio N° 10:



Eje horizontal: Clases: yi

Gráfico de bastones (ni; hi)

La suma de las ni= 8+ 11+ 17 + 15 +4

$\sum$ni= 55

n=55

A partir del gráfico, completar la tabla. Expresar los valores en dos decimales efectuando el redondeo simétrico.

Si el tercer decimal es mayor o igual que 5 redondeamos hacia arriba:

8/55= 0,145 ≈ 0,15

Si el tercer decimal es menor que 5 redondeamos hacia abajo:

36/55= 0,654 ≈ 0,65

yi	ni	hi	Ni	Hi
1	**8**	8/55= 0,15	8	8/55= 0,15
2	11	11/55= 0,20	8+ 11= 19	19/55=0,35
3	17	17/55= 0,31	8+ 11+ 17= 36	36/55=0,65
4	15	15/55= 0,27	8+ 11+ 17+15= 51	51/55=0,93
5	4	4/55= 0,07	8+ 11+ 17+15+4= 55	55/55= 1

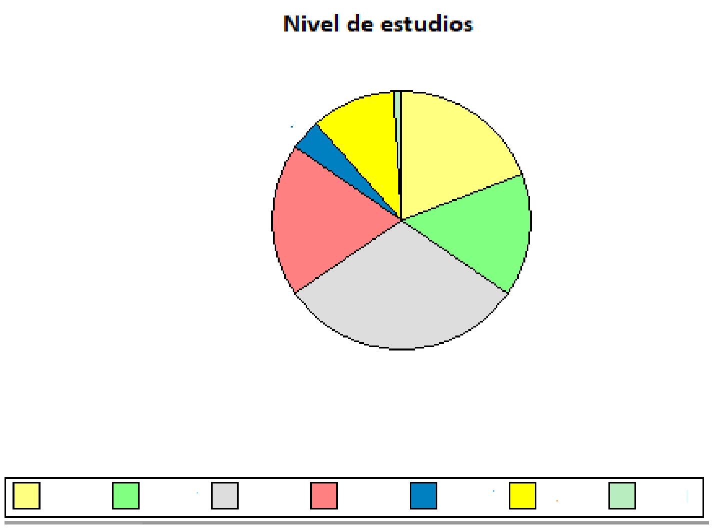
Ejercicio N° 11:

Dada la siguiente tabla:

Nivel de estudios	Cantidad de empleados
1-PC	32
2-PI	20
3-SC	16
4-SI	20
5-T	11
6-UI	1
7-UC	4

Nivel de estudios: Dato cualitativo ordinal.

Completar el siguiente gráfico:



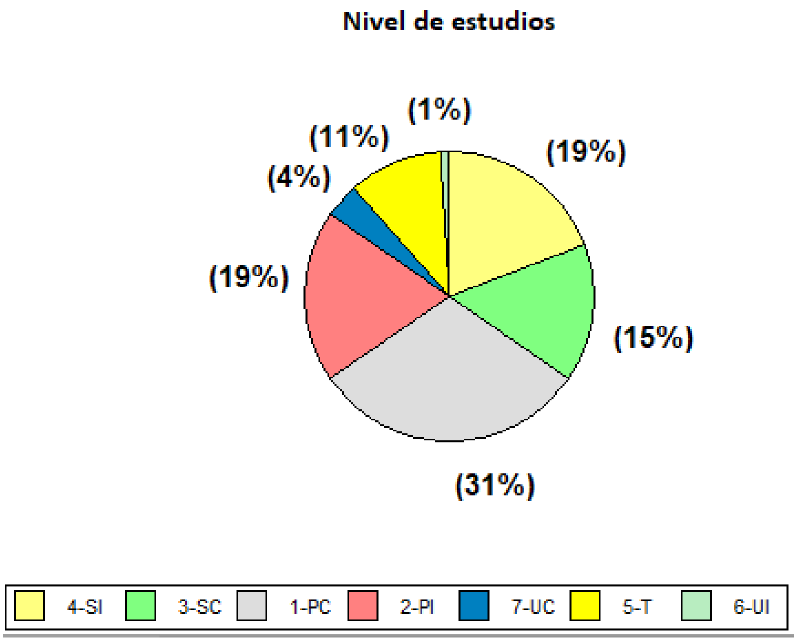
Fuente: Área de Estadística y Censos del Ministerio de Trabajo de la Pcia. X

Ubicar los siguientes valores y categorías:

11%	1%	19%	15%	31%	19%	4%
PI	PC	SI	SC	T	UI	UC

Respuesta:

Para completar el gráfico debemos tener en cuenta que a las mayores cantidades les corresponden los mayores porcentajes.



Fuente: Área de Estadística y Censos del Ministerio de Trabajo de la Pcia. X

Ejercicio N° 12:

Completar la siguiente tabla:

Inasistencias anuales de 55 empleados. Utilizar redondeo simétrico y dos decimales.

yi	ni	hi	Ni	Hi
1	**8**	8/55	8	8/55
2			**19**	
3		**0,31**		
4	**15**			**0,93**
5			55	

Respuesta:

yi	Ni	hi	Ni	Hi
1	**8**	0,15	8	0,15
2	19-8= 11	11/55= 0.20	**19**	19/55= 0,35
3	0,31*55= 17	**0,31**	19+ 17= 36	36/55= 065

4	**15**	15/55= 0,27	0,93*55= 51	**0,93**
5	55- 51= 4	4/55**=** 0,07	55	1

Ejercicio N° 13:

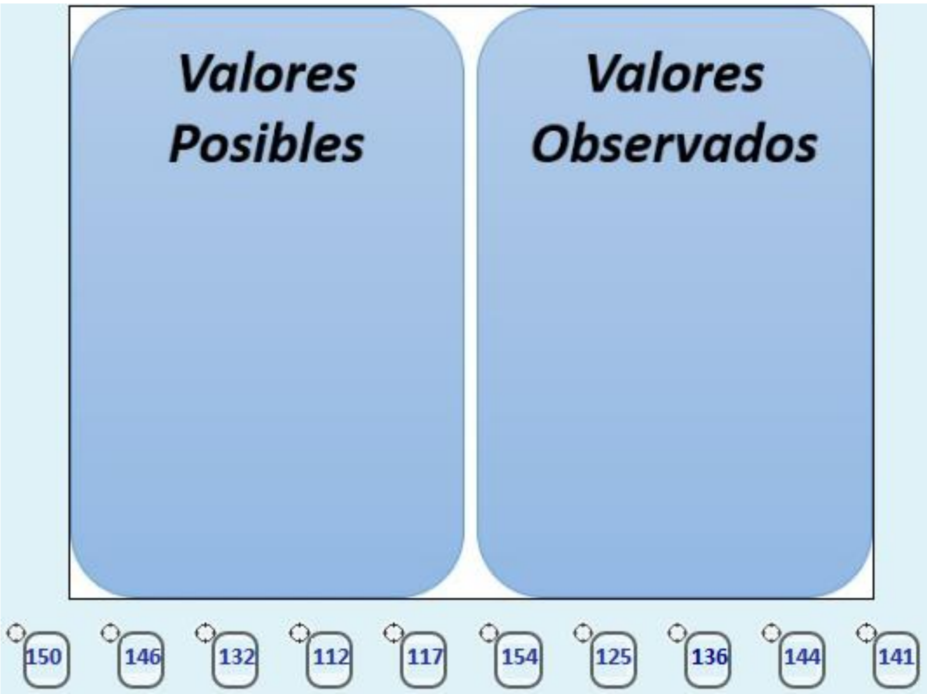
Salarios diarios de 50 empleados encuestados que trabajan en una cierta empresa:

Estos son los valores observados porque son los valores que se presentaron al recolectar los datos.

153	**123**	**129**	**132**	**147**	**138**	**137**	**134**	**131**	**147**
134	**148**	**125**	**139**	**146**	**145**	**148**	**135**	**152**	**128**
138	**138**	**122**	**146**	**137**	**151**	**145**	**124**	**132**	**138**
137	**146**	**138**	**146**	**140**	**137**	**129**	**126**	**117**	**136**
128	**127**	**146**	**144**	**137**	**128**	**142**	**141**	**136**	**138**

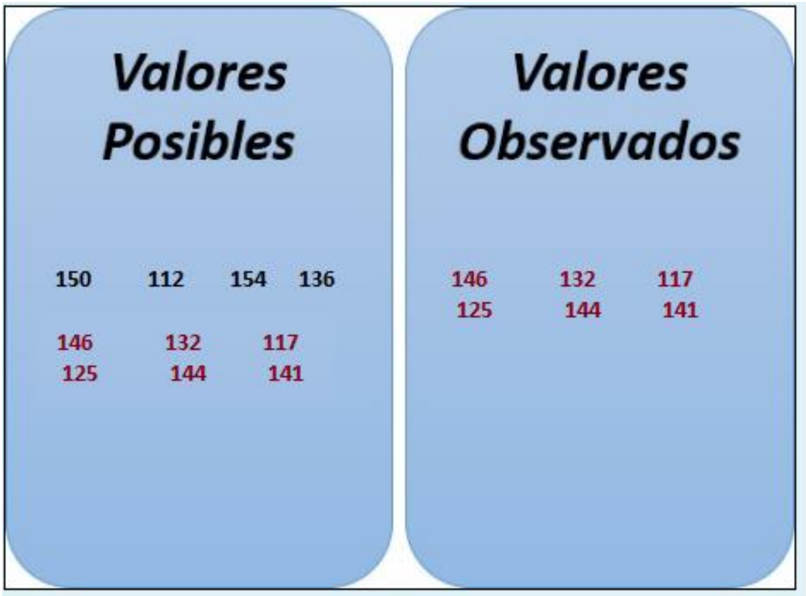
Los valores observados generalmente no incluyen a todos los valores posibles.

Existen otros valores posibles para la variable salario, pero no se presentaron al encuestar a los 50 empleados.



Respuesta:

153	**123**	**129**	**132**	**147**	**138**	**137**	**134**	**131**	**147**
134	**148**	**125**	**139**	**146**	**145**	**148**	**135**	**152**	**128**
138	**138**	**122**	**146**	**137**	**151**	**145**	**124**	**132**	**138**
137	**146**	**138**	**146**	**140**	**137**	**129**	**126**	**117**	**136**
128	**127**	**146**	**144**	**137**	**128**	**142**	**141**	**136**	**138**



Ejercicio N° 14:

Los siguientes valores corresponden al grosor de 50 piezas:

0,26	0,23	0,79	0,28	0,22	0,26	0,37	0,28	0,41	0,35
0,36	0,24	0,73	0,28	0,30	0,48	0,38	0,28	0,38	0,23
0,33	0,27	0,66	0,27	0,32	0,25	0,36	0,31	0,59	0,33
0,25	0,23	0,28	0,29	0,36	0,23	0,29	0,65	0,58	0,23
0,27	1,17	0,29	0,28	0,33	0,23	0,30	0,54	0,59	0,30

Señalar la tabla que agrupa correctamente los valores observados:

a) **Es correcta, abarca un total de 50 valores observados y dichos valores están comprendidos entre el límite inferior del primer intervalo y el límite superior del último intervalo**

LI	LS	MC	FA	FR	FAA	FRA
[0.19	0.39)	0.29	39	0.78	39	0.78
[0.39	0.59)	0.49	4	0.08	43	0.86
[0.59	0.79)	0.69	5	0.10	48	0.96
[0.79	0.99)	0.89	1	0.02	49	0.98
[0.99	1.19]	1.09	1	0.02	50	1.00

b) **El total de observaciones no es correcto (n=40 no es correcto) ya que hay un total de 50 observaciones, es decir que n=50**

LI	LS	MC	FA	FR	FAA	FRA
[0.19	0.39)	0.29	32	0.80	32	0.80
[0.39	0.59)	0.49	2	0.05	34	0.85
[0.59	0.79)	0.69	5	0.13	39	0.98
[0.79	0.99)	0.89	0	0.00	39	0.98
[0.99	1.19]	1.09	1	0.03	40	1.00

c) **El límite inferior del primer intervalo no es correcto. La tabla no abarca la totalidad de las observaciones debido a que los valores 0,22 y 0,23 quedan excluidos de la tabla**

LI		LS		MC	FA	FR	FAA	FRA
[	$\boxed{0.24}$	0.44	)	0.34	40	0.80	40	0.80
[	0.44	0.64	)	0.54	5	0.10	45	0.90
[	0.64	0.84	)	0.74	4	0.08	49	0.98
[	0.84	1.04	)	0.94	0	0.00	49	0.98
[	1.04	1.24	]	1.14	1	0.02	50	1.00

d) **El límite superior del último intervalo no es correcto. La tabla no abarca la totalidad de las observaciones debido a que el valor 1,17 queda excluido de la tabla**

LI		LS		MC	FA	FR	FAA	FRA
[	0.10	0.30	)	0.20	27	0.54	27	0.54
[	0.30	0.50	)	0.40	14	0.28	41	0.82
[	0.50	0.70	)	0.60	6	0.12	47	0.94
[	0.70	0.90	)	0.80	2	0.04	49	0.98
[	0.90	$\boxed{1.10}$	]	1.00	1	0.02	50	1.00

Ejercicio N° 15:

Gastos en cierto rubro de un grupo de familias que fueron seleccionadas para ser encuestadas:

351.8	266.6	344.8	381.2	258.4	393.8	344.8
395.6	311.2	381.4	345	285.2	256.8	358.8
258.6	314.4	215.6	257	335	262.2	384.6
262	344.6	260.2	390	392.4	271.8	389.8
265.8	211	331.2	233.4	393.4	311.6	393.6

Calcular:

Media

Mediana

Primer Cuartil

Tercer Cuartil

Varianza

Desviación Estándar

Coeficiente de Variación:

Coeficiente de Asimetría:

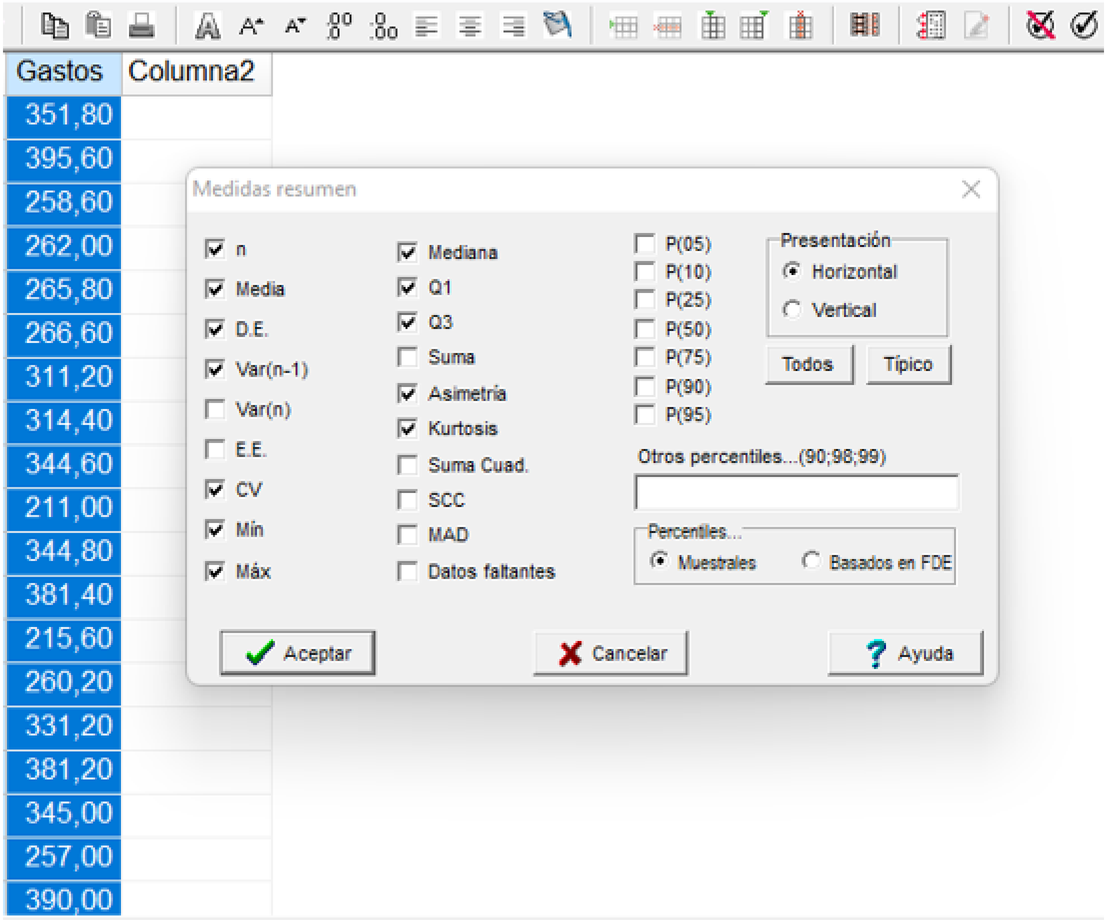
Mínimo:

Máximo:

A
Gastos
351.8
395.6
258.6
262
265.8
266.6
311.2
314.4
344.6
211
344.8
381.4
215.6
260.2
331.2
381.2
345
257
390
233.4
258.4
285.2
335
392.4
393.4
393.8
256.8
262.2
271.8

Infostat:

Estadísticas ___ Medidas Resumen



Medidas resumen

Resumen	Gastos
n	35,00
Media	318,67
D.E.	59,29
Var(n-1)	3515,47
CV	18,61
Mín	211,00
Máx	395,60
Mediana	331,20
Q1	262,00
Q3	381,40
Asimetría	-0,16
Kurtosis	-1,35

Var(n-1): Varianza muestral

Var: Varianza Poblacional

Coeficiente de curtosis (Kurtosis en Infostat)

Menor que 0 $\Longrightarrow$ Distribución Platicúrtica.

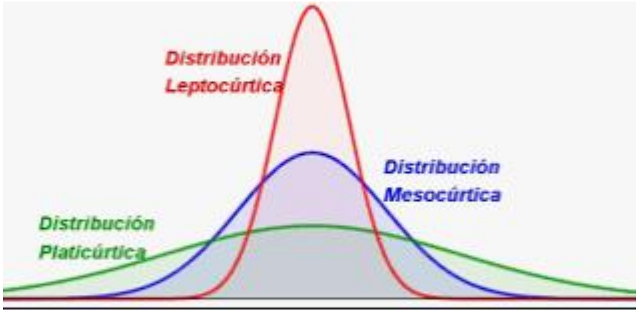
La distribución presenta un reducido grado de concentración con respecto a los valores centrales de la variable.

Igual que 0 $\Rightarrow$ Distribución Mesocúrtica

La distribución presenta un grado medio de concentración con respecto a los valores centrales de la variable.

Mayor que 0 $\Rightarrow$ Distribución Leptocúrtica

La distribución presenta un alto grado de concentración con respecto a los valores centrales de la variable.



Coeficiente de Curtosis= -1,35 Distribución Platicúrtica.

Medidas de Asimetría

Coeficiente de Asimetría o Coeficiente de Pearson

CP= $3.\dfrac{\textbf{Media}-\textbf{Mediana}}{\textbf{Desviación Standar}}$

(Media= 318,67) < (Mediana= 331,20)

Coeficiente de Asimetría o Coeficiente de Pearson= -0,16

(Media < Mediana) $\Rightarrow$ **CP < 0** $\Rightarrow$ **Distribución Asimétrica Izquierda**

Las mayores frecuencias corresponden a los valores de la variable más altos.

Las menores frecuencias corresponden a los valores de la variable más bajos.

Predominan los valores de la variable que son más altos.

Predominan los gastos más altos.