

Resumen de los contenidos básicos

Tema 1. Distribuciones de frecuencias unidimensionales



Carmen Trueba Salas
Lorena Remuzgo Pérez

DEPARTAMENTO DE ECONOMÍA

Este tema se publica bajo Licencia:
[Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



Contenidos

- 1.1 Conceptos básicos
- 1.2 Clasificación de variables
- 1.3 Distribución de frecuencias
- 1.4 Representaciones gráficas

1.1 Conceptos básicos

Población

Conjunto de individuos o elementos que tienen la propiedad o característica que se desea estudiar.

Muestra

Subconjunto representativo de la población.

Individuos (o unidades)

Elementos de la población que poseen la característica que se desea estudiar.

Variable estadística

Característica de los individuos de la población.

Datos (u observaciones)

Información que proporciona la observación de la variable en los individuos de la población.

1.2 Clasificación de variables

Tipo de datos (naturaleza)

Variables cualitativas o atributos

Los datos no son numéricos. Cada observación distinta de una variable cualitativa se denomina **modalidad o categoría**.

Variables cuantitativas

Los datos son numéricos. Cada observación distinta de una variable cuantitativa se denomina **valor**.

- **Discretas**

Los valores son números enteros.

- **Continuas**

Pueden tomar cualquier valor dentro de un intervalo.

1.2 Clasificación de variables

Tipo de escala

Escala nominal

Sus modalidades son mutuamente excluyentes y no es posible establecer ningún orden ni jerarquía entre ellas .

Escala ordinal

Sus modalidades son mutuamente excluyentes y sí es posible establecer algún orden o jerarquía entre ellas.

Escala de proporción (razón o cociente)

Un valor igual a cero indica la ausencia de la característica que se desea estudiar. Tiene sentido establecer proporciones entre sus valores.

Escala de intervalos

Un valor igual a cero no indica la ausencia de la característica que se desea estudiar. No tiene sentido establecer proporciones entre sus valores.

1.2 Clasificación de variables

Referencia temporal

Serie temporal

Las observaciones de la variable se recogen en diferentes periodos de tiempo.

Corte transversal o *cross-section*

Las observaciones de la variable hacen referencia a distintos individuos (o elementos) en un mismo periodo temporal.

Datos de panel

Se combinan referencias temporales y transversales.

1.3 Distribución de frecuencias

Permite presentar el conjunto de observaciones de una variable de manera ordenada:

- Distribución de frecuencias de una variable cualitativa
- Distribución de frecuencias de una variable cuantitativa (no agrupada en intervalos)
- Distribución de frecuencias de una variable cuantitativa (agrupada en intervalos)

1.3 Distribución de frecuencias

Elementos de la distribución de frecuencias de una variable cualitativa

Modalidades (a_i)	n_i	N_i	f_i	F_i
a_1	n_1	$N_1 = n_1$	$f_1 = n_1/N$	$F_1 = f_1 = N_1/N$
a_2	n_2	$N_2 = n_1 + n_2 = N_1 + n_2$	$f_2 = n_2/N$	$F_2 = f_1 + f_2 = F_1 + f_2 = N_2/N$
\vdots	\vdots	\vdots	\vdots	\vdots
a_i	n_i	$N_i = n_1 + \dots + n_i = N_{i-1} + n_i$	$f_i = n_i/N$	$F_i = f_1 + \dots + f_i = F_{i-1} + f_i = N_i/N$
\vdots	\vdots	\vdots	\vdots	\vdots
a_k	n_k	$N_k = n_1 + \dots + n_k = N_{k-1} + n_k = N$	$f_k = n_k/N$	$F_k = f_1 + \dots + f_k = F_{k-1} + f_k = N_k/N = 1$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$	

La distribución de frecuencias del atributo A viene dada por el conjunto de modalidades junto con sus correspondientes frecuencias absolutas (a_i, n_i) o relativas (a_i, f_i).

1.3 Distribución de frecuencias

Elementos de la distribución de frecuencias de una variable cuantitativa (no agrupada en intervalos)

Valores (x_i)	n_i	N_i	f_i	F_i
x_1	n_1	$N_1 = n_1$	$f_1 = n_1/N$	$F_1 = f_1 = N_1/N$
x_2	n_2	$N_2 = n_1 + n_2 = N_1 + n_2$	$f_2 = n_2/N$	$F_2 = f_1 + f_2 = F_1 + f_2 = N_2/N$
\vdots	\vdots	\vdots	\vdots	\vdots
x_i	n_i	$N_i = n_1 + \dots + n_i = N_{i-1} + n_i$	$f_i = n_i/N$	$F_i = f_1 + \dots + f_i = F_{i-1} + f_i = N_i/N$
\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_k	$N_k = n_1 + \dots + n_k = N_{k-1} + n_k = N$	$f_k = n_k/N$	$F_k = f_1 + \dots + f_k = F_{k-1} + f_k = N_k/N = 1$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$	

La distribución de frecuencias de la variable X viene dada por el conjunto de valores junto con sus correspondientes frecuencias absolutas (x_i, n_i) o relativas (x_i, f_i).

1.3 Distribución de frecuencias

Elementos de la distribución de frecuencias de una variable cuantitativa (agrupada en intervalos)

Intervalos ($L_{i-1} - L_i$)	n_i	N_i	f_i	F_i	Marcas de clase (x_i)	c_i	d_i
$L_0 - L_1$	n_1	N_1	f_1	F_1	$x_1 = (L_0 + L_1)/2$	$c_1 = L_1 - L_0$	$d_1 = n_1/c_1$
$L_1 - L_2$	n_2	N_2	f_2	F_2	$x_2 = (L_1 + L_2)/2$	$c_2 = L_2 - L_1$	$d_2 = n_2/c_2$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$L_{i-1} - L_i$	n_i	N_i	f_i	F_i	$x_i = (L_{i-1} + L_i)/2$	$c_i = L_i - L_{i-1}$	$d_i = n_i/c_i$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$L_{k-1} - L_k$	n_k	$N_k = N$	f_k	$F_k = 1$	$x_k = (L_{k-1} + L_k)/2$	$c_k = L_k - L_{k-1}$	$d_k = n_k/c_k$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$				

La distribución de frecuencias de la variable X viene dada por el conjunto de intervalos en los que se agrupan los valores junto con sus correspondientes frecuencias absolutas ($L_{i-1} - L_i, n_i$) o relativas ($L_{i-1} - L_i, f_i$).

1.4 Representaciones gráficas

Variables cualitativas

Diagrama de barras

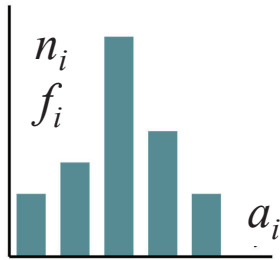
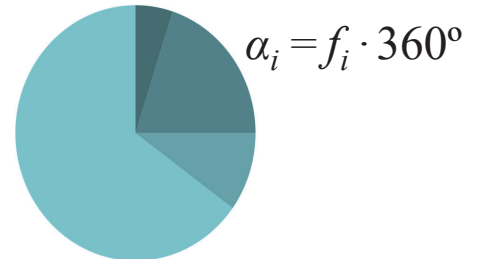


Diagrama de sectores



1.4 Representaciones gráficas

Variables cuantitativas

Diagrama de barras

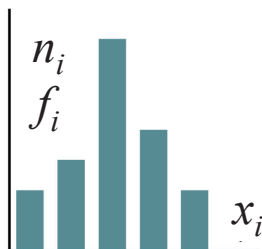
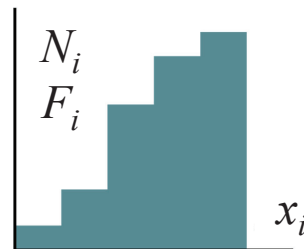
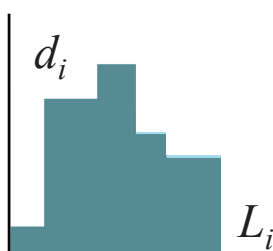


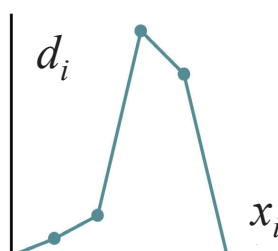
Diagrama de escalera



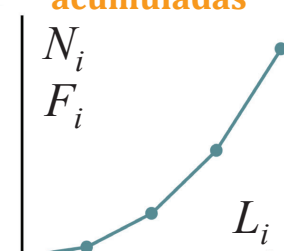
Histograma



Polígono de frecuencias



Polígono de frecuencias acumuladas



Datos no agrupados

Datos agrupados