

Tema 1

Estadística descriptiva:

Tipos de variables estadísticas,
tablas y sus gráficos

Curso 2017/18
Grado en biología sanitaria
Departamento de Física y Matemáticas
Marcos Marv Ruiz



Estadística descriptiva

- Tipos de datos (variables estadísticas)
- Representación de datos (análisis exploratorio)
- Resumir conjuntos de datos de números
 - Medidas de centralización
 - Medidas de dispersión
 - Medidas de posición

TIPOS DE VARIABLES ESTADÍSTICAS

	Sex	Class	Age	HR	Temp
1	Male	2	24	57	35.2
2	Male	2	24	106	35.4
3	Female	3	68	66	35.4
4	Female	3	68	57	35.5
5	Male	1	18	107	35.5
6	Male	2	68	75	35.5

Característica
que deseamos
estudiar en la
población



**VARIABLE
ESTADISTICA**

Cualitativas
Niveles

Nominal: no implica orden

Ordenada: no tiene sentido sumar,...

Cuantitativas

Discretas: valores aislados

Continuas: todos los valores entre
Cada par de datos

Sección 1.1 del libro

Un poco de notación: ¡No hay consenso!

Variable de estudio: X , (cuando hay más: X_1, X_2, \dots, X_k Y, Z).

- Tamaño de la muestra: n (a veces se usa N).
- Dada una muestra de n observaciones de la variable X ,

$$x_1, x_2, \dots, x_n \quad X = \{x_1, x_2, \dots, x_n\}$$

	Sex	Class	Age	HR	Temp
1	Male	2	24	57	35.2
2	Male	2	24	106	35.4
3	Female	3	68	66	35.4
4	Female	3	68	57	35.5
5	Male	1	18	107	35.5
6	Male	2	68	75	35.5

Ejemplo:

Var. cualitativa nominal $X = \text{“Género”}$

$$x_1 = \text{Male}, x_2 = \text{Male}, x_3 = \text{Female}, \dots$$

Var. cualitativa ordenada $X = \text{“Calificación”}$

$$x_1 = 2, x_2 = 2, x_3 = 3, \dots$$

Var. cuantitativa discreta $X = \text{“Frecuencia cardíaca”}$

$$x_1 = 57, x_2 = 106, x_3 = 66, \dots$$

Var. cuantitativa continua $X = \text{“Temperatura”}$

$$x_1 = 35.2, x_2 = 35.4, x_3 = 35.4, \dots$$

Frecuencias absolutas

Frecuencia absoluta de $X = x$: nº de veces que se repite el valor x en la muestra se suele denotar por f_i (hay quien usa F_i)

Ejemplo: En una muestra de 153 personas hay 68 mujeres y 32 hombres.
Variable cualitativa nominal $X = \text{“género”}$,

Valores:

$$x_1 = \text{mujer}, \quad x_2 = \text{hombre}$$

Frecuencias:

$$f_1 = 68, \quad f_2 = 32$$

La **tabla de frecuencias absolutas** recoge los valores que toma la variable junto con su frecuencia absoluta en la muestra

Género	f_i
Mujer	80
Hombre	73

Estado:	1	2	3
Frecuencia:	51	54	48

Frecuencias absolutas

Si los valores no se repiten, o sólo lo hacen unos pocos, no tiene sentido hacer tablas (recuentos) para resumirlos

Solución: AGRUPAR en clases

Ejemplo: Temperatura (C) de 153 personas

35.2 35.4 35.4 35.5 35.5 35.5 35.6 35.7 35.8 35.8 35.8 35.9 36.0 36.0 36.0, 36.1 36.1 36.1 36.2 36.2 36.2 36.3 ,
36.3 36.4 36.4 36.4 36.4 36.5 36.5 36.5, 36.5 36.6 36.6 36.6 36.7 36.7 36.7 36.7 36.7 36.7 36.7 36.7 36.7 36.7,
36.8, 36.8 36.8 36.8 36.9 36.9 36.9 36.9 37.0 37.0 37.0 37.0 37.0 37.1 37.1 37.1, 37.1 37.1 37.1 37.1 37.2 37.2,
37.2 37.3 37.3 37.3 37.3 37.3 37.1 37.1 37.1, 37.1 37.1 37.1 37.1 37.2 37.2 37.2 37.3 37.3 37.3 37.3 37.3 37.3,
37.3 37.4, 37.4 37.5 37.5 37.5 37.5 37.5 37.5 37.5 37.6 37.6 37.7 37.7 37.3 37.3 37.4, 37.4 37.5 37.5 37.5 37.5,
37.5 37.5 37.5 37.6 37.6 37.7 37.7 37.7 37.8 37.8, 37.8 37.8 37.9 37.9 38.1 38.1 38.7 39.5 41.0 37.7 37.8 37.8,
37.8 37.8 37.9, 37.9 38.1 38.1 38.7 39.5 41.0 37.7 37.8 37.8 37.8 37.8 37.9 37.9 38.1 38.1, 39.5 41.0 49.0

Frecuencias absolutas

35.2	35.4	35.5	35.6	35.7	35.8	35.9	36	36.1	36.2	36.3	36.4	36.5	36.6	36.7	36.8
1	2	3	1	1	3	1	3	3	3	2	4	4	3	10	4
36.9	37	37.1	37.2	37.3	37.4	37.5	37.6	37.7	37.8	37.9	38.1	38.7	39.5	41	49
4	5	14	6	14	4	14	4	7	12	6	6	2	3	3	1

Agrupados en 7 clases

[35.2,37.2]	(37.2,39.1]	(39.1,41.1]	(41.1,43.1]	(43.1,45.1]	(45.1,47]	(47,49]
71	75	6	0	0	0	1

Frecuencias relativas

La **frecuencia relativa** del valor $X = x$ en una muestra es el tanto por uno del valor x en la muestra:

$$f'_i = f_i / N$$

donde N es el tamaño de la muestra → **tabla de frecuencias relativas**

Ejemplo: Sobre los ejemplos anteriores

Género	f_i	f'_i
Mujer	80	0.52
Hombre	73	0.48

Absolutas

[35.2,37.2]	(37.2,39.1]	(39.1,41.1]	(41.1,43.1]	(43.1,45.1]	(45.1,47]	(47,49]
71	75	6	0	0	0	1

Relativas

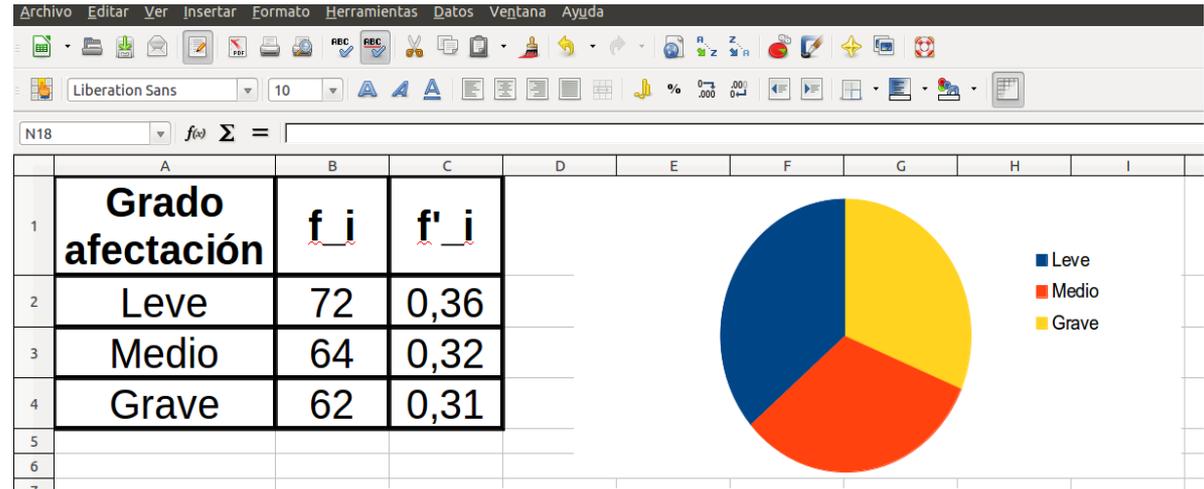
[35.2,37.2]	(37.2,39.1]	(39.1,41.1]	(41.1,43.1]	(43.1,45.1]	(45.1,47]	(47,49]
0.464052288	0.490196078	0.039215686	0.0	0.0	0.0	0.006535948

Diagramas de sectores

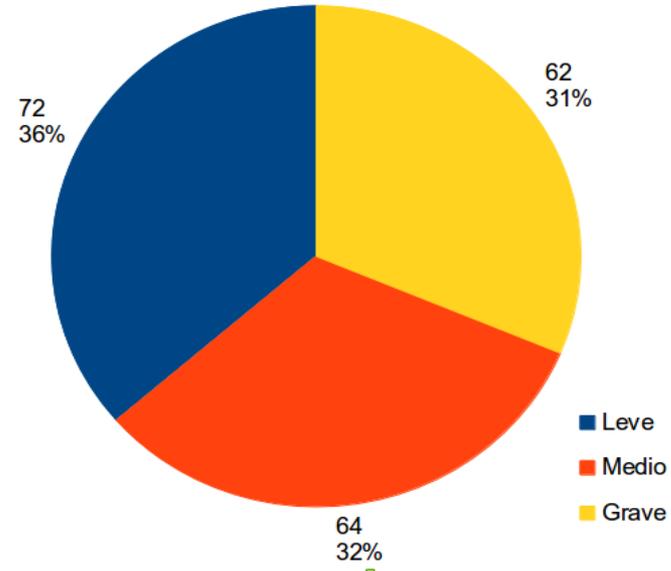
Se utilizan cuando queremos mostrar frecuencias (o porcentajes, recuentos,...) de variables cualitativas (categóricas).

Ejemplo: en una planta de un hospital hay 198 pacientes de distinta gravedad, como indica la tabla

¿Ves algo raro en la tabla?



- Se **desaconseja** su uso porque no siempre se visualizan claramente las diferencias.
- Si las tienes que usar, etiqueta cada sector con su valor



Diagramas con barras

Se utilizan para representar la frecuencia de los valores de una variable

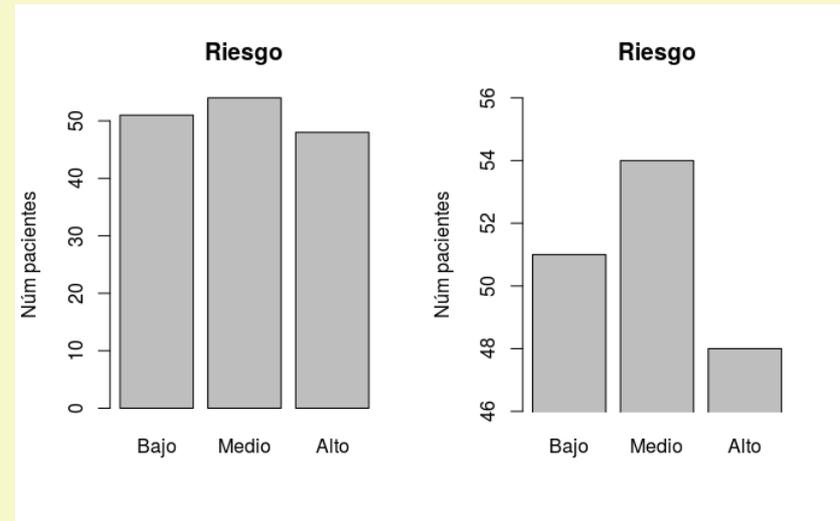
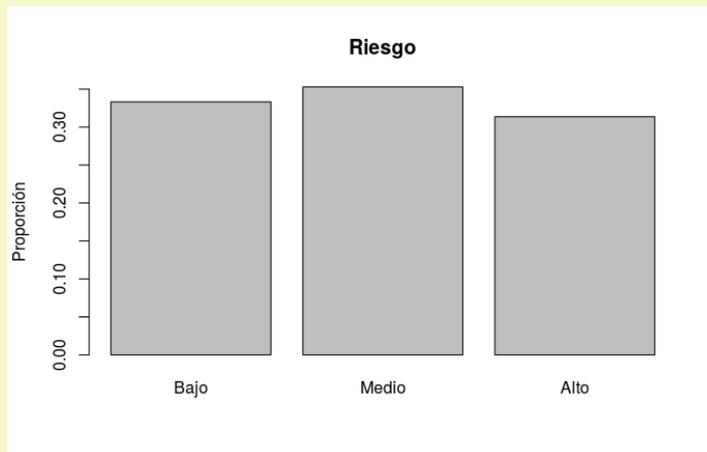
Diagrama de barras

Variables cualitativas, cuantitativas no agrupadas

- * Una barra para cada nivel (valor) de la categoría
- * La altura de cada barra es la frecuencia de ese nivel
- * Las barras (en general) no se tocan

Ejemplo: en un hospital hay 100 pacientes de distinta gravedad, como en la tabla

Estado:	Leve	Mod	Grave
Frec abs:	51	54	48
Frec rel:	0.33	0.35	0.32



Diagramas con barras

Se utilizan para representar la frecuencia de los valores de una variable

Histogramas

Variables cuantitativas agrupadas en clases (intervalos)

- * Los intervalos cubren el recorrido de la variable
- * Cada intervalo es la base de una barra: las barras se tocan
- * El área de cada barra es proporcional a la frecuencia de la clase

$$\text{altura} = \frac{\text{frecuencia}}{\text{base}} = \frac{\text{frecuencia del intervalo}}{\text{longitud del intervalo}}$$

Diagramas con barras

Se utilizan para representar la frecuencia de los valores de una variable

Histogramas

Variables cuantitativas agrupadas en clases (intervalos)

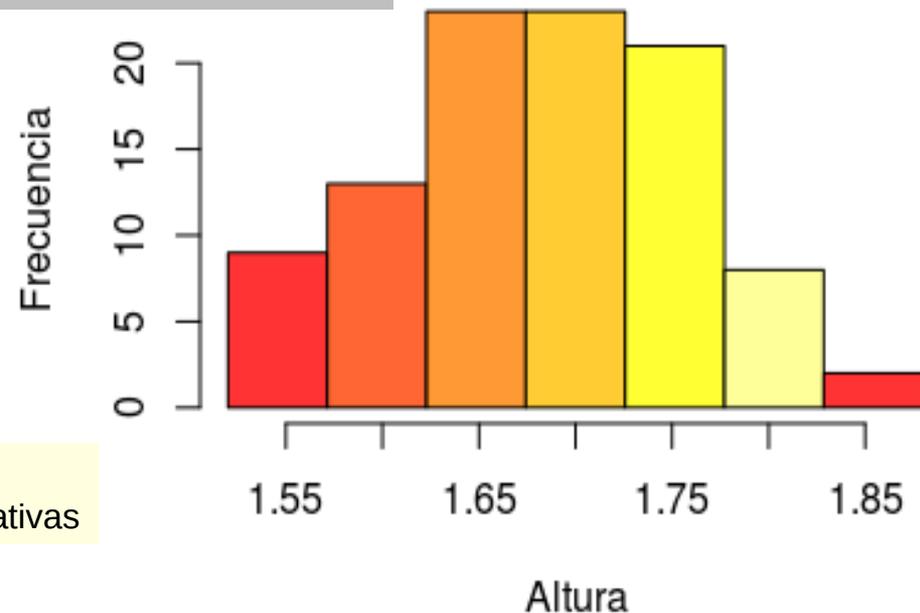
- * Los intervalos cubren el recorrido de la variable
- * Cada intervalo es la base de una barra: las barras se tocan
- * El área de cada barra es proporcional a la frecuencia de la clase

$$\text{altura} = \frac{\text{frecuencia}}{\text{base}} = \frac{\text{frecuencia del intervalo}}{\text{longitud del intervalo}}$$

Ejemplo: temperatura

[35.2, 35.7]	(35.7, 36.2]	(36.2, 36.7]	(36.7, 37.3]
8	13	23	23
(37.3,37.8]	(37.8,38.3]	(38.3,38.8]	
21	8	2	

Clases de la misma anchura -> eje vertical
frecuencias absolutas o relativas

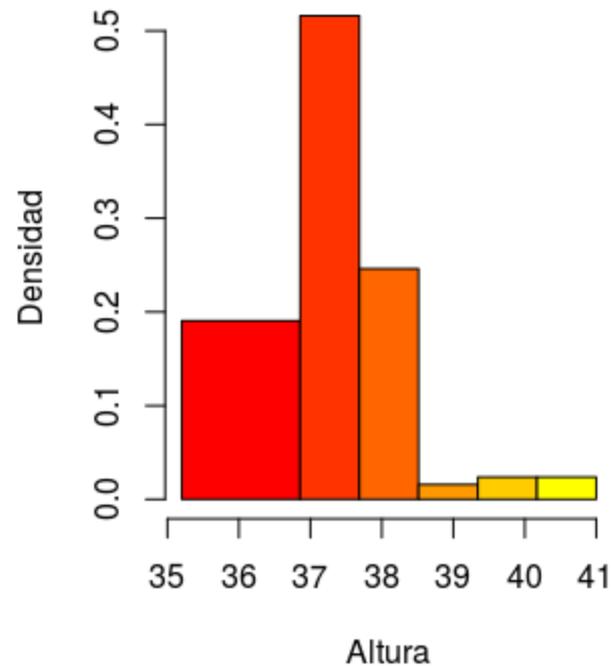
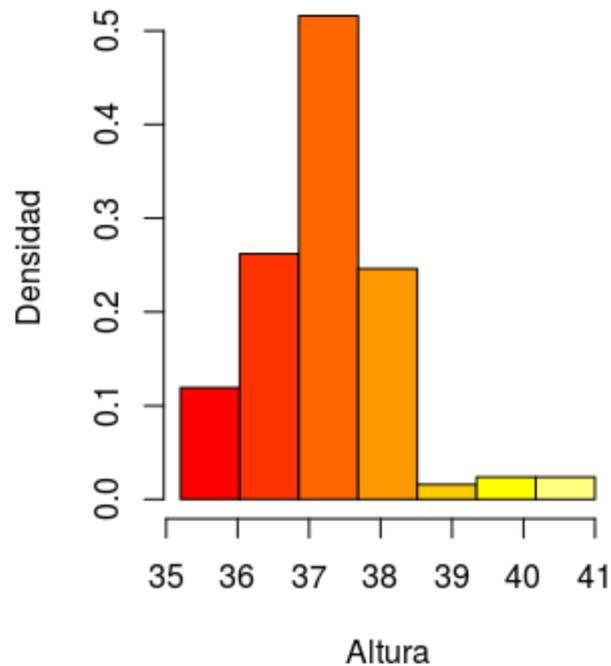


Diagramas con barras

Se utilizan para representar la frecuencia de los valores de una variable

Histogramas

Ejemplo: Clases de distinta anchura -> eje vertical frecuencias absolutas o relativas



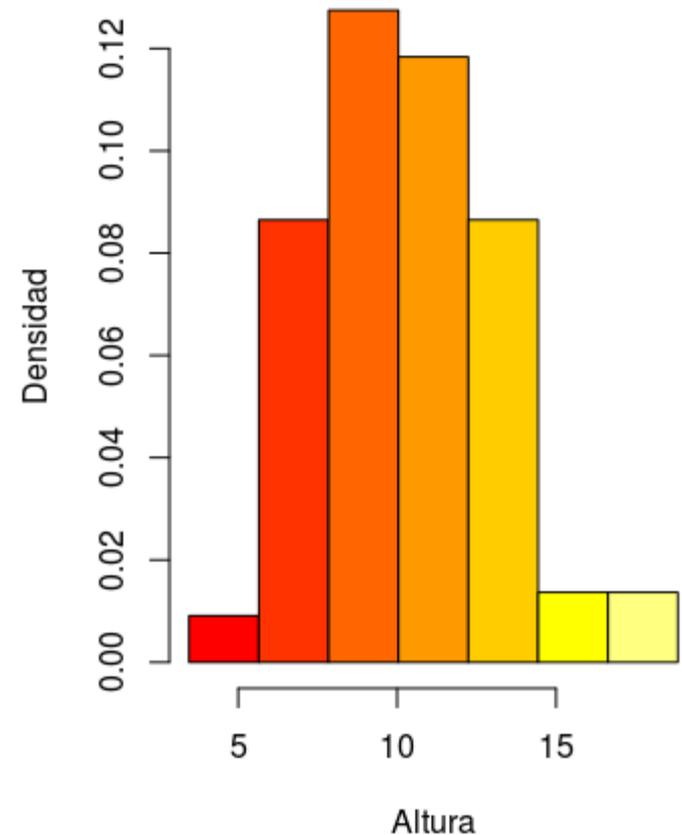
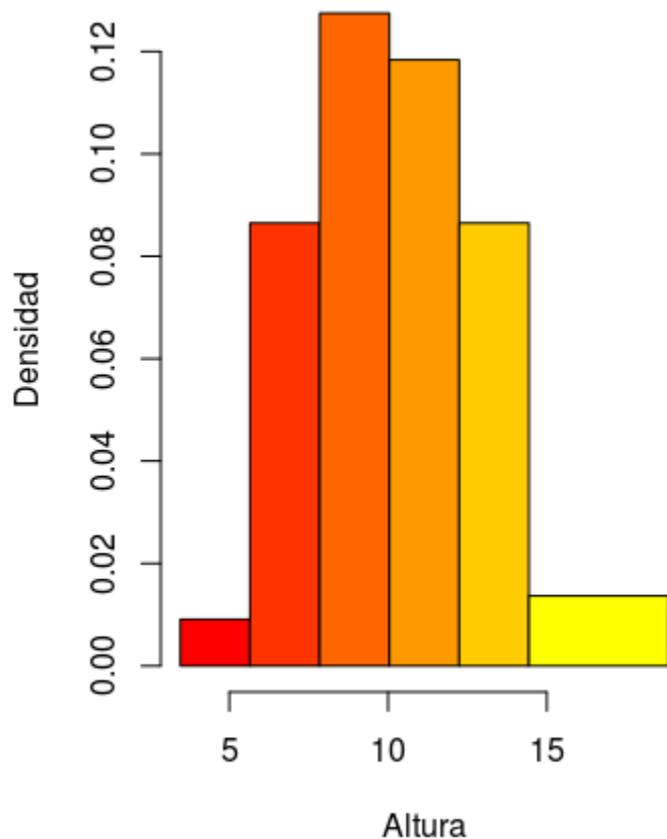
Observa el cambio de unidades en los ejes: el área total es 1

Diagramas con barras

Se utilizan para representar la frecuencia de los valores de una variable

Histogramas

Ejemplo: con otros datos, si unimos las dos últimas clases



La barra no es más alta; mantiene el área

Variables continuas, variables agrupadas en clases e histogramas

Qué hemos visto

Marco general

Tipos de variables

Variables discretas:

- Frecuencias absoluta y relativa. Tablas.
- Diagramas de barras y sectores.

Variables continuas:

- Agrupar en clases.
- Histogramas.

Leer sección 1.3 del libro: precisión y exactitud (imprescindible para los problemas)
Practica con el primer cuestionario