

TEMA 2. Tipos de datos y representaciones

Alicia Nieto Reyes

BIOESTADÍSTICA

Tipos de datos

Tenemos tres estudiantes del grado de enfermería

Variable: característica que puede diferir de un individuo, o entidad, a otro

Ejemplos de variables: altura, temperatura en la calle, estado de ánimo, ...

	María (individuo 1)	Pedro (individuo 2)	Ramón (individuo 3)
altura	1,69	1,94	1,75
hora de dormir	0:30	22:15	1
sexo	2	1	1
grupo de prácticas	D	A	C
clase social	3	2	2
nº de leucocitos por mm ³	5.000	9.000	8.000

Leyenda:

- sexo: masculino=1, femenino=2
- clase social: baja=1, media=2, alta=3

Clasificación de variables

	María	Pedro	Ramón
altura cuantitativa	1,69	1,94	1,75
hora de dormir cuantitativa	0:30	22:15	1
sexo cualitativa	2	1	1
grupo de prácticas cualitativa	D	A	C
clase social cualitativa	3	2	2
nº de leucocitos por mm ³ cuantitativa	5.000	9.000	8.000

Leyenda:

- sexo: masculino=1, femenino=2
- clase social: baja=1, media=2, alta=3

Clasificación de variables cuantitativas

	María	Pedro	Ramón
altura cuantitativa escala racional continua	1,69	1,94	1,75
hora de dormir cuantitativa escala de intervalos continua	0:30	22:15	1
nº de leucocitos por mm ³ cuantitativa escala racional discreta	5.000	9.000	8.000

Clasificación de variables cualitativas

	María	Pedro	Ramón
sexo cualitativa escala nominal	2	1	1
grupo de prácticas cualitativa escala nominal	D	A	C
clase social cualitativa escala ordinal	3	2	2

Leyenda:

- sexo: masculino=1, femenino=2
- clase social: baja=1, media=2, alta=3

Clasificación de variables

Nota que las variables

- cuantitativas pueden pasar a ser cualitativas; pero no al revés
- con escala ordinal pueden pasar a escala nominal; pero no al revés

	María	Pedro	Ramón
altura cuantitativa escala racional continua	1,69	1,94	1,75
altura cualitativa escala ordinal	2	3	2
clase social cualitativa escala ordinal	3	2	2
clase social cualitativa escala nominal	C	C	C

Leyenda:

- altura: baja=1, media=2, alta=3
- clase social: baja=1, media=2, alta=3
- clase social: con recursos=C, sin recursos=S

Variables:

- **Cuantitativas** miden cantidades
 - **escala racional** el cero tiene un significado físico
 - **escala de intervalos** el cero es arbitrario
 - **continuas** pueden tomar, de forma teórica, todos los valores intermedios
 - **discretas** no pueden tomar todos los valores intermedios
- **Cualitativas** miden cualidades
 - **escala ordinal** admiten una ordenación entre sus valores
 - **escala nominal** no admiten ordenación

Representación de datos

Datos de velocidad de una reacción enzimática (Treolar 1974)

- **concentración del sustrato:** en ppm (Partes por millón, es una unidad de medida que se refiere a los mg que hay en un kg de disolución)
- **velocidad:** en counts/min²; mide cuantos átomos de un material radioactivo se desintegran por min²
- **estado:** si se ha tratado la encima con Puromycin o no

individuos	concentración	velocidad	estado
1	0.02	76	tratado
2	0.02	47	tratado
3	0.06	97	tratado
4	0.06	107	tratado
5	0.11	123	tratado
6	0.11	139	tratado
7	0.22	159	tratado
8	0.22	152	tratado
9	0.56	191	tratado

Representación de datos

Datos de velocidad de una reacción enzimática (Treolar 1974)

individuos	concentración	velocidad	estado
10	0.56	201	tratado
11	1.10	207	tratado
12	1.10	200	tratado
13	0.02	67	no tratado
14	0.02	51	no tratado
15	0.06	84	no tratado
16	0.06	86	no tratado
17	0.11	98	no tratado
18	0.11	115	no tratado
19	0.22	131	no tratado
20	0.22	124	no tratado
21	0.56	144	no tratado
22	0.56	158	no tratado
23	1.10	160	no tratado

Datos de velocidad de una reacción encimática

estado

estado: si se ha tratado la encima con Puromycin o no
variable **cuantitativa** nominal

Tabla de frecuencias:

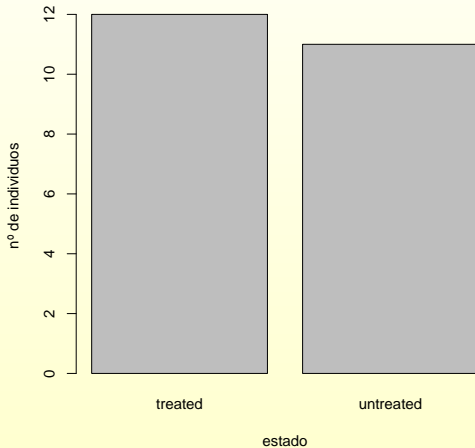
estado	nº de individuos
tratados	12
no tratados	11

Las variables **cuantitativas** y las **cuantitativas discretas** se representan mediante un diagrama de barras

Datos de velocidad de una reacción encimática

estado

Diagrama de barras



Datos de velocidad de una reacción encimática

Recomendación para la elección del número de clases en variables cuantitativas

Regla de Sturges:

- si el número de datos pertenezca al intervalo $[2^3, 2^4)$, el número de clases sugerido es 4
Nota que $2^3 = 8$ y $2^4 = 16$. Que el número de datos pertenezca al intervalo $[8, 16)$ es que puede ser 8 ó 9 ó 10 ó 11 ó 12 ó 13 ó 14 ó 15
- si el número de datos pertenezca al intervalo $[2^4, 2^5)$, el número de clases sugerido es 5
- si el número de datos pertenezca al intervalo $[2^5, 2^6)$, el número de clases sugerido es 6
- ...

Como tenemos 23 datos y $2^4 = 16$ y $2^5 = 32$, el número de clases sugerido es 5

Datos de velocidad de una reacción enzimática

concentración del sustrato

concentración del sustrato: en ppm (Partes por millón, es una unidad de medida que se refiere a los mg que hay en un kg de disolución)

variable **cuantitativa continua** racional

Tabla de frecuencias:

límites	punto medio	frecuencia absoluta	frecuencia relativa	porcentaje de observaciones
0 - 0.2	0.1	12	0.522	52.2
0.2 - 0.4	0.3	4	0.174	17.4
0.4 - 0.6	0.5	4	0.174	17.4
0.6 - 0.8	0.7	0	0	0
0.8 - 1	0.9	0	0	0
1 - 1.2	1.1	3	0.13	13

Datos de velocidad de una reacción encimática

concentración del sustrato

Histograma

