

## Estadística y probabilidad

La estadística y la teoría de la probabilidad son dos ramas de la matemática fuertemente relacionadas. Las dos estudian colecciones de datos, con perspectivas ligeramente diferentes, pero complementarias.

### Colecciones de datos

En el mundo real estamos rodeados de datos numéricos, muchas veces en cantidades tan grandes que cuesta hacerse una idea de cuánto. Su procesamiento adecuado es tan importante hoy en día que se ha creado el concepto de *big data* para agrupar todas las técnicas bajo una única denominación. Tanto la estadística como la teoría de la probabilidad tienen un papel determinante en el *big data*.

Las colecciones de datos se mantendrán en ordenadores usando el formato que resulte más adecuado (algún tipo de **base de datos**). Pero para estudiarlos en la enseñanza secundaria se te presentarán escritos uno tras otro, quizá colocados en una tabla para que los puedas ver bien.

### Valores y repeticiones

Cuando estudiamos una colección de datos lo primero que debemos estudiar es qué valores diferentes aparecen en ella. Un mismo valor puede repetirse cualquier número de veces y es importante reconocerlo.

#### Ejemplo 1

En una urbanización hay veinte casas. En cada casa vive una familia. Hacemos una encuesta casa por casa preguntando cuántos hijos hay en cada familia. Vamos anotando los datos obtenidos y al final tenemos una colección con veinte datos. Por ejemplo, podría ser así:

2	3	0	2	1	0	4	3	2	1	1	0	3	3	1	4	2	2	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Observamos que en la colección de datos solo hay cinco valores diferentes (0, 1, 2, 3 y 4) pero que cada uno de ellos aparece más de una vez.

#### Ejemplo 2

Usamos un dado con forma de hexaedro con las caras numeradas del 1 al 6; es decir, el típico dado de muchos juegos de azar. Lanzamos 25 veces el dado y vamos anotando los números que van saliendo. Por ejemplo, podríamos obtener esto:

3	1	2	1	5	4	5	5	6	2	6	3	3	3	6	3	2	4	6	4	1	3	3	5	4
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Observamos que en la colección de datos solo hay seis valores diferentes (1, 2, 3, 4, 5 y 6) pero que cada uno de ellos aparecen más de una vez.

### Diferencia en los dos ejemplos

Los dos ejemplos presentados permiten comenzar a hacerse una idea de la diferencia entre estadística y probabilidad.

- \* En el ejemplo (1) debemos comenzar por tomar los datos de la realidad y, **después** de obtenerlos, los estudiaremos. Este tipo de colecciones se estudian en estadística.
- \* En el ejemplo (2) podríamos estudiar qué resultados esperamos obtener **antes** de realizar los lanzamientos del dado. Este tipo de colecciones se estudian en teoría de la probabilidad.