

Clase 2

Variables estadísticas

2.1 Introducción

Como vimos en la clase anterior, la estadística realiza descripciones e inferencias acerca de las poblaciones y sus características.

Sobre los elementos de la población, o de la muestra, se miden características que recogen información del concepto en estudio.

Como estas características son diferentes en cada elemento se denominan **variables estadísticas**.

Por ejemplo, si se está interesado en averiguar sobre la actividad laboral en un grupo de individuos en Puerto Santa Cruz, se pueden medir variables como nivel de estudio, de satisfacción con su trabajo o de salario en los individuos empleados, o edad y tiempo de inactividad entre los desempleados.

Como podemos notar en el ejemplo anterior, las variables pueden tomar valores numéricos como en el caso de la edad, el tiempo de inactividad o el salario, o bien valores no numéricos como pueden ser el nivel de estudio o de satisfacción con su trabajo.

Estos distintos tipos de variables aleatorias nos llevan a la clasificación que veremos en esta clase.

2.2 Variable estadística

Definición

Una variable estadística es toda característica de los elementos de una población o muestra que se puede estudiar y que se caracteriza por adoptar distintos valores.

Por ejemplo, el color de pelo, las notas de un examen, la cantidad de memoria RAM, la marca de procesador de una PC, la estatura de una persona, las pulgadas de la pantalla del celular, son variables estadísticas.

Otro ejemplo, si tomamos la población de todas las personas que trabajan en informática, algunas variables estadísticas de dicha población serían: la cantidad de horas que pasan frente a la PC, la marca de teclado que usan, la marca de mouse que usa, las pulgadas del monitor que usan, el tipo de silla que usan, el tiempo de descanso que se toman, el nivel de satisfacción con su trabajo, el salario, etc.

Como vemos hay muchísimas variables estadísticas para una misma población o muestra. Básicamente es toda característica factible de ser consultada o medida.

La variedad de tipos de variable estadística nos lleva a su clasificación en cualitativas o cuantitativas; y dentro de las primeras en ordinales o nominales, mientras que las segundas se dividen a su vez en discretas o continuas, tal como se muestra en el siguiente cuadro.

Cualitativa	Ordinal
	Nominal
Cuantitativa	Discreta
	Continua

2.2.1 Variable cualitativa

*Aquella variable que mida un atributo o característica que se pueda clasificar en categorías se le denomina **cualitativa**. Generalmente la identificamos como aquella que se expresa de manera no numérica, mediante texto o letras.*

Son ejemplos de dichas variables cualitativas:

Variable	Valor que puede tomar
El color de ojos de una persona	Azul, verde, marrón, negro, celeste, grisáceo, etc.
El talle de los pantalones	XS, S, M, L, XL, XXL
La calificación de un examen en formato de letras	I, R, B, MB, S, E
El nivel de satisfacción de un cliente	Poco satisfecho, Satisfecho, Muy satisfecho
Los sabores de helado	Dulce de leche, frutilla, chocolate, vainilla, granizado, americana, etc.

En general, si observamos los valores que puede tomar la variable (indicados en azul), una variable cualitativa no toma valores numéricos.

En los ejemplos de la tabla hay una diferencia fundamental entre los tres del medio con respecto al primero y al último.

Dicha diferencia es que las variables correspondientes a el talle de los pantalones tienen un orden, es decir XS es menor que S, que es menor que M y así sucesivamente. Lo mismo con la calificación y con el nivel de satisfacción.

Sin embargo, el color de ojos de una persona no tiene un orden intrínseco, es decir azul no es menor, ni mayor, que verde. Lo mismo para los sabores de helado.

Esto deriva en la subclasificación de las variables cualitativas en ordinales o nominales.

Variable cualitativa ordinal: identifica características no numéricas que se pueden clasificar en grupos que admiten orden. Ejemplos de esto son el resultado de una entrevista (rechazado, normal y sobresaliente) o el premio en una prueba olímpica (oro, plata, bronce).

Variable cualitativa nominal: identifica características no numéricas que se pueden clasificar en grupos que no admiten un orden natural, por ejemplo, la ciudad de nacimiento o el color de los ojos.

2.2.2 Variable cuantitativa

Por supuesto, una de los tipos de variables aleatorias más usadas son aquellas que se expresan mediante valores numéricos.

*Una **variable cuantitativa** es aquella que expresa una característica que se puede medir o contar. Una característica distintiva de una variable cuantitativa es que con un grupo de dichas variables se pueden realizar operaciones aritméticas.*

Son ejemplos de variables cuantitativas:

Variable	Valor que puede tomar
La cantidad de alumnos de un curso.	0, 1, 2, 3, 4, ..., 30, 31, 32, ...
La cantidad de goles que hace un equipo de fútbol en un partido	0, 1, 2, 3, 4, 5, 6, 7, ...
El piso de un edificio	..., -2, -1, 0, 1, 2, 3, 4, 5, ...
La cantidad de litros de agua de un lago	$[0, +\infty)$
La profundidad que desciende un buzo en un lago de 300 m de profundidad.	$[-300 ; 0]$

En los ejemplos anteriores vemos una diferencia entre los tres primeros ejemplos y los dos últimos.

En los tres primeros los valores que pueden adoptar son discretos, es decir, en muchos casos pueden ser infinitos, pero entre dos valores cualesquiera hay una cantidad finita contable de otros valores.

En los dos últimos casos los valores que puede tomar la variable están representados por intervalos de números reales, los cuales tienen la característica de ser continuos; es decir, entre dos valores cualesquiera hay infinitos valores

reales. Notar por ejemplo que el intervalo correspondiente al último ejemplo, $[-300; 0]$, no se extiende hasta infinito, sin embargo, es continuo.

De esta diferenciación se desprende la clasificación de las variables cualitativas en discretas y continuas.

Variable cuantitativa discreta: es aquella que toma un número finito o infinito contable¹ de valores. El número de respuestas correctas en un examen de cinco preguntas $\{0,1,2,3,4,5\}$ o el número de hijos de una persona $\{0,1,2,3,\dots\}$ son ejemplos de estas variables.

Variable cuantitativa continua: es aquella que puede tomar infinitos valores numéricos en cualquier intervalo razonable de la variable. La estatura de un hombre adulto es una variable aleatoria continua. Nótese que existe una infinidad de posibilidades para la estatura en un intervalo arbitrario, por ejemplo, entre los 160 y 161 centímetros. Aunque las personas suelen responder con un número entero cuando se les pregunta por su estatura (en centímetros o en metros), **esta variable se puede medir con la precisión deseada utilizando el instrumento adecuado**, por lo tanto, la consideraremos continua.

Finalmente dejemos la siguiente tabla a modo de resumen:

		Los valores que toma la variable:
Cualitativa	Ordinal	Se pueden clasificar en grupos ordenados
	Nominal	Se pueden clasificar en grupos pero no hay un orden entre ellos
Cuantitativa	Discreta	Son finitos o contables
	Continua	Son infinitos en cualquier intervalo razonable

2.3 Ordenamiento y manejo de la información numérica

Este último apartado no es en realidad un concepto estadístico, pero si una metodología de trabajo que me pareció importante remarcar con el fin de evitar errores en los temas de las siguientes clases.

En general, cuando trabajemos con un conjunto de datos, conviene ordenarlos de menor a mayor y además colocando los valores que se repiten, es decir, si el dato 2 aparece 3 veces, colocarlo dicha cantidad de veces.

Esta manera de trabajar nos facilitará los cálculos que realizaremos, sobre todo aquellos valores de las medidas de posición y de dispersión, así como una correcta tabla de frecuencias y el trabajo con intervalos de clase.

Vamos a mostrar un ejemplo corto de cómo se debería ordenar un conjunto de datos.

¹ En términos matemáticos, un conjunto se dice contable si es posible construir una correspondencia entre dicho conjunto y los números naturales.

Ejemplo

En la siguiente tabla se muestran diferentes tiempos de respuesta, en minutos, que un escuadrón de bomberos demora en estar preparado para salir a atender una llamada:

10,5	11	15	8,5	9	7	18
12	20	8,5	12	7	10	11
19	17,5	21	9	12,5	15	12
15	17	17,5	12,5	11	8	13
10	15,5	18	14	15	8	9,25

Está claro que esta variable es “tiempo de demora en salir” y es de tipo cuantitativa, pues se expresa mediante valores numéricos.

Ahora bien, ¿es de tipo discreta o continua? Bueno, los valores son de tiempo y se podrían medir con la suficiente precisión en caso de ser necesario, adoptando infinitos valores de un intervalo continuo. Por lo tanto, es una **variable cuantitativa continua**.

El primer paso para trabajar con estos 35 datos, es ordenarlos de menor a mayor, incluyendo las repeticiones, de la siguiente manera:

7	7	8	8	8,5	8,5	9	9	9,25	10	10	10,5	11	11	11	12	12	12
---	---	---	---	-----	-----	---	---	------	----	----	------	----	----	----	----	----	----

12,5	12,5	13	14	15	15	15	15	15,5	17	17,5	17,5	18	18	19	20	21
------	------	----	----	----	----	----	----	------	----	------	------	----	----	----	----	----

De esa manera tenemos los 35 datos ordenados de menor a mayor. Notar también que se colocan las repeticiones, es decir, el 7 aparece dos veces, el 8 dos veces, el 11 tres veces, etc.

El ordenamiento se puede realizar también en Excel, colocando todos los datos en forma de una única fila, o columna, y seleccionando ordenar de menor a mayor.

Con esta metodología de trabajo, relativamente simple, lograremos varias cosas:

- Se nos facilitará contar la frecuencia de cada dato.
- Se nos facilitará calcular la media y la moda, porque muchas veces se nos olvida sumar la repetición de algún valor.
- Este ordenamiento es imprescindible para calcular el valor de la mediana.

Además, evitará que nos salteemos algún valor, con lo cual tendremos menos errores en los cálculos que iremos introduciendo a partir de la próxima clase.