

SOFTWARE ESTADÍSTICO PARA EL CÁLCULO DE LA MUESTRA Y LOS TIPOS DE MUESTREO

K, López. B, Sandoval. S, Urquizo.
 Universidad de las Fuerzas Armadas ESPE, Ecuador.
 Ingeniería en Software, Ciencias Exactas.
 [kalopez5, bsandoval, srurquizo] @espe.edu,ec

Resumen Gráfico



Resumen:

Con el fin de conseguir el tamaño de la muestra y seleccionar un tipo de muestreo adecuado se desarrolló un software capaz de realizar los cálculos de manera optimizada e interactiva.

Para ello se utilizó el lenguaje .Net en el IDE proporcionado por Windows, Visual Studio en el cual a partir de ventanas permitirá al usuario navegar por el Software. Para el efecto se utiliza como factor clave para estos cálculos los números aleatorios los cuales representan individuos mediante números aleatorios proporcionados por funciones propias del lenguaje.

Los resultados de muestreo se establecen según la técnica seleccionada estableciendo un muestreo aleatorio, simple, un sistemático y un muestreo estratificado, además de ello se utiliza fórmulas para la deducción del tamaño de la muestra.

Palabras clave: Muestreo, tamaño de la muestra, Interfaz, .NET.

Abstract

In order to achieve the sample size and select a suitable type of sampling, software was developed that was able to perform the calculations in an optimized and interactive way.

For this, the .Net language was used in the IDE provided by Windows, Visual Studio in which from windows it will allow the user to navigate the Software. For this purpose, random numbers are used as the key factor for these calculations, which represent individuals through random numbers provided by language functions.

The sampling results are established according to the selected technique by establishing a random, simple, systematic and stratified sampling, in addition to these formulas for the deduction of the sample size are used.

Keywords: Sampling, sample size, interactive, .NET.

1.0 INTRODUCCIÓN

El presente documento de investigación se ha enfocado en la estadística, la determinación de la muestra que puede corresponder a objetos, cosas u otro material físico, la selección correcta del tipo de muestreo y los elementos fundamentales que intervienen en el mismo.

El elemento puede ser una entidad simple como por ejemplo un ser humano o una entidad compleja algo en plural como una tribu que se denomina unidad [1]. *Brindándonos* un punto de partida y demostrando que la estadística está presente en nuestro entorno, es así como el análisis de la muestra permite realizar, extrapolar o suponer conclusiones a la población con un alto grado de certeza [2], ampliando al mismo tiempo el panorama para estudiar los tipos de muestreo y aplicarlo en el análisis existente.

Existen varios tipos de muestreo, el primero de ellos es el muestreo aleatorio simple (MAS), éste es un muestreo básico y fundamental debido a que, para la selección de una muestra, se puede utilizar la tabla de números aleatorios [3] y de ahí escoger cuantos valores necesitemos en forma aleatoria, asignándole igual probabilidad de elegir a cada elemento [3] y demostrando que es un tipo de muestreo elemental, pero a la vez importante en la toma de decisiones.

Siendo el muestreo aleatorio simple (MAS) una base fundamental para los otros tipos de muestreo se puede conocer en qué consiste el Muestreo Sistemático (MS). El muestreo sistemático tiene como propósito facilitar la selección de elementos cuando la población es un listado de muchos elementos, su cálculo se determina mediante la división del número de elementos de la población por el tamaño de la muestra, obteniendo un valor de k que a posteriori representará el intervalo fijo.

Una vez establecido el intervalo fijo, seleccionamos el punto de inicio en forma aleatoria dentro del primer intervalo y posterior iremos contando cada k -ésimo intervalo hasta llegar a culminar el número de muestras dentro del marco muestral, garantizando de esta manera tener una muestra distribuida acorde a los elementos que proporciona el caso a estudiar.

Se puede ver claramente que el enfoque del muestreo es obtener datos que puedan representar a la población, por lo que existe otro tipo de muestreo llamado muestreo estratificado (ME). [3] El muestreo estratificado contribuye con una mayor amplitud a la hora de obtener una muestra puesto que este tipo de muestro divide a la misma en Estratos que tengan características en común, por ejemplo, la edad, el nivel de estudio, el género, nivel socioeconómico, la nacionalidad, la religión, etc.

Una vez realizada esta división o separación en grupos o estratos de características en común, se calcula el número de muestras de cada grupo y se retoma un muestreo aleatorio simple para la selección de las mismas dentro del grupo. Al final debe coincidir la población total con la suma de las subpoblaciones en cada estrato, así como también debe coincidir el número total de las muestras de cada estrato con el total de las muestras asignadas.

2.0 EXPERIMENTACIÓN

2.1. Microsoft Visual Studio entorno de desarrollo Visual Basic

Para la fase experimental se utilizó el IDE Microsoft Visual Studio del cual se eligió el entorno de desarrollo Visual Basic donde se procedió a programar línea a línea teniendo en cuenta ciertos criterios de programación, además de la previa realización de Mockus con el fin de determinar la interfaz más amigable para el usuario (Docentes y estudiantes).

2.2. Creación y diseño de interfaces de SampligSoft

El diseño de interfaz debe cumplir ciertos requerimientos, pero el fundamental es ser una interfaz amigable para el usuario, se tienen 8 interfaces pensadas para un mejor manejo de los tipos de muestreo aleatorio, las cuales son:

2.2.1. Interfaz de menú principal

Se procedió con el diseño de la interfaz agregando Label para la colocación de títulos, además de PictureBox para colocar imágenes, se programó dos botones, el botón *CALCULO DE MUESTRA* está enlazado a otra interfaz para

seleccionar el tipo de población que se tiene como dato y de igual manera el botón *TIPOS DE MUESTREO* está programada para que nos lleve a otra interfaz para determinar el tipo de muestreo con el que se desea trabajar.



Figura 1. Interfaz menú principal

2.2.2. Interfaz del cálculo de muestras

Se procedió con el diseño de la interfaz agregando Label para la colocación de títulos, se programó dos botones, estos dos botones nos llevan a otra interfaz en la que nos muestra las siguientes opciones:



Figura 2. Interfaz de menú de muestras

2.2.3. Interfaz tamaño de la población conocida

Se procedió con el diseño de la interfaz agregando Label para la colocación de títulos, se programó un solo botón el cual tiene integrado el cálculo de la fórmula mostrada en el PictureBox, solo se debe tener en cuenta que el ingreso de datos en cada uno de los TextBox está validado para no permitir el ingreso de datos no numéricos, se muestra la comprobación con la fórmula [4] y la integración en el software.

$$n = \frac{N \times Z_a^2 \times p \times q}{d^2 \times (N - 1) + Z_a^2 \times p \times q} \quad 1)$$

Donde:

N es el tamaño de la población

Z es el nivel de confianza

p es la probabilidad de éxito

q es la probabilidad de fracaso

d es el error máximo admisible

$$n = \frac{92 \times (1.96^2) \times (0.9) \times 0.1}{0.05^2 \times (92 - 1) + (1.96^2) \times 0.9 \times 0.1}$$

$$n = 55 \text{ aprox.}$$

Poblacion Conocida

Cálculo del tamaño de la muestra conociendo el tamaño de la población

$$n = \frac{N \times Z_a^2 \times p \times q}{d^2 \times (N - 1) + Z_a^2 \times p \times q}$$

En donde, N = tamaño de la población Z = nivel de confianza, P = probabilidad de éxito, o proporción esperada Q = probabilidad de fracaso D = precisión (Error máximo admisible en términos de proporción).

INGRESO DE DATOS

Ingrese el tamaño de la población N :	92
Ingrese el nivel de confianza Z(Dato tabla):	1.96
Ingrese la probabilidad de éxito P :	0.9
Ingrese la probabilidad de fracaso Q :	0.1
Ingrese la precisión (error admisible) D :	0.05

Calcular

El resultado de la muestra es **55**

REGRESAR SamplingSoft

Figura 3. Interfaz tamaño de muestra conocido

2.2.4. Interfaz tamaño de la población desconocida

Cuando la población es infinita o mayor que 10000 se manifiesta que la población es desconocida. Se procede con el diseño de la interfaz agregando Label para la colocación de títulos, se programó un solo botón el cual tiene integrado el cálculo de la fórmula mostrada en el PictureBox, solo se debe tener en cuenta que el ingreso de datos en cada uno de los TextBox está validado para no permitir el ingreso de datos no numéricos, se muestra la comprobación con la fórmula [5] y la integración en el software.

$$n = \frac{Z_a^2 \times p \times q}{d^2} \quad (2)$$

Donde:

Z es el nivel de confianza

p es la probabilidad de éxito

q es la probabilidad de fracaso

d es el error máximo admisible

$$n = \frac{1.96^2 \times 0.9 \times 0.1}{0.05^2}$$

$$n = 138 \text{ aprox.}$$

Figura 4. Interfaz tamaño de muestra desconocido

2.2.5. Interfaz de menú de Tipos de Muestreo

Se procedió con el diseño de la interfaz agregando Label para la colocación de títulos, se programó dos botones, estos dos botones nos llevan a otra interfaz en la cual tenemos las opciones de Muestreo Aleatorio Simple, Muestreo Sistemático y Muestreo Estratificado [6]:



Figura 5. Interfaz de menú de tipos de muestreo

2.2.6. Interfaz de muestreo aleatorio simple

El muestreo aleatorio simple se basa en un conjunto de datos aleatorios de los cuales se seleccionarán muestras al azar. Se procedió con el diseño de la interfaz agregando Label para la colocación de títulos e instrucciones, además de TextBox en donde, primero ingresaremos en valor de nuestro marco muestral y posteriormente una muestra de dicho marco, además se programó dos botones, el botón *Observar Números* que nos muestra el marco muestral y el botón *Generar Números* que nos da la muestra de dicho marco muestral.

En la figura 6 se muestra el ingreso de 8 elementos que conforman el marco muestral y 3 números aleatorios que han sido generados de manera aleatoria.

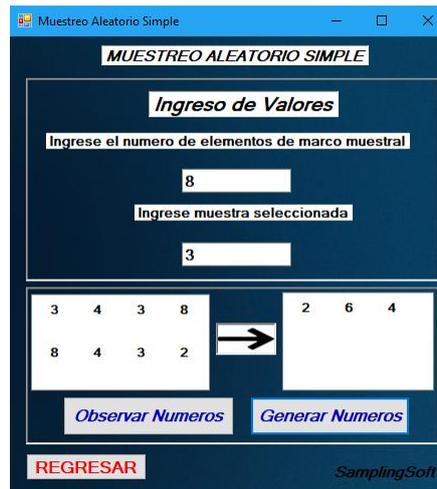


Figura 6. Interfaz de muestreo aleatorio simple

2.2.7. Interfaz de muestreo sistemático

Para la programación de este tipo de muestreo, se procedió con el diseño de la interfaz agregando Label para la colocación de títulos e instrucciones, además de TextBox en donde, primero ingresaremos en valor de nuestro marco muestral y posteriormente una muestra de dicho marco, además se programó dos botones, el botón *Calcular* que nos muestra el valor de cada k -ésimo término que será seleccionado y el botón *Generar Números* que indica la muestra según cada k -ésimo término seleccionado del marco muestral, además se visualiza la comprobación con la fórmula [7] y la integración en el software.

$$k = \frac{N}{n} \quad (3)$$

N es la población

n es la muestra

$$k = \frac{20}{5} = 4$$

Después de la obtención de k , se toman los números de la matriz que se presenta en la figura 8 y a partir del primer término se toman en cuenta los primeros 4 números y de estos se toma una sola muestra, siguiendo este proceso hasta el final del marco muestral.

En el ejemplo mostrado en la figura 7, en el costado izquierdo de la ventana se determina un marco muestral de 15 elementos y una muestra

seleccionada de 3. Se calcula el valor de k que da como resultado 5 y posterior a esto se genera los números aleatorios. En el costado derecho de la ventana se muestra el marco muestral en forma aleatoria y en la parte inferior derecha se muestra el primer quinto elemento elegido en forma aleatoria que es el 5, luego el siguiente quinto elemento que es el 11 y finalmente el siguiente quinto elemento que es el 6.

Figura 7. Interfaz de muestreo sistemático

2.2.8. Interfaz de muestreo estratificado

El muestreo aleatorio estratificado consiste en dividir la población en estratos a partir de ciertas características, fundamentando que entre los estratos es mayor la variabilidad, siendo éste un tipo de muestro que simplifica varios procesos. Es un muestreo que tiene un costo muy bajo y además es rápido de realizar y se utiliza para conocer la opinión de una determinada población.

Para la interfaz se procedió con su diseño agregando Label para la colocación de títulos e instrucciones, además de TextBox en donde, primero ingresaremos un valor de la muestra de dicho marco muestral, después nuestro botón *Generar*, nos genera la muestra aleatoriamente y por consiguiente en el *NumericUp*, seleccionamos la cantidad de estratos que queremos para la muestra, también seleccionaremos el rango de cada uno de los estratos, en el botón *Aceptar* está programada la fórmula de obtención de estratos.

Este muestreo trabaja primeramente eligiendo la cantidad de estratos que serán determinamos por rangos, en donde se da un rango mínimo y un rango máximo para cada estrato y después se da una muestra para cada uno de los estratos, donde se aplica la siguiente fórmula [8] para el cálculo de cada n por cada estrato.

$$n_k = \frac{N_k}{N_T} * n \tag{4}$$

n_k es la muestra a calcular

N_k es la poblacion de cada estrato

N_T es la poblacion total de todo el marco muestral

n es la muestra que se selecciona del estrato

En el ejemplo mostrado en la figura 8, se determina inicialmente el marco muestral [9], en este caso conformado por 20 elementos, luego se determina una muestra de 9. En el costado derecho de la ventana se determina el número de estratos a trabajar, en este caso se tomarán 3 estratos, luego se determina el rango mínimo y máximo para cada estrato y finalmente en la parte inferior derecha se muestra un resumen con las respuestas obtenidas [10].

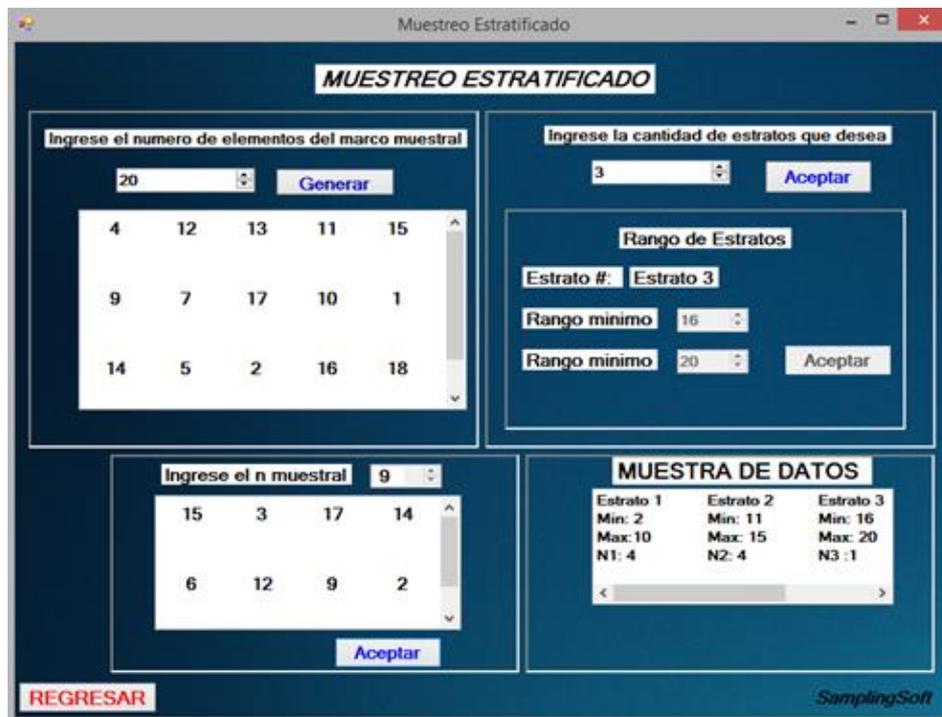


Figura 8. Interfaz de muestreo estratificado

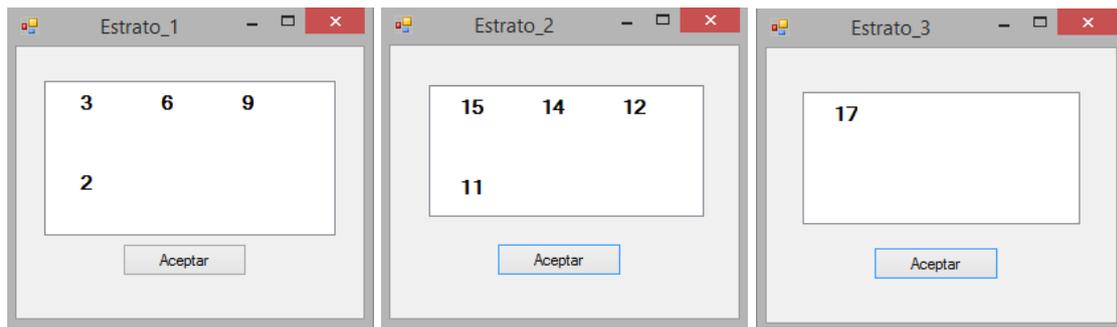


Figura 9. Interfaz de Estrato 1,2,3

3.0 RESULTADOS Y DISCUSION

Los resultados están representados en la Tabla1 y Tabla 2.

Tabla 1. Análisis de resultados obtenidos en el cálculo de la muestra

Tamaños de la muestra	Tamaño de la población	Nivel de confianza	Probabilidad de éxito	Probabilidad de fracaso	Error	Resultado (muestra aprox.)
Conociendo el tamaño de la población	92	1.96	0.9	0.1	0.05	55
Desconociendo el tamaño de la población	-	1.96	0.9	0.1	0.05	138

Tabla 2. Análisis de resultados obtenidos en el tipo de muestreo

Tamaños de la muestra	Número de elementos del marco mastral	Muestra	Estratos	Rango	Ni	Resultado
Muestreo Aleatorio Simple	8	3	-	-	-	{2,4,6}
Muestreo Sistemático	15	3	-	-	-	{5,11,6}
Muestreo Estratificado	20	9	3	2-10	N1=4	{3,6,9,2}
				11-15	N2= 4	{15,14,12,11}
				16-20	N3= 1	{17}

3.1 Análisis de resultados

El cálculo del tamaño de la muestra es un paso importante según el propósito de la investigación y se elige dependiendo del tamaño que represente la población.

En la tabla 1 se ingresó una población de tamaño 92, un valor de z de 1.96 correspondiente al 95% de nivel de confianza. Una probabilidad de éxito del

90% y un 10% de fracaso de que los datos no cumplan con las características para ser evaluados, finalmente un error del 5%. Con estos datos ingresados, el sistema calcula internamente mediante la fórmula ingresada y obtiene una muestra de 55 individuos, los cuales serán el conjunto que será estudiado.

En el caso de que no se conozca el tamaño de la población o el tamaño de la población sea muy alta, se utiliza la segunda opción de la misma tabla 1, en la que no se requiere del tamaño de la población, por lo que se omite el mismo y sí se ingresa un valor de z de 1.96 correspondiente al 95% de nivel de confianza. Una probabilidad de éxito del 90% y un 10% de fracaso de que los datos no cumplan con las características para ser evaluados, finalmente un error del 5%. Con estos datos ingresados, el sistema calcula internamente mediante la fórmula ingresada y obtiene una muestra de 138 individuos, los cuales serán el conjunto que será estudiado.

Para establecer resultados a partir de la selección del tipo de muestreo se puede utilizar el muestreo aleatorio simple, en la tabla 2 se refleja los datos del número total de elementos del marco muestral que en este caso es 8 y una muestra requerida de 3 individuos, el sistema calcula en forma aleatoria el siguiente resultado {2, 6, 4}, cumpliendo la definición de éste tipo de muestreo, la selección aleatoria de un número determinado de muestras.

Dentro del muestreo sistemático como se puede apreciar en la tabla 2 se ingresó 15 elementos del marco muestral y una muestra requerida de 3, generando un valor de $K=5$ calculado lo que permite que exista un intervalo de cada 5 elementos dentro de marco muestral, obtenido así la siguiente muestra {5,11,6}.

Cuando se selecciona el Muestreo Estratificado se debe prever que este muestreo necesita dividir la muestra por estratos para ello como se muestra en la tabla 2 se ingresa un número total de elementos del marco muestral que en este caso es 20 y una muestra requerida de 9.

Para el efecto se ha dividido en 3 estratos, se establecen rangos según la tabla tenemos números mayores o iguales que 2 pero menores que 10 para el rango 1, números mayores o iguales que 11 pero menores que 15 para el rango 2,

números mayores o iguales que 16 pero menores que 20 para el rango 3, terminando así el ingreso y procediendo a calcular, esto permitirá mostrar el resultado que se puede exponer en la figura 10, 11 y 12 como resultados los datos que fueron escogidos según el rango, para finalmente calcular las muestras de cada rango que serían {3,6,9,2} para el primer rango; {15,14,12,11} para el segundo rango; y {17} para el tercero, resultados que evidencian que para cada estrato se escogió cierta cantidad de individuos a partir de rangos establecidos para ser analizados.

4.0 CONCLUSIONES

- Se pudo concluir que la implementación de fórmulas sobre el cálculo de muestras fue aplicada efectivamente mediante la programación debido a que los datos que fueron ingresados de manera manual fueron comprobados posteriormente con la ayuda del software.
- Basándonos en la comparación de resultados, se concluye que la aplicación de un software estadístico resulta viable debido a la precisión de los resultados obtenidos.
- SamplingSoft es un software con carácter educativo creado para calcular el tamaño de la muestra según parámetros establecidos y seleccionar el tipo de muestreo adecuado para un análisis en particular.
- Se utilizó el IDE Visual Basic, con el fin de realizar una interfaz amigable para el usuario sea docente o estudiante y así mejorar la comprensión de temas fundamentales de la estadística referentes al muestreo y sus técnicas.

REFERENCIAS

- [1] Morphol, J. (2017). Técnicas de Muestreo sobre una Población a Estudio, 35(1):227-232.
- [2] Hammersley M, y Atkinson P. "El diseño de la investigación; problemas, casos y muestras". Etnografía. En: Hammersley M, y Atkinson P. Métodos de investigación. Barcelona: Paidós; 2001. p. 40-68.
- [3] Crespo Blanco C y Salamanca Castro A (2007). El muestreo en la investigación cualitativa, Madrid: fuden.
- [4] Fernández M, Virtudes A. Ruiz Fuentes N. Muestreo estadístico, 1era edición. Editorial Oviedo, 2004: 15-200.
- [5] Curso de estadística Aplicada a las Ciencias Sociales. Tema 8 Poblaciones y Muestras. Manual (tema 19) y Agresti (Cap. 2) URL disponible

<http://www.uclm.es/profesorado/imezo/estadistica/t8.pdf>. Fecha de acceso 1 de julio del 2011.

- [6] Salgado C. (El muestreo en investigación cualitativa). Universidad Autónoma Metropolitana – México.
- [7] Pizarro J, Rubio A and Lira G, (2001), (Optimización De Una Técnica De Muestreo In Situ: Análisis De Cu Y Fe En Sistemas Acuáticos), Recuperado de la web: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0366-16442001000300007.
- [8] Urbano, J. A., Fernández, D. B., Luis, M. I. I. G. G. N. M. I., & Santamaría, E. G. (2016, November). Modelación de escenarios mediante programación lineal y muestreo aleatorio simple. In Congreso Interdisciplinario de Ingenierías (p. 73).
- [9] Kleeberg-Hidalgo, F., & Ramos-Ramírez, J. C. (2009). Aplicación de las técnicas de muestreo en los negocios y la industria. Ingeniería Industrial, (027), 11-40.
- [10] Ramírez, J. R., Rodríguez, G. E. H., & Pérez-Salazar, C. (2014). Software libre que calcula el tamaño de muestra mediante métodos de muestreo probabilístico. Cuadernos de Educación y Desarrollo, (53).



Kevin Alexander López Díaz graduado en la Unidad educativa fiscomisional don Bosco de la Tola en Bachillerato en Informática. Actualmente estudiante de Universidad de las Fuerzas Armadas ESPE-L, cursando quinto semestre en Ingeniería en Software.



Bryan Lenin Sandoval Maiza se graduó en el Instituto Tecnológico Superior Bolívar de la ciudad de Ambato, en Bachillerato en Informática. Actualmente estudiante de la Universidad de las Fuerzas Armadas ESPE-L, cursando quinto semestre en Ingeniería en Software.



Santiago Rafael Urquiza Vaca Máster en Ingeniería en Matemática y Computación. Magister en Tecnología de la Información y Multimedia Educativa. Diplomado Superior en Gestión para el Aprendizaje Universitario. Ingeniero Electrónico.