

INSTITUCIÓN EDUCATIVA SAN LUCAS

GUÍA DE TRABAJO Nº 01

ÁREA: MATEMÁTICAS

ASIGNATURA: ESTADÍSTICA

TEMA: INTRODUCCIÓN A LA ESTADÍSTICA

GRADO: _____

GRUPO: _____

PROFESOR: WILLMAN ORTEGA MENA

ESTUDIANTES:

DEFINICIÓN DE ESTADÍSTICA

La estadística podría definirse como la ciencia que se encarga de recopilar, organizar, procesar, analizar e interpretar datos con el fin de deducir las características de una población objeto de estudio.

En otra definición se puede afirmar que la estadística en general es una rama de las matemáticas que trata de la recolección, organización, presentación, análisis e interpretación de datos numéricos con el fin de realizar una toma de decisiones más efectiva respecto a una situación de interés.

Esta rama de las matemáticas es una ciencia transversal, es decir, aplicable a una variedad de disciplinas, que van desde la física a las ciencias sociales, ciencias de la salud o el control de calidad. Además, tiene gran valor en actividades de negocio o gubernamentales, donde el estudio de los datos obtenidos permite facilitar la toma de decisiones o hacer generalizaciones.

DIVISIÓN DE LA ESTADÍSTICA

La estadística se divide en dos grandes ramas:

Estadística descriptiva: Es la rama de la estadística que se dedica a la descripción, visualización y resumen de datos originados a partir de los fenómenos de estudio. Los datos pueden ser resumidos numéricamente o gráficamente. Su objetivo es organizar y describir las características sobre un conjunto de datos con el propósito de facilitar su aplicación generalmente con el apoyo de gráficas, tablas o medidas numéricas.

La **Estadística descriptiva** trata de “describir” y analizar algunas características de los individuos de un grupo dado, sin extraer conclusiones para un grupo mayor.

Para este estudio, se siguen estos pasos:

- Selección de las características que interesa estudiar.
- Análisis de cada característica: diseño de la encuesta o del experimento y recogida de datos.
- Clasificación y organización de los resultados en tablas de frecuencias.
- Elaboración de gráficos, si conviene, para divulgarlos a un público amplio (no experto).
- Obtención de parámetros: valores numéricos que resumen la información obtenida.

Estadística inferencial: es la rama de la estadística que se dedica a la generación de los modelos, inferencias y predicciones asociadas a los fenómenos o situaciones que son objeto de estudio, teniendo en cuenta lo aleatorio y la incertidumbre en las observaciones. Se usa para modelar patrones en los datos y extraer inferencias acerca del fenómeno estudiado.

La **estadística inferencial** trabaja con muestras y pretende, a partir de ellas, “inferir” características de toda la población. Es decir, se pretende tomar como generales propiedades que solo se han verificado para casos

particulares. En este proceso hay que actuar con mucha cautela: ¿Cómo se elige la muestra?, ¿Qué grado de confianza se puede tener en el resultado obtenido? Etc.

Aplicaciones de la Estadística

La importancia de la Estadística radica en sus aplicaciones en varios campos de la ciencia, la tecnología y la sociedad. A continuación, te presento algunos ejemplos:

- **En Política:** Conocer la preferencia de los electores por un candidato para diseñar una campaña política.
- **En Gobierno:** Para planear programas sociales de salud, educación, servicios.
- **En Economía:** Para dar a conocer los índices económicos relativos a la producción, a la mano de obra, índices de precios para el consumidor. Para predecir las fluctuaciones del mercado bursátil, las tasas de interés, el índice de inflación, el costo de la vida. Determinación de las variaciones de precios y salarios. Determinar el poder adquisitivo de la población. Etc.
- **La Biología:** Evolución de las poblaciones. Estudios genéticos.
- **Los Deportes.** Conocer el rendimiento de un equipo o de los jugadores.
- **La Medicina:** Probar la eficiencia de un medicamento o producto médico. Realizar el diagnóstico de enfermedades. Planificar los programas de salud pública. Etc.

EERCICIOS

- 1) Investiga al menos otros cinco campos con ejemplos donde se aplique la estadística.
- 2) Describe al menos dos actividades cotidianas en las que hayas empleado la estadística

CONCEPTOS BÁSICOS

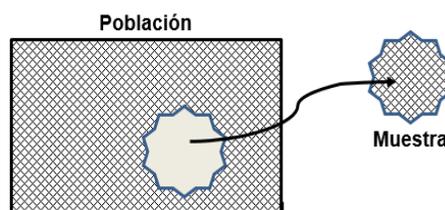
Para una mejor comprensión del estudio de la estadística a continuación presentamos ciertos conceptos que son de frecuente manejo.

Experimento Estadístico. Es una actividad planificada, cuyos resultados producen un conjunto de datos. Es el proceso mediante el cual una observación o medición es registrada. Por ejemplo, si quisiéramos conocer el promedio de edad de los habitantes de una determinada región, las acciones realizadas con este propósito serán nuestro experimento estadístico

Población: Conjunto de todos los individuos (personas, objetos, animales, etc.) que porten información sobre el fenómeno que se estudia. Por ejemplo, si estudiamos la edad de los habitantes en una ciudad, la población será el total de los habitantes de dicha ciudad.

Muestra: Es un subconjunto de la población, cuando es seleccionada de tal manera que cada individuo de la población ha tenido la misma oportunidad de hacer parte de ella, se dice que la muestra es representativa o una

Muestra Aleatoria. Por ejemplo, al elegir 30 personas al azar por cada barrio de la ciudad para saber sus edades, será representativo para conocer el promedio de edad de la ciudad.



Individuo: es cualquier elemento que porte información sobre el fenómeno que se estudia. Así, si estudiamos la altura de los niños de una clase, cada alumno es un individuo; si estudiamos la edad de cada habitante, cada habitante es un individuo.

Variable: Una variable es una característica observable que puede tener diferentes valores en los distintos elementos o individuos de una población o una muestra. Las variables pueden ser de dos tipos: Variables Cualitativas o atributos y Variables Cuantitativas

Variabes cualitativas o atributos: Son variables que no se pueden medir numéricamente (ejemplo: nacionalidad, color de la piel, sexo). Estas a su vez se pueden clasificar en **Nominales y Ordinales**.

- **Variabes cualitativas Nominales:** Son aquellas cuyos valores No se pueden ordenar, por ejemplo: sexo, grupo sanguíneo, nacionalidad, etc.
- **Variabes cualitativas Ordinales:** Son aquellas cuyos valores Si es posible ordenar, por ejemplo: grado de satisfacción, sensación de dolor, etc.

Variabes cuantitativas: Estas variables tienen valor numérico (ejemplo: edad, precio de un artículo, el dinero de la merienda, la altura de un edificio, etc.). Estas a su vez se pueden clasificar en **Discretas y Continuas**.

- **Variabes cuantitativas Discretas:** Son aquellas que sólo pueden tomar valores enteros (1, 2, 8, -4, etc.). Por ejemplo: número de hermanos (puede ser 1, 2, 3, ..., etc).
- **Variabes cuantitativas Continuas:** Estas variables pueden tomar cualquier valor real dentro de un intervalo. Por ejemplo, la velocidad de un vehículo puede ser 80,3 km/h.

En la siguiente tabla mostramos un resumen de los distintos tipos de variables y su utilidad

TIPO DE VARIABLE	SIRVE PARA
Cualitativa nominal	CLASIFICAR
Cualitativa ordinal	JERARQUIZAR
Cuantitativa discreta	CONTAR
Cuantitativa continua	MEDIR

Dato: Es cada uno de los valores que puede tener una variable, por ejemplo: En los ciudadanos de la población colombiana puede variar

- El grupo sanguíneo (A, B, AB, O)
- Su Nivel de felicidad (Muy feliz, Feliz, Moderadamente Feliz, Deprimido)
- Su estatura en centímetros (53,2; 57; 63,66; 128,9; ...; etc.)
- Número de miembros en la familia (2, 3, 4, ..., etc.)

EJERCICIOS

Resuelve de manera individual

1. Indica escribiendo al lado de cada variable cuáles son cualitativas (**Cuali**) y cuáles son cuantitativas (**Cuanti**):
 - a. Comida Favorita. (___)
 - b. Profesión que te gusta. (___)
 - c. Número de goles marcados por El Junior en el último partido. (___)

- d. Número de alumnos de tu escuela. (___)
 - e. El color de los ojos de tus compañeros de clase. (___)
2. Indica escribiendo al lado de cada variable si son discretas (**D**) o continuas (**C**).
- a. Número de acciones vendidas cada día en la Bolsa. (___)
 - b. Temperaturas registradas cada hora en un observatorio. (___)
 - c. Vida útil de un automóvil. (___)
 - d. El diámetro de las ruedas de varias bicicletas. (___)
 - e. Número de hijos de 50 familias. (___)
3. Clasifica las siguientes variables escribiendo al lado de ella (**CO**) si es cualitativa ordinal, (**CN**) si es cualitativa nominal, (**CD**) si es cuantitativas discretas o (**CC**) si es cuantitativa continua.
- a. Número de litros de agua contenidos en un depósito. (___)
 - b. Número de libros en un estante de librería. (___)
 - c. Suma de puntos obtenidos al lanzar de un par de dados. (___)
 - d. La profesión de una persona. (___)
 - e. El área de las distintas baldosas de un edificio. (___)
4. Escribe al lado de cada una de las siguientes afirmaciones (**V**) si es verdadero y (**F**) si es falso
- a. Una muestra aleatoria es aquella en la cual ciertos elementos tienen mayor posibilidad que otros de ser seleccionados. (___)
 - b. Una variable cuantitativa no se le puede asignar un valor numérico. (___)
 - c. La variable “nacionalidad de una persona” es cualitativa discreta. (___)
 - d. La estadística inferencial se ocupa de la colección y clasificación de información, de su resumen en cuadros y gráficos adecuados que resuman en forma apropiada la información captada. (___)

PROCESAMIENTO DE DATOS

Para un estudio estadístico es necesario recolectar información en distintos momentos y de diferentes fuentes para aspirar a una mejor toma de decisiones. Por ejemplo, una empresa que desea lanzar un nuevo producto, antes necesita recopilar información sobre la demanda, las preferencias de los clientes, los competidores, etc. En caso de que esto no suceda, el producto recién lanzado puede fracasar por diferentes razones.

Aunque los datos son un activo valioso para cualquier investigación, no sirven para nada si no se analizan o procesan. Para iniciar el procesamiento de los datos es importante luego que han sido recolectados por diferentes técnicas que trabajaremos con posterioridad, conocer y manejar ciertos conceptos y procedimientos de corriente uso para la Estadística.

INVESTIGACIÓN ESTADÍSTICA

Es toda operación orientada a la recopilación de información acerca de un fenómeno en particular, siguiendo procedimientos estandarizados y certificados.

El proceso de la investigación estadística implica una serie de etapas; pues lanzarse a investigar sin un criterio previo o preparación adecuada puede demandar más tiempo del programado.

La **formulación del problema** es el paso inicial del procedimiento de investigación, un correcto planteamiento o formulación del problema establece los límites de tiempo y espacio en los que se efectuara la investigación, lo cual sirve para optimizar el tiempo y los recursos empleados en la investigación.

Luego de plantear el problema, la aplicación de la estadística implica las siguientes etapas:

- **Recolección de información**
- **Organización de la información.**
- **Análisis de la información.**
- **Interpretación de resultados.**

RECOLECCIÓN DE INFORMACIÓN

En cualquier estudio estadístico, es de gran importancia el proceso de recogida, ordenación y presentación de los datos, para ello empleamos diversas técnicas y herramientas estadísticas.

Elección de la muestra.

Para la elección de la muestra distinguimos dos tipos fundamentales de muestreo:

- **Muestreo probabilístico (aleatorio):** En este tipo de muestreo, todos los individuos de la población pueden formar parte de la muestra, es decir que todos tienen la probabilidad de formar parte de la muestra. Por lo tanto, es el tipo de muestreo que deberemos utilizar en nuestras investigaciones, por ser el riguroso y científico.
- **Muestreo no probabilístico (no aleatorio):** En este tipo de muestreo, puede haber clara influencia de la persona o personas que seleccionan la muestra o simplemente se realiza atendiendo a razones de comodidad. No es un tipo de muestreo riguroso y científico, dado que no todos los elementos de la población pueden formar parte de la muestra. Por ejemplo, si hacemos una encuesta telefónica por la mañana, las personas que no tienen teléfono o que están trabajando, no podrán formar parte de la muestra.

Luego de elegida la muestra se procede a la obtención de los datos. Esto puede ser realizado mediante las diversas **Técnicas de recolección de datos**. Entre ellas podemos citar:

La observación: La observación es una técnica que consiste precisamente en observar el desarrollo del fenómeno que se desea analizar. Éste método puede usarse para obtener información cualitativa o cuantitativa de acuerdo con el modo en que se realiza.

- En investigación cualitativa permite analizar las relaciones entre los participantes gracias al análisis de sus comportamientos y de su comunicación no verbal.
- En investigación cuantitativa es útil para hacer seguimiento a la frecuencia de fenómenos biológicos o al funcionamiento de una máquina.

Por ejemplo, si se desea comprender los motivos de la deserción escolar, puede ser útil observar la forma en que se relacionan maestros y estudiantes. En este caso, puede aplicarse la técnica de observación en una clase cualquiera.

Al aplicar esta técnica con enfoque cualitativo, es necesario organizar las observaciones en categorías temáticas para poder darle un orden al análisis. Estas categorías deben relacionarse con la información obtenida a través de otras técnicas para tener mayor validez.

Las entrevistas: La entrevista es, en esencia, una conversación bien planificada. En ella, el investigador plantea una serie de preguntas o temas de debate a una o varias personas, con el fin de obtener información específica. Puede realizarse personalmente, por teléfono o de manera virtual. Sin embargo, en algunos casos es importante la interacción personal con el entrevistado, para poder tomar nota de la información que brinda la comunicación no verbal.

Por ejemplo, en una investigación que indaga sobre las causas de la deserción escolar en una institución, pueden aplicarse entrevistas. En este caso puede ser útil entrevistar a actores de la problemática como padres y estudiantes, al igual que funcionarios públicos para comprender mejor la problemática.

Según la **organización de una entrevista**, puede ser estructurada, semiestructurada o informal.

- Una **entrevista estructurada** es aquella en la cual el entrevistador tiene una lista de preguntas definidas previamente y se limita estrictamente a ellas.
- En la **entrevista semiestructurada** existe una guía de preguntas o temas generales de conversación. Sin embargo, el entrevistador puede desarrollar preguntas nuevas a medida que vayan surgiendo los temas de su interés.
- La **entrevista informal**, es aquella que no está guiada por una lista de preguntas determinadas. El entrevistador tiene claros los temas sobre los que quiere indagar y los introduce de manera espontánea en la conversación.

Los cuestionarios y las encuestas, son técnicas en las cuales se plantea un listado de preguntas cerradas para obtener datos precisos. Usualmente se usan en investigaciones cuantitativas, pero también pueden incluirse preguntas abiertas para permitir un análisis cualitativo.

Es una técnica muy extendida porque permite obtener información precisa de una gran cantidad de personas. El hecho de tener preguntas cerradas, permite calcular los resultados y obtener porcentajes que permitan un análisis rápido de los mismos. Además, es un método ágil, teniendo en cuenta que no requiere la presencia del investigador para realizarse. Puede hacerse masivamente por correo, a través de internet o vía telefónica.

Para continuar con el ejemplo de la deserción escolar, los cuestionarios pueden ser útiles para obtener información precisa de los estudiantes. Por ejemplo: edad, grado en el que abandona la escuela, motivos para abandonarla, etc.

En poblaciones pequeñas (por ejemplo, los trabajadores de una empresa) se pueden realizar encuestas censales, es decir, a todos los individuos de un colectivo, pero normalmente las encuestas se realizan mediante muestras estadísticamente representativas de una población mayor, para posteriormente extrapolar los resultados al conjunto de la población.

La encuesta se diferencia de otros métodos de investigación en que la información obtenida ya está de antemano preparada y estructurada. En este sentido, la encuesta presenta notables limitaciones, al restringir las posibilidades de obtener información a validar o refutar hipótesis previamente establecidas en el cuestionario, coartando el discurso del entrevistado, y sin que resulte posible saber si existe información relevante que no se tiene en cuenta. Además, al tratarse de entrevistas individuales, se pierde la riqueza de fenómenos que resultan de la interacción social.

El Cuestionario es el instrumento de la encuesta y es un instrumento de recogida de datos rigurosamente estandarizado que operacionaliza las variables objeto de observación e investigación.

Los **tipos de preguntas** que se realizan en un cuestionario se clasifican según la contestación que admitan:

- **Preguntas Abiertas:** (preguntas que sólo formulan la pregunta, sin establecer categorías de respuesta). Se deben utilizar muy poco en las encuestas porque después de la encuesta hay que cerrarlas y luego estandarizarlas.
- **Preguntas Cerradas:** Usualmente establecen sólo 2 alternativas de respuesta, "Si o No" y a veces NS/NR (no sabe o no responde) Se deben utilizar sólo para temas muy bien definidos que admiten estas alternativas como respuesta.

EJERCICIOS

Resuelve de manera individual

- 1) Escoja una problemática que esté afectando a su comunidad e investigue sobre la situación o el estado en que se encuentra.
- 2) Plantee una hipótesis sobre lo que cree que es causa de la problemática señalada.
- 3) Utilice una de las técnicas para la recolectar información sobre la problemática.

ORGANIZACIÓN, PRESENTACIÓN Y ANÁLISIS DE LA INFORMACIÓN

Para presentar los datos que van a ser procesados, comúnmente se utiliza una tabla de **Distribución de Frecuencias**

La frecuencia es una medida que sirve para comparar la aparición de un elemento X_i en un conjunto de elementos (X_1, X_2, \dots, X_N) .

Las frecuencias de cada elemento se pueden expresar tanto como **frecuencias absolutas** (número total de apariciones) como **frecuencias relativas** (proporción de apariciones).

Si los datos provienen de una variable cuantitativa o existen muchas categorías, los datos se suelen agrupar en **clases**. Cada clase, que tiene un intervalo o amplitud constante, está representada por su marca de clase, que es el punto medio de la misma.

Antes de continuar un concepto importante

Sumatoria o Notación Sigma

La sumatoria (llamada también notación sigma) es una operación matemática que se emplea para calcular la suma de muchos sumandos; se expresa con la letra griega sigma mayúscula Σ , y se representa así:

$$\sum_i^k X_i$$
$$\sum_i^k X_i = X_1 + X_2 + X_3 + \dots + X_k$$

Esta expresión se lee: "sumatoria de todos los X_i desde i igual a uno hasta k , donde i toma los valores desde 1 hasta k ".

- i es el valor inicial, llamado límite inferior.
- k es el valor final, llamado límite superior. Pero necesariamente debe cumplirse que: $i \leq k$

Conociendo esto podemos definir:

Frecuencia absoluta:

La frecuencia absoluta (n_i) de un valor X_i es el número de veces que el valor aparece en el conjunto de datos. La suma de las frecuencias absolutas de todos los elementos diferentes del conjunto o muestra de datos debe ser el número total de sujetos N . Si el conjunto tiene k números (o categorías) diferentes, entonces:

$$\sum_i^k n_i = N$$

$$\sum_i^k n_i = n_1 + n_2 + n_3 + \dots + n_k = N$$

EJEMPLO

El profesor Arnold tiene la lista de las notas en matemáticas de 30 alumnos de su clase de octavo. Las notas son las siguientes:

NOTAS DE MATEMÁTICAS DE GRADO 8º									
3,0	5,0	2,5	2,5	2,0	2,0	3,0	3,0	2,5	2,0
3,0	3,5	3,5	2,5	3,0	1,5	3,0	3,5	4,5	2,5
3,0	2,5	3,5	1,5	4,0	4,0	2,0	3,5	4,0	4,5

Se realiza el **recuento** de la variable que se estudia (notas) para ver el número de veces que aparece cada nota. Una vez realizado el recuento, se representan las **frecuencias absolutas** de cada una de las notas (n_i). Las frecuencias son: $n_1(1,5) = 2$; $n_2(2,0) = 4$; $n_3(2,5) = 6$; $n_4(3,0) = 7$; $n_5(3,5) = 5$; $n_6(4,0) = 3$; $n_7(4,5) = 2$; $n_8(5,0) = 1$

X_i	Frecuencia Absoluta (n_i)
1,5	2
2,0	4
2,5	6
3,0	7
3,5	5
4,0	3
4,5	2
5,0	1
Total	30

EJERCICIO

Indaga vía chat o telefónicamente entre tus hermanos, amigos y primos sus edades y realiza una tabla con sus frecuencias absolutas

Bibliografía

- <https://sites.google.com/site/339estadistica/05---etapas-del-proceso-estadistico>
- <http://www.monografias.com/trabajos30/estadistica-basica/estadistica-basica.shtml>
- http://fcps.uaq.mx/descargas/prope2014/estadistica/1/estadistica_conceptos_basicos.pdf
- <http://www.profesorenlinea.cl/matematica/Sumatoria.html>
- <http://www.eumed.net/libros-gratis/2007a/239/1c.htm>
- http://www.slideshare.net/winman73/savedfiles?s_title=investigacin-estadstica-15314208&user_login=ingricardoguevara
- <http://leolabor.blogspot.com/2009/08/que-es-investigacion-estadistica.html>
- <http://ligrfg.obolog.es/estadistica-se-divide-dos-ramas-56026>
- http://maralboran.org/wikipedia/index.php/Tipos_de_muestreos
- <https://www.lifeder.com/tecnicas-instrumentos-recoleccion-datos/>
- <https://www.universoformulas.com/estadistica/descriptiva/frecuencia/>
- Ardila Gutiérrez Víctor Hernando. Inteligencia Lógico – Matemática. Edit. Voluntad, Bogotá (Col.) 2003