

Método Lógico

Guía para elaborar un proyecto de investigación

Parte VIII

Dra. Merced Velázquez Quintana¹

En una investigación, el análisis de los datos inicia al terminar de recolectar la información (Tabla 1). Para este procedimiento sugerimos utilizar un programa estadístico computarizado como el Epi Info, desarrollado por la Organización Mundial de la Salud y el Centro de Control de Enfermedades de los Estados Unidos (CDC), distribuido en forma gratuita en la dirección www.cdc.gov. Este programa, que puede utilizarse en inglés o español, vie-

Tabla 1

ETAPAS DE UNA INVESTIGACIÓN

1. Elaboración del proyecto
2. Recolección de datos
3. Análisis estadístico
4. Elaboración del informe

ne con una guía completa para su uso.

En el análisis estadístico podemos distinguir la descripción, la inferencia y la estimación. Todos los estudios inician con la descripción estadística de cada variable estudiada; solo en los estudios no descriptivos se hace una correlación de variables para obtener la inferencia y la estimación (Tabla 2).

I. ESTUDIOS DESCRIPTIVOS

Utilizan el análisis de una sola variable, conocido como análisis univariado, el cual muestra conjuntos de datos a través de la estadística descriptiva. Los procedimientos estadísticos varían de acuerdo a la escala que utilizamos para medir la variable:

¹Neurologa, Maestra en Ciencias Médicas
Para correspondencia:
Ortiz Mena # 1815 Fraccionamiento, Las Águilas
Chihuahua, Chih. México CP 31230
Tel. y Fax (14) 10 - 34 - 49
mchvelazquez@infosel.net.mx

Tabla 2

TIPO DE ANÁLISIS DE ACUERDO AL DISEÑO

- I. Estudio descriptivo – análisis univariado.
- II. Casos y controles – medidas de riesgo y análisis multivariado.
- III. Cohortes – análisis de supervivencia, indicadores de riesgo y pruebas de hipótesis.
- IV. Ensayo clínico controlado – prueba de hipótesis.

1. Variables obtenidas en escala nominal.

Son valores expresados en categorías que no establecen un orden. Su descripción se hace en proporciones, porcentajes o razones, los cuales pueden presentarse en tablas simples o de contingencia y en gráficos de barras.

- **Proporción.** Es dividir la parte entre el todo. Se obtiene dividiendo el número de observaciones con determinada característica entre el número total de observaciones. Se expresa siempre como una fracción de la unidad.

- **Porcentaje.** Consiste en expresar la proporción como fracción de 100. Se obtiene multiplicando la proporción por 100.

- **Índice o tasa.** Consiste en expresar la proporción como fracción de unidad de población y tiempo. Las formas más frecuentemente utilizadas tienen la base de 100,000 habitantes por 1 año.

- **Razón.** Es dividir una parte entre la otra. Se obtiene dividiendo el número de observaciones con determinada característica entre el número de observaciones del mismo grupo sin dicha característica. Se expresa comparando con la unidad (Tablas 3 y 4).

2. VARIABLES OBTENIDAS EN ESCALA ORDINAL

Son valores de datos expresados en categorías que guardan un orden inherente, Ej. Escolaridad: Primaria, Secundaria y Preparatoria. Su descripción

se hace en porcentajes o proporciones; puede también utilizar la Mediana como medida de tendencia central y al Rango como medida de dispersión. Se muestran en tablas de contingencia o en gráficas de barras.

3. VARIABLES OBTENIDAS EN ESCALAS NUMÉRICAS

Son valores de datos expresados en números. Su resumen se expresa con el Rango seguido de una medida de tendencia central y una medida de dispersión (Tabla 5).

MEDIDAS DE TENDENCIA CENTRAL

Utilizadas sobre todo en variables numéricas. Sirven para resumir en un solo valor, considerado típico, los datos que se recolectan en numerosas observaciones. El programa Epi Info muestra en forma automatizada las tres medidas de tendencia central más utilizadas, que son la Media, la Mediana y la Moda.

1. Media (X). Resume variables numéricas. Es el promedio aritmético, que se obtiene con la suma de cada una de las observaciones dividida entre el número de observaciones realizadas.

2. Mediana (Md). Puede usarse en variables ordinales o numéricas. Es la observación central. Para obtenerla, se ordenan las observaciones de menor a mayor; en un número de observaciones impar, la Mediana es el valor que queda en el centro; si el número de observaciones es par, la Mediana se obtiene sumando los valores centrales y dividiéndolos entre 2 (correspondería a la media de los 2 valores centrales).

3. Moda (Mo). Es el valor que se presenta con mayor frecuencia en una serie de datos. Cuando ningún valor se repite, la serie no tiene moda; cuando varios valores se repiten con igual frecuencia, la serie tiene más de una moda, lo que se conoce como distribución bimodal, trimodal, etc.

4. Media Geométrica (GM). Es la raíz enésima del producto de las observaciones: el valor se multiplica por sí mismo y se le saca la raíz correspondien-

te, esto es, en dos observaciones, se saca raíz cuadrada, en tres, raíz cúbica, etc. Se usa sobre todo en mediciones hechas en escala logarítmica (Tabla 6).

MEDIDAS DE DISPERSIÓN.

Es el resumen de datos que muestra la amplitud o dispersión de las observaciones. El programa EPI Info muestra en forma automatizada el Valor Mínimo, el Valor Máximo, la Desviación estándar, la Varianza, la Percentil 25 y la Percentil 75.

1. Rango. Son los valores menor y mayor de las observaciones.

2. Recorrido. Es la diferencia entre el valor mayor y menor de una serie.

3. Desviación estándar (s). Es la medida de dispersión más utilizada, ya que se expresa en las unidades en que se midió la variable. Muestra la dispersión de los valores alrededor de la media: independientemente de la distribución de las observaciones, el 75% de los valores se agrupan a dos desviaciones estándar más y menos de la medida central. Se usa cuando la medida central se expresa con la Media.

• **Coefficiente de variación (CV).** Es una estandarización útil para comparar mediciones obtenidas con diferente escala. Se obtiene dividiendo la desviación estándar de las observaciones entre la medida central y multiplicando el resultado por 100. Se expresa como un porcentaje.

4. Varianza (S²). Expresa la dispersión de los valores en torno a la Media en unidades cuadradas.

5. Percentiles. Se usan para comparar un valor individual con una norma establecida. Para obtenerlas, se divide el rango entre 100, siendo una percentil cada uno de los límites obtenidos. La expresión "La talla de 1.20 m se encuentra en la percentil 95 de la edad", indica que 95% de las personas de esa edad miden 1.20 m o menos.

6. Rango intercuartílico. El rango se divide en cuatro partes o cuartiles, cuyos límites son la percentil 25, la Media y la percentil 75 (Tabla 7).

Tabla 3

DESCRIPCIÓN DE UNA VARIABLE MEDIDA EN ESCALA NOMINAL

Variable – Sexo, medido en escala nominal, con indicadores Femenino y Masculino.

Observaciones: Femenino – 200; Masculino – 100; Total – 300.

- **PROPORCIÓN** – Dividir la parte entre el todo. El sexo femenino se presenta en una proporción de 0.66.
- **PORCENTAJE** – Expresar la proporción como fracción de 100. El 66% de los participantes son de sexo femenino.
- **RAZÓN** – Se expresa comparando con la unidad. El sexo femenino se presenta a razón de 2:1.

Redacción: El resultado se expresa sólo en una de las tres formas.

Tabla 4
TABLA SIMPLE

CATEGORÍAS	FRECUENCIAS	PORCENTAJES
Agricultor	10	10%
Empleado	50	50%
Obrero	30	30%
Negocio Propio	20	20%

La tabla simple se utiliza para mostrar una variable nominal u ordinal.

Tabla 5

DESCRIPCIÓN DE UNA VARIABLE MEDIDA EN ESCALA NUMÉRICA

Variable – Edad, medida en escala numérica, con indicadores en años cumplidos.
Observaciones: Total – 300.

- **RANGO** – Edad mínima 15 años. Edad máxima 56 años.
- **MEDIA** – 35 años.
- **DESVIACIÓN ESTÁNDAR** – 7.

Redacción – El resultado se expresa con el Rango seguido de una de las medidas de tendencia central y una de las medidas de dispersión:

“La edad de los participantes fue de 15 a 56 años (35 ± 7)...”

Tabla 6

MEDIDAS DE TENDENCIA CENTRAL

Resumen que muestra el valor medio de los datos, mismo que se considera como típico del conjunto.

- **MEDIA (X)** – Para resumir variables numéricas.
- **MEDIANA (Md)** – Para variables numéricas y ordinales.
- **MODA (Mo)** – Cuando la variable numérica tiene dos o más valores que se presentan con mayor frecuencia.
- **MEDIA GEOMÉTRICA (GM)** – En variables numéricas medidas en escala logarítmica.

Cuando la Media y la Mediana son iguales, se utiliza la Media.

Esto indica que los datos tienen una distribución normal.

Cuando la Media y la Mediana difieren, se utiliza la Mediana.

Esto indica que los datos tienen una distribución sesgada.

El uso de la Moda indica que los datos tienen una distribución bimodal.

Tabla 7

MEDIDAS DE DISPERSIÓN

Muestran la amplitud que tienen los valores de las observaciones.

- **RANGO** – Expresa el valor mayor y menor de las observaciones.
- **RECORRIDO** – Diferencia entre el valor mayor y menor de una serie.
- **DESVIACIÓN ESTANDAR (s)** – Dispersión de los valores alrededor de la media.
- **VARIANZA (s²)** – Diseminación de los valores en torno a la media, expresado en unidades cuadradas, por lo que no se puede aplicar en términos de unidades originales.
- **COEFICIENTE DE VARIACIÓN (CV)** – Es una estandarización útil para comparar mediciones obtenidas con diferente escala.
- **PERCENTILES** – Se usan para comparar un valor individual con una norma establecida.
- **RANGO INTERCUARTÍLICO** – El rango se divide en cuatro partes o cuartiles.

LECTURAS RECOMENDADAS:

1. Hernández R, Fernández C, Baptista P. *Metodología de la investigación*. Mc Graw Hill. Segunda edición. México 1998.
2. Daniel W. *Bioestadística. Base para el análisis de las ciencias de la salud*. Editorial Limusa. 5ª edición. México 1996.
3. Hopkins K, Hopkins B, Glass G. *Estadística básica para las ciencias sociales y del comportamiento*. Prentice May. Tercera edición. México 1997.
4. Dawson-Saunders B, Trapp R. *Bioestadística Médica. Manual Moderno*. México 1993.

