



# Mi Universidad

## Ensayo

*Nombre del Alumno: Lidia Peralta Castellanos*

*Nombre del tema: Unidad I. Estadística descriptiva. Unidad II. Teoría de la probabilidad*

*Parcial: Primero*

*Nombre de la Materia: Tendencias y sistemas de salud en México*

*Nombre del profesor: DAE. MA. Cecilia Zamorano R.*

*Nombre de la Maestría: Administración en los servicios de salud.*

*Cuatrimestre: Primero*

*Pichucalco, Chiapas a 26 de Noviembre del 2022*

# **ESTADÍSTICA DESCRIPTIVA**

## **Procesamiento estadístico de datos**

El proceso estadístico es el conjunto de etapas o fases que deben completarse para realizar una investigación basada en información cuantitativa y obtener unos resultados fieles a la realidad estudiada, es una serie de pasos para garantizar que los resultados se adecuen a la realidad y poder tomar mejores decisiones.

## **Etapas**

### **Planteamiento del problema**

Aquí se sitúa el eje central sobre el que articular todo lo es una reflexión ordenada que va dando cuenta de una transición lógica del pensamiento desde el punto de vista metodológico.

### **Recolección de datos**

Es importante describir la organización que va a tener el investigador para poder obtener la información de la población del estudio, en ese sentido, es importante que en este apartado se describa la forma para obtener los permisos, el tiempo, lugar y fecha para la aplicación de instrumentos, entre otros.

El procedimiento para recolectar información es:

- Recolectar (a través de bases de datos o de encuestas personalizadas ).
- Sistematizar.
- Presentar.

### **Análisis de los datos**

Una vez planteado el problema, recolectados los datos y organizados podemos analizarlos de forma eficaz. Dependiendo del planteamiento del problema, se realizará un tipo de análisis u otro. El análisis de datos es la ciencia que se encarga de ex

aminar un conjunto de datos con el propósito de sacar conclusiones sobre la información para poder tomar decisiones, o simplemente ampliar los conocimientos sobre diversos temas.

### **Técnicas de análisis de datos**

-Análisis de datos cualitativo: Los datos cualitativos se presentan de manera verbal. Se basa en la interpretación.

-Análisis de datos cuantitativos: Los datos cuantitativos se presentan en forma numérica. Se basa en resultados tangibles.

### **Ventajas del análisis de datos**

- Capacidad para tomar decisiones de negocios más rápidas e informadas, respaldadas por hechos.
- Ayuda a las empresas a identificar problemas de rendimiento que requieren algún tipo de acción.
- Comprensión más profunda de los requisitos de los clientes, lo que, a su vez, crea mejores relaciones comerciales.
- Mayor conciencia del riesgo, permitiendo la implementación de medidas preventivas.

### **Interpretación de los datos**

El objetivo de la recopilación e interpretación es adquirir información útil y utilizable y tomar las decisiones más informadas posibles.

### **Distribuciones de frecuencia**

Las tablas de distribución de frecuencias se utilizan cuando se recolectan datos, con ellas se pueden representar los datos de manera que es más fácil analizarlos.

Permiten resumir los datos en una tabla que recoge:

Clase: es cada uno de los grupos en que se dividen los datos.

La frecuencia absoluta: es el número de veces que se repite cada dato. Cuando se agrupan los datos, es el número de datos que tiene cada clase.

La marca de clase: es el punto medio de la clase. Se obtiene dividiendo entre dos la suma de los valores extremos de cada clase.

El rango: es la diferencia entre el valor mayor y el valor menor en estudio de una distribución de datos.

La frecuencia relativa de un dato da información sobre qué parte de la población o de la muestra en estudio corresponde a la característica analizada. La frecuencia acumulada relativa es la frecuencia relativa total hasta el límite superior de cada clase.

Los límites de clase: son los valores que separan a una clase en particular de la anterior y de la siguiente.

### **Presentación gráfica**

Con la finalidad de facilitar la comprensión y el análisis tanto por parte de los mismos investigadores como de cara a mostrar la variabilidad de los datos y de donde salen las conclusiones al resto del mundo, es de gran utilidad emplear elementos visuales de fácil interpretación: las gráficas o gráficos.

Una gráfica es la representación en unos ejes de coordenadas de los pares ordenados de una tabla. Esta representación visual sirve de apoyo a la hora de mostrar y comprender de manera sintetizada los datos recabados durante la investigación, de manera que puede tanto los investigadores que llevan a cabo el análisis como otros puedan comprender los resultados y resulte sencillo utilizarlo como referencia.

### **Medidas de tendencia central**

Persiguen una serie de objetivos que justifican su existencia, sirven para conocer en qué lugar se ubica el elemento promedio, o típico del grupo. Además sirven para comparar e interpretar los resultados obtenidos con relación a los distintos valores observados.

Principales medidas de tendencia central:

Media: Es el valor promedio de un conjunto de datos numéricos, calculada como la suma del conjunto de valores dividida entre el número total de valores. Existen diferentes tipos de media y su elección tiene que ver principalmente con el tipo de dato que se calcula.

Mediana: Es un estadístico de posición central que parte la distribución en dos, es decir, que deja la misma cantidad de valores a un lado que a otro.

Moda: Es el valor que más se repite en una muestra estadística o población.

### **Medidas de dispersión**

Las medidas de dispersión indican qué tanto se dispersan o agrupan los datos con respecto a su media aritmética. Las medidas de dispersión son de dos tipos:

-Medidas de dispersión absoluta: como recorrido, desviación media, varianza y desviación típica, que se usan en los análisis estadísticos generales.

-Medidas de dispersión relativa: que determinan la dispersión de la distribución estadística independientemente de las unidades en que se exprese la variable.

### **Teorema de Chebyshev**

Es considerado una desigualdad probabilística, proporciona un límite superior a la probabilidad de que la desviación absoluta de una variable correspondiente o aleatoria, de su medida, excede un umbral dado.

El teorema recibe el nombre en honor al matemático ruso Pafnuty Chebyshev (también escrito como Chebychev o Tchebycheff) quien, a pesar de no ser el primero en enunciar dicho teorema, fue el primero en dar una demostración en el año 1867.

Este teorema explica que al menos  $1 - 1/k^2$  de datos de una muestra deben caer dentro de  $K$ , que es las desviaciones estándar de la media. En cualquier ejercicio o prueba, el  $K$  es un número real positivo mayor que uno. Es un conjunto de datos que se distribuye, o se encuentra en forma de curva de campana.

## **Regla empírica**

Esta regla sugiere que cada dato que se puede observar caerá bajo tres desviaciones estándar diferentes de la media en una distribución normal. También se le conoce como regla 68-95-99.7 o la regla de los tres dígitos.

De acuerdo con esta regla el 68% de los datos caerán en la primera desviación estándar, el 95% caerán en la primera y la segunda desviación y el 99.7% de los datos caerán en las tres desviaciones.

Si tenemos una distribución normal de los datos en un gráfico en el eje x, la curva de campana estará en el centro. La primera desviación estándar incluye la mitad positiva y la mitad negativa. Podemos obtener rápidamente una estimación de probabilidad aproximada de los datos mediante la regla 68-95-99.7. Se puede utilizar este método como simple prueba cuando la población de los datos es normal. Sin embargo, si la población de los datos no es normal, se puede utilizar este método como una prueba de normalidad.

## **TEORÍA DE LA PROBABILIDAD**

Se ocupa de asignar un cierto número a cada posible resultado que pueda ocurrir en un experimento aleatorio, con el fin de cuantificar dichos resultados y saber si un suceso es más probable que otro.

La probabilidad como concepto existe desde hace miles de años. Hay evidencias históricas, que nos indican que la primera civilización era capaz de construir dados de 4 caras trabajando con huesos. Más tarde los juegos de azares se hicieron populares en Egipto, Grecia y Roma. Concretamente fue Gerolamo Cardano quién escribió en 1553 un tratado sobre el juego de dados, aunque su obra fue publicada hasta 1663.

### **Enfoques de probabilidad**

a) Enfoque clásico de la probabilidad: Permite determinar valores de probabilidad antes de ser observado el experimento por lo que se le denomina enfoque a priori.

Es aplicado cuando todos los resultados son igualmente probables y no pueden ocurrir al mismo tiempo.

b) Enfoque de frecuencias relativas: Permite determinar la probabilidad con base en la proporción de veces que ocurre un resultado favorable en cierto número experimentos. No implica ningún supuesto previo de igualdad de probabilidades. Se le denomina también enfoque empírico debido a que para determinar los valores de probabilidad se requiere de la observación y de la recopilación de datos. También se le denomina a posteriori, ya que el resultado se obtiene después de realizar el experimento un cierto número de veces.

c) Enfoque subjetivo de la probabilidad; Se diferencia de los dos enfoques anteriores, debido a que tanto el enfoque clásico como el de frecuencia relativa producen valores de probabilidad objetivos. Señala que la probabilidad de un evento es el grado de confianza que una persona tiene en que el evento ocurra, con base en toda la evidencia que tiene disponible, fundamentado en la intuición, opiniones, creencias personales y otra información indirecta. No depende de la repetitividad de ningún evento y permite calcular la probabilidad de sucesos únicos y se da el caso de que ocurra o no esa única vez. Debido a que el valor de la probabilidad es un juicio personal, al enfoque subjetivo se le denomina también enfoque personalista.

### **Espacio muestral**

El espacio muestral es una parte del espacio probabilístico. Como su propio nombre indica, está formado por los elementos de la muestra. Al contrario, el espacio probabilístico engloba todos los elementos. Incluso aunque no salgan recogidos en la muestra.

Dependiendo de cómo sea este conjunto, los espacios muestrales pueden ser:

- Espacio muestral discreto finito. Consta de un número finito de elementos, por ejemplo, lanzar un dado.
- Espacio muestral discreto infinito. Consta de un número infinito numerable de elementos, por ejemplo, lanzar un dado hasta que salga un cinco.

- Espacio muestral continuo. Consta de un número infinito no numerable de elementos, por ejemplo, todas las medidas posibles de espárragos extraídos aleatoriamente de una población.

### **Eventos simples y complejos**

Un Evento es un resultado particular de un experimento aleatorio. En términos de conjuntos, un evento es un subconjunto del espacio muestral.

Un evento simple es un evento con un solo resultado y un evento compuesto es un evento con más de un resultado.

La probabilidad de un evento es la frecuencia con que se espera que ocurra. Cuando todos los resultados posibles de un experimento son igualmente probables, la probabilidad es la relación entre el tamaño del espacio de eventos (los resultados en el evento) y el espacio muestral (todos los posibles resultados del experimento).

### **Leyes de probabilidades**

REGLA DE LA ADICION: La regla de la adición o regla de la suma establece que la probabilidad de ocurrencia de cualquier evento en particular es igual a la suma de las probabilidades individuales, si es que los eventos son mutuamente excluyentes, es decir, que dos no pueden ocurrir al mismo tiempo.

REGLA DE LA MULTIPLICACION: La regla de la multiplicación establece que la probabilidad de ocurrencia de dos o más eventos estadísticamente independientes es igual al producto de sus probabilidades individuales.

La regla de Laplace establece que la probabilidad de ocurrencia de un suceso imposible es 0. La probabilidad de ocurrencia de un suceso seguro es 1.

### **Tablas de contingencia**

Consiste en crear al menos dos filas y dos columnas para representar datos categóricos en términos de conteos de frecuencia. Esta herramienta, que también se conoce como tabla cruzada o como tabla de dos vías, tiene el objetivo de representar en un resumen, la relación entre diferentes variables categóricas.

La tabla permite medir la interacción entre dos variables para conocer una serie de información oculta de gran utilidad para comprender con mayor claridad los resultados de una investigación.

Los objetivos de la tabla de contingencia son los siguientes:

- Ordenar la información recolectada para un estudio cuando los datos se encuentran divididos de forma bidimensional, esto significa a que se relaciona con dos factores cualitativos.
- El otro objetivo de la tabla de contingencia es analizar si hay una relación entre las variables cualitativas, ya sean dependientes o independientes.

### **Teorema de Bayes**

El Teorema de Bayes enunciado por el matemático inglés Thomas Bayes (1702-1761) es un sistema de cálculo de probabilidades, pero hecho de forma inversa a cómo se calculan habitualmente. Tiene en cuenta la información que conocemos que se ha producido en determinado entorno con determinados factores para saber cuáles de esos factores han producido esas consecuencias.

La fórmula del Teorema de Bayes es:

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$

Donde B es el suceso que conocemos, A el conjunto de posibles causas, excluyentes entre sí, que pueden producirlo y, por tanto, P(A/B) son las posibilidades a posteriori, P(A) las posibilidades a priori y P(B/A) la posibilidad de que se dé B en cada hipótesis de A.

## BIBLIOGRAFÍA

- Montgomery, D. C y Runger, R. (2008). Probabilidad y Estadística Aplicadas a la Ingeniería. 2da. Edición. Limusa Wiley. Mexico.
- Rincón. L. (2010). Curso Elemental de Probabilidad y Estadística. 1ra Edición. Circuito Exterior de CU. México D.F.
- Walpole, R. E., Myers, R. H., Myers, S. L. y Ye, K. (2007). Probabilidad y Estadística para Ingeniería y Ciencias. 8va. Edición. Pearson. Mexico
- I. Espejo Miranda, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, A. M. Rodríguez Chía, A. Sánchez Navas, C. Valero Franco. (2006). Estadística Descriptiva y Probabilidad. 3ra. Edición. Servicio de Publicaciones de la Universidad de Cádiz.
- Devore, J. (2008). Probabilidad y Estadística para Ingeniería y Ciencias. 7ma. Edición. Cengage Learning Editores. México
- Canavos, G. Probabilidad y Estadística. Métodos y Aplicaciones.
- Daniel, W. (2009). Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. Limusa Wiley. México.
- Mendenhall, W., Wackerly, D. y Scheaffer, R. (1994). Estadística Matemática con Aplicaciones. Grupo Editorial Iberoamericana. México.
- José Francisco López, 28 de Febrero, 2019. Teoría de la probabilidad.
- ALEA, V. et al. (2006) Estadística Aplicada a les Ciències Econòmiques i Socials. Barcelona: Edicions McGraw-Hill EUB.
- CANAVOS, G. (2008) Probabilidad y Estadística. Aplicaciones y Métodos. México: McGraw-Hill.
- DURA PEIRÓ, J. M. y LÓPEZ CUÑAT, J.M. (2006) Fundamentos de Estadística. Estadística Descriptiva y Modelos Probabilísticos para la Inferencia. Madrid: Ariel Editorial.
- ESCUDER, R. y SANTIAGO, J. (2010) Estadística aplicada. Economía y Ciencias Sociales. Valencia: Tirant lo Blanch.
- FERNÁNDEZ CUESTA, C., y FUENTES GARCÍA, F. (2015) Curso de Estadística Descriptiva. Teoría y Práctica. Madrid: Ariel.
- MARTÍN PLIEGO, F. (1994) Introducción a la Estadística Económica y Empresarial. (Teoría y Práctica) Madrid: AC.
- Paul L. Meyer. Probabilidad y Aplicaciones Estadísticas. S.A. ALHAMBRA MEXICANA.
- URIEL, E. y MUÑIZ, M. (1988) Estadística Económica y Empresarial. Teoría y ejercicios. Madrid: AC.
- Anderson D., Sweeney D., Williams T. Estadística para la administración y economía. Décima edición. Cengage Learning. 2008.
- Berenson M., Levine D., Krehbiel T. Estadística para administración. Segunda edición. Prentice Hall. 2000.
- Meyer P. Probabilidad y Aplicaciones estadísticas. Edición revisada. Addison Wesley Logman. 1998.
- José Francisco López, 07 de Octubre, 2019. Medidas de tendencia central.