



**MATERIA: TENDENCIAS Y SISTEMAS DE SALUD  
EN MEXICO**

**UNIDAD I ESTADÍSTICA DESCRIPTIVA**

**UNIDAD II TEORÍA DE LA PROBABILIDAD**

**DOCENTE: DRA. MARÍA CECILIA ZAMORANO  
RODRÍGUEZ**

**ALUMNO: LIC. AURELIO MAGAÑA GARCIA**

**FECHA: 02/11/2024**

## PROCESAMIENTO ESTADÍSTICO DE DATOS

Este es un proceso de manejo de datos, es una manera de tratar los datos para sacar conclusiones. Consta de pasos que son fundamentales seguir correctamente para evitar datos erróneos al momento de dar resultados. El procesamiento de datos de forma estadística ayuda a calcular mejor las probabilidades u objetivos que queremos alcanzar con el objetivo de prevenir pérdidas en el área en la que estemos aplicando este procedimiento. Una de las principales características del método estadístico es que facilita interpretar un conjunto de datos y extraer conclusiones de él. Además, permite inferir los resultados obtenidos en el análisis estadístico de una muestra de datos a toda la población estadística. Este procedimiento consta de etapas específicas que ayudan a agilizar el logro de lo que se está buscando, las cuales son las siguientes:

**PLANTEAMIENTO DEL PROBLEMA:** Es una descripción clara y concisa de un problema o desafío que requiere ser abordado, sirve como guía para la resolución de problemas y toma de decisiones, ayudando a equipos e individuos a definir el enlace de su trabajo. Para darle paso al problema a solucionar se debe realizar los siguientes cuestionamientos, ¿Dónde?, ¿Cuándo? Y ¿Cómo?, estos serán la base que darán rumbo a la búsqueda de las soluciones deseadas.

**RECOLECCIÓN DE DATOS:** Es la etapa del proceso que nos permite recabar información que ayude a alcanzar los objetivos. Es importante definir el tipo de muestreo o la forma de obtener la información requerida, una de las más comunes es la entrevista, la cual se utiliza para recabar información verbal, a través de preguntas.

**ORGANIZACIÓN DE DATOS:** De acuerdo a la información obtenida se realiza el análisis de los datos de acuerdo a la variable del estudio y se determina si es cuantitativo o cualitativo.

**PRESENTACIÓN:** Una vez obtenida la información a través de las diferentes técnicas e instrumentos y luego de organizar, revisar y clasificar los datos recolectados estos deben ser presentados en cuadros y gráficos estadísticos para hacer el análisis e interpretación correspondiente.

**ANÁLISIS E INTERPRETACIÓN DE DATOS:** Este proceso se utiliza para entender e interpretar los datos para tomar decisiones basadas en información real y tangible.

## DISTRIBUCIONES DE FRECUENCIAS

La distribución de frecuencias es una tabla en la que se agrupan los diferentes valores de una muestra en filas y en cada columna se muestra un tipo de frecuencia de cada valor. Por lo tanto, la distribución de frecuencias sirve para mostrar todos los tipos de frecuencias de un conjunto de datos. Se pueden elaborar tablas de distribución de frecuencias para datos no agrupados y para datos agrupados. Estas últimas se utiliza cuando se tienen muchos datos. En concreto, una distribución de frecuencias incluye la frecuencia absoluta, la frecuencia absoluta acumulada, la frecuencia relativa y la frecuencia relativa acumulada. Una de las características de las distribuciones de frecuencias es que son muy útiles para resumir una muestra estadística tanto de una variable cuantitativa como de una variable cualitativa. Para elaborar tablas de distribuciones de frecuencia se debe tener en cuenta lo siguiente: Cuando hay muchos datos se agrupan en clases. Esto consiste en agrupar los datos en una distribución de frecuencias, que puede definirse como una ordenación o arreglo de datos en clases o categorías que muestran para cada una de ellas, el número de elementos que contiene, denominada frecuencia.

**FRECUENCIA ABSOLUTA:** Es el número de veces que aparece un valor en un conjunto de datos. Es decir, la frecuencia absoluta es la cantidad de veces que se repite un resultado. Por ejemplo, si en una encuesta siete personas han respondido que su color favorito es el rojo, entonces la frecuencia absoluta del color rojo es 7. Por lo tanto, la suma de las frecuencias absolutas de todos los valores es igual al número total de datos de la muestra estadística. En general, se utiliza la letra  $f$  con el subíndice  $i$  para representar la frecuencia absoluta del valor  $i$ , de modo que el símbolo de la frecuencia absoluta es  $f_i$ .

**FRECUENCIA RELATIVA:** es una medida que indica el número de veces que aparece un valor en una muestra de datos en forma de proporción o porcentaje. En concreto, la frecuencia relativa es igual a la frecuencia absoluta dividido por el número total de datos. Por ejemplo, si la frecuencia absoluta de un valor es 15 y en total hay 100 datos, la frecuencia relativa de ese valor es 0,15 ( $15/100=0,15$ ). En general, se utiliza el símbolo  $h_i$  para representar la frecuencia relativa. Aunque aún no existe un consenso en la comunidad estadística y, por tanto, puede que veas la frecuencia relativa representada con otro símbolo.

FRECUENCIA ACUMULADA: Es la suma de las frecuencias absolutas  $f$ , desde la menor hasta la que corresponde a un determinado valor de la variable. A su vez la frecuencia absoluta es la cantidad de veces que aparece una observación en el conjunto de datos. Evidentemente, la variable de estudio debe ser ordenable. Y como la frecuencia acumulada se obtiene sumando las frecuencias absolutas, resulta que la frecuencia acumulada hasta el último dato, debe coincidir con el total de ellos. De lo contrario hay un error en los cálculos. Por lo general, la frecuencia acumulada se denota como  $F_i$  (o a veces  $n_i$ ), para distinguirla de la frecuencia absoluta  $f_i$  y es importante añadir una columna para ella en la tabla con la que se organizan los datos, conocida como tabla de frecuencias. De esta forma se facilita, entre otras cosas, llevar la cuenta de cuántos datos se contaron hasta determinada observación. A  $F_i$  se la conoce también como frecuencia acumulada absoluta. Si se divide entre el total de datos, se tiene la frecuencia acumulada relativa, cuya suma final debe ser igual a 1.

## PRESENTACIÓN GRAFICA

Es una representación gráfica de un conjunto de datos estadísticos. Es decir, un gráfico estadístico permite resumir una muestra de datos de manera visual. Por lo tanto, un gráfico estadístico sirve para analizar una muestra de datos visualmente. En estadística, los diagramas son muy útiles porque permiten extraer conclusiones de un conjunto de datos sin necesidad de hacer cálculos. La representación gráfica de datos estadísticos tiene como objetivo ofrecer una visión de conjunto del hecho sometido a investigación, de una manera más directa y perceptible que la mera presentación de los datos numéricos. Por ello, las representaciones gráficas son un medio eficaz para análisis de las estadísticas ya que las regularidades se recuerdan y aprecian con más facilidad cuando se examinan gráficamente. Pero no dejan de ser un método auxiliar, ya que la interpretación de un diagrama siempre es subjetiva. También se pueden usar como avance o para llamar la atención del experimentador sobre los elementos más destacables de los datos o poner de manifiesto alguna relación inesperada. También son útiles para expresar los resultados en forma visible después de haber realizado un análisis. una vez hechas dichas observaciones hace falta tener en cuenta aspectos como la frecuencia, la media, la moda o la dispersión de los datos obtenidos. A un nivel estadístico y matemático, denominados gráfica a aquella representación visual a partir de la cual pueden representarse e interpretarse valores generalmente numéricos. De entre las múltiples informaciones extraíbles de la observación de la gráfica podemos encontrar la existencia de relación entre variables y el grado en que se da, las frecuencias o la proporción de aparición de determinados valores. 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, como información a tener en cuenta o como punto de contraste ante la realización de nuevas investigaciones y metaanálisis.

Existen diferentes tipos de graficar para representar valores o cantidades dependiendo de la función que se le determine.

**GRAFICO DE BARRAS:** Uno de los gráficos mas comunes y utilizados, en este se presentan los datos utilizando barras en ejes cartesianos que indican los valores, se emplea para representar la frecuencia de diferentes condiciones o variables.

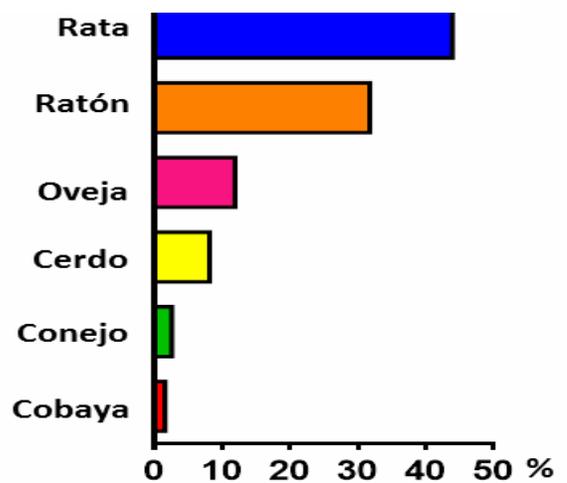
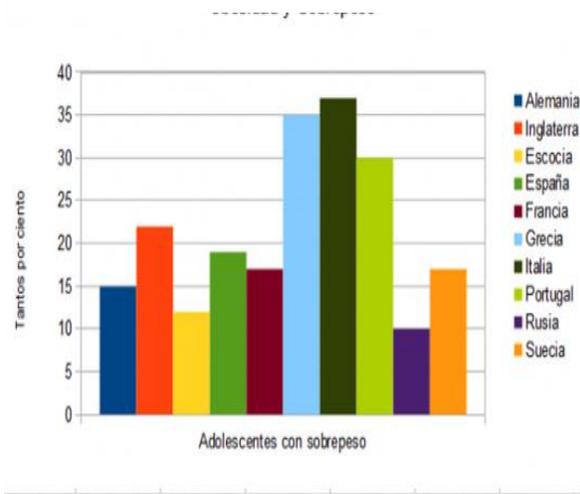
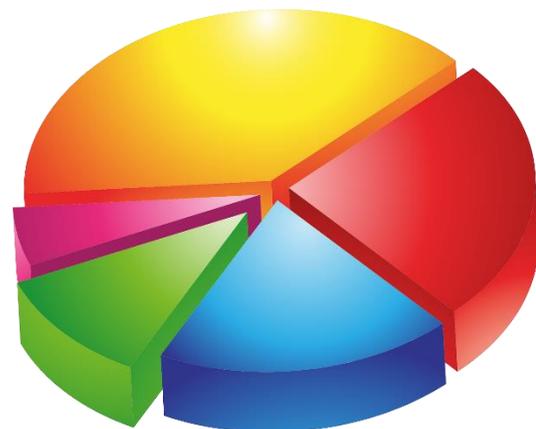
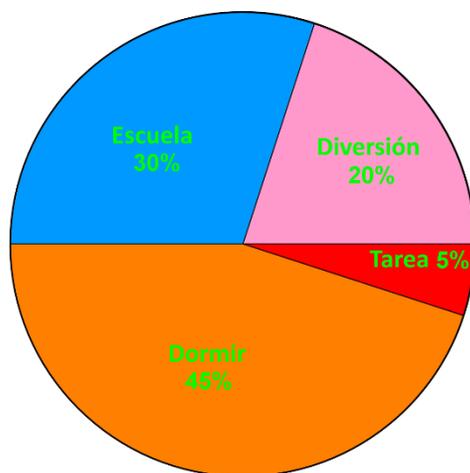
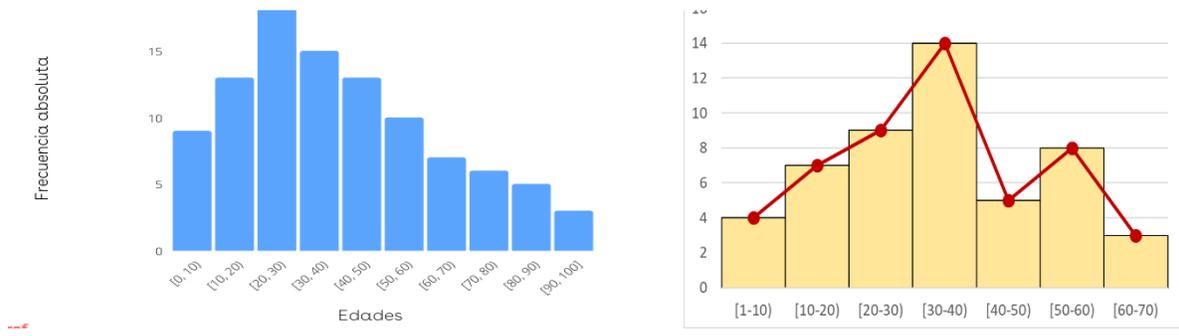


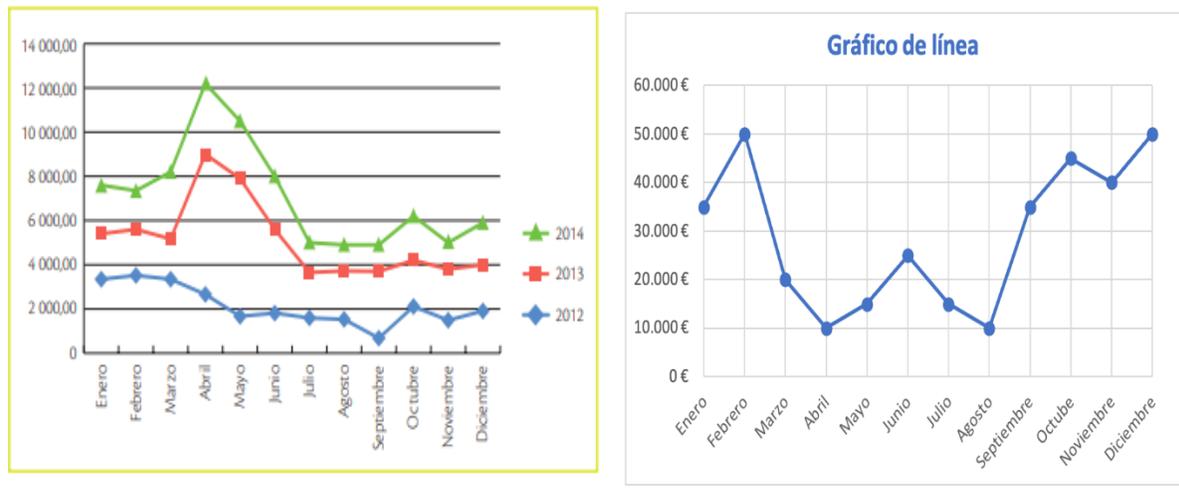
GRAFICO CIRCULAR O POR SECTORES: Se utiliza para representación de datos mediante la división de un círculo en partes que sean necesarias y el tamaño de cada división dependerá del valor a representar. Este tipo de gráfico o diagrama es habitual cuando se está mostrando la proporción de casos dentro del total, utilizando para representarlo valores porcentuales.



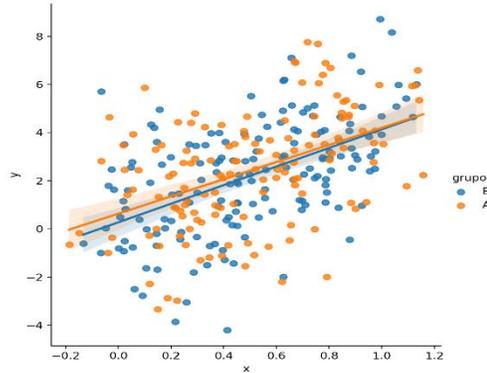
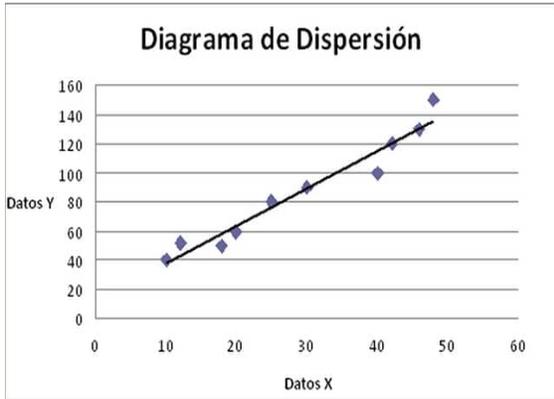
**HISTOGRAMA:** Tiene mucha similitud al gráfico de barras, este tipo de gráfico resulta más importante y fiable, es tipo de diagrama estadístico en el que se representa un conjunto de datos estadísticos mediante barras rectangulares, de manera que cada barra del histograma es proporcional a su frecuencia correspondiente. Los histogramas sirven para representar gráficamente variables continuas, como por ejemplo el peso de una muestra estadística. Además, un histograma permite visualizar rápidamente la forma que tiene una distribución.



**GRÁFICOS DE LÍNEAS:** También llamado diagramas de líneas, se utiliza para representar gráficamente una serie temporal. Consiste en líneas que indica la evolución de uno de los datos a lo largo del tiempo. Suelen representar los períodos de tiempo en el eje X del gráfico de líneas, por otro lado, el eje Y corresponde al valor de la variable que va fluctuando según el período.

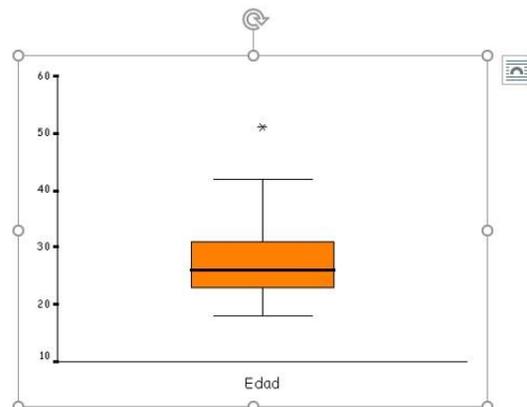
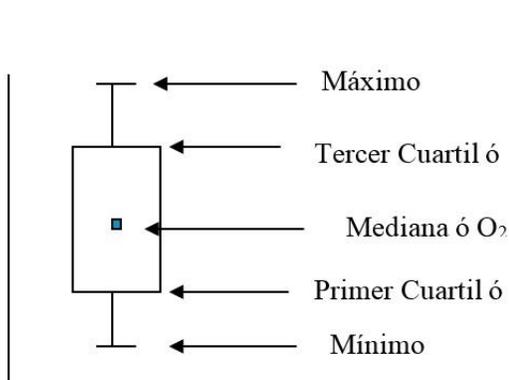


**GRAFICO DE DISPERSIÓN:** Tipo de gráfico en el cual mediante los ejes cartesianos se representa en forma de puntos todos los datos obtenidos mediante la observación. Los ejes x e y muestran cada uno los valores de una variable dependiente y otra independiente o dos variables de la que se esté observando si presentan algún tipo de relación.

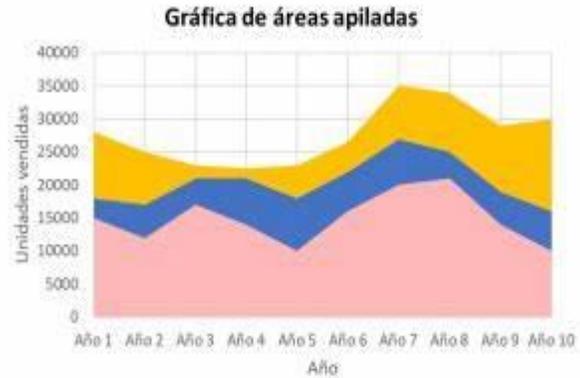


**GRAFICO DE CAJA Y BIGOTES:** Se utilizan para observar la dispersión de datos y como estos agrupan sus valores. Se parte del cálculo de los cuartiles, los cuales son los valores que permiten dividir los datos en cuatro partes iguales. Los llamados bigotes serían la representación gráfica de los valores extremos. Este gráfico es útil a la hora de evaluar intervalos, así como de observar el nivel de dispersión de los datos a partir de los valores de los cuartiles y los valores extremos.

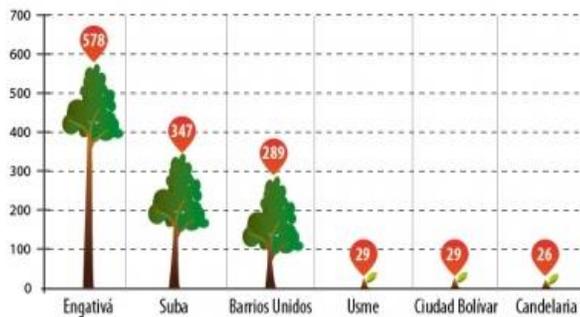
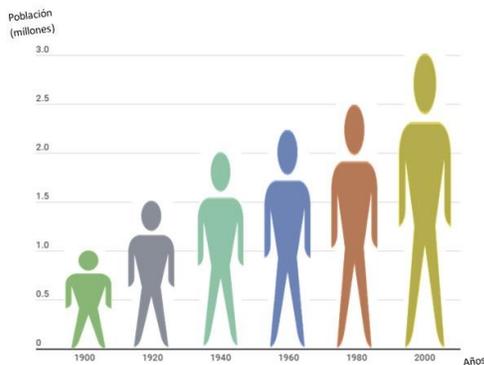
### Gráfica de Cajas y Bigotes.



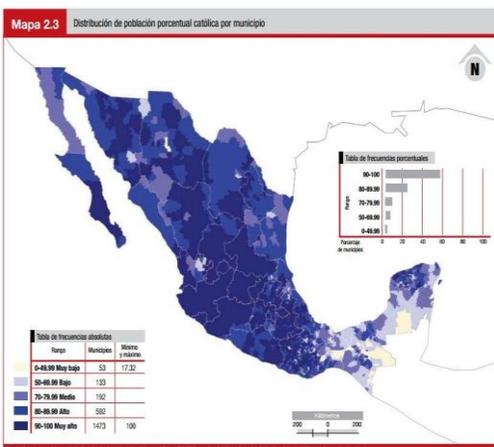
**GRAFICO DE ÁREAS:** A través de él se pueden medir y comparar los valores de diferentes muestras.



**PICTOGRAMA:** Su función es similar al grafico de barras en cuanto a representación de frecuencias, para representación se utiliza elementos propios del tema investigado.

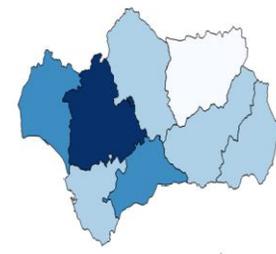


**CARTOGRAMA:** Indican las zonas geográficas en las que aparece con mayor o menor frecuencia un determinado valor de una variable.



Renta Per Capita y Poblacion en Andalucia

CARTOGRAMA  
 Menos de 16000 € / Anuales  
 16000 - 17000 € / Anuales  
 17000 - 18000 € / Anuales  
 Mas de 18000 € / Anuales



Renta Per Capita en Andalucia

ANDALUCIA\_PROVINCIA\_DATOS  
 Menos de 16000 € / Anuales  
 16000 - 17000 € / Anuales  
 17000 - 18000 € / Anuales  
 Mas de 18000 € / Anuales

## MEDIDAS DE TENDENCIA CENTRAL

son métricas estadísticas que indican el valor central de una distribución. Es decir, las medidas de tendencia central sirven para encontrar un valor representativo del centro de un conjunto de datos. Las medidas de tendencia central más utilizadas son la media, la mediana y la moda. Las medidas de tendencia central, como es obvio, persiguen una serie de objetivos que justifican su existencia. En primer lugar, las medidas de tendencia central sirven para conocer en qué lugar se ubica el elemento promedio, o típico del grupo. Imaginemos que queremos saber qué grupo de música es el favorito de la clase. Para ello, podemos utilizar la moda. Asimismo, las medidas de tendencia central sirven para comparar, así como para interpretar los resultados obtenidos con relación a los distintos valores. También, las medidas de tendencia central sirven para comparar e interpretar el valor de una misma variable en distintas ocasiones. Este tipo de medidas sirven para comparar los resultados con otros grupos, atendiendo a estas mismas medidas de tendencia central. Se distinguen dos clases principales de valores promedio:

- Las medidas de posición centrales: medias (aritmética, geométrica, cuadrática, ponderada), mediana y moda.
- Las medidas de posición no centrales: entre las que destacan especialmente los cuantiles.

**MEDIA ARITMÉTICA:** La media aritmética es un valor central característico de un conjunto de datos estadísticos. Para calcular la media aritmética se suman todos los valores y se divide entre el número total de datos. Además, la media aritmética es uno de los principales indicadores que se utilizan para hacer un estudio estadístico de una muestra.

**MEDIANA:** Es el valor del medio de todos los datos ordenados de menor a mayor. Es decir, la mediana divide todo el conjunto de datos ordenados en dos partes iguales.

**MODA:** Es el valor del conjunto de datos que tiene una mayor frecuencia absoluta, es decir, la moda es el valor que más se repite de un conjunto de datos. Por lo tanto, para calcular la moda de un conjunto de datos estadísticos basta con contar el número de veces que aparece cada dato en la muestra, y el dato más repetido será la moda.

## MEDIDAS DE DISPERSIÓN

Las medidas de dispersión pueden definirse como los valores numéricos cuyo objeto es analizar el grado de separación de los valores de una serie estadística con respecto a las medidas de tendencia central consideradas.

Las medidas de dispersión son unas métricas estadísticas que indican la dispersión de un conjunto de datos. Es decir, las medidas de dispersión se utilizan para evaluar cuánto de dispersos están los datos de una muestra. Las medidas de dispersión son fundamentales para comprender la variabilidad de un conjunto de datos.

Estas medidas nos permiten evaluar cuánto se desvían los valores individuales de un conjunto de datos con respecto a su valor central o promedio. En este artículo, vamos a desentrañar las principales medidas de dispersión utilizadas en estadística, incluyendo el rango, la varianza, la desviación estándar y más. Además, discutiremos su importancia y aplicaciones en diversos campos.

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. Se trata de parámetros más técnicos y utilizados en estudios específicos, y entre ellas se encuentran los coeficientes de apertura, el recorrido relativo, el coeficiente de variación (índice de dispersión de Pearson) y el índice de dispersión mediana.

## TEOREMA DE CHEBYSHEV

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.

En el estudio de la teoría de la probabilidad ocurre que, si se conoce la función de distribución de una variable aleatoria  $X$ , se puede calcular su valor esperado —o esperanza matemática  $E(X)$ — y su varianza  $\text{Var}(X)$ , siempre y cuando dichas cantidades existan.

Es decir, conociendo  $E(X)$  y  $\text{Var}(X)$  no necesariamente se puede obtener la función de distribución de  $X$ , por lo cual cantidades como  $P(|X|>k)$  para algún  $k>0$ , son muy difíciles de obtener. Pero gracias a la desigualdad de Chebyshev es posible hacer una estimación de la probabilidad de la variable aleatoria. El teorema de Chebyshev nos dice que, si tenemos una variable aleatoria  $X$  sobre un espacio muestral  $S$  con una función de probabilidad  $p$ , y si  $k>0$ .

Acotamiento de probabilidades: Esta es la aplicación más común y se utiliza para dar una cota superior para  $P(|X-E(X)|\geq k)$  donde  $k>0$ , solo con la varianza y la esperanza de la variable aleatoria  $X$ , sin conocer la función de probabilidad.

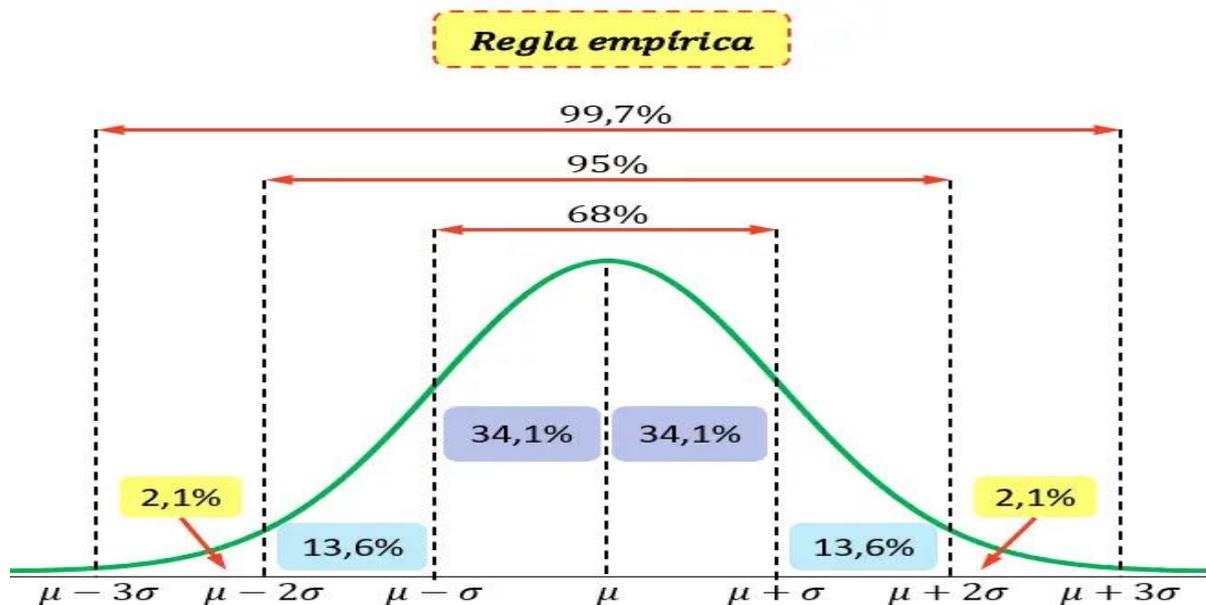
Tamaño de muestra: En términos de la varianza, la desigualdad de Chebyshev nos permite encontrar un tamaño de muestra  $n$  que es suficiente para garantizar que la probabilidad de que  $|\bar{X}_n-\mu|\geq k$  ocurra sea tan pequeña como se desee, lo cual permite tener una aproximación a la media.

## REGLA EMPÍRICA

La regla empírica, a la que también se le conoce como la regla 68,5-95-99,7, constituye una manera útil de analizar datos estadísticos. Sin embargo, solo funciona para una distribución normal (la campana de Gauss) y solo es posible producir estimaciones. Una regla empírica es el resultado de la experiencia práctica y la observación de la vida real. Por ejemplo, se puede saber que especies de aves se pueden observar en ciertos lugares en cada época del año y de esa observación se puede establecer una “regla” que describa los ciclos de vida de estas aves. En estadística, la regla empírica hace referencia a la forma de agrupación de las observaciones alrededor de un valor central, la media o promedio, en unidades de desviación estándar. La regla empírica es una manera rápida y práctica de analizar datos estadísticos, siendo cada vez más confiable en la medida en que la distribución se acerca a la simetría.

Así pues, la regla empírica establece que:

- El 68% de los valores se encuentran a una desviación estándar de la media.
- El 95% de los valores se encuentran a dos desviaciones estándar de la media.
- El 99,7% de los valores se encuentran a tres desviaciones estándar de la media.



## TEORÍA DE LA PROBABILIDAD

Es un conjunto de reglas y propiedades que sirven para calcular la probabilidad de un fenómeno aleatorio. Así pues, la teoría de la probabilidad permite saber qué resultado de un experimento aleatorio es más probable que ocurra.

De modo que la teoría de la probabilidad son un conjunto de leyes que nos permiten determinar cuánto de probable es que ocurra un fenómeno aleatorio. Se define como cálculo de probabilidad al conjunto de reglas que permiten determinar si un fenómeno ha de producirse, fundando la suposición en el cálculo, las estadísticas o la teoría. El valor más pequeño que puede tener la probabilidad de ocurrencia de un evento es igual a 0, el cual indica que el evento es imposible, y el valor mayor es 1, que indica que el evento ciertamente ocurrirá. Dos o más eventos son independientes cuando la ocurrencia o no-ocurrencia de un evento no tiene efecto sobre la probabilidad de ocurrencia del otro evento (o eventos). Un caso típico de eventos independiente es el muestreo con reposición, es decir, una vez tomada la muestra se regresa de nuevo a la población donde se obtuvo.

Es una distribución de probabilidad continua que es tanto simétrica como mesocúrtica. La curva que representa la distribución de probabilidad normal se describe generalmente como en forma de campana.

En teoría de probabilidades, un suceso (o evento) es cada uno de los posibles resultados de un experimento aleatorio. Por lo tanto, la probabilidad de un suceso es un valor que indica cuánto de probable es que ocurra un resultado. En cualquier caso, sin desviarnos del concepto de teoría de la probabilidad, diremos que está formada por un conjunto de técnicas que nos permiten asignar un número a la posibilidad de que un evento ocurra.

## ENFOQUES DE PROBABILIDAD

La probabilidad, mide la frecuencia con la cual se obtiene un resultado en oportunidad de la realización de un experimento sobre el cual se conocen todos los resultados posibles gracias a las condiciones de estabilidad que el contexto supone de antemano. Expresado matemáticamente, es igual al número de formas que un evento específico puede ocurrir, dividido por el número total de posibles eventos. La probabilidad está absolutamente inmersa en nuestro día a día como parte integrante de una sociedad y comunidad determinada, ya que en el análisis de riesgos y en el comercio de materias primas, la probabilidad, tiene una incidencia y una importancia vital.

Tipos de enfoques de Probabilidad:

**CLÁSICO:** Los resultados de un experimento son igualmente viables, es decir, tienen teóricamente las mismas posibilidades de ocurrir. La probabilidad clásica es un número entre 0 y 1. Cuanto más probable de que ocurra un evento, mayor será la probabilidad clásica, por contra, cuanto menos probable sea de que suceda un evento, menor será el valor de la probabilidad clásica. A diferencia de otros tipos de probabilidades, no hace falta hacer ningún experimento para hallar la probabilidad clásica de un evento, sino que se trata de un cálculo teórico.

**RELATIVO:** La probabilidad de que un evento suceda se determina observando eventos similares en el pasado. Este método utiliza la frecuencia relativa de las presentaciones pasadas de un evento como una probabilidad. Determinamos qué tan frecuente ha sucedido algo en el pasado y usamos esa cifra para predecir la probabilidad de que suceda de nuevo en el futuro.

**SUBJETIVO:** Se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que se tenga disponible. Esa evidencia puede presentarse en forma de frecuencia relativa de presentación de eventos pasados o puede tratarse simplemente de una creencia meditada.

## ESPACIO MUESTRAL

Conjunto de todos los resultados posibles que pueden surgir de un experimento aleatorio. Comprender el espacio muestral es esencial para el análisis de situaciones donde la incertidumbre y el azar juegan un papel crucial. Desde lanzar una moneda hasta realizar encuestas en una población, el espacio muestral proporciona la base para calcular probabilidades y evaluar riesgos. El espacio muestral se puede definir como el conjunto total de todos los resultados posibles provenientes de un experimento aleatorio. Cada uno de estos resultados es conocido como un evento. Se simboliza comúnmente como  $E$  o  $\Omega$ . Comprender el espacio muestral es crucial para la recolección de datos y la ejecución de análisis estadísticos.

La principal razón de esto es que el espacio muestral permite identificar todas las posibilidades que pueden surgir de un experimento. Esto, a su vez, facilita la estimación de probabilidades, la que es fundamental para tomar decisiones informadas. Por lo tanto, cualquier análisis estadístico debe comenzar delimitando adecuadamente el espacio muestral. Dentro de un espacio muestral, se pueden identificar subconjuntos que son conocidos como sucesos. Un suceso es cualquier colección de uno o más resultados del espacio muestral. Por ejemplo, si estamos lanzando un dado, un suceso específico puede ser obtener un número impar. Este suceso se puede denotar como  $B = \{1, 3, 5\}$ , que son los resultados favorables dentro del espacio muestral.

El espacio muestral es un aspecto esencial en la probabilidad y la estadística, proporcionando un contexto claro para el análisis y la toma de decisiones informadas. Desde simples experimentos como lanzar una moneda o un dado, hasta aplicaciones complejas en finanzas y análisis de riesgos, el espacio muestral ofrece una estructura que nos ayuda a entender la incertidumbre. Al dominar el concepto de espacio muestral, los individuos pueden realizar evaluaciones más efectivas en diversas áreas, lo que resulta en una comprensión más profunda de la naturaleza del azar y la probabilidad.

## EVENTOS SIMPLES Y COMPLEJOS

Un evento es una colección particular de resultados. Los eventos usualmente son descritos usando una característica común de los resultados. Un evento compuesto, también llamado suceso compuesto, es un conjunto de posibles resultados de un experimento aleatorio. Por lo tanto, un evento compuesto es un conjunto de eventos simples y un subconjunto del espacio muestral. Un evento simple (o suceso simple) es un solo resultado de un experimento aleatorio, en cambio, un evento compuesto (o suceso compuesto) es un conjunto de dos o más posibles resultados. Es decir, un evento compuesto es una combinación de eventos simples. Por ejemplo, en el experimento del lanzamiento de un dado, obtener la cara con el número 1 se trata de un evento simple. Por contra, sacar un número menor que 6 es un evento compuesto por cinco eventos simples (1, 2, 3, 4 y 5). 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). La probabilidad de un evento  $E$  normalmente se escribe  $P(E)$ . Es una práctica común en probabilidad, como en las fracciones en general, simplificar una probabilidad en sus términos más bajos para que sea más fácil tener idea de qué tan grande es. A menos que exista una razón para no hacerlo, expresaremos todas las probabilidades en sus términos más bajos. La probabilidad nos ayuda a comprender la posibilidad de que ocurra un evento. La probabilidad generalmente se escribe como una fracción, decimal o porcentaje con los eventos favorables escritos como numerador y la cantidad total de resultados escritos como denominador. La probabilidad simple expresa la probabilidad de que ocurra un evento y, a menudo, se expresa visualmente usando monedas, dados, canicas o ruleta. La probabilidad compuesta describe las posibilidades de que ocurra más de un evento separado, por ejemplo, lanzar cara a una moneda y sacar un 7 de una baraja de cartas estándar. Los problemas de probabilidad compuesta a menudo se resuelven utilizando el cálculo matemático de multiplicar los resultados de los diversos resultados de probabilidad simple, usando una lista para mostrar los diferentes resultados, o usando un diagrama de árbol para modelar los diversos resultados. Conocer la probabilidad simple y compuesta le ayuda a dar los primeros pasos hacia una mayor comprensión de la probabilidad y la estadística.

## LEYES DE PROBABILIDAD

La probabilidad estadística es una rama de las matemáticas y la estadística que se enfoca en el estudio y la cuantificación de la incertidumbre en los eventos y fenómenos del mundo real. Se utiliza para analizar y comprender la aleatoriedad y la variabilidad en los datos y en los procesos naturales o sociales. Existen diversas formas como método abstracto, como la teoría de Dempster y la teoría de la relatividad numérica, esta última con un alto grado de aceptación si se toma en cuenta que disminuye considerablemente las posibilidades hasta un nivel mínimo ya que somete a todas las antiguas reglas a una simple ley de relatividad. La probabilidad de un evento se denota con la letra  $p$  y se expresa en términos de una fracción y no en porcentajes, por lo que el valor de  $p$  cae entre 0 y 1. Por otra parte, la probabilidad de que un evento no ocurra equivale a 1 menos el valor de  $p$  y se denota con la letra  $q$ .

Regla de la adición: 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 multiplicación: 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, es decir,  $P(A) = 1$ . Para aplicar la regla de Laplace es necesario que los experimentos den lugar a sucesos equiprobables, es decir, que todos tengan o posean la misma probabilidad.

Se puede decir razonablemente que el descubrimiento de métodos rigurosos para calcular y combinar los cálculos de probabilidad ha tenido un profundo efecto en la sociedad moderna. Por consiguiente, puede ser de alguna importancia para la mayoría de los ciudadanos entender cómo se calculan los pronósticos y las probabilidades, y cómo contribuyen a la reputación y a las decisiones, especialmente en una democracia.

## TABLAS DE CONTINGENCIA

Una tabla de contingencia es un conjunto de filas y columnas que sirve para clasificar los datos recopilados. En estadística, las tablas de contingencia se usan para calcular probabilidades y analizar la información recopilada. Hacer una tabla de contingencia es muy simple, primero tienes que realizar una investigación y recoger datos numéricos sobre el tema en cuestión. Después solamente tienes que ordenar los datos según su naturaleza. Lógicamente, debes tener muy claro las variables con las que quieres clasificar los datos para poder hacer una interpretación correcta. Principalmente, una tabla de contingencia sirve para poder analizar los datos de una variable categórica. Resulta complicado hacer un estudio estadístico de este tipo de variables, pero utilizando las tablas de contingencia se pueden ordenar los datos de manera clara y, en consecuencia, es más fácil interpretarlos. Además, las tablas de contingencia facilitan deducir las relaciones entre las variables de una investigación con solo observarlas, ya que presentan los datos de manera ordenada y limpia. 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. La tabla sólo mostrará los encuestados que respondieron ambas preguntas, lo que significa que las frecuencias mostradas pueden diferir de una tabla de frecuencias estándar. La tabla de contingencia generalmente se realiza en datos categóricos, es decir que se pueden dividir en grupos mutuamente excluyentes. Un ejemplo de datos categóricos es la región de ventas de un producto. Típicamente, la región se puede dividir en categorías como área geográfica (norte, sur, noreste, oeste, etc.) o estado. Es importante recordar que los datos categóricos no pueden pertenecer a más de una categoría. Uno de los principales usos de una tabla de contingencia es analizar la relación que existe entre los datos, las cuales no son fáciles de identificar. Esto permite realizar un estudio de mercado efectivo, examinando la información recolectada en una encuesta.

## TEOREMA DE BAYES

En la teoría de la probabilidad, el teorema de Bayes es una ley que sirve para calcular la probabilidad de un evento cuando se conoce información a priori sobre dicho suceso. En concreto, el teorema de Bayes relaciona matemáticamente la probabilidad del evento A dado el evento B con la probabilidad de B dado A. El teorema de Bayes tiene muchas aplicaciones, por ejemplo, se utiliza en medicina, en economía o en tecnología para calcular las probabilidades de algunos eventos que están condicionadas por otros eventos. Podemos calcular la probabilidad de un suceso A, sabiendo además que ese A cumple cierta característica que condiciona su probabilidad. Para calcular la probabilidad tal como la definió Bayes en este tipo de sucesos, necesitamos una fórmula. La fórmula se define matemáticamente como: Donde B es el suceso sobre el que tenemos información previa y A(n) son los distintos sucesos condicionados. En la parte del numerador tenemos la probabilidad condicionada, y en la parte de abajo la probabilidad total. En cualquier caso, aunque la fórmula parezca un poco abstracta, es muy sencilla. Para demostrarlo, utilizaremos un ejemplo en el que en lugar de A (1), A (2) y A (3), utilizaremos directamente A, B y C. Las aplicaciones del teorema de Bayes son muchas, algunas de ellas son:

**Pruebas médicas:** el teorema de Bayes se usa frecuentemente en medicina para determinar la probabilidad de acierto de las pruebas diagnósticas. Por ejemplo, en el caso de una prueba de detección del VIH, se puede utilizar el teorema para calcular la probabilidad de que una persona tenga realmente el virus dado un resultado positivo en la prueba.

**Análisis financiero:** en finanzas se usa el teorema de Bayes para calcular la probabilidad de que ciertos eventos económicos, como un aumento o una bajada en el valor de las acciones, ocurran dado un conjunto de variables económicas.

**Estudio de mercado:** el teorema de Bayes permite determinar, por ejemplo, la probabilidad de que una persona compre un producto cuando ha visto un anuncio de ese producto.

**Pronóstico del tiempo:** los modelos meteorológicos también utilizan el teorema de Bayes para encontrar la probabilidad de que dada una predicción meteorológica en función de los datos observados esta se acabe cumpliendo. De esta forma se mejora la precisión de las previsiones climática

## BIBLIOGRAFÍA

- 1.- ALEA, V. et al. (2006) Estadística Aplicada a les Ciències Econòmiques y Sociales. Barcelona: Edició McGraw-Hill EUB.
- 2.- DURA Peiró, J. M. y López CUñAT, J.M. (2006) Fundamentos de Estadística. Estadística Descriptiva y Modelos Probabilísticos para la Inferencia. Madrid: Ariel Editorial.
- 3.- FREEDMAN, D., et al. (2015) Estadística. Barcelona: A. Bosch Ed.
- 4.- FREIXA, M., et al. (2012) Análisis exploratorio de datos: Nuevas técnicas estadísticas. Barcelona: PPU.
- 5.- -MARTÍN PLIEGO, F. y RUIZ-MAYA, L. (1995) Estadística II: Inferencia. Madrid: AC
- 6.- -MENDENHALL, W., et al. (1994) Estadística Matemática con Aplicaciones. México: Grupo Editorial Iberoamérica.
- 7.- -Seymour Lipschutz Ph.D. 2000 Problemas Resueltos de Matemática Discretas. MCGRAW-HILL.
- 8.- TANUR, J. (1992) La Estadística, una Guía de lo Desconocido. Madrid: Alianza Editorial.
- 9.- URIEL, E. y MUñIZ, M. (1988) Estadística Económica y Empresarial. Teoría y ejercicios. Madrid: AC.
- 10.- Anderson D., Sweeney D., Williams T. Estadística para la administración y economía. Décima edición. Cengage Learning. 2008