CAMPUS TAPACHULA

PSU – 314/2012

MAESTRÍA EN ADMINISTRACIÓN EN SISTEMAS DE SALUD.

MATERÍA: TENDENCIAS Y SISTEMAS DE SALUD EN MÉXICO.

DOCENTE: DRA. MARIA CECILIA ZAMORANO RODRÍGUEZ.

TEMA: ENSAYO DE LA UNIDAD I ESTADÍSTICA DESCRIPTIVA Y LA UNIDAD II TEORÍA DE LA PROBALIDAD.

ALUMNO: LIC. JOEL ANTULIO GÓMEZ KELLER.

CUATRIMESTRE 1 VÍA ONLINE.

TAPACHULA DE CÓRDOVA Y ORDOÑEZ, CHIAPAS A 31 DE OCTUBRE DEL 2024.

INTRODUCCIÓN.

La estadística es una rama de las matemáticas que se encarga de recolectar, organizar, presentar, analizar e interpretar datos para la toma de decisiones. Esta ciencia se divide en dos grandes áreas: la estadística descriptiva y la estadística inferencial. La primera se enfoca en resumir y describir las características de los datos recopilados, mientras que la segunda se orienta a hacer inferencias o predicciones sobre una población a partir de una muestra. Además, la teoría de la probabilidad proporciona las bases matemáticas para el análisis de la incertidumbre y la toma de decisiones en situaciones donde los resultados no son deterministas.

La estadística es una disciplina fundamental en el análisis de datos, ya que facilita la comprensión y síntesis de grandes volúmenes de información. Esta se divide en estadística descriptiva y estadística inferencial. La primera se enfoca en organizar, resumir y describir datos mediante gráficos y medidas numéricas, mientras que la segunda se encarga de realizar inferencias sobre una población con base en una muestra.

Por otro lado, la teoría de la probabilidad proporciona una base matemática para analizar fenómenos aleatorios y cuantificar la incertidumbre. Ambas áreas son fundamentales para la toma de decisiones en diversos campos como la economía, las ciencias sociales, la ingeniería y la investigación científica.

INSTRUCCIONES: Elabore un ensayo con las siguientes características.

ENSAYO: El ensayo es un tipo de texto que analiza, evalúa o interpreta un tema determinado, ya sea de manera oficial o libre. Su principal característica es que se trata de un texto en el que el escritor cuenta con total libertad para organizar el contenido y la información. Es, además, considerado un género literario, como puede ser la poesía, la ficción o el drama. Todos los ensayos suelen presentar una estructura bastante clara, la cual se articula en torno a una introducción, un desarrollo y una conclusión.

CRITERIOS A EVALUAR.

- Presentación o portada.
- Originalidad en la redacción y las ideas, los trabajos copiados y bajados de internet serán rechazados.
- Concordancia con los temas de la antología.
- Contenido, 2 cuartilla como mínimo, 4 máximo.
- Uso de mayúsculas y minúsculas.
- Cuidar ortografía.
- Tipo de letra: Arial.
- Tamaño de letra: 12 para títulos y 11 para texto Interlineado: 1.5 Márgenes: Superior: 2.5, Inferior: 2.5, Izquierdo: 3, Derecho: 2.5
- Incluir citas, fuentes y referencias.
- Guardar y subir formato en PDF.

NO SE CONCEDERAN PRORROGAS.

RESUMEN.

En resumen, la combinación de estadística descriptiva y teoría de la probabilidad es esencial para convertir datos en conocimiento útil, guiando decisiones en áreas tan diversas como la ciencia, la economía, la ingeniería y las ciencias sociales. Estas disciplinas ofrecen herramientas sólidas para enfrentar desafíos complejos y comprender fenómenos inciertos, facilitando la toma de decisiones fundamentadas y racionales.

La estadística descriptiva se encarga de recolectar, organizar, presentar, analizar e interpretar datos, permitiendo resumir grandes volúmenes de información en formatos claros y comprensibles. A través del procesamiento estadístico de datos, que incluye la recolección, organización y presentación gráfica, se facilita la interpretación de la información mediante herramientas como tablas, histogramas, diagramas de barras o gráficos circulares.

Las distribuciones de frecuencias ayudan a estructurar los datos en categorías o intervalos, mostrando la cantidad de veces que ocurre cada valor. Las medidas de tendencia central media, mediana y moda; ofrecen un valor representativo que describe el comportamiento general del conjunto de datos. A su vez, las medidas de dispersión rango, varianza y desviación estándar; que permiten evaluar la variabilidad o dispersión de los datos respecto a su centro.

Por otro lado, la teoría de la probabilidad proporciona un marco para analizar eventos aleatorios e incertidumbre. Los enfoques de probabilidad ya sea de frecuentista, clásico y subjetivo, ayudan a cuantificar la posibilidad de que ocurra un evento. El espacio muestral y los eventos simples y compuestos son conceptos fundamentales para entender las combinaciones posibles en experimentos aleatorios. Las leyes de probabilidad como la adición y multiplicación y además las tablas de contingencia permiten analizar la relación entre eventos y calcular probabilidades condicionadas.

En el teorema de Bayes es una herramienta esencial para actualizar probabilidades cuando se dispone de nueva información. Finalmente, el teorema de Tchebyshev y la regla empírica ofrecen formas de estimar la dispersión de los datos, ya sea en distribuciones generales o normales.

En conjunto, la estadística descriptiva y la probabilidad proporcionan herramientas poderosas para sintetizar información y manejar la incertidumbre, facilitando la toma de decisiones fundamentadas en una amplia variedad de disciplinas.

ENSAYO.

1. Procesamiento Estadístico de Datos.

1.1 El procesamiento de datos comprende varias fases esenciales:

Recolección de datos: Es la etapa en la que se obtienen los datos mediante encuestas, censos, o registros. Es crucial que la recolección se haga de manera precisa para evitar sesgos.

Organización de los datos: Consiste en estructurar la información en tablas o cuadros estadísticos, agrupando los valores para facilitar su análisis.

Presentación gráfica: Permite visualizar los datos de forma clara e intuitiva mediante herramientas como histogramas, gráficos de barras y diagramas de pastel.

Análisis e interpretación: En esta fase se extraen conclusiones y patrones relevantes a partir de los datos organizados, facilitando la toma de decisiones.

1.2 Distribuciones de Frecuencias.

Las distribuciones de frecuencias organizan los datos en intervalos o categorías, mostrando la cantidad de veces que aparece cada valor o grupo de valores. Estas se clasifican en:

Frecuencia absoluta: Número de veces que aparece un valor específico.

Frecuencia relativa: Proporción que representa cada valor respecto al total.

Frecuencia acumulada: Suma de las frecuencias absolutas hasta cierto punto en la distribución.

Este tipo de organización permite resumir datos complejos y proporcionar una visión clara de su distribución.

Gráfico de barras.

El gráfico de barras utiliza columnas o barras para representar valores discretos o categóricos. Cada barra representa una categoría y su altura refleja la magnitud de los datos.

Uso: Comparar categorías o valores.

Ventajas: Fácil interpretación y comparación.

Desventajas: No es adecuado para mostrar tendencias continuas.

Ejemplo: Comparación de ventas mensuales por productos.

Gráfico circular o por sectores.

Un gráfico circular divide un círculo en secciones, donde cada segmento representa una parte proporcional del total.

Uso: Mostrar proporciones o porcentajes.

Ventajas: Visualmente atractivo.

Desventajas: No es adecuado para comparaciones precisas entre categorías con valores muy similares.

Ejemplo: Distribución del presupuesto en distintas áreas de una empresa.

Histograma.

El histograma es similar al gráfico de barras, pero se utiliza para mostrar la distribución de frecuencias de datos numéricos continuos. Las barras se agrupan en intervalos.

Uso: Analizar la distribución de datos.

Ventajas: Ideal para identificar patrones como la simetría o la asimetría en los datos.

Desventajas: La elección del número de intervalos puede afectar la interpretación.

Ejemplo: Distribución de edades de los estudiantes en una universidad.

Gráfico de líneas.

Este gráfico conecta puntos mediante líneas para mostrar cambios a lo largo del tiempo o en secuencias ordenadas.

Uso: Mostrar tendencias o evolución de datos.

Ventajas: Ideal para visualizar cambios a lo largo del tiempo.

Desventajas: No adecuado para comparar categorías aisladas.

Ejemplo: Evolución de la temperatura durante el año.

Gráfico de dispersión.

Este gráfico representa puntos en un plano cartesiano para mostrar la relación entre dos variables.

Uso: Identificar correlaciones o patrones.

Ventajas: Ayuda a detectar relaciones lineales o no lineales.

Desventajas: Puede ser difícil de interpretar con muchos puntos superpuestos.

Ejemplo: Relación entre el peso y la altura de un grupo de personas.

Gráfico de caja y bigotes (Boxplot).

Este gráfico muestra la distribución de los datos a través de sus cuartiles, indicando la mediana y posibles valores atípicos.

Uso: Resumir y comparar distribuciones de datos.

Ventajas: Identifica outliers y dispersión.

Desventajas: Puede ser difícil de interpretar sin conocimiento previo.

Ejemplo: Análisis de los tiempos de entrega de diferentes repartidores.

Gráfico de áreas.

Este tipo de gráfico es una variación del gráfico de líneas, donde las áreas bajo las líneas se rellenan, enfatizando la magnitud de los datos.

Uso: Mostrar tendencias acumulativas o proporciones.

Ventajas: Visualmente claro para representar acumulaciones.

Desventajas: Puede ser confuso si hay muchas áreas superpuestas.

Ejemplo: Evolución del mercado de diferentes productos a lo largo del tiempo.

Pictograma.

El pictograma utiliza íconos o imágenes para representar valores, de manera proporcional a las cantidades que se desean mostrar.

Uso: Visualizar datos de forma creativa y atractiva.

Ventajas: Fácil de entender para el público en general.

Desventajas: Puede perder precisión si los íconos no se escalan correctamente.

Ejemplo: Representación del número de vehículos vendidos utilizando íconos de automóviles.

Cartograma.

Un cartograma es un mapa donde el tamaño de las áreas (como países o regiones) se distorsiona según una variable específica.

Uso: Mostrar datos geográficos de forma proporcional a variables cuantitativas.

Ventajas: Permite ver patrones geográficos de forma clara.

Desventajas: La distorsión puede dificultar la lectura del mapa.

Ejemplo: Cartograma que muestra la población de los países ajustando el tamaño de cada territorio.

1.3 Presentación Gráfica.

La presentación gráfica es fundamental para interpretar los datos de manera eficiente. Las herramientas más utilizadas son:

Histogramas: Representan datos continuos en intervalos, permitiendo visualizar la forma de la distribución.

Gráficos de barras: Utilizados para datos discretos o categóricos.

Gráficos circulares: Muestran proporciones de diferentes categorías.

Polígonos de frecuencia: Unen los puntos medios de las clases en una curva que describe la distribución.

1.4 Medidas de Tendencia Central.

Estas medidas resumen un conjunto de datos con un valor representativo que indica su "centro". Las principales son:

Media: Promedio aritmético de los datos.

Mediana: Valor central en un conjunto de datos ordenados.

Moda: Valor o valores que ocurren con mayor frecuencia.

Estas medidas ofrecen una síntesis numérica que facilita la comprensión del comportamiento de los datos.

$$\overline{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_n x_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum f_i x_i}{\sum f_i} \text{ con } i = 1, 2, \ldots, n$$

$$\overline{x} = \frac{p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum p_i x_i}{\sum p_i} \text{ con } i = 1, 2, \ldots, n$$

1.5 Medidas de Dispersión.

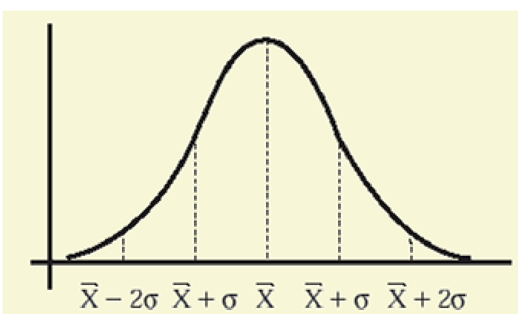
Las medidas de dispersión reflejan la variabilidad de los datos alrededor de la media. Entre las más importantes se encuentran:

Rango: Diferencia entre el valor máximo y mínimo.

Varianza: Promedio de las desviaciones al cuadrado respecto a la media.

Desviación estándar: Raíz cuadrada de la varianza, que mide la dispersión en las mismas unidades que los datos.

Estas medidas complementan las de tendencia central al proporcionar información sobre la heterogeneidad de los datos.



$$V = \frac{f_1(x_1 - \overline{x})^2 + f_2(x_2 - \overline{x})^2 + \ldots + f_n(x_n - \overline{x})^2}{f_1 + f_2 + \ldots + f_n} =$$
$$= \frac{\sum f_i(x_i - \overline{x})^2}{\sum f_i} \text{ con } i = 1, 2, 3, \ldots, n.$$

Por su parte, la desviación típica, simbolizada por s, se define sencillamente como la raíz cuadrada de la varianza:

$$\sigma = \sqrt{V}$$

Por lo tanto, se tiene que:

$$\sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

1.6 Teorema de Tchebyshev y Regla Empírica.

Teorema de Tchebyshev.

Este teorema establece que, para cualquier distribución, al menos 1− 1/k2 de los datos se encuentran dentro de K desviaciones estándar de la media, donde k >1. Esto es útil para garantizar una estimación de dispersión en datos que no siguen una distribución normal.

Regla Empírica

Aplica únicamente a datos con distribución normal. Según esta regla:

El 68% de los datos se encuentran dentro de 1 desviación estándar de la media.

El 95% dentro de 2 desviaciones estándar.

El 99.7% dentro de 3 desviaciones estándar.

Demostración.

Primero notemos lo siguiente:

$$E(S_n) = \frac{E(X_1) + E(X_2) + \cdots + E(X_n)}{n} = \frac{n\mu}{n} = \mu$$

Como X1, X2,…, Xn son independientes, se deduce que:

$$Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$$

Por lo tanto, es posible afirmar lo siguiente:

$$Var(S_n) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{Var(X_1 + X_2 + \cdots + X_n)}{n^2}$$

$$\frac{Var(X_1 + X_2 + \cdots + X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

1.7 Regla empiridca.

Es una herramienta fundamental para el análisis de datos, permitiendo describir, interpretar y predecir fenómenos en distintas disciplinas. Entre las diversas herramientas utilizadas en estadística, la regla empírica o regla del 68-95-99.7 es esencial para entender cómo se distribuyen los datos en una curva normal. Esta regla proporciona una forma intuitiva de interpretar los datos en relación con su media y su desviación estándar. En este ensayo, exploraremos en qué consiste la regla empírica, su importancia, y algunas de sus aplicaciones prácticas.

Además se aplica a distribuciones normales, también conocidas como curvas de Gauss o campanas de probabilidad. Esta regla establece que, en una distribución normal:

El 68% de los datos se encuentra dentro de una desviación estándar de la media.

El 95% de los datos se encuentra dentro de dos desviaciones estándar de la media.

El 99.7% de los datos se encuentra dentro de tres desviaciones estándar de la media.

Esto significa que, en condiciones normales, la mayor parte de los valores observados estarán relativamente cerca de la media, mientras que valores más extremos serán cada vez menos frecuentes.

La importancia que tiene la Regla Empírica es útil porque permite a los investigadores y analistas identificar qué tan dispersos están los datos en torno a su media. Además, facilita detectar anomalías o datos atípicos (outliers), ya que los valores que se encuentren más allá de tres desviaciones estándar son inusuales en una distribución normal. De esta manera, se convierte en una referencia rápida para evaluar la variabilidad de los datos y la probabilidad de ciertos eventos.

2.1 Teoría de la Probabilidad.

La probabilidad es una rama de la matemática que estudia la incertidumbre de eventos aleatorios. Existen tres enfoques principales:

Frecuentista: Define la probabilidad como la proporción de veces que ocurre un evento en una serie de experimentos repetidos.

Clásico: Calcula la probabilidad como el cociente entre los casos favorables y el total de casos posibles, suponiendo que todos son igualmente probables.

Subjetivo: Considera la probabilidad como una medida del grado de certeza o creencia personal sobre la ocurrencia de un evento.

El enfoque clásico de la administración se centra en la eficiencia, la división del trabajo y la racionalización de los procesos. Una de sus principales ideas es que las organizaciones funcionan mejor si se estructuran como sistemas mecánicos, donde cada persona tiene un rol claramente definido.

El enfoque de frecuencia relativa es una de las formas fundamentales para entender la probabilidad. Se basa en la observación empírica de cuántas veces ocurre un evento en un conjunto grande de ensayos o experimentos repetidos. A diferencia de enfoques más subjetivos o teóricos, esta perspectiva pone énfasis en la repetición y la recolección de datos como fuente de información. En este ensayo, explicaremos los principios del enfoque de frecuencia relativa, su importancia, un ejemplo práctico, y las limitaciones que pueden presentarse al aplicarlo.

Ejemplo del enfoque de frecuencia relativa.

Imaginemos un ejemplo simple: queremos calcular la probabilidad de que al lanzar una moneda justa salga cara. En lugar de asumir que la probabilidad es exactamente 0.5, realizamos una serie de lanzamientos y contamos las veces que obtenemos cara.

Supongamos que realizamos el experimento 100 veces y obtenemos cara en 48 ocasiones. La frecuencia relativa de obtener cara será:

$$\text{Frecuencia relativa} = \frac{48}{100} = 0.48$$

Si repetimos el experimento más veces, por ejemplo, 1,000 lanzamientos, y obtenemos cara en 502 ocasiones, la nueva frecuencia relativa será:

$$\text{Frecuencia relativa} = \frac{502}{1000} = 0.502$$

Observamos que, a medida que aumentamos el número de lanzamientos, la frecuencia relativa tiende a aproximarse a 0.5, que es el valor teórico de la probabilidad para una moneda justa. Este es un ejemplo clásico del enfoque de frecuencia relativa.

2.1.1 Enfoque de probabilidad.

La probabilidad es una disciplina que busca medir la incertidumbre en distintos fenómenos. A lo largo del tiempo, diferentes enfoques han sido desarrollados para definir el concepto de probabilidad, cada uno con su propia perspectiva sobre cómo se deben interpretar y calcular los eventos aleatorios. Los tres enfoques más comunes son el clásico, frecuencia relativa y subjetiva. Este ensayo explorará en qué consiste cada enfoque, su aplicación en la vida real y las limitaciones que pueden presentar.

Enfoque clásico de la probabilidad.

El enfoque clásico, también conocido como enfoque a priori, fue formulado en el siglo XVIII y se basa en la idea de que todos los resultados posibles de un experimento son igualmente probables. La probabilidad de un evento se define como la proporción entre los casos favorables y los casos totales. La fórmula es:

$$P(A) = \frac{\text{Casos favorables}}{\text{Casos posibles}}$$

Este enfoque es útil en situaciones donde los resultados son simétricos y bien definidos, como el lanzamiento de una moneda o los sorteos de cartas. Por ejemplo, la probabilidad de obtener un número par al lanzar un dado es:

$$P(\text{Número par}) = \frac{3}{6} = 0.5$$

Enfoque de frecuencia relativa.

Este enfoque se basa en la observación empírica. La probabilidad se define como el valor al que tiende la frecuencia relativa de un evento conforme el experimento se repite muchas veces. Es útil en situaciones donde se pueden realizar ensayos repetidos para recopilar datos, como el análisis de calidad en una línea de producción.

Por ejemplo, si lanzamos una moneda 100 veces y obtenemos 47 caras, la frecuencia relativa será:

$$P(\text{Cara}) = \frac{47}{100} = 0.47$$

A medida que se realicen más lanzamientos, esta probabilidad tenderá a acercarse al valor teórico (0.5).

Enfoque subjetivo de la probabilidad.

Este enfoque se basa en la opinión o creencia personal sobre la ocurrencia de un evento. A diferencia de los enfoques anteriores, no depende de resultados simétricos ni de datos experimentales. Se utiliza en situaciones donde no es posible realizar repeticiones, como estimar la probabilidad de que una persona gane una elección o que llueva mañana.

Por ejemplo, una persona podría estimar que la probabilidad de que su equipo favorito gane el próximo partido es del 70%, basándose en su conocimiento de los jugadores y su desempeño reciente.

2.1.2 Espacio Muestral.

El espacio muestral es el conjunto de todos los resultados posibles de un experimento aleatorio. Un evento puede ser:

Simple: Cuando se compone de un solo resultado.

Compuesto: Cuando está formado por varios resultados del espacio muestral.

También el espacio muestral es el conjunto de todos los resultados posibles de un experimento aleatorio. Se representa con la letra SSS u Ω\Omega. Cada uno de los elementos del espacio muestral se llama evento elemental. Dependiendo del experimento, el espacio muestral puede ser finito (con un número limitado de posibles resultados) o infinito (cuando los resultados forman un conjunto infinito).

Por ejemplo:

En el lanzamiento de una moneda, el espacio muestral es:

$$S = \{Cara, Cruz\}$$

En el lanzamiento de un dado, los posibles resultados son:

$$S = \{1, 2, 3, 4, 5, 6\}$$

El espacio muestral puede ser clasificado según la naturaleza del experimento:

Espacio muestral discreto: Cuando los resultados son contables, como el lanzamiento de una moneda o de un dado.

Espacio muestral continuo: Cuando los resultados posibles forman un intervalo infinito, como la temperatura registrada en un día o el tiempo que toma completar una carrera. En estos casos, se utilizan intervalos numéricos o funciones de densidad para describir el espacio.

2.1.3 Experimentos simples y complejos.

Los experimentos aleatorios constituyen la base de esta disciplina, y pueden clasificarse en simples y complejos según la cantidad de resultados y la estructura del fenómeno que se estudia. Un entendimiento claro de ambos tipos de experimentos permite abordar problemas probabilísticos con mayor precisión. En este ensayo, se explicará qué son los experimentos simples y complejos, se identificarán sus diferencias, y se analizarán ejemplos prácticos de cada uno.

Un experimento simple es aquel en el que se realiza una única acción o prueba, generando un solo resultado posible por cada ejecución. Estos experimentos tienen espacios muéstrales pequeños y son fáciles de analizar.

Ejemplo: El lanzamiento de una moneda es un experimento simple, ya que tiene solo dos resultados posibles: cara o cruz.

Otro ejemplo: Al lanzar un dado, el espacio muestral es:

$$S = \{1, 2, 3, 4, 5, 6\}$$

Cada resultado es independiente y no requiere mayor estructura para ser evaluado.

Características:

1. Pocos resultados posibles.
2. Sencillo de modelar y resolver.
3. No involucra combinaciones complejas entre varios eventos.

Experimento complejo.

Un experimento complejo consiste en la combinación de varios experimentos simples o en la realización de múltiples etapas, cada una generando resultados. El espacio muestral en un experimento complejo puede ser mucho más amplio, ya que se consideran todas las combinaciones posibles de los experimentos individuales.

Ejemplo: El lanzamiento de dos dados es un experimento complejo porque involucra dos etapas. El espacio muestral es:

$$S = \{(1,1), (1,2), (1,3), \ldots, (6,6)\}$$

En total, hay 36 combinaciones posibles.

Otro ejemplo: Extraer dos cartas de una baraja sin reposición también es un experimento complejo, ya que cada extracción afecta las probabilidades de la siguiente. La cantidad de combinaciones posibles es mayor y el cálculo de probabilidades se vuelve más complicado.

Características:

1. Puede involucrar varias etapas o pruebas sucesivas.

2. El espacio muestral es más extenso y requiere análisis más detallados.
3. La probabilidad total depende de las interacciones entre eventos parciales.

2.1.4 Leyes de Probabilidad.

Las leyes de probabilidad ofrecen reglas matemáticas que describen cómo se combinan las probabilidades de eventos individuales para formar probabilidades más complejas. Estas leyes facilitan el análisis de experimentos aleatorios y proporcionan una base rigurosa para tomar decisiones en escenarios inciertos.

Las principales leyes de la probabilidad son:

Ley de adición: La probabilidad de que ocurra al menos uno de dos eventos es la suma de sus probabilidades individuales menos la probabilidad de su intersección.

La ley aditiva se utiliza para calcular la probabilidad de que ocurra al menos uno de varios eventos. Esta ley es especialmente útil cuando los eventos son mutuamente excluyentes (es decir, no pueden ocurrir al mismo tiempo). La ley se expresa así:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Si los eventos son mutuamente excluyentes, entonces P (A∩B)=0P(A \cap B) = 0P (A∩B)=0, y la fórmula se simplifica a:

$$P(A \cup B) = P(A) + P(B)$$

Ley de multiplicación: La probabilidad conjunta de dos eventos independientes es el producto de sus probabilidades individuales.

La ley multiplicativa se aplica para calcular la probabilidad conjunta de varios eventos independientes. La fórmula general es:

$$P(A \cap B) = P(A) \times P(B)$$

Si los eventos son dependientes, la probabilidad conjunta se calcula considerando la influencia de un evento sobre el otro:

$$P(A \cap B) = P(A) \times P(B|A)$$

Ejemplo: Si lanzamos dos monedas, la probabilidad de que ambas monedas caigan en cara es:

$$P(\text{Cara en 1.ª moneda}) = \frac{1}{2}, \quad P(\text{Cara en 2.ª moneda}) = \frac{1}{2}$$

Entonces:

$$P(\text{Ambas caras}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Las tablas de contingencia organizan los datos en dos o más categorías, permitiendo analizar la relación entre variables y calcular probabilidades condicionales.

2.1.4 Tablas de contingencia.

Una tabla de contingencia es una matriz que organiza las frecuencias absolutas o relativas de dos o más variables categóricas, permitiendo observar cómo se distribuyen los valores simultáneamente. Se compone de filas y columnas que representan cada categoría de las variables, mientras que las celdas contienen el número de observaciones que corresponden a las combinaciones específicas de las categorías.

Ejemplo básico: Si queremos analizar la relación entre el género (hombre, mujer) y la preferencia por un tipo de bebida (café, té), la tabla de contingencia podría ser:

	Café	**Té**	**Total**
Hombre	30	20	50
Mujer	25	25	50
Total	55	45	100

Esta tabla muestra cuántos hombres y mujeres prefieren café o té, así como los totales por fila y columna.

En una tabla de contingencia se pueden calcular varios tipos de frecuencias:

Frecuencia absoluta: Es el número de veces que ocurre una combinación específica de las variables. En el ejemplo anterior, la frecuencia absoluta de hombres que prefieren café es 30.

Frecuencia marginal: Corresponde a la suma de las frecuencias de una fila o columna. Por ejemplo, la frecuencia marginal de hombres es 50.

Frecuencia relativa: Es la proporción de cada celda respecto al total de observaciones, calculada como:

$$\text{Frecuencia relativa} = \frac{\text{Frecuencia absoluta}}{\text{Total general}}$$

Ventajas y limitaciones de las tablas de contingencia.

Ventajas:

Simplicidad: Son fáciles de construir e interpretar.

Versatilidad: Pueden utilizarse en una amplia variedad de disciplinas, como ciencias sociales, salud, y negocios.

Análisis de relaciones: Facilitan el análisis de asociaciones entre variables categóricas.

Limitaciones:

Datos categóricos: Las tablas de contingencia solo funcionan con datos categóricos, lo que limita su aplicación en variables numéricas continuas.

Tamaño: A medida que aumentan las categorías de las variables, las tablas pueden volverse muy grandes y difíciles de interpretar.

Correlación espuria: La asociación encontrada entre dos variables puede deberse a otros factores no considerados en la tabla.

Con una tabla de contingencia también se pueden calcular probabilidades condicionales. Sigamos con el ejemplo del consumo de café y té según género:

	Café	Té	Total
Hombre	30	20	50
Mujer	25	25	50
Total	55	45	100

¿Cuál es la probabilidad de que una persona consuma café dado que es hombre?

La probabilidad condicional se calcula como:

$$P(\text{Café}|\text{Hombre}) = \frac{30}{50} = 0.6$$

Esto significa que, si seleccionamos a un hombre al azar, hay un 60% de probabilidad de que prefiera café.

2.1 6 Teorema de Bayes.

El teorema de Bayes permite calcular la probabilidad de un evento condicionado a que otro haya ocurrido. Es una herramienta valiosa en situaciones donde es necesario actualizar la probabilidad con nueva información. Su expresión general es:

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

Donde P(A/B) es la probabilidad de A dado B, y P (B) es la probabilidad total de B.

P (A/B): Probabilidad posterior de AAA, es decir, la probabilidad de AAA dado que BBB ha ocurrido.

P (B/A) P (B/A) P (B/A): Probabilidad de que ocurra BBB si se sabe que AAA es cierto.

P (A) P(A) P(A): Probabilidad previa o inicial de AAA.

P (B) P (B) P (B): Probabilidad de que ocurra BBB, sin importar si AAA ocurrió o no.

El teorema de Bayes refleja el concepto de probabilidad condicional y nos permite actualizar nuestras creencias a medida que recibimos nueva información. En otras palabras, la probabilidad inicial que se tiene sobre un evento (probabilidad a priori) cambia en función de la evidencia observada, lo que da lugar a la probabilidad posterior.

Por ejemplo, en medicina, si inicialmente se conoce la probabilidad de que una persona tenga una enfermedad (a priori), esta probabilidad puede ser actualizada con los resultados de un examen diagnóstico (evidencia), obteniendo así la probabilidad posterior de que esa persona esté realmente enferma.

Supongamos que queremos determinar la probabilidad de que una persona tenga una enfermedad rara dado que ha dado positivo en una prueba.

La probabilidad a priori de tener la enfermedad es P (E)=0.01P (E) = 0.01P (E)=0.01 (es decir, 1% de la población tiene la enfermedad).

La prueba es 99% precisa, lo que significa que P (Positivo|E)=0.99P (\text Positivo/E) = 0.99P (Positivo|E)=0.99.

Sin embargo, también existe una probabilidad de falso positivo: P (Positivo/E)=0.05P (\text Positivo \neg E) = 0.05P (Positivo/E)=0.05 (es decir, 5% de las personas sin la enfermedad dan positivo).

Queremos calcular la probabilidad posterior de que la persona tenga la enfermedad dado que ha recibido un resultado positivo: P (E/Positivo) P (E | \text Positivo) P (E/Positivo).

Aplicando el Teorema de Bayes:

$$P(E|\text{Positivo}) = \frac{P(\text{Positivo}|E) \cdot P(E)}{P(\text{Positivo})}$$

Primero, calculamos P (Positivo) P (\text Positivo) P (Positivo), que es la probabilidad total de obtener un resultado positivo:

$$P(\text{Positivo}) = P(\text{Positivo}|E) \cdot P(E) + P(\text{Positivo}|\neg E) \cdot P(\neg E)$$

$$P(\text{Positivo}) = 0.99 \cdot 0.01 + 0.05 \cdot 0.99 = 0.0594$$

Ahora aplicamos el teorema de Bayes:

$$P(E|\text{Positivo}) = \frac{0.99 \cdot 0.01}{0.0594} = 0.1667$$

Por lo tanto, la probabilidad de que la persona tenga la enfermedad, dado que obtuvo un resultado positivo, es aproximadamente 16.67%. Este resultado ilustra que, aunque la prueba sea bastante precisa, la probabilidad de tener la enfermedad sigue siendo baja debido a la baja prevalencia inicial de la misma.

CONCLUSIÓN PERSONAL.

En estos temas aprendimos que podemos centrarnos en la importancia de estos conocimientos para el análisis de datos, la toma de decisiones y la interpretación de información en contextos variados. Así como la estadística descriptiva son los conceptos de estadística descriptiva son fundamentales para entender y procesar información cuantitativa. Desde la recolección y organización de datos hasta su análisis, este enfoque permite transformar datos en información útil. Las distribuciones de frecuencias, la presentación gráfica y las medidas de tendencia central y dispersión ayudan a comprender la estructura de los datos, su centralidad y su variabilidad. Además, el uso de herramientas como el teorema de Tchebyshev y la regla empírica facilita la interpretación de la dispersión y la distribución de los datos, siendo útiles tanto en contextos académicos como en aplicaciones prácticas.

Además la teoría de la probabilidad es complementa en el análisis estadístico al permitir evaluar y predecir la ocurrencia de eventos bajo condiciones de incertidumbre. La comprensión de conceptos como el espacio muestral, los eventos simples y compuestos, así como las leyes de probabilidad y el teorema de Bayes, proporciona una base sólida para analizar situaciones en las que el resultado no es completamente predecible. También la teoría de la probabilidad se aplica ampliamente en áreas como la ciencia, la economía y la ingeniería para calcular riesgos, entender patrones y realizar inferencias sobre poblaciones mayores basadas en muestras.

En conjunto, la estadística descriptiva y la teoría de la probabilidad permiten tanto el análisis efectivo de datos como el entendimiento de fenómenos inciertos, aportando una estructura para organizar, interpretar y tomar decisiones informadas.

CONCLUSIÓN.

Aprendimos que la estadística descriptiva y la teoría de la probabilidad son esenciales para el análisis de datos y la toma de decisiones. La primera proporciona herramientas para resumir, organizar y presentar datos, mientras que la segunda permite gestionar la incertidumbre en situaciones aleatorias. El uso combinado de estas disciplinas permite interpretar fenómenos complejos de manera más eficiente y fundamentar decisiones en evidencia cuantitativa. Comprender estos conceptos es clave en múltiples áreas del conocimiento, desde la investigación científica hasta el mundo empresarial.

Además son pilares fundamentales para analizar datos y tomar decisiones informadas en diversos campos. La estadística descriptiva proporciona herramientas para recolectar, organizar, presentar, y analizar datos, permitiendo transformar grandes volúmenes de información en conocimientos comprensibles. A través de las distribuciones de frecuencias, la presentación gráfica y las medidas de tendencia central, se sintetizan los datos, facilitando la identificación de patrones y tendencias.

El procesamiento estadístico de datos es una etapa clave que abarca desde la recolección hasta la interpretación, asegurando que los resultados sean precisos y útiles para la toma de decisiones. Las medidas de dispersión complementan las de tendencia central al revelar la variabilidad en los datos, proporcionando una imagen más completa.

Por otro lado, la teoría de la probabilidad permite gestionar la incertidumbre en situaciones donde los resultados no son deterministas. Los diferentes enfoques de probabilidad, el espacio muestral, los eventos simples y compuestos, y las leyes de probabilidad proporcionan un marco matemático para analizar fenómenos aleatorios. Herramientas como las tablas de contingencia y el teorema de Bayes son esenciales para entender relaciones entre variables y actualizar probabilidades con información nueva.

El teorema de Tchebyshev y la regla empírica ofrecen formas prácticas de interpretar la dispersión de los datos, ya sea para distribuciones generales o normales, lo que refuerza la importancia de comprender tanto los patrones centrales como la variabilidad de los datos.

Bibliografías.

1. ALEA, V. et al. (2006) Estadística Aplicada a les Ciencias Econòmiques y Sociales. Barcelona: Edicions McGraw-Hill EUB.
2. CANAVOS, G. (2008) Probabilidad y Estadística. Aplicaciones y Métodos. México: McGraw-Hill. 3.-DURA Peirón, J. M. y López Cuña, J.M. (2006) Fundamentos de Estadística. Estadística Descriptiva y Modelos Probabilísticos para la Inferencia. Madrid: Ariel Editorial.
3. ESCUDER, R. y SANTIAGO, J. (2010) Estadística aplicada. Economía y Ciencias Sociales. Valencia: Tirant lo Blanch.
4. Fernández CUESTA, C., y FUENTES García, F. (2015) Curso de Estadística Descriptiva. Teoría y Práctica. Madrid: Ariel.
5. Martínez-González, M.A.; Faulin, F.J. y Sánchez, A. (2006). Bioestadística amigable, 2ª ed. Díaz de Santos, Madrid.
6. Martín PLIEGO, F. (1994) Introducción a la Estadística Económica y Empresarial. (Teoría y Práctica) Madrid: AC.
7. Martín PLIEGO, F. y RUIZ-MAYA, L. (1995) Estadística I: Probabilidad. Madrid: AC.
8. Martín PLIEGO, F. y RUIZ-MAYA, L. (1995) Estadística II: Inferencia. Madrid: AC.
9. Martín Guzmán, P. y Martín PLIEGO, F. (1985) Curso Básico de Estadística Económica. Madrid: AC.
10. MENDENHALL, W., et al. (1994) Estadística Matemática con Aplicaciones. México: Grupo Editorial Iberoamérica.
11. MONTIEL, A.M., RIUS, F. y Barón, F.J. (1997) Elementos Básicos de Estadística Económica y Empresarial. Madrid: Prentice Hall.
12. Kai Lai Chung. Elementary Proability Theory with Stochastic Processes. Springer-Verlag New York Inc.
13. Kenneth. Rosen .Matemáticas Discretas y sus Aplicaciones. S.A.MCGRAW HILL/INTERAMERICANA DE ESPAÑA.
14. Paul L. Meyer. Probabilidad y Aplicaciones Estadísticas. S.A. ALHAMBRA MEXICANA.
15. Seymour Lipschutz Ph.D. 2000 Problemas Resueltos de Matemática Discretas. McGRAW-HILL.
16. https://es.wikihow.com/usar-la-regla-emp%C3%ADrica

17. http://www.stat119review.com/more-material/normal-distribution/empirical-rule/solving empirical-rule-questions
18. https://www.khanacademy.org/math/probability/normal-distributions-a2/normal distributions-a2ii/v/ck12-org-normal-distribution-problems-empirical-rule
19. http://www.stat119review.com/more-material/normal-distribution/empirical-rule/solving empirical-rule-questions
20. https://www.youtube.com/watch?v=T7-eeg6rhjY
21. http://www.stat119review.com/more-material/normal-distribution/empirical-rule/solving empirical-rule-questions
22. https://www.khanacademy.org/math/probability/normal-distributions-a2/normal distributions-a2ii/v/ck12-org-normal-distribution-problems-empirical-rule
23. https://www.nku.edu/~statistics/212_Using_the_Empirical_Rule.htm