

ESTADISTICA DESCRIPTIVA Y TEORIA DE LA PROBABILIDAD (ENSAYO)

PAOLA JAZMIN MONZON HERNANDEZ

Asignatura: Tendencias y sistemas de salud en México

08 de septiembre 2023

Introducción

La estadística descriptiva es un conjunto de técnicas numéricas y gráficas para describir y analizar un grupo de datos, se dedica a los métodos de recolección, descripción, visualización y resumen de datos originados a partir de los fenómenos de estudio. A través del tema se introducirán algunas técnicas descriptivas básicas, que permitirán realizar la descripción de datos, pretendiendo abordar de manera resumida cada uno de los conceptos para la comprensión e interpretación del tema, también analizaremos algunos conceptos elementales de la teoría de la probabilidad la cual estudia los fenómenos o experimentos aleatorios que surge como herramienta utilizada por los nobles con la finalidad de ganar en los juegos de aquella época, con el tiempo las técnicas matemáticas se perfeccionaron y encontraron usos muy diferentes a las que fueron creadas. En la actualidad surgen nuevas metodologías que permiten aumentar el uso de la computación en el estudio de las probabilidades, disminuyendo los márgenes de error en los cálculos.

UNIDAD I Estadística Descriptiva

El proceso estadístico es el conjunto de etapas que deben completarse para realizar una investigación basada en información cuantitativa, es necesario llevar a cabo cada una de las etapas dando como resultado final un efecto fiable en la culminación del estudio, pues de lo contrario se obtendrán conclusiones erróneas que nos llevarán a toma de malas decisiones.

Este proceso está compuesto por 5 fases.

- 1. Planteamiento del problema:** Sitúa el eje central y responde a las siguientes preguntas: ¿Qué necesito estudiar y por qué?, ¿Dónde?, ¿Cuándo? y ¿Cómo?
- 2. Recolección de datos:** Se establece el tipo, tamaño de muestreo y el tipo de recolección de datos (se lleva a cabo mediante base de datos como entrevistas, cuestionarios, encuestas personalizadas, inspección de registros y observación).
- 3. Organización de datos:** Una vez adquiridos todos los datos queda unificarlos y organizarlos por separado en datos cualitativos y cuantitativos.
- 4. Análisis de los datos:** El análisis de datos consiste en someter los datos a la realización de operaciones, esto se hace con la finalidad de obtener conclusiones precisas que nos ayuden a alcanzar nuestros objetivos.

5. Interpretación de datos: Si la interpretación es errónea, entonces las decisiones tendrán un efecto no deseado.

Después de realizar cada una de las etapas, se realiza la distribución de frecuencias por medio de tablas en la que se dispone la modalidad de la variable por filas. En las columnas se establece el número de ocurrencias por cada valor, permitiendo la reducción de datos mediante el agrupamiento en frecuencias, lo que nos facilita su análisis. Para facilitar la comprensión y la variabilidad de los datos que se están obteniendo mediante la investigación, es de gran utilidad emplear elementos visuales de fácil interpretación, las representaciones gráficas nos lo permiten con solo verlas, a través de estas se entiende la naturaleza de los datos, se observan sus características e incluso es más simple formular conclusiones sobre el comportamiento de una población o muestra. Para elaborar cualquier representación gráfica es preciso tomar en cuenta el tipo de representación gráfica para alcanzar su objetivo con mayor sencillez, indicando correctamente el título, el origen, las escalas, las claves explicativas y las unidades de medida.

Las medidas de tendencia central son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos, las utilizadas son: media (serie de valores que da un resultado producido al sumar todos ellos y dividir la suma por el número total de valores), mediana (valor de la variable que ocupa la posición central) y moda (valor de la variable que más se repite), estas tendencias contribuyen a conocer de forma aproximada el comportamiento de una distribución estadística. Mientras que las medidas de dispersión se definen como 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, pretenden resumir en un solo valor la dispersión que tiene un conjunto de datos. Las medidas de dispersión más utilizadas son: Rango de variación (diferencia entre el mayor valor de la variable y el menor valor de la variable), Varianza (medida para caracterizar la dispersión de una distribución o muestra), Desviación estándar (es la raíz cuadrada de la varianza), Coeficiente de variación (desviación estándar de la muestra expresada como porcentaje de la media muestral).

Dos conceptos muy relacionados en estadística son el teorema de Chebyshev y la regla empírica, ya que ambos se usan para calcular la probabilidad de intervalos de confianza. El teorema de Chebyshev, es considerado una desigualdad probabilística, proporciona un límite superior a la probabilidad, se emplea para medir la dispersión de los datos para

cualquier distribución determinando la posibilidad de que un valor se encuentre dentro de un intervalo de confianza, también se usa para demostrar otros teoremas estadísticos, tales como la ley de los grandes números (establece que para k , las variables aleatorias independientes e idénticamente distribuidas es la media muestral). En tanto que la regla empírica constituye una manera útil de analizar datos estadísticos, pues define el porcentaje de valores de una distribución normal ya que 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 y 99,7% de los valores se encuentran a tres desviaciones estándar de la media. La diferencia entre el teorema de Chebyshev y la regla empírica es que el teorema de Chebyshev se puede usar en cualquier tipo de distribución, en cambio, la regla empírica solo es válida para una distribución normal. el uso del teorema de Chebyshev es más amplio, sin embargo, la regla empírica proporciona resultados más precisos para una distribución normal.

UNIDAD II Teoría de la Probabilidad

La existencia de fenómenos o experimentos no determinados, donde el conocimiento de las condiciones en las que éstos se desarrollan no garantiza los resultados, hace indispensable el uso de una función que asigne niveles de certeza a cada uno de los desenlaces del fenómeno, disminuyendo de este modo los márgenes de error en los cálculos y ahí es donde aparece la probabilidad. La teoría de la probabilidad es un modelo matemático que se ocupa de analizar los fenómenos aleatorios; esto implica la contraposición respecto de los fenómenos ya determinados, que son aquellos en los cuales el resultado del experimento que se realiza produce un resultado único y previsible, que se repetirá la cantidad de veces que éste vuelva a hacerse, siempre y cuando se respeten las mismas condiciones.

Existen tres tipos de enfoques de Probabilidad:

- Clásico: Probabilidad priori antes de que el experimento se realice.
- Relativo: Probabilidad a posteriori después de que el experimento se realizó.
- Subjetivo: Donde se expresa un grado meramente personal de creencia de la posibilidad de que un acontecimiento específico de un experimento aleatorio suceda.

En cualquier experimento aleatorio se corre el riesgo de obtener variabilidad de resultados, por ello existe un conjunto formado por todos los posibles resultados elementales de un experimento aleatorio al cual se le denomina espacio muestral, dependiendo de cómo sea

este conjunto, los espacios muestrales pueden ser de tres tipos: 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) y espacio muestral continuo (consta de un número infinito no numerable de elementos).

Cuando trabajamos con probabilidad, una acción aleatoria o serie de acciones se llama experimento. Un resultado es la consecuencia de un experimento, y un evento es un resultado particular de un experimento aleatorio. En términos de conjuntos, un evento es un subconjunto del espacio muestral, a continuación, se describe dos tipos; el evento simple es aquel suceso que está formado por un único resultado del espacio muestral, por ejemplo, lanzar un dado existen 6 resultados posibles, sacar un 1 sería un evento simple y el evento compuesto es un evento con más de un resultado, por ejemplo, lanzar un dado de 6 lados y sacar un número par: 2, 4, y 6. Cuando lanzamos muchas veces un dado de 6 lados, no debemos esperar que un resultado ocurra más frecuentemente que otro.

La probabilidad constituye un importante parámetro en la determinación de las diversas casualidades obtenidas tras una serie de eventos esperados dentro de un rango estadístico, para ello es necesario conocer los tres métodos para calcular probabilidades como son la regla de la adición o suma que instituye 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; la regla de la multiplicación, nos indica que la probabilidad de ocurrencia de dos o más eventos estadísticamente independientes es igual al producto de sus probabilidades individuales y la regla de Laplace, establece que la probabilidad de ocurrencia de un suceso imposible es 0 y la de ocurrencia de un suceso seguro es 1, es decir, $P(A) = 1$. De esta manera proseguimos a ordenar los datos en la tabla de contingencia la cual consiste en crear al menos dos filas y dos columnas para representar datos categóricos en términos de conteos de frecuencia, tiene el objetivo de 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, analiza si hay una relación entre las variables cualitativas, ya sean dependientes o independientes, representando los datos de manera resumida.

El teorema de Bayes es utilizado para calcular la probabilidad de un suceso, teniendo información de antemano sobre ese suceso. Podemos calcular la probabilidad de un suceso

A, sabiendo además que ese A cumple cierta característica que condiciona su probabilidad. El teorema de Bayes entiende la probabilidad de forma inversa al teorema de la probabilidad total. El teorema de la probabilidad total hace inferencia sobre un suceso B, a partir de los resultados de los sucesos A. Por su parte, Bayes calcula la probabilidad de A condicionado a B. El teorema de Bayes ha sido muy cuestionado. Lo cual se ha debido, principalmente, a su mala aplicación. Ya que, mientras se cumplan los supuestos de sucesos disjuntos y exhaustivos, el teorema es totalmente válido.

Conclusión

A través de este ensayo se logró comprender que la estadística descriptiva tiene como objetivo resumir la evidencia encontrada en una investigación de manera sencilla y clara para su interpretación. Consta de representaciones mediante el uso de tablas o cuadros, figuras o gráficas para facilitar la comprensión y la variabilidad de los datos que se están obteniendo mediante el estudio, siendo necesario el uso de elementos visuales de fácil interpretación con la finalidad de señalar tendencias y comparaciones, por medio del proceso estadístico surgen incógnitas de fenómenos o experimentos no determinados que nos llevan a estudiar la teoría de la probabilidad, concluyendo con que esta nace para responder las interrogantes sobre el azar y eventos futuros, siendo que esta hipótesis se basa a través de teoremas o principios aplicados a poblaciones o muestras para resolver los enigmas que se presenten.

Referencias Bibliográficas

1. Avello Martínez, R., & Seisdedo Losa, A. (2017). El procesamiento estadístico con R en la investigación científica. *MediSur*, 15(5), 583-586.
2. Benitez, M. D., Iranzo, J. L., Isla, F., & Sarrion, M. D. (2013). *Estadística descriptiva*. Madrid: McGraw-Hill.
3. Pértega Díaz, S., & Pita Fernández, S. (2001). Representación gráfica en el análisis de datos. *Cad Aten Primaria*, 8, 112-117.
4. Quevedo, F. (2011). Medidas de tendencia central y dispersión. *Medwave*, 11(03).
5. Batanero, C., López-Martín, M. D. M., González-Ruiz, I., & Díaz-Levicoy, D. (2015). Las medidas de dispersión en el estudio de la inferencia estadística.
6. Romero Domínguez, A. (2021). Propuesta de estrategia de monitoreo transaccional anti lavado de activos empleando el método de Ward y el teorema de Chebyshev.
7. López, A. (2016). Usos del programa R en la enseñanza de la estadística.
8. Llinás, H. (2015). *Introducción a la teoría de la probabilidad*. Universidad del Norte.
9. Nieves Hurtado, A., & Domínguez Sánchez, F. C. (2009). *Probabilidad y estadística para ingeniería un enfoque moderno (No. 519 N5.)*. México: McGraw Hill.
10. Bolivar, L. M. A. V., Saavedra, C. X. C., & Grajales, R. J. R. (2020). Espacio muestral en la teoría de la probabilidad. *Covalente*, 2(1).
11. Inzunza, S. (2017). Conexiones entre las aproximaciones clásicas y frecuencial de la probabilidad en un ambiente de modelación computacional. *Avances de Investigación en Educación Matemática*, 11, 69-86.
12. Panteleeva, O. V. (2005). *Fundamentos de probabilidad y estadística*. UAEM.
13. López-Roldán, P., & Fachelli, S. (2015). Análisis de tablas de contingencia. *Metodología de la investigación social cuantitativa*, cap-III.
14. Martínez, M. C. P., & García, J. A. P. (2009). El planteamiento de problemas y la construcción del teorema de Bayes. *Enseñanza de las Ciencias. Revista de investigación y experiencias didácticas*, 27(3), 331-342.