

Ensayo sobre estadística descriptiva y teoría de la probabilidad

04/09/2023

Maestría en administración en sistemas de salud
Jesús Antonio García salvador
Tendencias y sistemas de salud en México
Docente: María Cecilia Zamorano Rodríguez
Universidad del sureste

y jklzxcvbnmqwertyu
iopasdf rtyuio
pasdfgh yuiopa
sdfghjk opasd
fghjklzx asdfg
hijklzxcvbnmqwertyuionpasdfghj
klzxcvbnmqwertyuionpasdfghjklz
xcvbnmqwertyuionpasdfghjklzxc
vbnmqwertyuionpasdfghjklzxcvb
nmqwertyuionpasdfghjklzxcvbn
mqwertyuionpasdfghjklzxcvbnm
qwertyuionpasdfghjklzxcvbnmqw
ertyuionpasdfghjklzxcvbnmqwert
vionpasdfghijklzxcvbnmqwertvii

Antes de comenzar un proyecto ¿Qué tipo de proceso se debe realizar para evitar tomar malas decisiones?

Antes de comenzar un proyecto, independientemente del espacio del contexto es necesario aplicar un proceso estadístico y ¿Qué es un proceso estadístico? es un conjunto de etapas que deben completarse para poder realizar una investigación basada en información cuantitativa de tal manera que podamos obtener resultados apegados a la realidad en la que queremos aplicarla.

La primera etapa del procesamiento estadístico es el planteamiento del problema, para ello es necesario preguntarse ¿Qué necesito estudiar? Y ¿Por qué? Después de tener la idea es necesario preguntarse el espacio del contexto ¿Dónde? Estos pueden ser desde un lugar físico, hasta un campo de estudio, movimiento literario o corriente de pensamiento. Después vendrá ¿Cuándo? Que es el tiempo en que lo quiero aplicar y por último el modo ¿Cómo? El método a realizarlo. La segunda etapa del proceso estadístico consiste en la recolección de datos, en esta etapa se debe establecer el tipo, el tamaño de la muestra y el tipo de recolección de datos, los métodos utilizados son entrevistas, cuestionarios, inspección de registros, hechos y la observación. La tercera etapa viene después de haber recolectado la información que necesitábamos, es aquí donde se deben organizar los datos previamente obtenidos, seguido de su cuarta etapa que consiste en analizar la información obtenida con el propósito de sacar conclusiones y llegar a la interpretación de los datos la cual es la última etapa para poder llegar a una conclusión satisfactoria que cubra los diferentes ángulos y que sostengan el problema planteado inicialmente, dicho lo anterior se puede aplicar a diferentes ámbitos que van desde abrir un negocio, la mercadotecnia, recursos humanos, estudios sociales y estudios clínicos entre mencionar algunos.

Una manera más fácil de analizar los datos después de su obtención es mediante las tablas de distribución de frecuencias las cuales se puede utilizar para datos no agrupados y para datos agrupados pero para ello es necesario primero conocer algunos términos importantes que serán muy comunes en su utilización más adelante como las marcas de clase (X_i) son utilizados para separar datos entre clases como clasificaciones entre edades, medidas o variantes, para la realización de una tabla de frecuencias es necesario contar con datos que nos ayudaran a llegar a resultados factibles como son la frecuencia absoluta(f_i), Frecuencia acumulada(F_i), frecuencia relativa(h_i), frecuencia relativa acumulada(H_i) y al momento de calcular el intervalo se utiliza la fórmula de STURGES $1+3.3\text{Log}(N)$ siendo (N) el número total de los datos.

Para dar una presentación visual de los valores o datos obtenidos lo ideal sería utilizar una representación gráfica de un estudio como podrían ser las ventas mensuales de una tienda, el porcentaje de edad de los colaboradores en una empresa, accidentes laborales ocurridos en un cierto tiempo determinado, el número de coches que vende una concesionaria mensualmente, bimestrales, trimestrales o anuales, comparando las estadísticas de ventas entre ellas para obtener un resultado comparativo y saber que decisiones o medidas tomar en base a los resultados, estas y muchas otras aplicaciones se les puede dar a las representaciones gráficas. Es importante conocer con qué tipo de representaciones graficas podemos contar, siendo el más común y utilizado el grafico de barras para representar datos de manera lineal, el grafico circular o de pastel, se lleva a cabo mediante la división de un circulo el cual representa un 100% y se va fragmentando sus porcentajes en base a los datos representados, el histograma siendo muy parecido a la gráfica de barras en el se observa un rango de valores, que además podrían llegar a reflejar intervalos de diferentes longitudes. Ello permite observar no solo la frecuencia sino también la dispersión de un continuo de valores, lo que a su vez puede ayudar a inferir la probabilidad, la gráfica lineal muestra los cambios a través del tiempo, mediante una línea gráfico de dispersión o gráfico xy es un 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. Gráfico de caja y bigotes tienden a utilizarse de cara a observar la dispersión de los datos y cómo éstos agrupan sus valores, se parte del cálculo de los cuartiles, los cuales son los valores que permiten dividir los datos en cuatro partes iguales. El Gráfico de áreas en este se observa de manera semejante lo que ocurre con los gráficos de líneas, la relación entre variable dependiente e independiente. Inicialmente se hace una línea que une los puntos que marcan los diferentes valores de la variable, a través de él se pueden medir y comparar los valores de diferentes muestras. El cartograma resulta de utilidad en el terreno de la epidemiología, indicando las zonas o áreas geográficas en las que aparece con mayor o menor frecuencia un determinado valor de una variable, un ejemplo más reciente puede ser la expansión del Covid-19 durante su trayectoria por todo el mundo.

La media aritmética, la moda o la mediana, ayudan a conocer de forma aproximada el comportamiento de una distribución estadística y conforman las medidas de tendencia central, mientras que las medidas de dispersión están conformadas por rango, varianza, desviación y coeficiente de variación.

¿Cuál es la función del teorema de Chebyshev? El teorema de Chebyshev es uno de los resultados más importantes que han existido dentro de la probabilidad ya que permite estimar la probabilidad de un evento, a pesar de que el teorema de Chebyshev fuera formulado por primera vez por el francés Irénée-Jules Bienaymé, el teorema recibe este nombre porque fue el ruso Pafnuty Chebushev quien lo probó en 1867. La regla empírica sirve para una distribución normal y solo es posible producir estimaciones, es necesario contar la media y la desviación estándar de los datos para poder estimar cuantos de los datos se encuentra dentro de un determinado rango.

Pierre de Fermat fue un matemático del siglo XVII al que se apodó príncipe de los aficionados, puesto que en realidad no era matemático sino que estudió Derecho en la ciudad de Toulouse, donde ejerció de juez durante el reinado de Luis XIV. Fue uno de los fundadores de la Geometría Analítica (la combinación de la Geometría con el Cálculo y el Álgebra); que describió en su obra *Ad locus planos et solidos isagoge*. La correspondencia con Blaise Pascal permitió que ambos cofundaran y asentaran las bases de la Teoría de la Probabilidad. La probabilidad es útil para estudiar fenómenos aleatorios y realizar predicciones en situaciones cercanas y cotidianas.

Los conocimientos probabilísticos son necesarios para estudiar experimentos aleatorios que se requieran en otros campos de estudio (Nilsson, 2013).

Los eventos mutuamente excluyentes y eventos no excluyentes de dos eventos o más quiere decir que si no pueden ocurrir simultáneamente. Es decir, la ocurrencia de un evento impide automáticamente la ocurrencia del otro evento. Un ejemplo de evento es excluyente sucede al lanzar una moneda solo puede ocurrir que salga cara o sello pero no los dos a la vez. 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, siguiendo el anterior ejemplo quiere decir que al lanzar dos veces la misma moneda no tiene por qué afectar el segundo lanzamiento, ya que son eventos independientes, de salir cara en el primer lanzamiento no quiere decir que pueda repetirse, al igualmente lo contrario que con sello, por eso es explícitamente un evento independiente. Dos o más eventos serán dependientes cuando la ocurrencia o no-ocurrencia de uno de ellos afecta la probabilidad de ocurrencia del otro, un ejemplo sería que al caminar una persona le metiéramos el pie, tiene probabilidades de que se desarrolle un evento dependiente lo que ocasionaría que la persona que camina termine cayéndose, siendo explícitamente condicional al evento anterior.

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, atendiendo a determinadas condiciones, 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. 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 de importancia vital

El espacio muestral es el conjunto de todos los resultados posibles individuales de un experimento aleatorio existen diferentes tipos, el espacio muestral discreto finito consta de un número finito de elementos, por ejemplo lanzar un dado. El espacio muestral discreto infinito consta de un número infinito numerable de elementos, por ejemplo lanzar un dado hasta que salga un cinco. El espacio muestral continuo consta de un número infinito no numerable de elementos.

El teorema de bayes es utilizado para calcular la probabilidad de un suceso mediante su fórmula, 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 conclusión la estadística descriptiva y la probabilidad se mantienen en un estado de constante de interrelación, comenzando desde el proceso de recolección de datos para poder llegar a resultados tangibles, plasmando esos resultados de manera visual, utilizando diferentes métodos como las tablas de frecuencias para poder analizar la información y el uso de los diferentes teoremas como el de chebyshev y el teorema de bayes utilizados dentro de la estadística y la probabilidad para así acercarse a lo más cercano a poder predecir eventos y sucesos que podrían ocurrir dentro de un proyecto, empresa o estudio implementándola en muchos más ámbitos, dando un enfoque matemático inclusive a situaciones cotidianas de la vida diaria.

Bibliografía

https://www.probabilidadyestadistica.net/teorema-de-chebyshev/#google_vignette

<https://economipedia.com/definiciones/distribucion-de-frecuencias.html>

<https://www.youtube.com/watch?v=cK1RFoihGOc>

<https://www.youtube.com/watch?v=rU9C3Fs16hc>

https://www.ine.es/explica/docs/historia_personajes.pdf

Nilsson, P. (2013). Challenges in seeing data as useful evidence in making predictions on the probability of a real-world phenomenon. *Statistics Education Research Journal*, 12(2), 71-83.

<https://doi.org/10.52041/serj.v12i2.305>

<https://www.youtube.com/watch?v=Fi6G48j0IZ4>