



EL DIAGRAMA DE TALLO Y HOJAS COMO ESTRATEGIA EN LA CONSTRUCCIÓN DE UNA DISTRIBUCIÓN DE FRECUENCIAS PARA DATOS AGRUPADOS

Ferney Anturi y Alirio Quesada
Universidad de la Amazonia (Colombia)
fantury@uniamazonia.edu.co, Alirio51@hotmail.com

En la presente ponencia se busca mostrar de forma precisa los resultados obtenidos al adoptar el Diagrama de tallo y hojas como estrategia metodológica para agrupar datos en clases o intervalos de clase. Dicha propuesta ha tenido su desarrollo dentro del espacio académico denominado 'Estadística y probabilidad' de la Licenciatura en matemáticas y física de la Universidad de la Amazonia. Con éste se busca hacer un llamado a incluir en el aula de clase, métodos, estrategias, ayudas didácticas que contribuyan a un clamor pertinente que ha venido tomando fuerza en lo relacionado con la alfabetización estadística.

PALABRAS CLAVE

Estrategia, diagrama de tallo y hojas, datos agrupados, alfabetización Estadística

INTRODUCCIÓN

El surgimiento de los lineamientos curriculares para matemáticas a finales de los años noventas, generó al interior de las prácticas escolares una serie de cambios en la clase, pues el maestro debe dar razón del pensamiento aleatorio, situación que motiva al estudio y reflexión sobre los abordajes que se deben dar en el aula de matemáticas.

En la Universidad de la Amazonia, buscando contribuir a lo arriba expuesto, hemos constituido un grupo de docentes y estudiantes de la Licenciatura en matemáticas y física que tiene como prioridad aportar propuestas de experiencias encaminadas a mejorar el estudio de los objetos matemáticos (en estadística), concebidos y estudiados desde su estructura conceptual. Se pretende, además, generar una verdadera cultura estadística apuntando prioritariamente a la nueva propuesta denominada 'alfabetización estadística'.

Siempre que se quiere agrupar datos en intervalos de clase, el profesor, dada su experiencia, usualmente recurre bien sea a la fórmula de Sturges o al número deseado de clases; a esto se suma el seguimiento estricto a los libros de texto. Nuestra práctica diaria posibilita generar nuevas propuestas, las que nos ha permitido notar que en el caso particular de la elaboración de un diagrama de tallo y hojas, en la mayoría de las veces el docente sólo lo emplea para hacer un análisis de datos y conocer la forma que toma la serie; esto nos permite afirmar que el mencionado diagrama es utilizado en un porcentaje mínimo, dejando de lado toda la enseñanza adicional que éste nos ofrece y que como educadores debemos aprovechar, pues la experiencia que hemos podido



adquirir, desde el abordaje del espacio 'Estadística y probabilidad' durante los últimos dos años, así lo muestran.

MARCO DE REFERENCIA

La estadística es una ciencia que se remonta al inicio de las sociedades humanas, como una necesidad para la realización de actividades elementales, desde una simple colección de datos hasta la interpretación de los mismos. Las primeras sociedades que se formaron en la historia, necesitaban saber cuántos habitantes formaban cada tribu, la cantidad de bienes poseían, etc. de esa manera planificaban mejor la asignación de los recursos; luego se podría decir, que la estadística permite recolectar, organizar, resumir y analizar datos, a partir de los cuales se deducen conclusiones que validen decisiones adecuadas.

Cuando hacemos investigación necesariamente debemos tener presente que sin la estadística no sería posible llegar a las conclusiones, descripciones, caracterizaciones o inferencias y por ende no podríamos tomar decisiones; luego es un compromiso ineludible de los maestros propiciar espacios adecuados para el logro de los objetivos que se plantean en el momento de la planeación de las clases buscando siempre que nuestros estudiantes se vinculen plenamente a los procesos que al respecto se dan en las instituciones educativas. Es por lo anterior que consideramos de mucha importancia que desde la planeación misma se generen dinámicas y procedimientos, que permitan a futuro hacer posible acceder a conceptos y definiciones que antes se consideraban demasiado axiomáticas.

Es por esta razón que nuestras prácticas escolares deben hoy en día contener un alto grado de innovación y estar dotadas de un tratamiento disciplinar y didáctico, a fin de que nuestros estudiantes reconozcan que el aprendizaje es en definitiva un proceso en el cual convergen personas comprometidas y direccionadas hacia el logro de unos objetivos precisos y comunes a los intervinientes del mencionado proceso. La presente propuesta retoma el legado del ingeniero químico y matemático Jhon Wilder Tukey, quien dio su aporte a la estadística con varias de las gráficas más usadas en el análisis exploratorio de datos y a quien adicionalmente se le asigna la introducción de las modernas técnicas para estimar el espectro de las series temporales. En 1965, en un artículo conjunto con J. W. Cooley, publicado en la revista *Mathematics of Computation*, introdujo el algoritmo de la *Transformada rápida de Fourier* (FFT), fundamental para crear el procesamiento digital de datos. El *Análisis Exploratorio de Datos* o EDA (Exploratory Data Analysis), es una nueva aproximación a la estadística básicamente gráfica, con un conjunto de técnicas de exploración de datos estadísticos basadas en el uso de gráficos. Destacan los gráficos 'Box-and-Whisker Plot' (Diagrama de caja y bigotes), el 'Stem-and-Leaf Diagram' (Diagrama de tallo y hojas), los 'Radiogramas' (rootograms) y los Diagramas de ajuste.

El diagrama de tallo y hojas, a pesar de no ser un gráfico definitivo para la presentación de datos, es fácil y rápido para realizar a mano; con él se puede caracterizar desde una primera perspectiva el comportamiento que siguen los datos, también es de gran utilidad para representar un conjunto de datos cuantitativos, este tipo de representación presenta similitudes con el histograma en cuanto que



proporciona información del recorrido de la distribución de datos en estudio, muestra la ubicación de la mayor concentración de mediciones y revela la presencia o ausencia de simetría. El diagrama de tallo y hojas tiene ventajas sobre el histograma; otra ventaja, es que esta representación elimina el paso de los datos originales a clases.

Algunas veces, la utilización del primero o de los dos primeros dígitos de los datos puntuales como tallos no proporcionan suficientes tallos como para permitir detectar la forma de su distribución. Una manera de solucionar esto es utilizar tallos dobles. Es decir, utilizar cada tallo dos veces: una vez para trazar las hojas inferiores y otra vez para trazar las hojas superiores. Muchos de los procedimientos estadísticos que se desarrollan suponen que la variable independiente estudiada posea al menos una distribución aproximadamente normal, para la cual el diagrama de tallos y hojas tiene forma de campana, por lo tanto, los diagramas de tallos y hojas dan una idea de la localización de los datos y de la forma de la distribución. Esta técnica funciona bien para los conjuntos de datos que no tienen una dispersión muy grande.

ASPECTOS METODOLÓGICOS

Desde el espacio académico ‘Estadística y probabilidad’ de la Licenciatura en matemáticas y física de la Universidad de la Amazonia, se pudo desarrollar la propuesta atendiendo a los requerimientos de la misma, y se llevó a cabo en dos fases.

La primera fase consistió en hacer el desarrollo normal del tema ‘Diagrama de tallo y hojas’ desde la clase tradicional, para lo cual se propuso la siguiente tarea en clase a través de equipos de trabajo.

El director nacional de tránsito, preocupado por la velocidad a la que los conductores manejan en un determinado tramo de la carrera entre el Florencia y El Doncello, se propuso recolectar información tendiente a la toma de decisiones. Para este fin, midió la velocidad de 45 conductores en este tramo en un determinado día encontrando la siguiente información expresada en kilómetros por hora (Km/h)

15 32 45 46 42 39 68 47 18

31 48 49 56 52 39 48 69 61

44 42 38 52 55 58 62 58 48

56 58 48 47 52 37 64 29 55

38 29 62 49 69 18 61 55 49

Se plantea como objetivo generar la discusión entre los estudiantes para conseguir la forma más adecuada para organizar la información en una tabla (distribución de frecuencias), donde se incluyan entre otros conceptos, el orden, la amplitud y el número de clases.

La segunda fase consistió en el desarrollo de ésta actividad ‘Construcción de la distribución de frecuencias’, pero mediante la utilización de la hoja de cálculo Excel siguiendo un proceso lógico, ayudado de las herramientas que ofrece está que a continuación mostramos con mayor detalle.



DESARROLLO DE LA PROPUESTA

El diagrama de tallo y hojas, a pesar de no ser un gráfico definitivo para la presentación de datos, es fácil y rápido para realizar a mano, con él se puede caracterizar desde una primera perspectiva el comportamiento que siguen los datos, también es de gran utilidad para representar un conjunto de datos cuantitativos, este tipo de representación presenta similitudes con el histograma en cuanto que proporciona información del recorrido de la distribución de datos en estudio, muestra la ubicación de la mayor concentración de mediciones y revela la presencia o ausencia de simetría. Cabe mencionar que el diagrama de tallo y hojas tiene ventajas sobre el histograma, porque conserva la información que puede arrojar las mediciones individuales, situación que se pierde en los intervalos del histograma; otra ventaja es que ésta representación elimina el paso de los datos originales a clases, lo que hace notar que se puede construir en el proceso de marcaje de los elementos en estudio.

Fueron los estudiantes de la Licenciatura en matemáticas y física de la Universidad de la Amazonia los participantes y con los cuales se procedió a desde el espacio académico 'Estadística y probabilidad' hacer el desarrollo de la presente propuesta participativa y con ayuda de la hoja de cálculo Excel en la cual se pudo construir la distribución de frecuencias y el cálculo de todas las medidas de centralización, dispersión, forma y posición relativa.

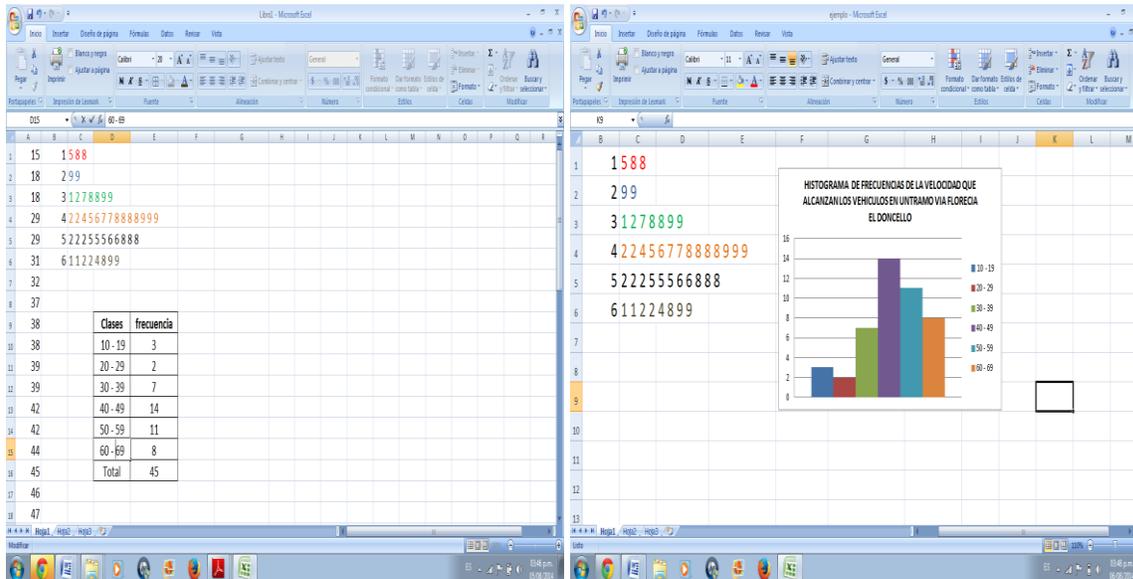
Inicialmente haremos un diagnóstico de cómo se encuentran los estudiantes con respecto a los conocimientos de estadística y el pensamiento estocástico, una vez realizado el análisis pasaremos a solucionar dificultades para luego pasar al manejo de la hoja de cálculo Excel, con el propósito de que al terminar la unidad el estudiante adquiera conocimientos y destrezas que puedan ser de gran utilidad en su vida cotidiana ya que como sabemos el conocimiento estocástico está creciendo cada vez más rápido y se hace necesario su comprensión y su razonamiento para la toma de decisiones en nuestra vida cotidiana. Paso seguido se hará una evaluación con el grupo de los resultados obtenidos en relación a los caminos de aprendizaje y las dificultades encontradas en el desarrollo, en tal sentido los desarrollos de la actividad se resumen a continuación.

Desde lo cotidiano y de acuerdo a los temas abordados, uno de los equipos de estudiantes llega a formalizar la siguiente distribución de frecuencias, adicionando que toman una amplitud de 10 para cada intervalo con seis intervalos.

Intervalos de Clase	f_i (Frecuencias)
10-19	3
20-29	2
30-39	7
40-49	14
50-59	11
60-69	8
TOTAL	45



Esta construcción generó controversia entre los demás estudiantes que construyeron su distribución de forma distinta, pero que una vez desarrollaron la misma actividad usando Excel aceptaron la construcción, en este sentido nos permitimos mostrar un archivo elaborado por uno de los equipos que inicialmente no había llegado a la anterior. Dichas construcciones nos permiten una vez más ratificar que la estrategia es válida y que si es acompañada por la utilización de la hoja de cálculo Excel se logran mejores resultados.



CONCLUSIONES

Se puede confirmar a través del trabajo práctico en la Licenciatura en matemáticas y física de la Universidad de la Amazonia, que el diagrama de tallo y hojas, más allá de servir como herramienta para hacer análisis exploratorio de datos, se convierte en una estrategia apropiada en el abordaje de la temática correspondiente a la construcción de distribuciones de frecuencias para datos agrupados en intervalos de clase.

Para los estudiantes de la Licenciatura en matemáticas y física de la Universidad de la Amazonia, la metodología utilizada propició un mejor desempeño desde el desarrollo mismo de la asignatura, puesto que generó una dinámica diferente a la clase usual, dado que esta actividad mostró la importancia que implica incluir la herramienta Excel en la clase de estadística.