



SEPARADOR 5

PENSAMIENTO ALEATORIO Y SISTEMAS DE DATOS



Carlos Octavio Gómez Tabares
I. C. Presbítero Camilo Torres Restrepo - Medellín

Gustavo Gallego Girón
I. E. José Felix Restrepo - Medellín

José Wilde Cisneros
I. E. Andrés Bello - Bello

Liliana Catrillón Giraldo
C. E. R. El Porvenir - Yolombó



PENSAMIENTO ALEATORIO Y SISTEMAS DE DATOS

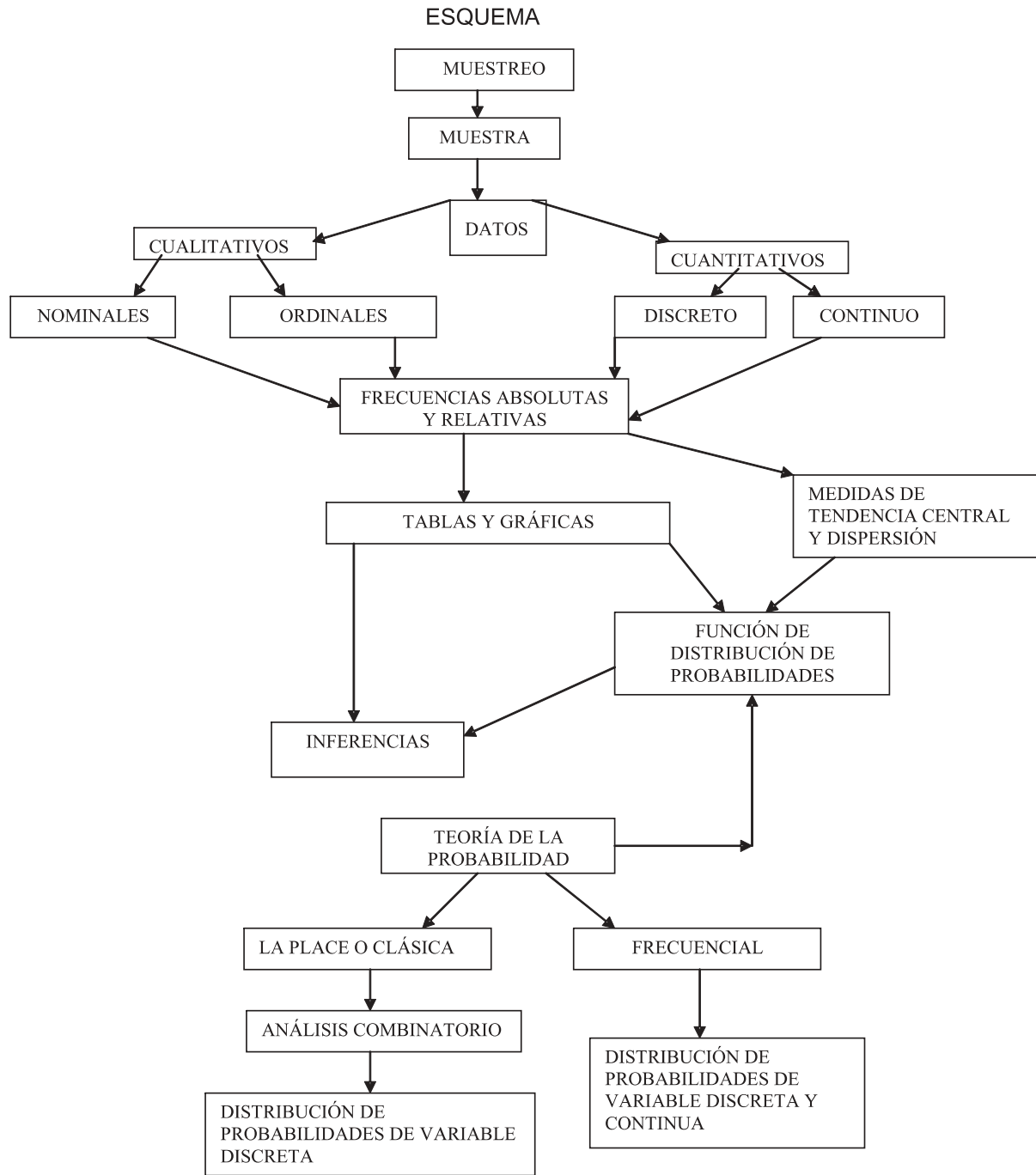
INTRODUCCIÓN

“Una tendencia actual en los currículos de matemáticas es la de favorecer el desarrollo del pensamiento aleatorio, el cual ha estado presente a lo largo de este siglo, en la ciencia, en la cultura y aún en la forma de pensar cotidiana. La teoría de la probabilidad y su aplicación a los fenómenos aleatorios, han construido un andamiaje matemático que de alguna manera logra dominar y manejar acertadamente la incertidumbre. Fenómenos que en un comienzo parecen caóticos, regidos por el azar, son ordenados por la estadística mediante leyes aleatorias de una manera semejante a como actúan las leyes determinísticas sobre otros fenómenos de las ciencias. Los dominios de la estadística han favorecido el tratamiento de la incertidumbre en ciencias como la biología, la medicina, la economía, la psicología, la antropología, la lingüística..., y aún más, han permitido desarrollos al interior de la misma matemática” (MEN, 1998).

La cita anterior manifiesta la necesidad de introducir el pensamiento aleatorio y los conceptos de sistemas de datos en los planes de área de Matemáticas, los cuáles no se han desarrollado en la mayoría de los currículos en el país. Ahora, las directrices del ministerio que incluyen no sólo los lineamientos curriculares sino también los estándares y las competencias exigen introducirlos en los planes desde el grado primero hasta el undécimo. Por eso es necesario, definir criterios básicos que permitan al educador de cada nivel desarrollar y afianzar sus acciones pedagógicas relacionadas con el pensamiento aleatorio y los sistemas de datos, enmarcadas dentro de las exigencias de los documentos rectores emanados por el ministerio de educación.

Esta propuesta inicialmente presenta un esquema general en el cual se muestran los conceptos fundamentales del pensamiento aleatorio, y una breve descripción de ellos. Luego se definen y explican unos posibles ejes temáticos mediante los cuales se organizan los estándares por niveles de grado.

A continuación se presentan esquemas generales de cada uno de estos ejes, con una somera explicación y finalmente se proponen unas situaciones problemas que sirven para desarrollar el pensamiento estadístico en el aula.



Para desarrollar ciertas investigaciones o estudios se requiere plantear sobre qué características de determinados objetos, cosas, personas o fenómenos se hará el proceso de recolección y análisis de datos, es decir, se define una *población* de estudio y las variables a estudiar. Como usualmente es casi imposible, por razones económicas o logísticas, obtener todos los datos de la población se debe seleccionar un conjunto de ésta para realizar el estudio, el cual debe ser lo más representativo posible de esa población y se le denomina *muestra*. Para que se cumpla la representatividad se hace un estudio previo de la selección de la muestra que se denomina *muestreo*, que es un proceso para



garantizar que, con la menor incertidumbre posible, los datos que se tomarán de la muestra representan o dan razón muy aproximada del verdadero comportamiento de la población. En la educación básica y media no se desarrolla la teoría del muestreo debido a que requiere de conceptos muy elaborados de la Estadística y las Matemáticas, fundamentalmente de probabilidad e inferencia.

En resumen, partiendo de un proceso serio y previo, el muestreo, se define una muestra que es representativa de la población para obtener un conjunto de datos que darán cuenta de forma aproximada de alguna de sus características.

Así, los datos a obtener nos darán información importante. Para lograr esto se requiere organizarlos, ya que usualmente se toman en desorden por razones operativas. La organización de los datos se puede hacer por medio de tablas, pictogramas, diagramas y gráficas, principalmente. Este proceso depende del tipo de datos que se está analizando, lo que hace indispensable, por razones pedagógicas, dividir su estudio.

Tipos de datos: Los datos pueden ser cualitativos o cuantitativos. Los primeros se refieren a eventos que agrupan cualidades no cuantificables como color, sexo, etc. También se pueden agrupar en cualidades que son susceptibles de estratificar, las que se denominan ordinales. Ejemplo: Estrato económico, las notas de calificación (D, I, A, S, E), los rangos militares. Los datos cuantitativos tienen la propiedad de ser cardinales, es decir, que además de un orden preestablecido se pueden comparar y relacionar entre sí. Se clasifican en discretos y continuos. Discretos cuando en un intervalo de la recta numérica todos los valores posibles son finitos o puntuales, ejemplos: número de hijos, la edad, la estatura, los precios. Y continuos cuando los valores posibles son todos los de la escala numérica real, ejemplo: rapidez de un móvil, la medida de las hipotenusas de los triángulos rectángulos inscritos en una semicircunferencia.

Si los datos son nominales se pueden agrupar en una tabla de doble entrada donde a cada fila le corresponde el nombre del dato en la primera columna, en la segunda columna se anota el número de veces que aparece el dato en la muestra, lo que se denomina *frecuencia absoluta*. En la tercera columna la relación entre la frecuencia absoluta y el tamaño de la muestra que es el número total de datos. Esta relación es la *frecuencia relativa*. Se puede agregar una cuarta columna con el porcentaje de participación de cada uno de los datos en la muestra. A esta tabla se le llama *tabla de distribución de frecuencias*. Con los datos organizados así se pueden elaborar gráficas como tortas, barras y pictogramas.

Cuando los datos son ordinales, a la tabla se le puede agregar dos columnas: una donde se van acumulando las frecuencias absolutas al pasar de una fila a la siguiente obteniéndose las *frecuencias absolutas acumuladas* y otra donde se acumulan las relativas y se calculan las *frecuencias relativas acumuladas*. Además de las gráficas anteriores se pueden elaborar los histogramas para cada frecuencia.

Cuando los datos son cuantitativos de tipo discreto, si el rango (diferencia entre el valor máximo y el mínimo de la muestra) o cantidad de valores diferentes es relativamente pequeña, por ejemplo las edades de un grupo de tercero diurno, se elabora la tabla tal como para los datos ordinales. Pero si el rango es grande, como por ejemplo las edades de un grupo de tercero de un nocturno, se debe trabajar una primera columna con los datos agrupados o por intervalos y una segunda columna donde se anota el valor o punto medio del intervalo, a lo que se denomina *marca de clase*. El resto de las columnas iguales para las cuatro diferentes frecuencias. Para ambos casos se pueden elaborar las mismas gráficas y, además, diagramas de cajón.



Si los datos son de tipo cuantitativo continuo se elabora la tabla como el caso anterior. Además de las gráficas anotadas y los diagramas de cajón aquí se pueden elaborar polígonos de frecuencias absolutas, relativas, absolutas acumuladas y relativas acumuladas. Las segundas se pueden ajustar a una curva que se aproxima a una función de distribución de frecuencias y las cuartas a otra curva que se denomina ojiva y de ella se puede definir una función de densidad de probabilidad. De ahí se pueden definir modelos probabilísticos que permiten hacer inferencias a partir de la muestra acerca de la población.

Pero para lograr hacer este tipo de análisis se requiere formular una teoría de la probabilidad: las diferentes definiciones de probabilidad, axiomas y teoremas y las distribuciones discretas como la binomial, Poisson, hipergeométrica y multinomial y las continuas como la normal, la chi-cuadrada, etc. y los tests para pruebas de hipótesis.

EJES TEMÁTICOS

Para organizar los estándares del pensamiento aleatorio y sistemas de datos se definieron unos grandes ejes temáticos que permiten agruparlos, por niveles de grado, los cuáles son: Organización de datos, medidas de posición y variabilidad y probabilidad e inferencia. Cada uno de los estándares se ubicó en uno o varios de los ejes según su afinidad temática.

El primer eje temático, la **Organización de datos** abarca temas relacionados con los diferentes procedimientos, técnicas y enfoques para organizar, recolectar y analizar un conjunto de datos obtenidos de una muestra, para que se le dé o tengan sentido dentro de su contexto y realizar las inferencias de acuerdo con ello, sin olvidar las diferentes formas de representación basados en la distribución de frecuencias.

Como ejemplo desarrollaremos una práctica que permite visualizar la eficiencia que tiene la Estadística para los fenómenos que, estudiados de otra manera, es casi imposible determinar sus regularidades. Esta consiste en determinar cual es la vocal más usada en nuestro idioma. Para comenzar, la población de estudio referida serán todas las vocales del idioma Español que es infinita. Pero podemos determinar una muestra a partir de una página completa de un libro cualesquiera, no es un buen ejemplo de muestreo, pero la experiencia nos ha mostrado que es una buena muestra para los propósitos de la clase. Por tanto, pediremos a cada estudiante que lleve a clase una copia de la página de un texto cualquiera. Es importante que la página esté completa de texto.

La toma de la muestra se podrá hacer realizando un listado de todas las vocales del texto en el orden en que aparecen y luego señalando con un color para cada vocal.

Continuando con el ejercicio de las vocales, los datos obtenidos de una manera real, se pueden organizar en la siguiente tabla:

Para el caso de las vocales se considera como variables nominales.

Vocal	ni	fi
a	324	0.31
e	308	0.294
i	122	0.117
o	194	0.185
u	98	0.094
Σ	1046	1.00

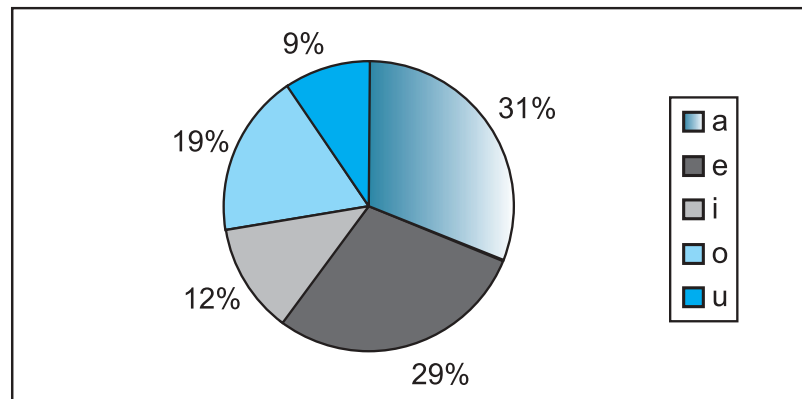
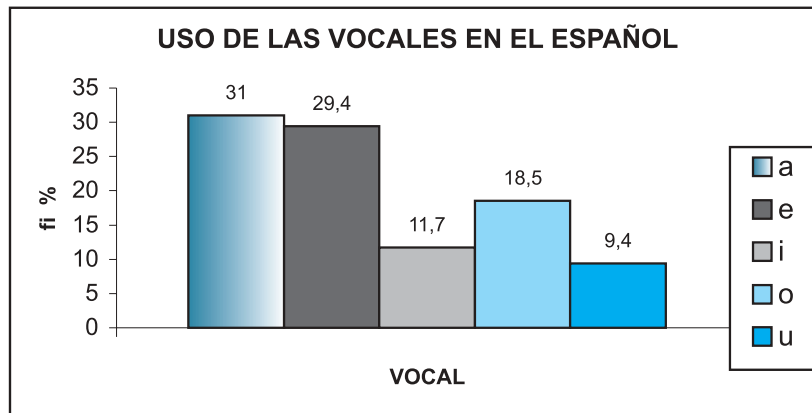


donde:

ni: Es la frecuencia absoluta, es decir, el valor que indica cuantas veces está el dato en la muestra.

fi: Es la frecuencia relativa, o sea, el valor proporcional que le corresponde al dato de la muestra. Si se desea puede agregar otra columna de **fi** en términos de porcentaje.

Los datos de la tabla anterior se pueden representar por medio de gráficas como barras o tortas.



Los temas del eje de **Las medidas de posición y variabilidad** son: media, mediana, moda, rango, varianza, desviación, etc. teniendo en cuenta que estos se obtienen a partir del conjunto de datos. Se comparan y se toman los valores que sean más representativos y permitan hacer inferencias. Se debe cubrir los diversos procedimientos estadísticos que se refieran a medidas de posición y variabilidad.

La media aritmética o promedio se calcula sumando cada uno de los datos, X_i y dividiendo el resultado por el número total de datos o tamaño de la muestra, N . Así:

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$$



La varianza S^2 , es el promedio del cuadrado de todas las distancias de cada uno de los datos respecto a la medida, o sea, se calcula realizando cada diferencia $(X_i - \bar{X})$, se eleva esta al cuadrado y se suman con las siguientes, así:

$$S^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}$$

La desviación S , equivale a la raíz cuadrada de la varianza: $S = \sqrt{S^2}$

Continuando con la práctica de las vocales se procede a analizar toda la información del grupo. Primero se recogen todos los resultados de la f_i en porcentaje. Como ejemplo, se muestra a continuación 32 datos de la vocal i tomados de un grupo de décimo:

12.4 10.5 9.5 19.5 15.9 19.1 14.6 15.8 20.0 13.7 15.0 13.3 18.0 15.0 19.0 15.0 17.6 10.4
15.6 17.0 21.0 12.0 15.0 14.4 12.4 13.7 9.4 1.6 14.6 15.6 13.7 17.9

Para fines más prácticos el resultado 1.6 debería revisarse o descartarse, pero los cálculos se hicieron con todos los datos. Aplicando las fórmulas se obtuvo:

$\bar{X} = 14.6 \%$, $S^2 = 14.67$, $S = 3.83$. Parece ser una varianza relativamente baja, lo que indica que la media de esta muestra está cercana a la de la población.

Los temas del eje de **probabilidad e inferencia** son: definiciones de probabilidad, teoría de la probabilidad, variables aleatorias, funciones de distribución de probabilidad, modelación de muestras, etc.

Organizando los datos anteriores se obtiene la siguiente tabla y gráfica.

Intervalo	x_i	n_i	f_i	N_i	F_i	$f_i \%$
1-3.99	2.5	1	0.03	1	0.03	3
4-7.99	6.5	2	0.06	3	0.09	6
8-17.99	10.5	3	0.09	6	0.18	9
12-15.99	14.5	17	0.54	23	0.72	54
16-19.99	18.5	8	0.25	31	0.97	25
20-23.99	22.5	1	0.03	32	1	3
		32				100

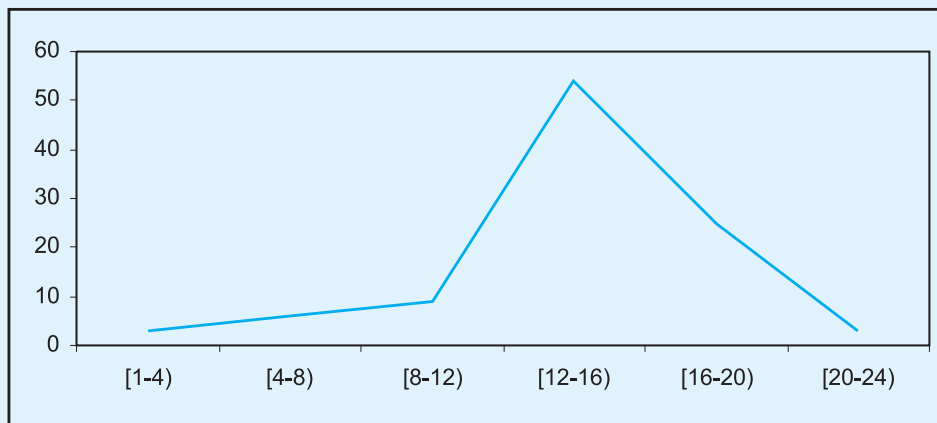
Donde: x_i : es el punto medio de cada intervalo y se denomina *marca de clase*

N_i : es la frecuencia absoluta acumulada

F_i : frecuencia relativa acumulada



Con los datos organizados se pueden construir histogramas o polígonos, tal como se muestra en la siguiente gráfica, donde se representa el polígono de frecuencias relativas de los datos de la vocal i.

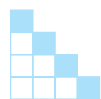


Esta grafica se puede aproximar a una curva y a una distribución de frecuencias estandarizada, lo que permite, entonces hacer inferencias respecto a la media, por ejemplo.

A continuación se presenta una propuesta para la distribución del área de Matemáticas para trabajar el pensamiento aleatorio o estadístico y el sistema de datos según los estándares, niveles y ejes temáticos

Ejes por grado	1° A 3°
Organización de datos	<ol style="list-style-type: none"> 1. Clasificar y organizar la presentación de datos (relativos a objetos reales o eventos escolares) de acuerdo con cualidades o atributos. 4. Representar datos relativos a su entorno usando objetos concretos, pictogramas y diagramas de barras 8. Resolver y formular preguntas que requieran para su solución coleccionar y analizar datos del entorno próximo.
Medidas de posición y variabilidad	<ol style="list-style-type: none"> 2. Interpretar cualitativamente datos referidos a situaciones del entorno escolar. 3. Describir situaciones o eventos a partir de un conjunto de datos. 5. Identificar regularidades y tendencias en un conjunto de datos.
Probabilidad e inferencia	<ol style="list-style-type: none"> 6. Explicar desde su experiencia la posibilidad o imposibilidad de ocurrencia de eventos cotidianos. 7. Predecir si la posibilidad de ocurrencia de un evento es mayor que la de otro.

Ejes por grado	4° A 5°
Organización de datos	<ol style="list-style-type: none"> 1. Representar datos usando tablas y gráficas (de barras, diagramas de líneas, diagramas circulares). 2. Comparar diferentes representaciones del mismo conjunto de datos.
Medidas de posición y variabilidad	<ol style="list-style-type: none"> 3. Interpretar información presentada en tablas y gráficas (de barras, diagramas de líneas, diagramas circulares). 5. Comparar y describir la distribución de un conjunto de datos. 6. Usar e interpretar la mediana (promedio).
Probabilidad e inferencia	<ol style="list-style-type: none"> 4. Hacer conjeturas y poner a prueba predicciones acerca de la posibilidad de ocurrencia de eventos.



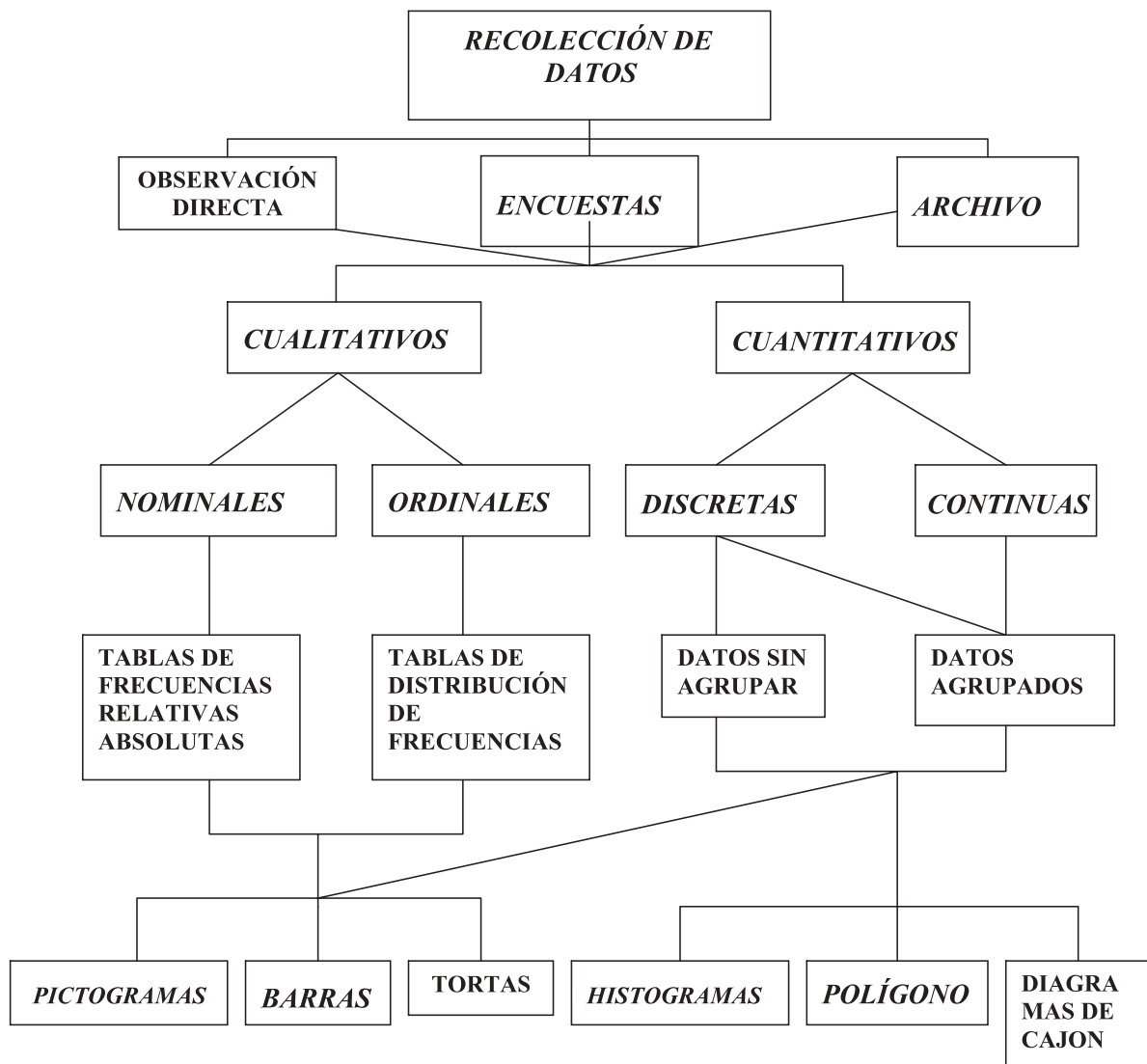
Ejes por grado	6° A 7°
Organización de datos	<ol style="list-style-type: none"> 2. Reconocer relación entre un conjunto de datos y su representación. 3. Usar representaciones gráficas adecuadas para presentar diversos tipos de datos (diagramas de barras, diagramas circulares). 7. Resolver y formular problemas a partir de un conjunto de datos presentados en tablas, diagramas de barras, diagramas circulares.
Medidas de posición y variabilidad	<ol style="list-style-type: none"> 1. Comparar e interpretar datos provenientes de diversas fuentes (prensa, revistas, televisión, experimentos, consultas, entrevistas). 4. Usar medidas de tendencia central (media, mediana, moda) para interpretar el comportamiento de un conjunto de datos.
Probabilidad e inferencia	<ol style="list-style-type: none"> 5. Usar modelos (diagramas de árbol, por ejemplo) para discutir y predecir posibilidad de ocurrencia de un evento. 6. Hacer conjeturas acerca del resultado de un experimento aleatorio usando proporcionalidad y nociones básicas de probabilidad. 7. Resolver y formular problemas a partir de un conjunto de datos provenientes de observaciones, consultas y experimentos. 8. Predecir y justificar razonamientos y conclusiones usando información estadística.

Ejes por grado	8° A 9°
Organización de datos	<ol style="list-style-type: none"> 1. Reconocer que, diferentes maneras de presentar la información, pueden dar origen a distintas interpretaciones.
Medidas de posición y variabilidad	<ol style="list-style-type: none"> 3. Interpretar conceptos de media, mediana y moda. 4. Seleccionar y usar algunos métodos estadísticos adecuados según el tipo de información. 7. Reconocer tendencias que se presentan en conjuntos de variables relacionadas.
Probabilidad e inferencia	<ol style="list-style-type: none"> 2. Interpretar analítica y críticamente información estadística proveniente de diversas fuentes (prensa, revistas, televisión, experimentos, consultas, entrevistas). 6. Resolver y formular problemas seleccionando información relevante en conjuntos de datos provenientes de fuentes diversas (prensa, revistas, televisión, experimentos, consultas, entrevistas). 5. Comparar resultados experimentales con probabilidad matemática esperada. 8. Calcular probabilidad de eventos simples usando métodos diversos (listados, diagramas de árbol, técnicas de conteo). 9. Usar conceptos básicos de probabilidad (espacio muestral, evento, independencia...).

Ejes por grado	10° A 11°
Organización de datos	
Medidas de posición y variabilidad	<ol style="list-style-type: none"> 1. Comparar estudios provenientes de medios de comunicación. 4. Describir tendencias que se observan en conjuntos de variables relacionadas. 5. Interpretar nociones básicas relacionadas con el manejo de información (como población, muestra, variable, estadígrafo y parámetro). 6. Usar comprensivamente algunas medidas de centralización, localización, dispersión y correlación (percentiles, cuartiles, centralidad, distancia, rango, varianza, covarianza y normalidad).
Probabilidad e inferencia	<ol style="list-style-type: none"> 2. Justificar inferencias provenientes de los medios o de estudios diseñados en el ámbito escolar. 3. Diseñar experimentos aleatorios (de las ciencias físicas, naturales o sociales) para estudiar un problema o pregunta. 7. Interpretar conceptos de probabilidad condicional e independencia de eventos. 8. Resolver y formular problemas usando conceptos básicos de conteo y probabilidad (combinaciones, permutaciones, espacio muestral, muestreo aleatorio, muestreo con reemplazamiento). 9. Proponer inferencias a partir del estudio de muestras probabilísticas.



EJE ORGANIZACIÓN DE DATOS



En el eje organización de datos se utilizan técnicas como la clasificación, organización, representación y modelos a partir de una presentación de datos que pueden obtenerse mediante:

1. Observación directa: En la solución de un problema de investigación, se acude a la fuente de la que se desea obtener la información, para ello se utilizan instrumentos como conversatorios o toma de datos por medio de apuntes de algún fenómeno que se esta observando, por ejemplo preguntar por los precios de algunos artículos en diferentes tiendas. Otro ejemplo de observación directa es anotar el número de infracciones de tránsito que ocurren en un sitio determinado en intervalos de tiempo.
2. Encuestas: La encuesta es un procedimiento utilizado en la investigación de mercados para obtener información mediante preguntas dirigidas a una muestra de individuos representativa de la población o universo, de forma que las conclusiones que se obtengan puedan generalizarse al



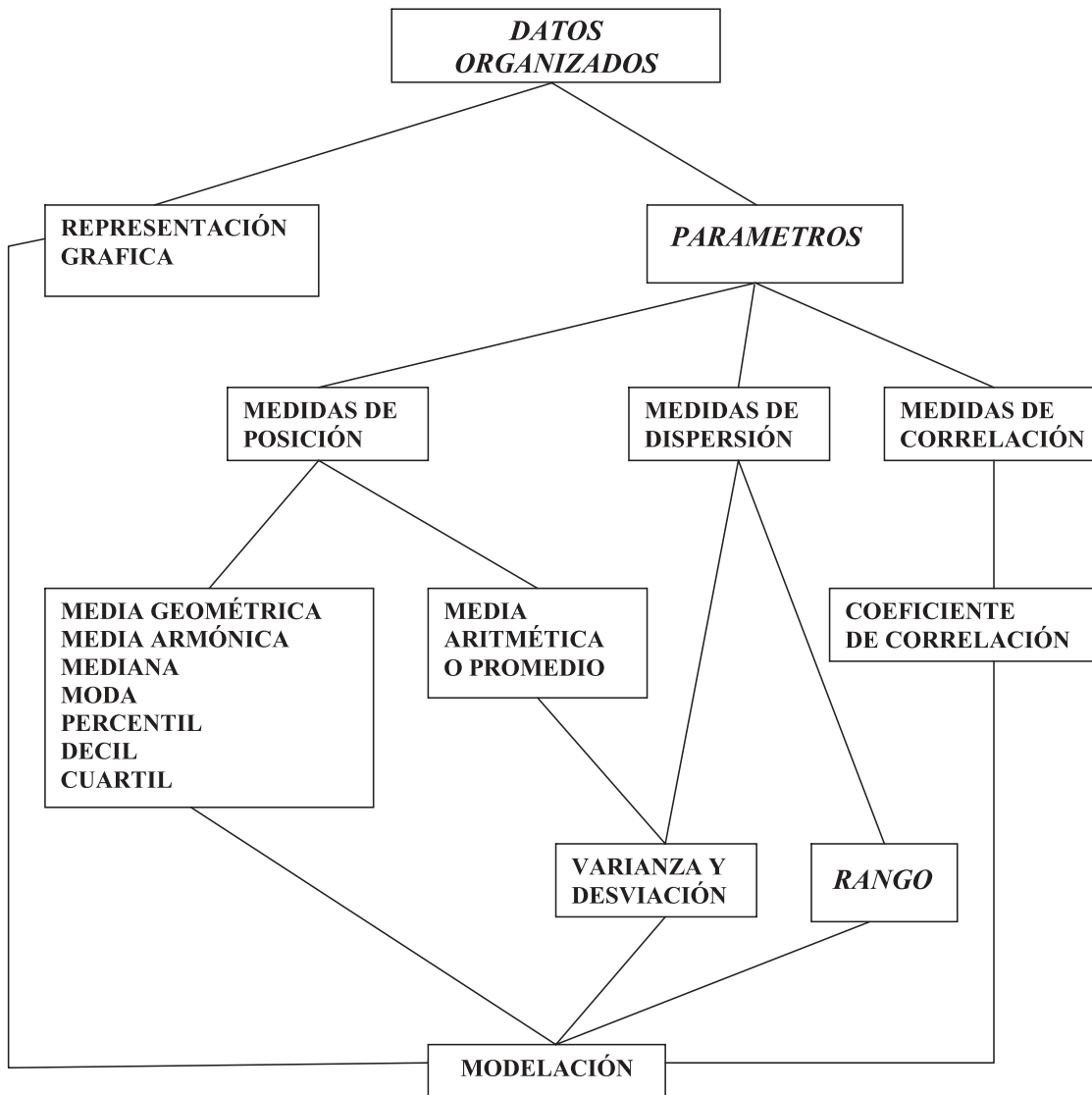
conjunto de la población siguiendo los principios básicos de la inferencia estadística, ya que la encuesta se basa en el método inductivo, es decir, a partir de un número suficiente de datos podemos obtener conclusiones a nivel general.

La principal ventaja de la encuesta frente a otras técnicas es su versatilidad o capacidad para recoger datos sobre una amplia gama de necesidades de información.

3. Archivo: Esta técnica consiste en utilizar datos que fueron recogidos con anterioridad por otra persona o un grupo de investigadores. Por ejemplo tomar datos del Anuario Departamental, informes del Dane o Sistema Nacional de Pruebas entre otros.

Los datos recogidos mediante los instrumentos anteriores se dividen en cualitativos y cuantitativos, y responden a las explicaciones dadas sobre estos mismos conceptos.

EJE MEDIDAS DE POSICIÓN Y VARIABILIDAD

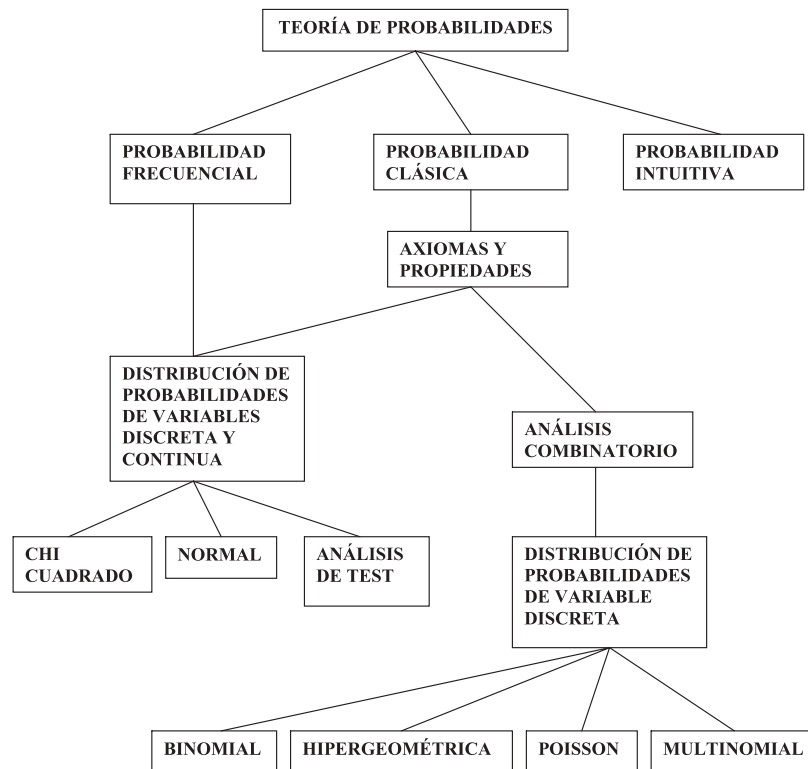




Las medidas de posición y variabilidad responden a lo que se va a realizar y para qué, entre ellas encontramos:

1. Representación gráfica: Tiene como objetivo darle un manejo eficiente a la información obtenida, de tal forma que el lector o quien la observe visualice fácilmente lo que expresan los datos. Por ejemplo para la distribución de variables categóricas, se puede utilizar el diagrama de bloques (rectángulos separados que representan cada uno una categoría de la variable). Si lo que se desea es comparar la frecuencia de una de las clases con la frecuencia de las demás clases o con el total de observaciones, el gráfico más apropiado es el diagrama circular (círculo dividido en tantos sectores como categorías tenga la variable). En el caso de una variable discreta se utiliza el diagrama de barras (varias barras paralelas separadas, las cuales representan una categoría). En el caso de que la variable sea cuantitativa continua, se utiliza un histograma (bloques o rectángulos adyacentes donde cada uno representa la clasificación de la variable).
2. Los parámetros: Son descripciones que se hacen de un conjunto de datos numéricos, entre ellos tenemos: Medidas de posición, que son aquellas que contribuyen a dar la imagen de la correspondiente distribución; ya que determinan valores especiales del conjunto de datos por el lugar que ocupan dentro de ella y entre ellas se encuentran media, mediana, moda, percentil, cuartil, media geométrica y media armónica. Medidas de dispersión; las cuales hacen referencia a que tan separados están entre sí los diferentes valores que asume la variable con respecto a una medida de posición que generalmente es la media. Entre ellas están la varianza, rango y la desviación estándar. Medidas de correlación, son las que establecen relación entre dos variables aleatorias llamadas variable dependiente asociada a la independiente. Entre ellas están el coeficiente de correlación.

EJE PROBABILIDAD E INFERENCIA





La probabilidad y la inferencia conforman el tercer eje temático enmarcado en la Teoría de Probabilidades. Existen fundamentalmente tres tipos de probabilidades: la Frecuencial, la Clásica y la Intuitiva.

Tanto la probabilidad clásica como la frecuencial se construyen sobre axiomas y propiedades de la distribución de probabilidades de variables discretas y continuas, así como también sobre el análisis combinatorio, siendo las distribuciones de probabilidades más usadas la chi-cuadrado, la normal y el análisis de test.

El análisis combinatorio se utiliza para el estudio de la distribución de probabilidades de variable discreta, siendo las más usadas la binomial, la hipergeométrica, poisson y multinomial.

La probabilidad intuitiva de un evento se entiende como el grado de creencia o confianza que un individuo coloca en la ocurrencia de cierto evento, basándose para ello en la evidencia de que dispone. No obstante en la solución de algunos problemas de riesgo o incertidumbre, no basta determinar que existe una probabilidad de ocurrencia de un evento, es necesario interpretarla y utilizar esa información para tomar decisiones que lleven a la consecución de un objetivo, como por ejemplo ganar o minimizar las pérdidas en un juego.



ACTIVIDADES METODOLÓGICAS

DESCRIPCIÓN

La búsqueda de respuestas a preguntas que sobre el mundo físico se hacen los niños resulta ser una actividad rica y llena de sentido si se hace a través de recolección y análisis de datos. Decidir la pertinencia de la información necesaria, la forma de recogerla, de representarla y de interpretarla para obtener las respuestas lleva a nuevas hipótesis y a exploraciones muy enriquecedoras para los estudiantes. Estas actividades permiten además encontrar relaciones con otras áreas del currículo y poner en práctica conocimientos sobre los números, las mediciones, la estimación y estrategias de resolución de problemas.

En la tarea de buscar y recoger datos es importante mantener claros los objetivos, las actitudes, los intereses que la indujeron, prever qué tipos de respuestas se pueden encontrar, las dificultades que podrían presentarse, las distintas fuentes como consultas, entrevistas, encuestas, observaciones, la evaluación de su veracidad, distorsiones, sesgos, lagunas, omisiones y la evaluación de la actitud ética de quien recoge los datos y su responsabilidad social (MEN, Lineamientos Curriculares de Matemáticas).

Algunos de los conceptos involucrados en las actividades o situaciones problema que vamos a presentar son: muestra, población, ordenamiento de datos, valor mínimo (menor), valor máximo (mayor), moda, media aritmética (promedio), distribución de frecuencias, porcentajes y gráficas. Además los relacionados con otros pensamientos, como por ejemplo con el pensamiento métrico, en la utilización del metro para la medición de las estaturas de los estudiantes; así se puede aprovechar la actividad para recordar algunos conceptos fundamentales involucrados en las medidas de longitud y del pensamiento numérico como: La estructura decimal, múltiplos, submúltiplos y notación decimal.

El objetivo de las actividades es presentar a los docentes ejemplos de exploración de datos que puedan utilizar con sus estudiantes. Un proyecto de análisis de datos, a partir de una clase o de un experimento por ejemplo, servirá para abordar los principales contenidos sobre análisis de datos en cualquier nivel de enseñanza, mostrar ejemplos de cuestiones que requieran el uso de conceptos y técnicas estadísticas, describir algunas dificultades previsibles de los estudiantes con los mismos y sugerir criterios para el trabajo en clase con ellos. La metodología del taller alternará el trabajo en grupo y discusión de los docentes y el resumen de algunos puntos claves por el docente.



SITUACIONES PROBLEMA

OBJETIVOS RELACIONADOS CON LOS PROCEDIMIENTOS ESTADÍSTICOS

1. Definir una muestra.
2. Ordenar adecuadamente datos.
3. Representar gráficamente los grupos de datos.
4. Utilizar los métodos estadísticos para realizar análisis de datos y posibles inferencias.

OBJETIVOS DIDÁCTICOS

Las siguientes actividades están enmarcadas dentro del siguiente propósito:

“Explorar e interpretar los datos, relacionarlos con otros, conjeturar, buscar configuraciones cualitativas, tendencias, oscilaciones, tipos de crecimiento, buscar correlaciones, distinguir correlación de causalidad, calcular correlaciones y su significación, hacer inferencias cualitativas, diseños, pruebas de hipótesis, reinterpretar los datos, criticarlos, leer entre líneas, hacer simulaciones, saber que hay riesgos en las decisiones basadas en inferencias” (Vasco C. citado en lineamientos).

PRESENTACIÓN DE ACTIVIDADES O SITUACIONES PROBLEMA A DESARROLLAR POR LOS ESTUDIANTES EN EL AULA

Situación problema uno: *¿Existe relación entre la estatura y el peso?*

Es una actividad a desarrollar por todo el grupo en el aula de clase, pero cada estudiante debe consignar los datos, las operaciones y resultados en su cuaderno.

Material

- Los propios del aula de clase.
- La lista de los estudiantes.
- Los estudiantes.
- Un metro.
- Una balanza.

Desarrollo de la actividad

Solicitar a los estudiantes que propongan métodos para medir la estatura y el peso de cada uno de los compañeros de clase. La discusión nos debe llevar al metro y al kilogramo (fuerza) de peso, aprovechando para recordar los conceptos fundamentales involucrados en ellos, como estructura decimal, múltiplos, submúltiplos, notación de mediciones menores y mediciones mayores.

Luego se les induce a organizar los datos y realizar sus representaciones gráficas.



Preguntas

1. ¿Cuál es el compañero más bajo del grupo?
2. ¿Cuál es el compañero más alto del grupo?
3. ¿Cuál es la estatura que más se repite?
4. ¿Si sumas todas las estaturas y las divides por el número de alumnos del grupo que número obtienes? ¿Qué significado le puedes dar a este número?
5. ¿Si sumas el valor más bajo con el más alto y los divides por dos, qué número obtienes? ¿Tiene alguna relación con el número obtenido en la pregunta anterior?
6. ¿Habrá alguna otra forma de presentar la tabla de valores?
7. Representa estos valores gráficamente.
8. Describe un rango en el cual esté el 70% de los alumnos.
9. Si tomas el dato de la pregunta 5, ¿Cuál es el porcentaje de estudiantes que están entre 5 cm menos y 5 cm más de ese dato?

Situación problema dos: Un fabricante de teclados y máquinas de escribir ha descubierto que en su producto, unas teclas se desgastan más rápidamente que otras. Él está interesado en realizar un estudio sobre cuáles de las teclas debe reforzar en sus productos.

Posibilidad de generar situaciones de aprendizaje . Lo usual es trabajar sobre varios textos , ya que se pretende obtener frecuencias absolutas y relativas sobre el uso de las vocales. Estos conjuntos de datos son obtenidos por los mismos estudiantes, mediante la realización de una lectura con sus compañeros sobre temas diversos, incluyendo temas relacionados con otras áreas o publicaciones.

Las preguntas o cuestionamientos que se le realizan pueden ser entre otras:

Grados 1° A 3°

¿Cuál es la vocal que más se repite?. En este caso que debe hacer el fabricante de teclados?

Grados 4° Y 5°

El ejercicio anterior se repite para los grados 4° y 5°, profundizando en cuanto a consonantes, artículos, y palabras más utilizadas por cualquier autor. Puede incluir en este momento una página o una hoja que el niño escoja.

Grados 6° y 7°

Se ahonda un poco más en el ejercicio y con la ayuda del profesor de Español se llega a las palabras que incluyen 3 vocales, el artículo más usado etc. para ello debe el docente proponer 3 ó 4 hojas ya que este ejercicio se torna más especial.



Grados 8° A 11°

No solamente se lleva el anterior record sino que se ahonda en diptongos, triptongos. Además el comportamiento de las gráficas en este nivel debe ser porcentual y por grados.

Existen vocales juntas? Cuáles son? Cuales se repiten más? Este tipo de situaciones permite correlacionar con otras áreas como humanidades y lengua castellana.

Representar gráficamente las anteriores situaciones.

Fuerte apoyo en representaciones gráficas: “Una idea fundamental del análisis exploratorio de datos es que al usar representaciones múltiples de los datos se convierte en un medio de desarrollar nuevos conocimientos y perspectivas. Esto puede ejemplificarse al pasar de tablas a gráficos, de lista de números a representaciones como la del “tronco”, reduciendo los números a una variedad discreta en un mapa estadístico para facilitar la exploración de la estructura total, construyendo gráficos, como el de la “caja” que hace posible la comparación de varias muestras”. (Biehler [4],pg.2). (Citado por Godino y Batanero 2002).

Situación problema tres: ¿Cuál es la influencia de la luz solar en el crecimiento de una planta?

El educador debe inducir al estudiante en este nivel al cultivo de una planta en lugares diferentes bajo las mismas condiciones en cuanto a la forma de abonarlas con el mismo producto, regarlas con la misma cantidad de agua, además a tomar medidas cada 3 días y representarlas en tablas. El ejercicio debe hacerse por un mes o dos (según la conveniencia). Con lo anterior se realizan preguntas como: cuál crece más rápido. Qué sucedería si las épocas de siembra son diferentes. Cuáles se pueden sembrar al mismo tiempo. En dónde no debe sembrarse. Bajo cuáles condiciones entre otras. Lo anterior se da para los grados 1° a 3°.

Para los siguientes grados 4° a 7°, la actividad ya debe conllevar a una situación casi problemática donde prácticamente, el estudiante es quien lleva la propia iniciativa para resolver la actividad. Además se le pide que grafique la situación de acuerdo a los datos de la tabla y al tipo de variable a estudiar.

Para los grados superiores la actividad en realidad se debe convertir en un proceso investigativo, ya que interviene en dicho proceso otros tipos de variables que pueden ser controlados como es el uso de insecticidas además es importante recalcarle al estudiante que la planta a sembrar debe ser de su propio medio como el entorno donde se mueve el estudiante y la posibilidad o no de este tipo de producto, otra variable a tener en cuenta es el tiempo de cosecha, los insectos, tipo de terreno etc. Se debe recordar el trabajo completo para los grados 10° y 11° con porcentajes y además con análisis predictivos.

Situación problema cuatro: Organización de un campeonato deportivo.

En concordancia con el departamento de Educación Física y Deportes, se le propone a los educadores formalizar los diferentes torneos interclases con la ayuda de los diferentes grupos.



Esta actividad para los grados de 1° a 3° se les induce para que se organice un campeonato con tres equipos. Los estudiantes escogen su deporte favorito y con base en ello, deben formar los equipos con diferentes camisetas, jugar todos contra todos en dos vueltas.

En los grados 4° y 5° se les propone 4 equipos bajo las condiciones anteriores y estando en la mitad del campeonato y realizar la tabla de partidos posibles, se les realiza preguntas como: En la próxima fecha que equipo queda de primero. Cuál va a ser el goleador, además predecir el equipo ganador de la fecha (del día).

A partir de los grados 6° a 11° se debe aumentar el número de equipos y la distribución se hará de acuerdo al grado con el objetivo de ir dificultando el conteo, además la organización del respectivo campeonato debe ser tal que deben entregar planillas desde el inicio del torneo.

A este nivel, el tema de probabilidad se hace más interesante ya que se les pregunta cuál es la posibilidad de que determinado grado gane el campeonato. O que el goleador del torneo sea de un grado específico. Cuál puede ser el goleador del equipo etc.



BIBLIOGRAFÍA

Lineamientos Curriculares de Matemáticas, MEN, 1998

Estándares Curriculares de Matemáticas, MEN, 2003

Carmen Batanero y Juan D. Godino, Estocástica y su Didáctica para maestros, proyectos edumat-maestros, Granada 2002.