REPRESENTACIONES GEOMÉTRICAS Y DINÁMICAS DE CONCEPTOS DE ESTADÍSTICA DESCRIPTIVA

MARLEIDY LORENA GARZÓN RODRÍGUEZ

LEIDY JASBLEIDY JEREZ MOSQUERA

UNIVERSIDAD PEDAGÓGICA NACIONAL

FACULTAD DE CIENCIA Y TECNOLOGÍA

LICENCIATURA EN MATEMÁTICAS

BOGOTÁ D.C.

2014

REPRESENTACIONES GEOMÉTRICAS Y DINÁMICAS DE CONCEPTOS DE ESTADÍSTICA DESCRIPTIVAS

MARLEIDY LORENA GARZON
CÓDIGOS: 2008140039
C.C: 1024508875
LEIDY JASBLEIDY JEREZ MOSQUERA
CÓDIGOS: 2008240038
C.C: 1024504678

Director:
Benjamín Sarmiento
Docente Departamento de Matemáticas

UNIVERSIDAD PEDAGÓGICA NACIONAL
FACULTAD DE CIENCIA Y TECNOLOGÍA
LICENCIATURA EN MATEMÁTICAS
BOGOTÁ D.C.
2014

1.Información General	
Tipo de documento	Trabajo de Grado
Acceso al documento	Universidad Pedagógica Nacional. Biblioteca Central
Título del documento	Representaciones geométricas y dinámicas de conceptos de estadística descriptiva
Autor(es)	Garzón Rodríguez, Marleydi Lorena; Jerez Mosquera, Lady Jasbleidy
Director	Sarmiento Lugo, Benjamín
Publicación	Bogotá. Universidad Pedagógica Nacional, 2014. 89p.
Unidad Patrocinante	Universidad pedagógica Nacional
Palabras Claves	Representaciones geométricas de conceptos de estadística descriptiva

2.Descripción
Este trabajo de grado presenta un recurso virtual como fuente modeladora de graficos estadísticos y de la recopilación de construcciones geométricas que representan diferentes conceptos de Estadística descriptiva. El documento consta de dos capítulos: en el primer capítulo se presenta la introducción, los objetivos y el marco conceptual. En el segundo capítulo se presentan subcapítulos de las construcciones recopiladas para los diferentes temas de estadística descriptiva; el primer subcapítulo exhibe un conjunto de representaciones dinámicas de los respectivos

gráficos estadísticos de acuerdo al tipo de variable, el siguiente expone las construcciones geométricas que representan los diferentes conceptos de las medidas de localización y las relaciones que se dan entre las mismas. En el tercer subcapítulo se presentan las construcciones que ilustran el significado desde una perspectiva geométrica de las medidas de dispersión, en el cuarto subcapítulo se enseña una construcción que muestra la relación entre las medidas de localización dadas según la medida de forma de la distribución. El último subcapítulo explica geométricamente el significado de la recta de regresión lineal.

Cada subcapítulo contiene la manera como se diseñó cada representación y una breve explicación de cómo se debe usar la construcción con una sugerencia de cómo utilizarla en clase.

Finalmente presentamos la recopilación de las construcciones en un C.D como recurso virtual, estas construcciones contienen la representación geométrica o dinámica, y junto a ella se podrá conocer el modo de uso por medio de una sugerencia para clase.

3.Fuentes

- Fernández, F. y Sarmiento B. (2009). Curso básico de estadística, introducción al análisis de datos. Bogotá, Colombia. Editorial CARGRAPHIS S.A.
- Triola, M. Estadística. (2004). México. Editorial Pearson. 614p
- Gonzáles, M. Estadística aplicada. (2009). Madrid, España. Editorial Díaz De Santos. 759 p.
- Murray R. Spiegel, John J. Schiller, R. Probabilidad y estadística. (2010). México. Editorial McGraw- Hill. 425 p.
- Ross, M Sheldon. Introducción a la Estadística. (2007). Barcelona- Bogotá- Buenos Aires- Caracas- México. Editorial REVERTÉ S. A.
- Martínez, C. Estadística básica aplicada. (2006). Santa fe de Bogotá: ECOE Ediciones. 386 p.

- MONTGOMERY, Douglas C. & RUNGER, George C. (1997). Probabilidad y Estadística aplicadas a la Ingeniería. México: McGraw Hill.
- Johnson, R. Kuby, P. ESTADÍSTICA ELEMENTAL. (2008). México. Editorial impresiones S.A
- Cuervo, C. Edilberto. Estadística Matemática. (2009). Bogotá, Colombia. Editores Litográficos Ltda.

4.Contenidos

MARCO CONCEPTUAL

En el desarrollo de este marco conceptual especificamos el significado de cada concepto a partir de la estadística. A continuación presentamos los títulos de los diferentes contenidos de este capítulo.

- Variables y clasificación de las variables.
  Para este ítem se presenta la definición de cada tipo de variables y sus subdivisiones.
- Escalas de medición.
  En este ítem se definen las escalas de medición de acuerdo a los diferentes tipos de variable
- Tablas de frecuencia
  Las tablas de frecuencia se presentan junto a los gráficos estadísticos, los cuales están agrupados de acuerdo al tipo de variable con el que corresponde cada uno.
- Diagramas según el tipo de variable:
  - Para Variables Cualitativas.
  - Para Variables Cuantitativas Discretas.
  - Para Variables Cuantitativas Continuas
- Análisis numérico.
  En este análisis numérico se presentan los conceptos y las relaciones dadas entre las medidas de tendencia central, de localización, de dispersión y de forma.
  Medidas de tendencia central.

  Medidas de localización.

  Medidas de dispersión.

  Medidas de forma.

- Regresión lineal.
  En este ítem se exhibe el método usado para encontrar la recta de mayor ajuste a

una distribución de puntos.

REPRESENTACIONES DINÁMICAS DE TABLAS Y GRÁFICOS.

En esta sección se presentan las construcciones dinámicas y geométricas con su respectiva interpretación.

- Representaciones geométricas de tablas y gráficos.
- Construcciones de medidas de localización.
- Construcciones de medidas de dispersión.
- Construcciones de medidas de forma.
- Construcción de recta regresión lineal.

CONCLUSIONES.
BIBLIOGRAFIA.

5.Conclusiones
En las conclusiones se presenta una breve descripción de los aportes que obtuvimos en el proceso de construcción de este recurso virtual, y otras que deducimos mediante el desarrollo del trabajo de grado. • De acuerdo al primer objetivo: *Recopilar construcciones geométricas sobre temas de Estadística Descriptiva.* Se concluye que en los textos se encuentran las ilustraciones de los gráficos estadísticos, aunque están no son dinámicas, y también es posible encontrar la construcción de mediana, moda y comparación de medias para datos agrupados. Estas representaciones muestran cómo llegar a las formulas por medio de semejanza de triángulos. • *Diseñar representaciones geométricas para ilustrar los diferentes conceptos de Estadística descriptiva.* Se concluye que por medio de diferentes elementos geométricos se pueden realizar construcciones geométricas interactivas para conceptos de estadística descriptiva • *Elaborar un C.D con la recopilación de las ilustraciones geométricas sobre los conceptos de Estadística descriptiva.* Se concluye que un recurso bibliográfico como este será una herramienta de clase que permitirá corroborar los procesos hechos en “lápiz y papel”

• La geometría y la estadística son dos ciencias que permiten hacer construcciones que las relacionan y por medio de las propiedades de los diferentes elementos geométricos se hace una lectura del concepto desde otra perspectiva. • La construcción de un recurso que relacione dos ciencias nos permite identificar la relación entre ellas y así mostramos que no son estudios independientes, si no que por el contrario podemos observar cómo se enlazan para el desarrollo de diferentes temas.

Elaborado por:	Garzón Rodríguez Marleydi Lorena y Jerez Mosquera Leidy Jasbleidy
Revisado por:	Sarmiento Lugo, Benjamín

Fecha de elaboración del Resumen:	21	07	2014

TABLA DE CONTENIDO

INTRODUCCIÓN

El presente trabajo surge a partir de la necesidad de generar una alternativa, que brinde elementos teóricos los cuales nos permitan desde la geometría acceder a diferentes conceptos de Estadística Descriptiva, razón por la cual se hace una recopilación de las construcciones geométricas y dinámicas; estas nos muestran una perspectiva distinta para las tablas de frecuencia, diagramas; medidas de tendencia central, dispersión, de forma y regresión lineal.

Por medio de las representaciones geométricas se muestra un camino distinto para interpretar diferentes conceptos, y a partir de las propiedades de cada construcción se logra deducir algunas fórmulas dadas en estadística para encontrar el valor de diferentes medidas.

Esta es una fuente asequible que recopila representaciones geométricas y dinámicas, las cuales permiten que el estudiante interactúe con un medio tecnológico, por medio del cual se identifica la fortaleza de este tipo de herramientas. De acuerdo a lo anterior construimos este recurso bibliográfico dirigido a docentes de matemáticas y de otras áreas relacionadas con los conceptos de Estadística descriptiva, para ser desarrollado en diversos contextos académicos como la escuela, colegio y universidad; los talleres que diseñen los docentes con base en este recurso deben estar acorde con el plan de estudios del área.

El trabajo de grado está compuesto de dos capítulos y un recurso virtual en C.D. En el primer capítulo se presenta la introducción, objetivos y se describe el marco conceptual donde se explica el significado e interpretación de cada tema desde la Estadística descriptiva. El segundo capítulo contiene las construcciones geométricas, con su respectiva interpretación y la explicación del uso de cada construcción, dando una sugerencia para generar una propuesta de taller o una actividad de observación que permitan identificar las relaciones entre diferentes conceptos y medidas a partir de las representaciones geométricas contenidas en los applets.

En el C.D se encuentran los applets de construcciones geométricas y dinámicas como recurso virtual con el que es posible interactuar; para identificar los cambios o relaciones de conceptos estudiados y representados con cada construcción, donde utilizarán las herramientas de la información y la comunicación que les permite confirmar relaciones o conceptos vistos desde la teoría, ya que de esta manera les es posible visualizar varios eventos que son susceptibles de cambios sobre la construcción identificando así relaciones entre diferentes elementos.

El aporte de este material es también de tipo pedagógico ya que se muestra como una herramienta para trabajar en clase de una manera distinta a lo que se acostumbra a desarrollar para estos conceptos. Desde nuestra experiencia como practicantes de la licenciatura en matemáticas y en el desarrollo de este trabajo de grado identificamos la fortaleza de este tipo de herramientas, por ejemplo: en la construcción de comparación de medias se presenta el caso de la desigualdad $\frac{2ab}{a+b} < \sqrt{ab} < \frac{a+b}{2} < \sqrt{\frac{a^2+b^2}{2}}$ donde a es distinto de b, pero cuando a=b se establece que: $\frac{2ab}{a+b} = \sqrt{ab} = \frac{a+b}{2} = \sqrt{\frac{a^2+b^2}{2}}$.

OBJETIVOS

Objetivo general

Elaborar un recurso bibliográfico[1] que contenga construcciones geométricas sobre temas de Estadística Descriptiva.

Objetivos específicos

- Recopilar las construcciones geométricas sobre medidas de tendencia central, dispersión y forma.

- Diseñar representaciones geométricas para ilustrar algunos conceptos de Estadística descriptiva.

- Elaborar un C.D con la recopilación de las ilustraciones geométricas para algunos conceptos de Estadística descriptiva

[1] Se entiende recurso bibliográfico como un conjunto de documentos que han pasado por un proceso el cual permite que estos estén disponibles para un mayor número de individuos, quienes estén interesados en la información contenida. Esta herramienta es un conjunto de metadocumentos de acuerdo a la definición dada por Ranganathan (Matemático y bibliotecario de India, 1973).

MARCO CONCEPTUAL

ESTADÍSTICA DESCRIPTIVA

Se le llama ESTADÍSTICA DESCRIPTIVA a la ciencia que se ocupa del proceso de describir y analizar una población por medio de un conjunto sistemático de procedimientos utilizados para observar y describir numéricamente dicha muestra representativa de la población en estudio, sin dar conclusiones o inferencias de tipo general para esta. En sus estudios solo se habla de datos tomados y conocidos. Se describen las características de la muestra resaltando tendencias, agrupaciones, promedios frente a datos numéricos y se representa los datos reunidos de manera gráfica y tabulada, los gráficos y las tablas deben ser acordes al tipo de información recolectada.

3. VARIABLES

Una variable es el dato en estudio que cambia de valor entre los diferentes individuos de una muestra representativa de la población que se quiere describir. De acuerdo a los diferentes datos obtenidos por medio de una encuesta se conocen dos tipos de variables, las cualitativas y las cuantitativas.

3.1 CLASIFICACIÓN DE LAS VARIABLES

3.1.1 Variables Cualitativas

Las variables cualitativas son elementos de una población que pueden ser descritos mediante palabras a las que usualmente se le llaman atributos los cuales agrupan la información obtenida de la población en grupos determinados por una de las clases dada por la variable, estas no representan cantidad, si no que da una cualidad de la población por ejemplo: Género, Estrato socioeconómico. Estas variables pueden ser de tipo nominal: cuando no existe una jerarquía entre ellas; u ordinales si es posible establecer un orden entre los posibles valores que puede tomar la variable.

3.1.2 Variables cuantitativas.

Las variables cuantitativas son elementos cuantificables que representan cantidades ejemplo: número de hijos, número de empleados, peso estatura, salarios, etc.

Estas pueden ser de tipo discreto o continuo

3.1.3 Variables cuantitativas discretas

Este tipo de variables toman valores enteros iguales o mayores a 0. Ejemplo: número de estudiantes de un curso, número de propiedades, etc. En estos casos no se puede decir que hay 30 niños y medio, o que se 3 y un cuarto de propiedades, por eso solo se admiten valores enteros.

3.1.4 Variables cuantitativas continuas

Este tipo de variable permite infinitos valores entre los números enteros, ya que en estos casos si es posible hablar de cantidades racionales no enteras como media libra, un metro y setenta centímetros etc. Para estos valores es posible establecer intervalos. Ejemplo: altura, presión, peso, etc.

3.2. ESCALAS DE MEDICIÓN

Existen cuatro tipos de escalas de medición, presentadas a continuación.

3. 2.1 Escala Nominal: Las unidades que pertenecen a esta escala se agrupan en clases excluyentes según determinada propiedad, con lo que se define cada partición sobre el conjunto de tales unidades. Ejemplo: Al estudiar el porcentaje de desempleo de un país y agrupar la población en dos grupos excluyentes de acuerdo a su género: masculino y femenino.

3.2.2 Escala Ordinal: Son los datos que tienen una jerarquización entre ellos, algunas veces son numéricos pero estos no representan cantidades tangibles. Por ejemplo: nivel de estudios o estrato socioeconómico.

3.2.3 Escala de intervalos: Los valores se pueden organizar en intervalos, los cuales tienen la misma longitud, estos valores pueden utilizasen para hacer comparaciones. El valor del cero en esta escala no es un cero absoluto sino arbitrario ya que no representa un conjunto vacío de la magnitud medida, este es el límite inferior de la primera clase. Ejemplo la escala de temperatura centígrada, en la cual se pueden hacer comparaciones de cuánto tiene demás una comparada con la otra, o la magnitud del intervalo, sin embargo no será posible hacer afirmaciones como: 40° es el doble de 20°, por no tener el cero absoluto.

3.2.4 Escala de razón: Esta es el nivel de medición más completo, ya que además de cumplir con las mismas propiedades de la escala de intervalos, posee el cero absoluto que representa la ausencia completa de magnitud para la variable de medida. En una escala de razón los valores se pueden comparar, ordenar, comprobar la igualdad de diferencia y la igualdad de razones y cocientes. Ejemplos: longitud, peso, distancia, etc.

3.3 TABLAS DE FRECUENCIA

La estructuración y ordenación de estos datos se realiza en una primera instancia en tablas de frecuencia, que son distintas según el tipo de datos recogidos, porque para cada clase de variable (dato) hay una tabla de frecuencia que le corresponde. Por lo tanto a continuación se especifica las características de cada tabla según el tipo de variable.

3.3.1 **Para variables cualitativas**

Si el dato analizado de la población y sus posibles valores calificativos son atributos medibles o calificados con palabras; es decir son características cualitativas de una población, las observaciones de estos se deben presentar en una tabla de dos columnas y m filas, siendo m los posibles valores de la variable y n el tamaño de la muestra.

Categoría	N_i
A_1	n_1
A_2	n_2
.	.
A_m	N_m

Tabla 1

En la tabla anterior A_1, A_2,.., A_m son los posibles valores que puede tomar la variable, por ejemplo si se midiera el número de personas que tienen x tipo de sangre, una seria O+, otra O-, otra A+, y así sucesivamente de acuerdo a la cantidad de valores posibles para la variable; y los N_i son las diferentes cantidades en que se da cada valor de la categoría, las cuales se corresponden con cada atributo, por ejemplo: 25 personas son O+, 10 personas son O-, 20 personas son A+ etc.

3.3.2 Para variables cuantitativas discretas

Si los valores observados corresponden a una variable cuantitativa discreta, para construir una tabla de frecuencias absolutas se colocan en la primera columna los valores observados en orden ascendente C_1, C_2,…, C_r y en la segunda columna las frecuencias absolutas correspondientes a cada valor.

Categorías	*Frecuencia Absoluta*
C_1	F_1
C_2	F_2
.	.
C_m	F_m

Tabla 2

3.3.3 Para variables cuantitativas continuas

En este caso se agrupan los datos en intervalos y se asignan a cada intervalo una frecuencia absoluta igual al número de observaciones que caen dentro de él.

Las tablas de frecuencias absolutas, relativas absolutas acumuladas y relativas acumuladas en este caso son de la siguiente forma:

Clases	Marcas	Frecuencias	Fre. Relativas	Frecuencia. Acumulada
$[L_0, L_1)$	X_1	f_1	—	
$[L_1, L_2)$	X_2	f_2	—— *100	
:	:	:	:	:
$[L_{i-1}, L_i)$	X_i	f_i	——	
:	:	:	:	
$[L_{k-1}, L_k)$	X_k	f_k	—— *100	

Tabla 3

Los extremos de estas clases se llaman límites reales porque coincide el extremo superior de cada una de ellas con el extremo inferior de la siguiente (L_0, L_1, L_2,…L_{n-1}, L_n). Con los limites reales no hay ambigüedad al hacer el recuento de los datos, ya que por ser semiabiertos no se incluyen los extremos en el conteo. En la segunda columna están las marcas de clase, lo que se define como punto medio del intervalo, (X_0, X_1, X_2,…X_K $=\frac{L_{n-1}+L_n}{2}$). En la tercera columna están las frecuencias absolutas correspondientes a cada intervalo, en la cuarta la frecuencia relativa porcentual (acumulada), y finalmente la última columna contiene las frecuencias acumuladas.

3.4. DIAGRAMAS SEGÚN EL TIPO DE VARIABLE

De acuerdo al tipo de variable y a la cantidad de datos se puede escoger el grafico adecuado para representar dicha información. Por lo tanto a continuación se muestran los gráficos con los cuales es posible representar el estudio de una variable cualitativa en una población y luego los gráficos utilizados para representar la información recolectada de variables cuantitativas.

3.4.1 Variables Cualitativas

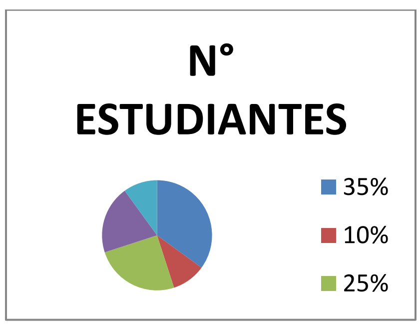
Para representar información de variables cualitativas, se debe tener en cuenta la cantidad de datos recolectados, al elegir el diagrama con el cual se ilustrara la información obtenida y el tipo de dato, por ejemplo si las frecuencias de los diferentes valores de la variable son porcentajes entonces el diagrama más adecuado es un gráfico de sectores.

- **Gráficos de Sectores**

Sobre un círculo que simboliza el total de la población se dibujan sectores circulares cuyo ángulo central es el proporcional a la frecuencia absoluta de la modalidad que representa. Ejemplo:

SEMESTRE	N° ESTUDIANTES
PRIMERO	35%
SEGUNDO	10%
TERCERO	25%
CUARTO	20%
QUINTO	10%

Tabla 4



Gráfica 1

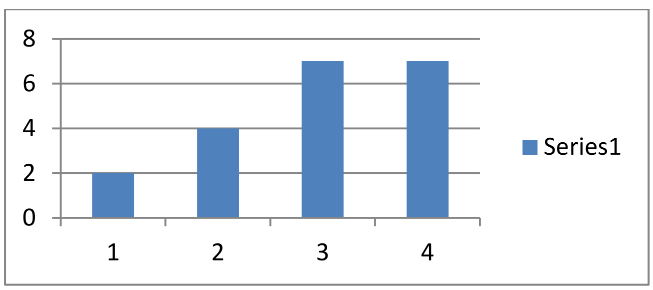
- **Diagrama de Columnas**

Es un diagrama formado sobre un plano cartesiano, sobre el eje de las abscisas se divide en tantos segmentos como la cantidad de posibles valores que pueda tomar la variable observada, los cuales deben ser de la misma longitud pero no consecutivos (esta es la diferencia grafica entre el diagrama de columnas y el histograma). En el eje de las ordenadas se representan las frecuencias absolutas, dibujando rectángulos que tienen como base los segmentos (de la abscisas) y alturas iguales a las frecuencias absolutas correspondientes, o proporcionales a ellas.

Ejemplo:

Estrato Socio económico	Frecuencia
0	2
1	4
2	7
3	7
4	4

Tabla 5



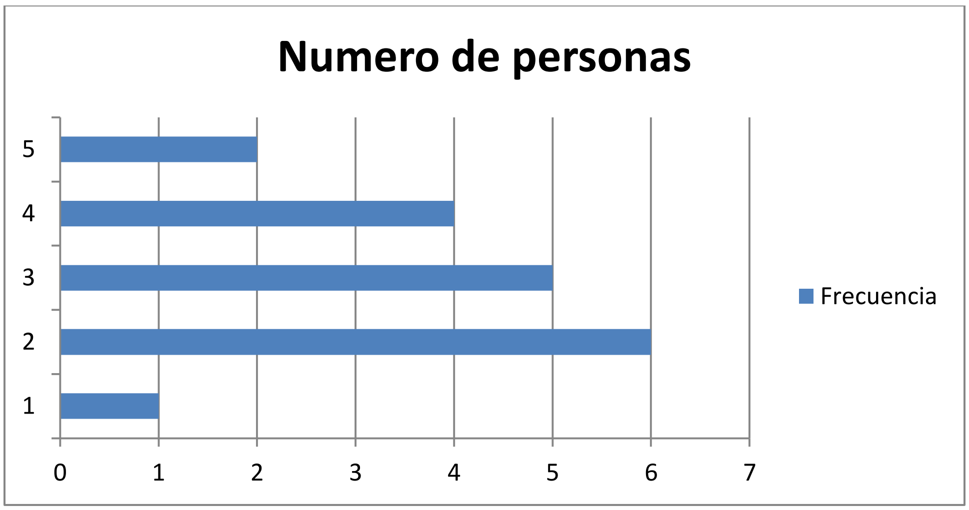
Gráfica 2

- **Diagrama de Barras**

Es un diagrama con una estructura similar al diagrama de columnas, por lo tanto es posible que para el tipo de información que sea adecuado el de Diagrama de columnas este también lo sea. Se construye sobre un plano cartesiano, en el cual se ubican sobre el eje de las ordenadas tantos segmentos como el número de valores que tenga la variable observada, los cuales deben ser de la misma longitud pero no consecutivos.

En el eje de las abscisas se representan las frecuencias absolutas, dibujando rectángulos que tienen como altura los segmentos (de la abscisas) y bases que miden tanto como las frecuencias absolutas correspondientes a cada uno, o proporcionales a ellas.

Ejemplo: Representamos la misma información de la tabla 1 con un diagrama de barras



Gráfica 3

- **Pictogramas**

Se tienen al lado de cada modalidad unos dibujos o figuras alusivas al carácter representado cuyas dimensiones, o número de veces que se repite el dibujo, son proporcionales a la frecuencia absoluta de correspondiente modalidad.

AÑOS	MILLONES DE HABITANTES
1930	62
1950	78
1970	101
1990	128

Tabla 6

Gráfica 4

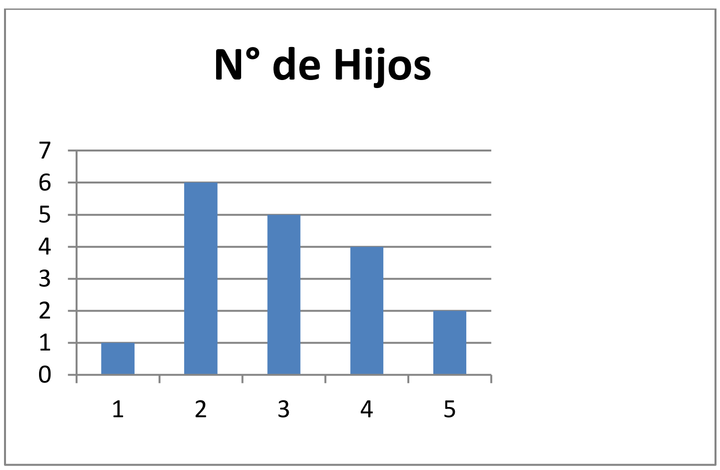
3.4.2 **Para Variables Cuantitativas Discretas**

- **Diagrama de Columnas**

Las alturas de las barras son directamente proporcionales a las frecuencias absolutas o relativas. Sus segmentos de base no son continuos al igual que para la variable cualitativa.

N° de Hijos	Frecuencia
0	2
1	4
2	7
3	7
4	4

Tabla 7



Gráfica 5

- **Diagrama de tallos y hojas**

Es para graficar pocos datos. Dados los datos numéricos de dos cifras, separamos la primer cifra de cada número y se ubican a la izquierda de la línea vertical que separa las cifras, y de esta manera se forma el tallo, luego la segunda cifra deberá ir a la derecha de la línea vertical frente a su primer cifra correspondiente, formando así la hoja, por ejemplo si tenemos 11, 12, 15, 16, 19 el uno hará parte del tallo y el 1,2,5,6,9 la primer hoja ubicada frente al uno, estas segundas cifras se ubican de manera horizontal.

Edades: 11,12, 11, 23, 21,24,14,11,32,13,33, 34, 35, 36, 42, 43, 41, 44, 52, 53, 54,57

Tabla 8

5	2 3 4 7
4	1 2 3
3	2 3 4 5
2	1 3
1	1 1 2 3

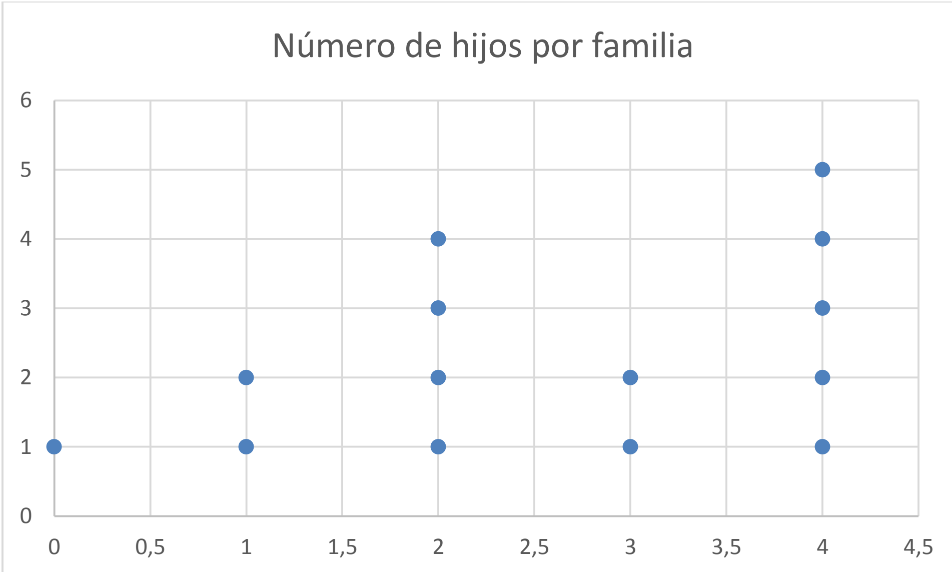
Gráfica 6

- **Diagrama de puntos**

Este gráfico se hace sobre una recta numérica con valores discretos la recta tiene un valor mínimo dado por el mismo conjunto de datos, al igual que el valor máximo que es el tope, la frecuencia absoluta de cada valor es igual a la cantidad de puntos ubicados en línea recta sobre el valor.

N° de Hijos	Frecuencia
0	1
1	2
2	4
3	2
4	5

Tabla 9



Gráfica 7

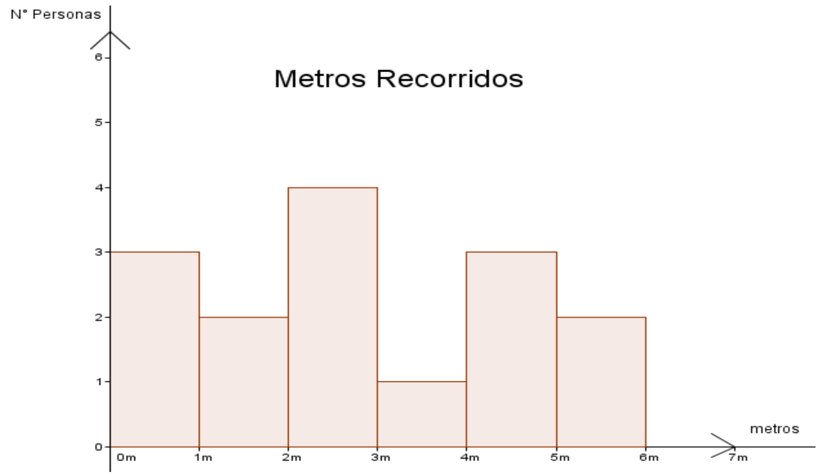
3.4.3 **Para Variables Cuantitativas Continuas**

- **Histogramas de frecuencias**

Se representa sobre un plano cartesiano, en el eje de las abscisas los límites reales de las clases, obteniendo de esta manera los segmentos de base de igual tamaño y consecutivos. Sobre cada una de las clases se traza un rectángulo de altura igual a la frecuencia absoluta sobre la amplitud de este, a lo que se le denomina la frecuencia media por unidad de amplitud.

Metros recorridos	N° de personas
0 - 1	3
1 - 2	2
2 - 3	4
3 - 4	1
4 - 5	3
5 - 6	2

Tabla 10



Gráfica 8

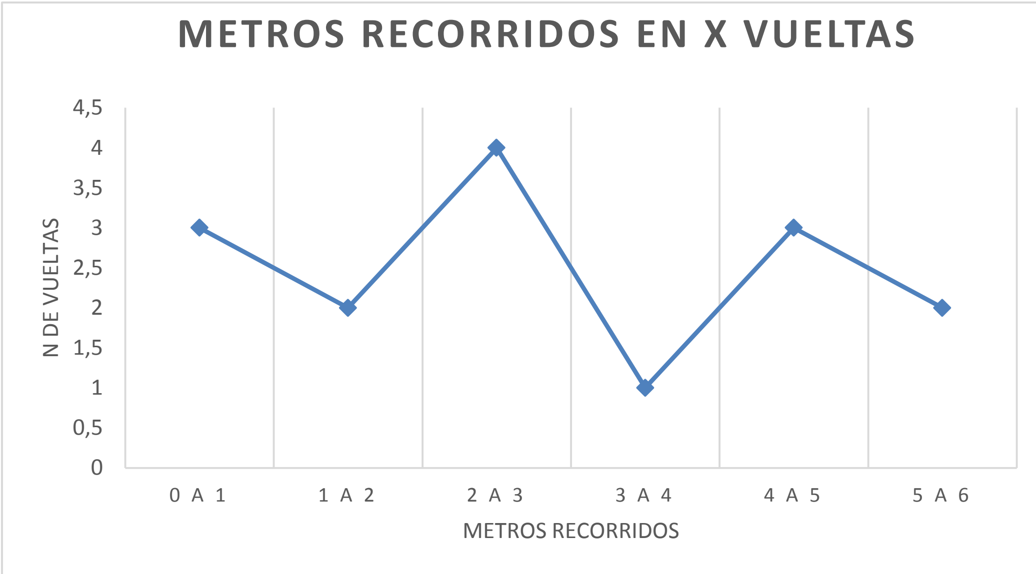
- **Polígono de frecuencia**

Se señalan en las bases superiores de los rectángulos del histograma, siendo estos los puntos medios y se dibujan dos clases más de la misma amplitud, una anterior a la primera y otra a continuación de la última, a las que se les asigna frecuencia cero, es decir, se dibujó los puntos, $\boldsymbol{x_i}, f_i$, siendo $\boldsymbol{x_i}$ las marcas de clase y f_i la frecuencia del intervalo correspondiente a la marca de clase; luego de tener los puntos estos se unen con segmentos de línea recta formando así el polígono.

Ejemplo: La representación gráfica por medio de un polígono de frecuencias para la tabla 11, se presenta en la gráfica 9.

Metros recorridos	N° de vueltas
0 a 1	3
1 a 2	2
2 a 3	4
3 a 4	1
4 a 5	3
5 a 6	2

Tabla 11



Gráfica 9

3.5. ANÁLISIS NUMÉRICO

Las variables cuantitativas permite además un resumen numérico de la muestra es decir obtener a partir de los datos observados un número reducido de valores característicos, que son los estadísticos o estadígrafos.

3.5.1 Medidas de tendencia central

3.5.1.1 **Media aritmética:** representa el promedio de un conjunto de datos agrupados o no agrupados.

a) si los datos no están agrupados en intervalos. La media aritmética se halla así:

$$\mathbf{x} = \frac{{}^{\mathbf{r}}_{\mathbf{j=1}}\, \mathbf{x_i n_i}}{\mathbf{n}}$$

Siendo $\mathbf{x_i}$ los diferentes valores de la variable cuantitativa discreta y $\mathbf{n_i}$ sus frecuencias absolutas.

b) Si los datos están agrupados en intervalos, se debe hallar así:

$$\mathbf{x} = \frac{{}_{\mathbf{j=1}}^{\mathbf{k}} \mathbf{x_i n_i}}{\mathbf{n}}$$

Siendo $\mathbf{x_i} = \frac{L_{i-1}+L_i}{2}$, la marca de clase del intervalo (L_{i-1}, L_i), y $\mathbf{n_i}$ la frecuencia absoluta de cada intervalo correspondiente a $\mathbf{x_i}$

3.5.1.2 Media geométrica: si todos los valores de la variable son positivos, se define su media geométrica como:

$$\mathbf{m_g} = \sqrt[\mathbf{n}]{\mathbf{x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}}}$$

3.5.1.3 **Media aritmética ponderada:** se define la media global de las s muestras como una media ponderada, considerando como pesos los tamaños de las s muestras:

$$\mathbf{x} = \frac{{}_{\mathbf{j=1}}^{s} \mathbf{x_i n_i}}{{}_{\mathbf{j=1}}^{s} \mathbf{n_i}}$$

Pues así es como se obtiene en el numerador le suma de los valores de las s muestras y en el denominador el total de valores recogidos. Donde cada valor tiene una importancia distinta.

3.5.1.4 **Media armónica:** si todos los valores de la variable son positivos, se define la media armónica del siguiente modo:

$$\mathbf{m_a} = \frac{{}_{\mathbf{j=1}}^{\mathbf{r}} \mathbf{n_i}}{{}_{\mathbf{j=1}}^{\mathbf{r}} \frac{\mathbf{1}}{\mathbf{x}} \mathbf{n_i}}$$

Es por tanto $\frac{1}{m_a} = \frac{{}_{i=1}^{r}\frac{1}{x_i}n_i}{{}_{i=1}^{r}n_i}$ es decir, la media armónica es el recíproco de la media aritmética de los valores de las variables.

3.5.1.5 media cuadrática: se define como la raíz cuadrada de la media aritmética de los cuadrados de los valores de la variable, es decir:

$$\mathbf{m_c} = \overline{\frac{{}_{i=1}^{r} x_i^2 n_i}{{}_{i=1}^{r} n_i}}$$

3.5.2 Medidas de localización.

3.5.2.1 **Mediana:** La mediana es la medida de tendencia central que separa la distribución en dos partes de igual frecuencia, es decir, la mediana es el valor que deja tantos valores de la variable por encima como por debajo de ella, cuando los datos están ordenados en orden creciente o decreciente. Se designará por $\mathbf{M_e}$.

Si los datos no están agrupados en intervalos:

Si el número de datos es impar, la medida es el valor central, es decir, se ubica en el lugar $\frac{n}{2} + 1$, parte entera de $\frac{n}{2}$ más una unidad al organizarse los datos de manera ascendente.

3.5.2.2 **Cuantiles**

Los cuantiles, también llamados centiles, son valores que dividen la distribución en partes de igual frecuencia, como lo son los cuartiles, los deciles y los percentiles.

3.5.2.3 **Los cuartiles**: son tres valores $\mathbf{Q_1, Q_2}$ $\mathbf{y}$ $\mathbf{Q_3}$ que dividen la distribución en cuatro partes de igual frecuencia.

Colocados los datos en orden creciente, el primer cuartel deja por debajo la cuarta parte de los datos, el segundo cuartel de la mitad y el tercer cuartel de las tres cuartas partes. Por tanto, el segundo cuartel coincide con la mediana: $\mathbf{Q_2 = M_e}$.

Si los datos no están agrupados en intervalos:

Para calcular, $\mathbf{Q_j}$ cuartil – j ésimo. Si $\frac{j_n}{4}$, no es entero, $\mathbf{Q_j}$ es el valor que ocupa el lugar $\frac{j*n}{4} + \mathbf{1}$, parte entera de $\frac{j*n}{4}$ más la unidad. Si $\frac{j*n}{4}$ es entero, $\mathbf{Q_j}$ es la media aritmética de los dos valores que ocupa los lugares $\frac{j*n}{4}$ y $\frac{j*n}{4} + \mathbf{1}$.

Si los datos están agrupados en intervalos:

Para determinar $\mathbf{Q_j}$, el cuartil j-esimo, primero se determina la clase que lo contiene que es aquella cuya frecuencia absoluta acumulada es la primera mayor o igual a $\frac{\mathbf{jn}}{\mathbf{4}}$. Si esta clase es $\mathbf{L_{i-1}, L_i}$ que se denomina Q_j como en el caso de la mediana obteniéndose:

$$Q_j = L_{i-1} + \frac{j\frac{n}{4} - N_{i-1}}{n_i} c_i$$

Siendo L_{i-1}: el límite inferior de la clase i, n representa el cuartil que se va a hallar, N_{i-1}: se refiere a la frecuencia acumulada hasta el intervalo anterior, n_i:es la frecuencia absoluta para dicho intervalo y c_i: es el tamaño del intervalo.

3.5.2.4. **Deciles:** los deciles son nueve valores que dividen la distribución en 10 partes de igual frecuencia. Colocados los datos en orden creciente D_j el decil j- esimo, deja por debajo las $\frac{j}{10}$ partes de las observaciones.

Para los datos agrupados en intervalos, primero se determina la clase $L_{i-1} - L_i$ que lo contiene y, como en los casos anteriores, se obtiene:

$$D_j = L_{i-1} + \frac{j\frac{n}{10} - N_{i-1}}{n_i} c_i$$

Siendo L_{i-1}: el límite inferior de la clase i, n representa el decil que se va a hallar, N_{i-1}: se refiere a la frecuencia acumulada hasta el intervalo anterior, n_i:es la frecuencia absoluta para dicho intervalo y c_i: es el tamaño del intervalo.

3.5.2.5 **Centiles o percentiles:** Son noventa y nueve valores. El percentil P_j es el valor de la variable tal que el j% de las observaciones son inferiores o iguales a él. Para datos agrupados en intervalos, primero se determina la clase que lo contiene, si esta es $L_{i-1} - L_i$ como en los casos anteriores, se obtiene:

$$P_j = L_{i-1} + \frac{j\frac{n}{100} - N_{i-1}}{n_i} c_i$$

Siendo L_{i-1}: el límite inferior de la clase i, n representa el percentil que se va a hallar, N_{i-1}: se refiere a la frecuencia acumulada hasta el intervalo anterior, n_i:es la frecuencia absoluta para dicho intervalo y c_i: es el tamaño del intervalo.

Si los datos están agrupados en intervalos: en primer lugar, se determina la clase modal que es la que tiene frecuencia máxima, o las clases modales si hay más de una clase con la máxima frecuencia, o más de una que cumpla la condición de que su frecuencia no es inferior a la de las clases anterior y posterior a ella.

Dentro de la clase modal, la moda es el punto al que le correspondería la frecuencia máxima suponiendo que el aumento de la frecuencia $n_i - n_{i-1}$ de la clase anterior a la clase modal, y la disminución de esta clase de la frecuencia de esta clase a la siguiente $n_i - n_{1+1}$ se distribuyen uniformemente.

3.5.3 Medidas de Dispersión

3.5.3.1 **Rango medio.**

Es la diferencia entre el mayor y el menor de los datos. Sólo se usa para datos no agrupados en intervalos.

3.5.3.2 **Desviación media**

Es la media aritmética de los valores absolutos de las desviaciones de la media.

$$\mathrm{Dm} = \frac{{}^{k}_{i=1}(x_i - x)}{n}$$

3.5.3.3 **Desviación estándar.**

Es la medida aritmética de las diferencias entre los datos y la media

$$s = \overline{\frac{{}^{k}_{i=1}(x_i - x)^2 f_i}{n}}$$

3.5.3.4 **Varianza**

Es la media aritmética de los cuadrados de las diferencias entre la media y el dato.

$$s^2 = \frac{{}^{k}_{i=1}(x_i - x)^2 f_i}{n}$$

3.5.4 **Medidas de forma**

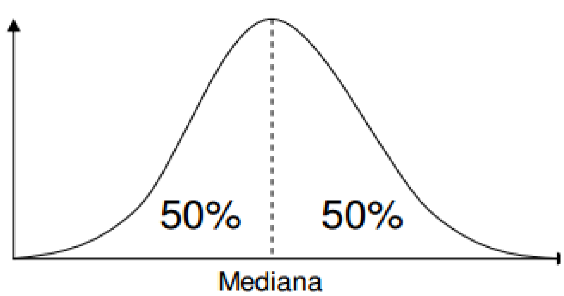
Las medidas de forma permiten comprobar si una distribución de frecuencia tiene características especiales como simetría, asimetría, nivel de concentración de datos y nivel de apuntamiento que la clasifiquen en un tipo particular de distribución.

Las medidas de forma son necesarias para determinar el comportamiento de los datos y así, poder adaptar herramientas para el análisis probabilístico.

3.5.4.1 **Simetría**

Al dividir una distribución de frecuencia mediante la mediana, ambas áreas resultantes son iguales, es decir, los datos se distribuyen de la misma forma y el área abarcada por ambos lados es equivalente (50% de los datos se encuentran distribuidos en ambas secciones).

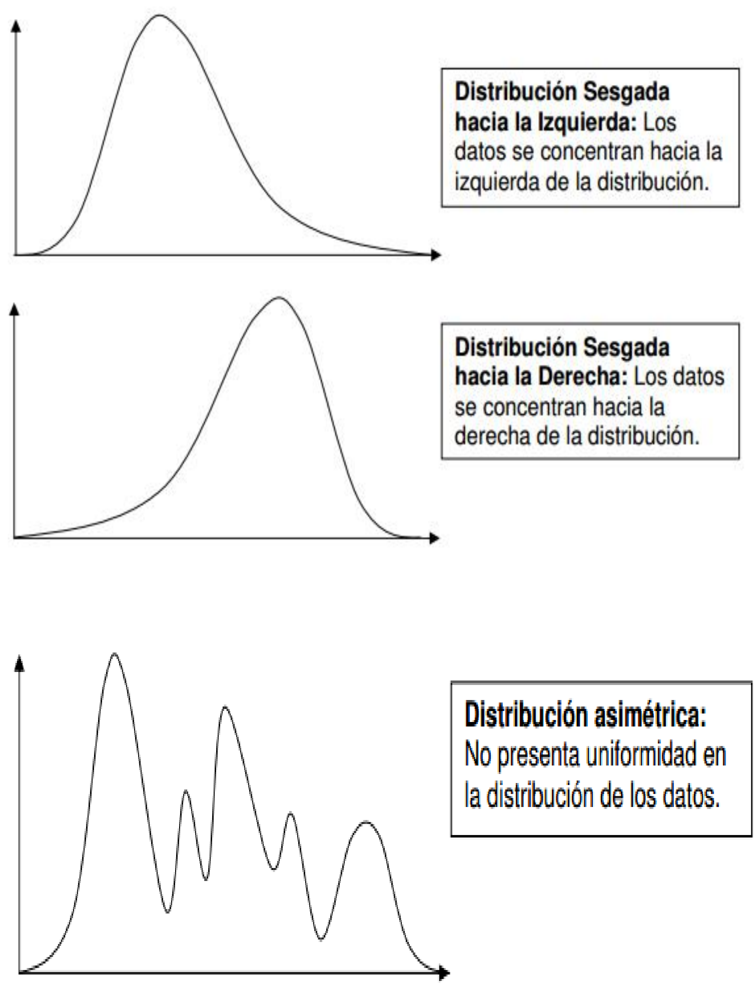
Esta medida es característica de una distribución normal; Esta distribución está definida por dos parámetros: la media μ y la desviación típica σ. Se presenta mediante una curva simétrica conocida como campana de Gauss. Esta distribución nos da la probabilidad de que al elegir un valor, éste tenga una medida contenida en unos intervalos definidos. Esto permitirá predecir de forma aproximada, el comportamiento futuro de un proceso.



Gráfica 10

3.5.4.2 **Asimetría**

Los datos no se distribuyen de forma uniforme y similar en las áreas que dan como resultado al dividir la distribución de frecuencia por la mediana.



Gráfica 11

3.5.4.2.1 **Coeficiente de asimetría de Fisher**

Este coeficiente mide el grado de asimetría de la distribución con respecto a la media. Para datos no agrupados se soluciona con la siguiente ecuación donde X_i representa el valor de cada dato en la distribución, X, μ representa la media de la distribución y se diferencian porque μ es para una distribución normal, al igual que la desviación estándar la cual se representa con: s, ó σ.

$$As = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{s} \right)^3 \quad o \quad As = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^3$$

.

Los resultados de este coeficiente se interpretan así:

• (As = 0): Se acepta que la distribución es Simétrica, es decir, existe aproximadamente la misma cantidad de valores a los dos lados de la media. Este valor es difícil de conseguir por lo que se tiende a tomar los valores que son cercanos ya sean positivos o negativos (± 0.5).

• (As > 0): La curva es asimétricamente positiva por lo que los valores se tienden a reunir más en la parte izquierda que en la derecha de la media.

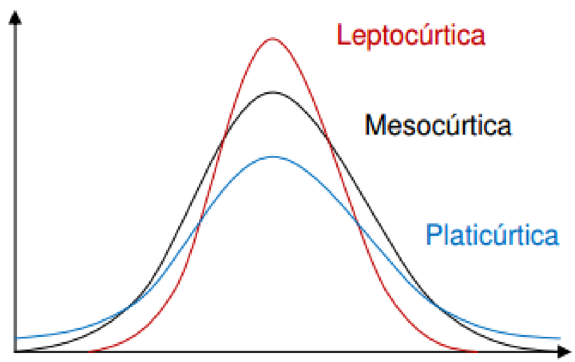
• (As < 0): La curva es asimétricamente negativa por lo que los valores se tienden a reunir más en la parte derecha de la media.

Desde luego entre mayor sea el número (Positivo o Negativo), mayor será la distancia que separa la aglomeración de los valores con respecto a la media.

3.5.4.3 Curtosis

Indica que tan apuntada o achatada se encuentra una distribución respecto a un comportamiento normal (distribución normal). Si los datos están muy concentrados hacia la media, la distribución es leptocúrtica (curtosis mayor a 0). Si los datos están muy dispersos, la distribución es platicúrtica (curtosis menor a 0).

El comportamiento normal exige que la curtosis sea igual a 0 (distribución mesocúrtica).



Gráfica 12

La fórmula empleada para calcular la Curtosis en datos desagrupados es:

$$Curtosis = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{s} \right)^4 \right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

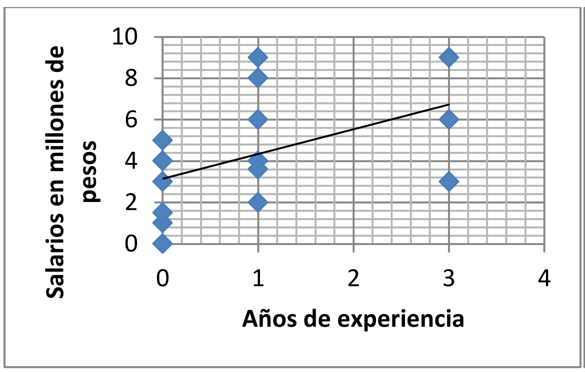
En esta formula X_i representa el valor de cada dato en la distribución, X representa la media de la distribución y la desviación estándar se representa con s.

3.6 Regresión Lineal

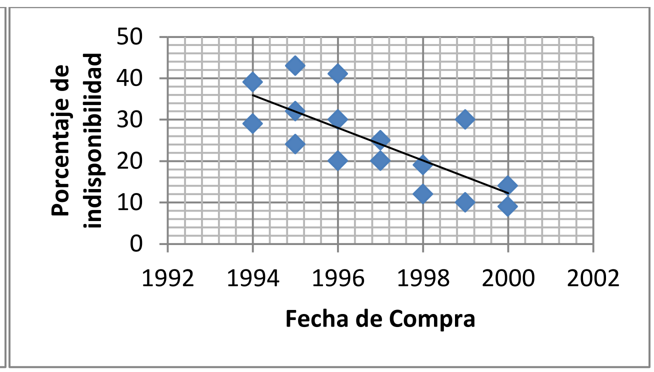
La regresión lineal es la estimación de una variable dependiente a partir de una variable independiente con la cual está relacionada. Este tipo de relación es una correlación simple, se presenta entre dos variables.

3.6.1 **Correlación Lineal**

Si X, Y denotan las variables que se relacionan, un diagrama de dispersión muestra la localización de los puntos (X, Y) en un mismo sistema de coordenadas rectangulares. La correlación es lineal cuando los puntos muestran un grado de relación igual a 1 de acuerdo al coeficiente de correlación, como en la *figura 13 y 14*.



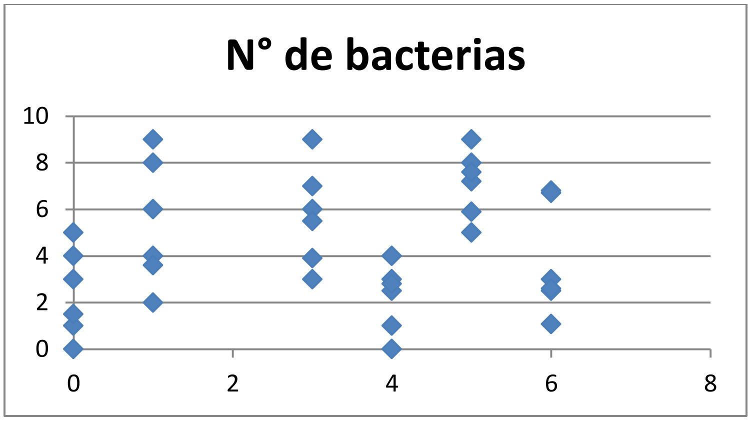
Gráfica 13



Gráfica 14

Si Y incrementa al incrementar X, la correlación se denomina positiva, o directa (gráfica 13). En el caso contrario, cuando Y disminuye, mientras X aumenta se denomina correlación negativa o inversa (gráfica 14).

Si no se evidencia una relación entre las variables se dice que no hay correlación o que no están correlacionadas, como en la gráfica 15.



Gráfica 15

3.6.2 **Medidas de correlación**

Las medidas de correlación nos permiten determinar qué tan ajustada es la recta para un conjunto de datos en un diagrama de dispersión. Entre estas tenemos:

3.6.2.1 **Error típico**

Se tiene un valor estimado Y, para valores de X dados, esta es una medida de dispersión alrededor de la recta de regresión de Y sobre X, esta medida de dispersión está determinada dada por:

$$s_{Y.X} = \sqrt{\frac{(Y - Y_{est})^2}{N}}$$

Si definimos la recta de regresión en X, calculando de manera análoga el error típico del valor estimado para X sobre Y, se define como:

$$s_{X.Y} = \sqrt{\frac{(X - X_{est})^2}{N}}$$

De lo anterior podemos concluir que: $s_{Y.X} \neq s_{X.Y}$

3.6.2.2 **Variación explicada y no explicada**

La *variación total* de Y se define como: $(Y - Y)^2$, es decir la suma de los cuadrados de las desviaciones de los valores de Y de su media Y. Esto se puede definir como:

$$(Y - Y)^2 = (Y - Y_{est})^2 + (Y_{est} - Y)^2$$

- $(Y - Y)^2$ este término representa las variaciones de los valores con respecto a la media de la muestra, $(Y - Y_{est})^2$: esta mide la variación de los valores con respecto a la recta de regresión y $(Y_{est} - Y)^2$ mide la variación de los valores estimados respecto a la media, luego la variación de los valores con respecto a la media es igual a la variación con respecto a la recta más la variación de los valores estimados con respecto a la media.

3.6.2.3 **Coeficiente de correlación:**

Se define como la razón de la variación explicada a la variación total; si la variación explicada es cero, es decir, la variación total es toda no explicada, la razón es cero. Si la variación no explicada es cero, es decir la variación total es toda explicada, la razón es uno. En casos donde la variación se compone de una variación explicada y otra no explicada, la razón se encuentra entre cero y uno, puesto que la razón es siempre positiva, entonces la denotamos como r^2. La cantidad r se llama *coeficiente de correlación* y está dado por:

$$r = \pm \sqrt{\frac{(Y_{est} - Y)^2}{(Y - Y_{est})^2}}$$

3.6.3 **RECTAS DE REGRESIÓN DE MÍNIMOS CUADRADOS**

Se define como el problema de bondad, el cual permite hacer una prueba de ajuste identificando de esta manera la línea recta que ilustra la relación entre dos variables, explicando de esta manera el tipo de relación entre estas. Por medio de la recta de regresión dada por los mínimos cuadrados es posible explicar este tipo de relaciones, ya que esta es generada por la suma de los cuadrados de la diferencia procurando que esta sea mínima, luego será posible identificar el grado de relación. La recta de regresión de mínimos cuadrados de Y sobre X es: $Y = a_0 + a_1X$; donde a_0 y a_1 se obtienen de las siguientes ecuaciones normales:

$$Y = a_0N + a_1 \quad X \longrightarrow a_0 = \frac{Y \quad X^2 - (\quad X)(\quad XY)}{N \quad X^2 - (\quad X)^2}$$

$$XY = a_0 \quad X + a_1 \quad X^2 \longrightarrow a_1 = \frac{N \quad XY - (\quad X)(\quad Y)}{N \quad X^2 - (\quad X)^2}$$

Y: Sumatoria de los valores de Y.

a_0: Corte con el eje Y.

a_1: Valor de la pendiente.

X: Sumatoria de los valores de X.

XY: Sumatoria de los productos XY.

X^2: Sumatoria de los cuadrados de X.

4. REPRESENTACIONES GEOMÉTRICAS Y DINÁMICAS

En este capítulo se recopilan 16 representaciones con las cuales se ilustran 26 conceptos de estadística descriptiva, entre las cuales hay 8 representaciones dinámicas, que ilustran los gráficos estadísticos tales como: Diagrama de columnas, Pictogramas, Diagrama de barras, Diagrama circular, Diagrama de tallos y hojas, Diagrama de puntos, Histograma y Polígono de frecuencias; y las otras 8 son representaciones geométricas, 2 en descartes, las cuales son: Mediana y Moda para datos agrupados y las otras 6 en Geogebra, entre las cuales tenemos: Comparación de medias (Media Aritmética, Media Geométrica, Media Armónica, Media Cuadrática); Media Ponderada; Rango Medio; Desviación estándar, Varianza, y Asimetría.

REPRESENTACIONES DINÁMICAS

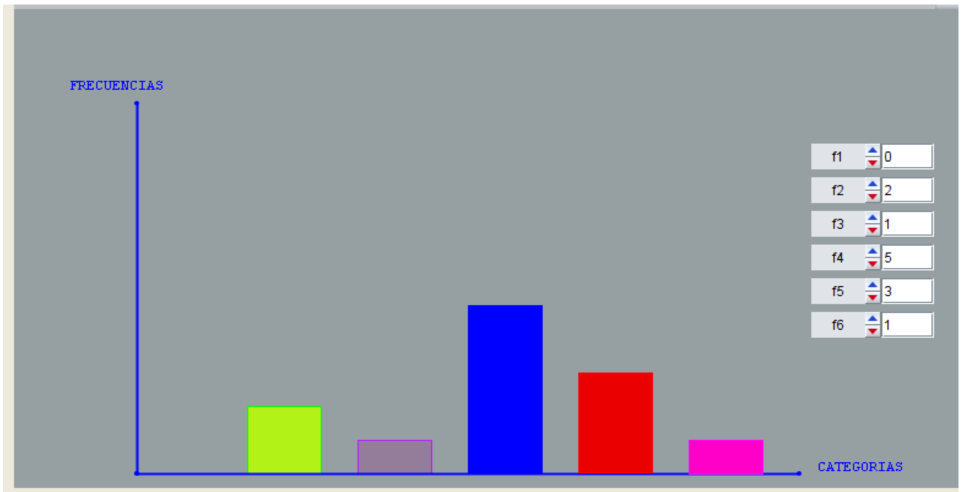
Para los diferentes gráficos estadísticos se realiza una representación dinámica en las cuales se muestra una explicación del tema, el diseño de cada construcción y el modo de uso.

Para la construcción de representaciones dinámicas de los gráficos estadísticos se utilizó el programa Descartes, en el cual en su configuración inicial de espacio se elige la opción, E1 (espacio 1), para trabajar en dos dimensiones.

4.1 Diagrama de Columnas

El diagrama de Columnas es utilizado para representar información de una muestra donde los datos son variables de tipo cuantitativo discreto o cualitativo tanto nominal como ordinal, este gráfico permite representar la frecuencia que presenta cada dato.

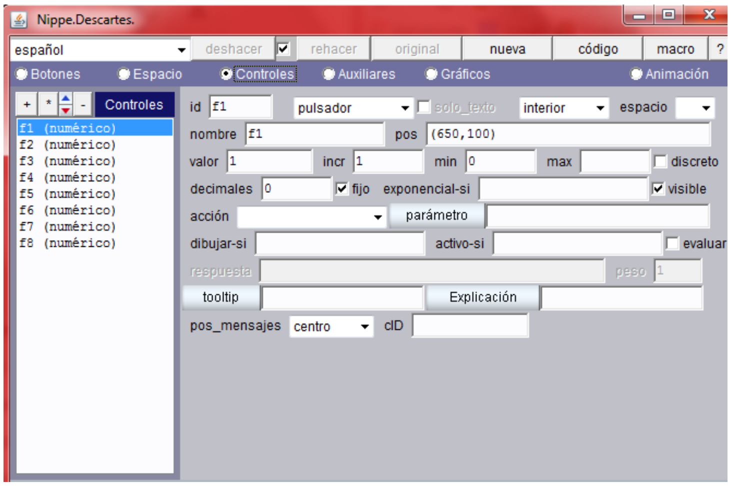
La importancia del diagrama radica en que es una forma de representar gráficamente un conjunto de datos o valores, con el cual el estudiante podrá visualizar el cambio producido en la altura al variar el valor de la frecuencia.



Gráfica 16

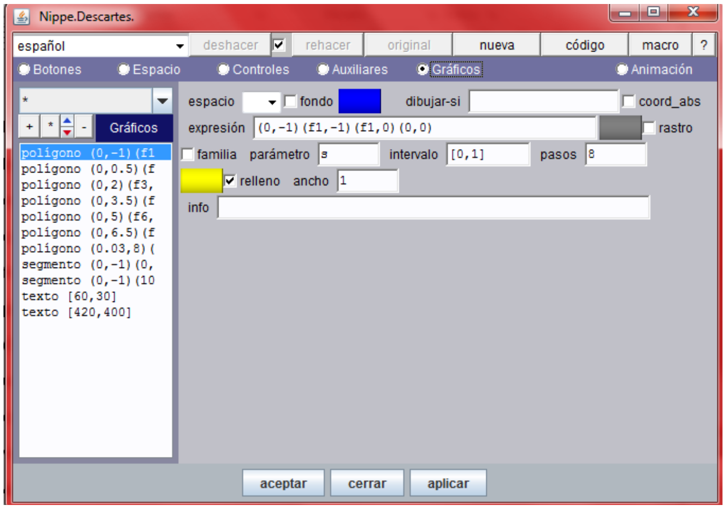
3.1.5 Diseño de la construcción:

Para iniciar la construcción se dirige al botón llamado config que se encuentra ubicado en el extremo superior derecho; ya que en esta opción se configura la ilustración, allí se elige la opción controles que se encuentra en la parte superior izquierda, luego se pulsa en el signo + que se encuentra en el costado izquierdo para agregar un control nuevo, para finalizar la construcción del control se pulsa sobre el icono aceptar en una nueva ventana que aparece. Luego se configura el botón con las siguientes características: se escoge pulsador (el usuario podrá cambiar la frecuencia del dato al dar clic sobre el control), se deja en la opción interior, se le da el nombre f_1, con un incremento de 1 en 1, coordenadas (0,0) y luego se incrementa en escala de 30 en 30 al ubicar los demás controles. El valor mínimo es 0, en número de decimales escribimos 0, y finalmente se da click sobre el botón aplicar que se encuentra en la parte inferior de la ventana, de esta manera tenemos el primer control con nombre f_1 en la pantalla del programa, después se realiza el mismo procedimiento para crear los demás controles, tales como: f_2, f_3, f_4, f_5 y f_6.



Gráfica 17

Después para construir las columnas se dirige a la opción gráficos que se encuentra en la parte superior de la ventana, se pulsa en el símbolo +, para agregar un gráfico, en este caso tomamos la opción polígono, al cual le damos las siguientes características: con coordenadas (0,0) (2,0) $(2,f_1)$ $(0,f_1)$ (0,0), se construye el primer rectángulo de frecuencia f_1 y se escoge el color deseado para rellenarlo, luego para el segundo polígono se cambia el color del relleno y se dejan las siguientes coordenadas (3,0) (5,0) $(5,f_2)$ $(3,f_2)$ (3,0) para construir los otros rectángulos se va incrementando de 1 en uno en la coordenada de la abscisa para hacer las demás columnas y completar la construcción dinámica.



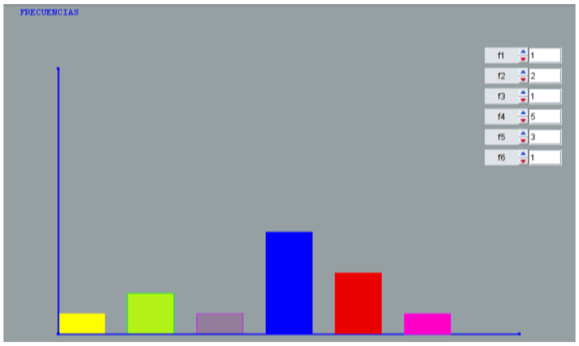
Gráfica 18

Al cambiar el valor de los controles se puede observar como varia la altura de cada rectángulo.

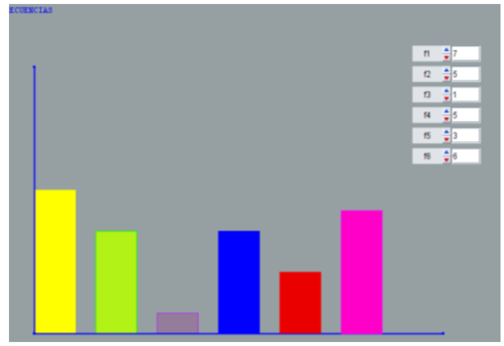
4.1.2 Modo de uso:

Al modificar los valores de la frecuencia se puede identificar los cambios de altura que presentan los rectángulos, manera en la que relacionan la altura con el número de veces que se da el dato en un conjunto, para visualizar dichos cambios se sugiere al docente implementar en el aula ejercicios donde el estudiante realice un diagrama de barras primero en el cuaderno y luego con la ayuda del applet explore la herramienta, cambiando los valores de las frecuencias, para que de esta manera identifique todas las posibles situaciones que se pueden presentar con un mismo evento, además de la que el trazó sobre su cuaderno.

Ejemplo: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 19

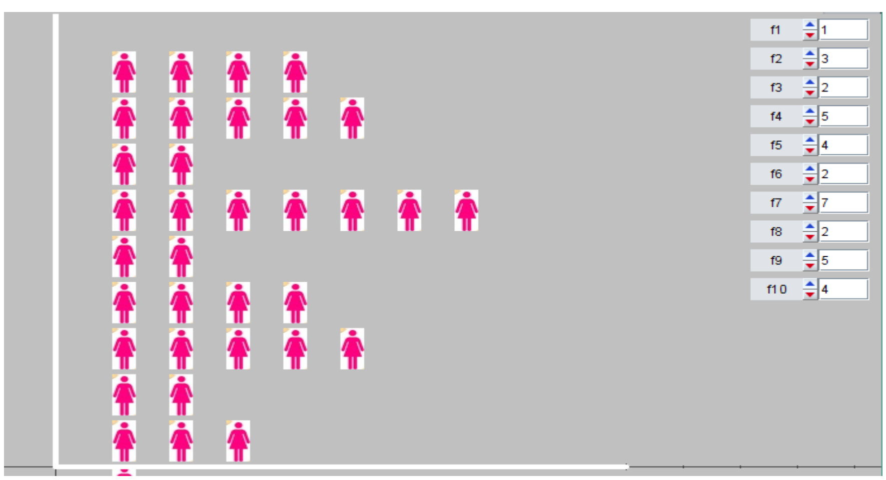


Gráfica 20

3.2 Pictogramas

El Pictograma es un gráfico que consiste en utilizar iconos que representan la variable analizada. En cada columna o fila se ubican tantos iconos como frecuencia que tenga el dato.

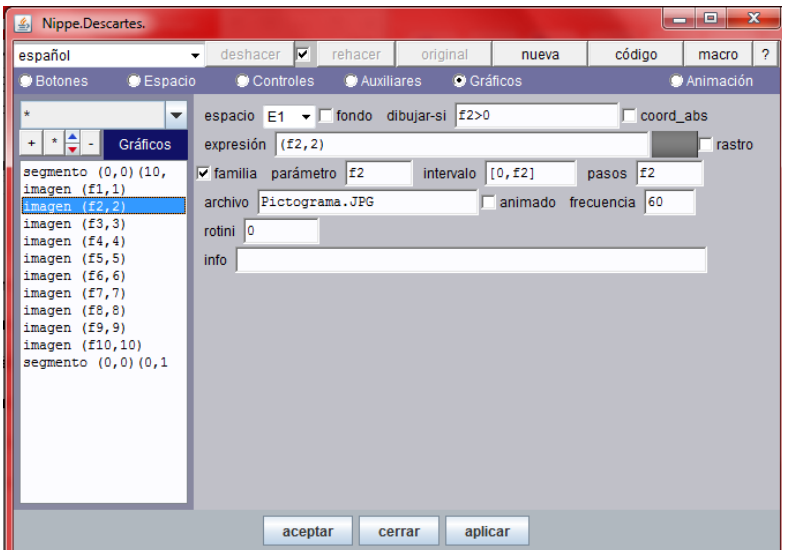
La importancia de este tipo de grafica radica en que le permite al estudiante observar el gráfico con dibujos alusivos al carácter que se está estudiando y cuyo número de iconos es equivalente a la frecuencia que representan; lo cual le permite un mayor acercamiento al problema.



Gráfica 21

3.2.1 Diseño de la construcción:

Para la elaboración de este diagrama se construyeron los controles de manera similar que en el diagrama de columnas, se le asignan los nombres desde f_1 hasta f_{10}, para la siguiente parte nos dirigimos a la opción gráficos que se encuentra en la parte superior de la ventana, se pulsa en el símbolo +, para agregar un gráfico, en este caso es una imagen donde las coordenadas son $(1, f_1)$ hasta $(10, f_{10})$, y en la opción de archivo se agrega la imagen que debe estar en la misma carpeta de la construcción.

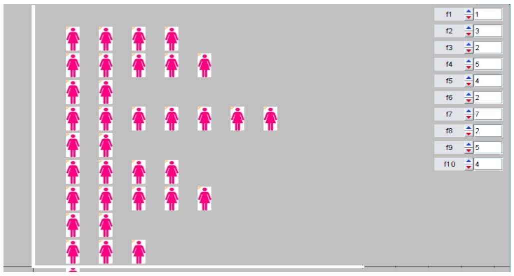


Gráfica 22

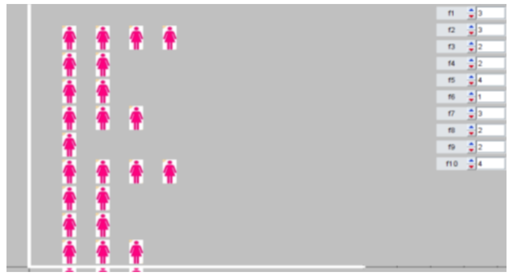
3.2.2 Modo de uso:

Al modificar los valores de la frecuencia se puede identificar los cambios de altura presentados por las figura que relacionan la altura con el número de veces que se da el dato en un conjunto, para visualizar dichos cambios se sugiere al docente diseñe un taller de exploración con el applet el cual el estudiante analice, los diferentes cambios que se pueden presentar, como por ejemplo el número de niños nacidos por día en un hospital etc., (se deja como sugerencia para el docente).

Ejemplo: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



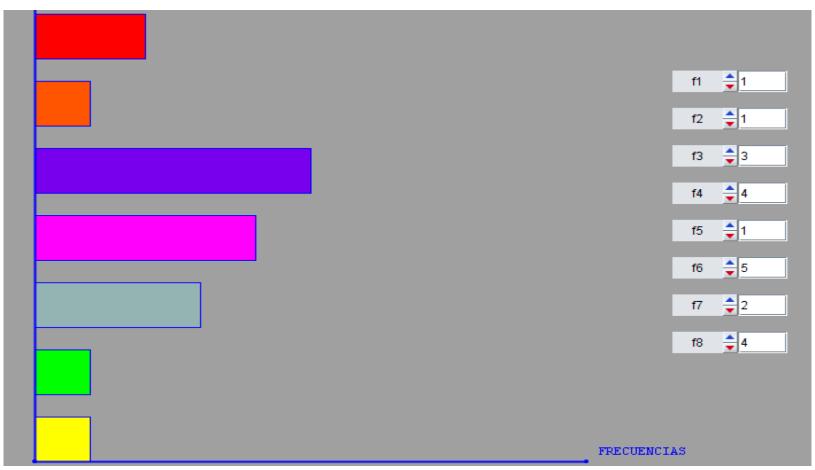
Gráfica 23



Gráfica 24

3.3 Diagrama de Barras

El diagrama de barras es utilizado para representar información de una muestra donde los datos son variables de tipo cualitativo o cuantitativo discreto; este gráfico permite representar la frecuencia con la que se presenta cada dato. Particularmente, el tamaño de la base de cada columna es equivalente a la frecuencia de la categoría representada, mientras que la altura es igual para todas las columnas, las columnas salen desde el eje vertical y crecen o decrecen de manera horizontal, entre ellas existen espacios del mismo tamaño representando que no existen valores decimales entre los mimos



Gráfica 25

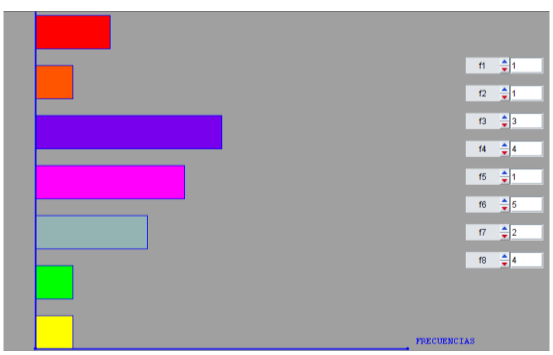
3.3.1 Diseño de la construcción

La construcción se realiza de manera similar que el diagrama de columnas, pero con la diferencia que los rectángulos se construyen en el eje vertical.

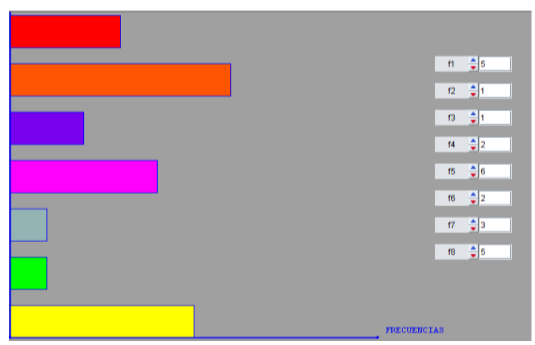
3.3.2 Modo de uso:

Para esta construcción se sugiere al docente implementar en el aula ejercicios como el siguiente: Solicitar al estudiante, que realice una encuesta de 10 preguntas a un número máximo de 30 personas sobre temas que llamen su atención como: deporte, música, comida, televisión etc. Solicitándole el estudiante organizar la información y representarla mediante diagramas de barras y con ayuda del applet.

Ejemplo: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 26

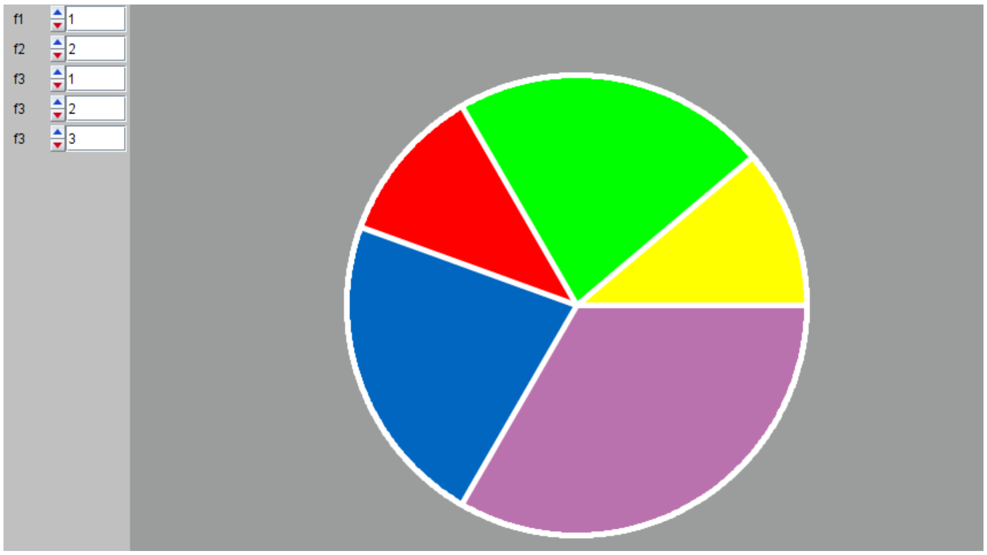


Gráfica 27

3.4 Circular

El gráfico circular es útil para representar proporciones de distintas clases dentro de una muestra. La muestra es representada por el área total de un círculo y cada una de las clases que la componen, por un sector de éste. El ángulo de cada sector es a 360° como cada frecuencia porcentual es al 100%, o como cada f_n es el total de los datos.

Este gráfico es útil para representar datos que sean variables cualitativas o cuantitativas; donde la información organizada en la tabla presente frecuencias porcentuales.

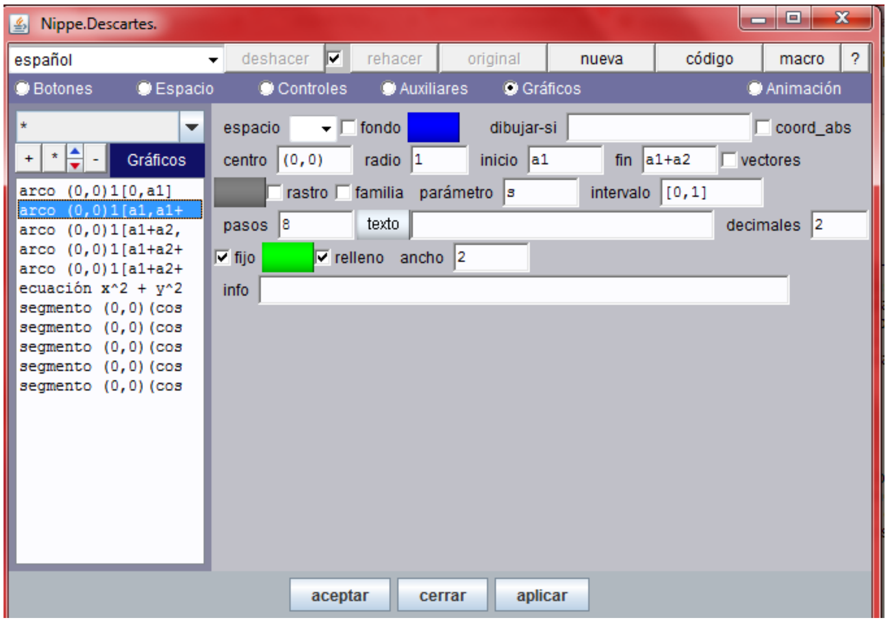


Gráfica 28

3.4.1 Diseño de la construcción:

Para la construcción de este diagrama se ubicaron cinco controles a la izquierda de la pantalla, los cuales se realizan de igual manera que en las anteriores construcciones, desde f_1 hasta f_5, luego se crean elementos auxiliares, como: $N\ (f_1 + f_2 + \cdots + f_4)$, que es igual al total de datos obtenidos, $a_i = \frac{360 * f_i}{N}$ es igual a la longitud del arco de la $i - esima$ sección circular, de acuerdo al valor de la frecuencia. Luego se construye la representación gráfica dibujando una circunferencia con $x^2 + y^2 = 1$; con centro en (0,0); y en esta misma se dibujan segmentos con coordenadas p_i: (0,0),

p_f: $\cos\frac{a_i*\pi}{180}, \sin\frac{a_i*\pi}{180}$, dividiendo de esta manera al círculo en sectores proporcionales a las frecuencias que corresponde a cada uno.

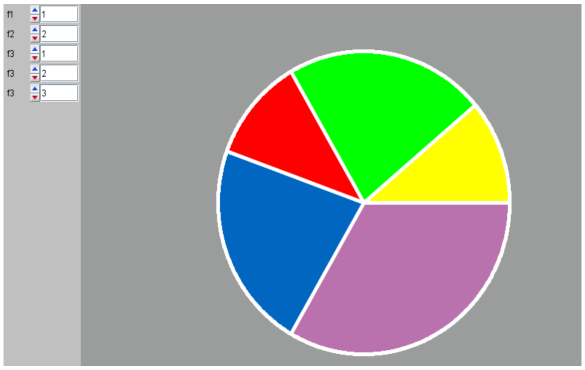


Gráfica 29

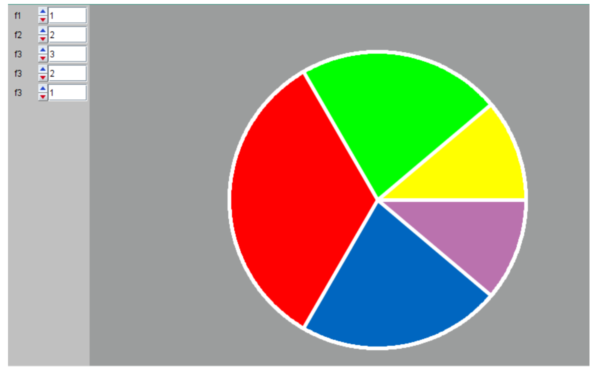
3.4.2 Modo de uso:

Se le sugiere al docente hacer un taller de exploración donde el estudiante caracterice el diagrama de acuerdo a lo observado e identificado en la ilustración, entregándole a cada uno de los estudiantes, una situación problema que sea posible representar con el diagrama circular. El estudiante deberá identificar que existe una relación directa ente el valor de la frecuencia y el sector circular con el que se corresponde; ya que el sector circular es al área total del círculo, como la frecuencia de cada dato es al total de datos.

Ejemplo: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 30



Gráfica 31

3.5 Tallo y hojas

El diagrama tallo y hojas permite obtener simultáneamente una distribución de frecuencias de la variable y la representación gráfica. Con este se representan variables de tipo cuantitativo discreto, el conjunto representado debe tener pocos datos para que sea fructífera su utilidad. Para construirlo se debe separar en cada dato el último dígito de la derecha (que constituye la hoja) del bloque de cifras restantes (que formará el tallo).

Gráfica 32

3.5.1 Modo de uso

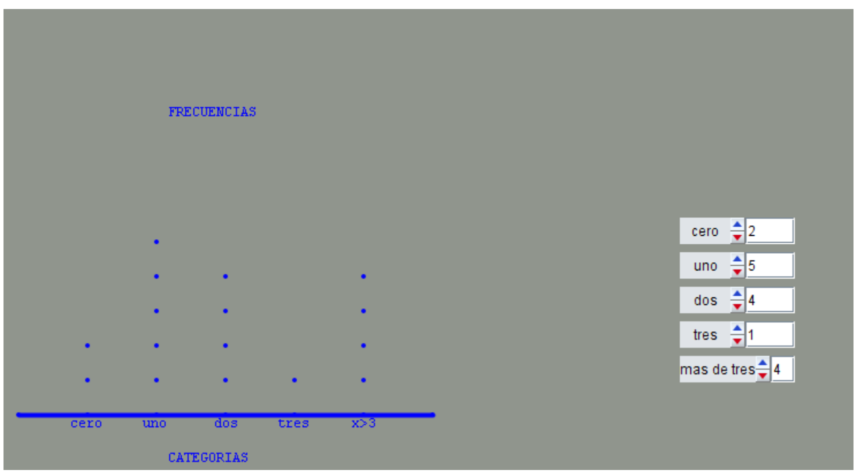
Para el diagrama de tallos y hojas se sugiere ejercicios donde el docente, trabaje con números decimales, darle una serie de datos al estudiante y solicitarle que por medio del ejemplo que se muestra en el Applet el mismo estudiante los ubique en un diagrama de tallos y hojas.

En esta representación será posible identificar que la cantidad de puntos ubicados verticalmente sobre cada dato es igual al valor de la frecuencia de dicho dato.

Esta representación ilustra la forma del diagrama.

3.6 Diagrama de Puntos

El diagrama de puntos se utiliza para representar pocos datos con variables cuantitativas discretas o cualitativas, se ubican tantos puntos en línea vertical sobre el dato representado, como sea el valor de la frecuencia del dato.



Gráfica 33

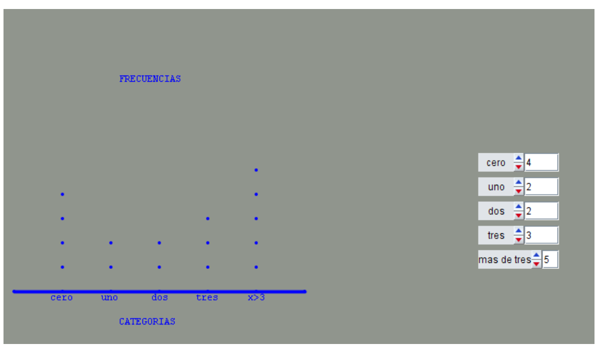
4.6.1 Diseño de la construcción:

En esta representación será posible identificar que la cantidad de puntos ubicados verticalmente sobre cada dato es igual al valor de la frecuencia de dicho dato, para la construcción de este diagrama construimos los controles de manera similar que los diagramas anteriores , se le asignan los nombres desde f_1 hasta f_5, se les asigna un texto a cada uno: cero, uno, dos, tres, más de tres, para la siguiente parte nos dirigimos a la opción gráficos que se encuentra en la parte superior de la ventana, se pulsa en el símbolo +, para agregar un gráfico, en este caso puntos que tiene por coordenadas (x, f) donde x es un numero cualquiera y f es valor de la frecuencia.

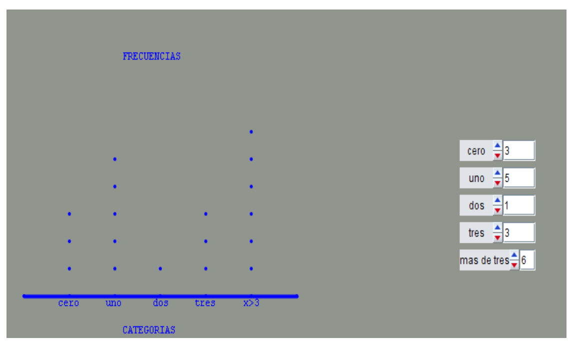
4.6.2 Modo de uso

Se sugiere al docente que para llevar este applet al aula de clases, ejercicios donde el estudiante adecue los datos al Apple y organice la información, desde el menor valor y el mayor, para la construcción del diagrama de puntos.

Ejemplo Gráfico: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 34



Gráfica 35

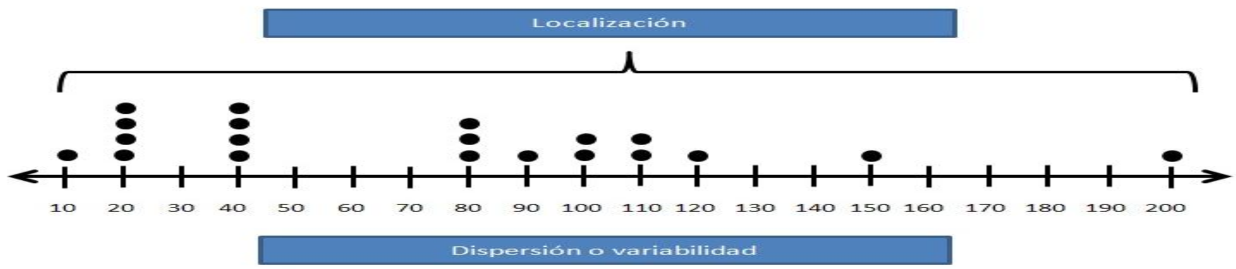
Ejemplo

La tabla siguiente muestra los datos de longitud en milímetros de un conjunto de cables que serán utilizados en un estudio de resistencia a la tensión:

Cable	Longitud	Cable	Longitud	Cable	Longitud	Cable	Longitud
1	20	6	40	11	40	16	40
2	80	7	20	12	110	17	200
3	110	8	20	13	120	18	10
4	100	9	90	14	20	19	100
5	80	10	80	15	40	20	150

Tabla 12

Trazar una línea horizontal desde el valor mínimo, seleccionar una escala y utilizando intervalos regulares, marcar la escala hasta que el valor máximo sea alcanzado.



Gráfica 36.1

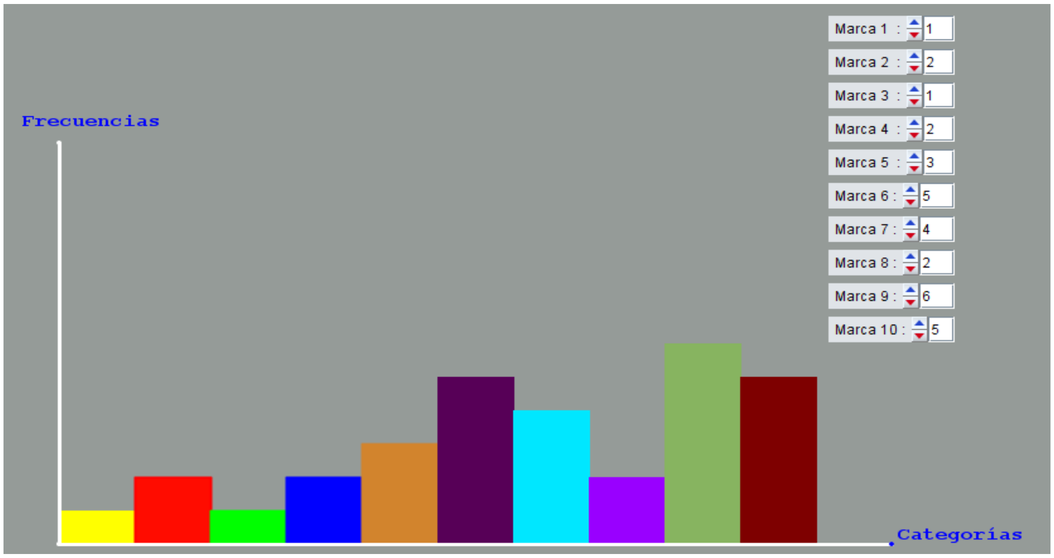
Para cada valor numérico presente en la tabla de datos, colocar un punto sobre la escala de valores en la recta numérica, cuando el valor numérico aparece más de una vez, apilar los puntos.

Gráfica 36. 2

3.7 Histograma

El Histograma es utilizado para representar información de una muestra donde los datos son variables de tipo cuantitativo continuo y permiten representar la frecuencia con la que se presenta cada dato, o representa daos agrupados en intervalos del mismo tamaño, el tamaño del intervalo es la base del triángulo, luego lo que varía en esta representación es la altura de cada uno. Particularmente, el tamaño de la altura de cada rectángulo es equivalente a la frecuencia de la categoría representada.

Esta representación es similar al diagrama de columnas, la diferencia radica en que en este si hay valores entre los datos, estos valores son reales, luego se hace un rectángulo por intervalo y no por cada dato.



Gráfica 37

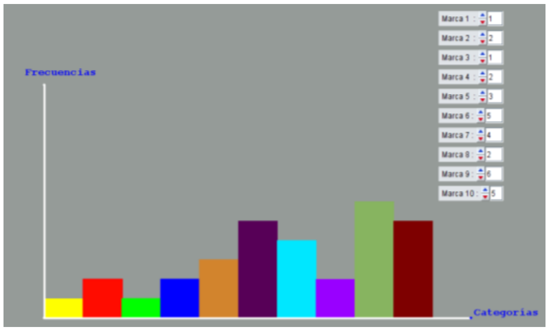
3.7.1 Diseño de la construcción:

La construcción se realiza de manera similar que el diagrama de columnas, pero con la diferencia que los rectángulos se construyen sin dejar espacios entre ellos, en este caso la coordenada del rectángulo siguiente empieza donde termina el anterior a él.

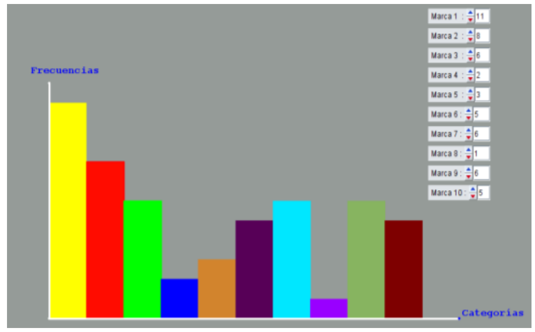
3.7.2 Modo de uso

Para esta grafica se sugiere al docente que le permita al estudiante experimenta con el Apple y que por medio de la experimentación, el estudiante conjeture una definición de histograma, y que proponga ejercicios donde los dalos se den por medio de intervalos.

Ejemplo Gráfico: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 38

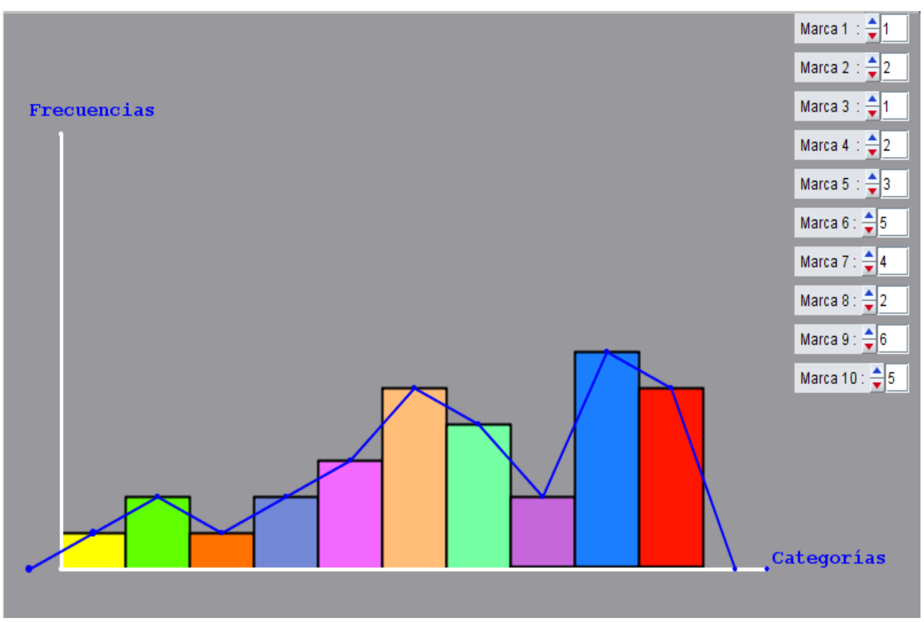


Gráfica 39

3.8 Polígono de frecuencias

Un polígono de frecuencias es utilizado para representar en forma geométrica la información obtenida de un histograma por medio de segmentos de recta que unen los puntos medios de los intervalos de clase adyacentes.

Para esta grafica se podrá observar que al modificar el valor de la frecuencia de cada marca cambiara la inclinación de los segmentos que unen los puntos medios de cada intervalo.



Gráfica 40

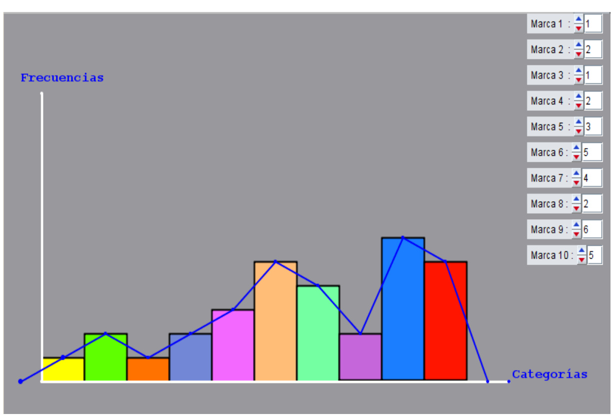
3.8.1 Diseño de la construcción:

Este gráfico se realiza a partir de la construcción de un histograma, ya que luego de tener el histograma encontramos la marca de clase (esta es el punto medio de la base de cada rectángulo), para cada intervalo y marcamos los puntos de coordenadas m_i, f_i donde m_i: es la marca de clase de algún intervalo, y f_i: es la frecuencia de dicho intervalo. Luego por medio de segmentos unimos estos puntos que tendrán las siguientes coordenadas: m_i, f_i , m_{i+1}, f_{i+1} .

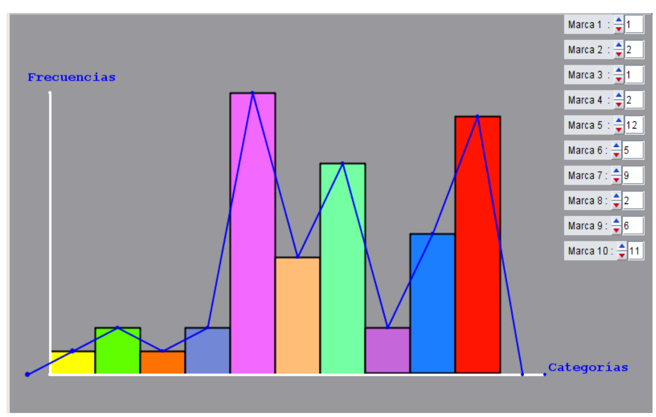
3.8.2 Modo de uso

Para esta ilustración se sugiere un taller donde por ejemplo: el estudiante realice una encuesta en su barrio de niños nacidos por mes, luego con la información recogida construir un polígono de frecuencias que represente los datos obtenidos; para esto la mitad del curso deberá hacerlo con lápiz y papel y la otra mitad modelarlo sobre el applet, luego comparar lo obtenido.

Ejemplo Gráfico: a continuación se muestra un ejemplo gráfico, al cambiar el valor de las frecuencias, a partir del grafico generado por los valores predeterminados.



Gráfica 41



Gráfica 42

4. REPRESENTACIONES GEOMÉTRICAS DE MEDIDAS DE LOCALIZACIÓN

En este capítulo estudiaremos las medidas de localización desde su debida representación geométrica. Estas construcciones se realizaron con software de matemáticas como lo son: Geogebra y Descartes.,

5.1 Desigualdad entre las medias

En un conjunto de datos siempre se presenta la misma relación de desigualdad entre: la media armónica, la geométrica, la aritmética y la cuadrática. Esta desigualdad es representada geométricamente a continuación.

$$\frac{2ab}{a+b} < \sqrt{ab} < \frac{a+b}{2} < \sqrt{\frac{a^2+b^2}{2}} .$$

H < G < $\bar{X}$ < Q,

Dónde:

H: Media armónica

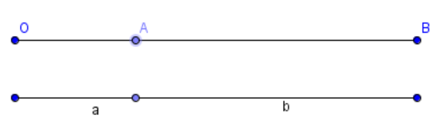
G: Media geométrica

X: Media aritmética

Q: Media cuadrática

Diseño y demostración de la construcción

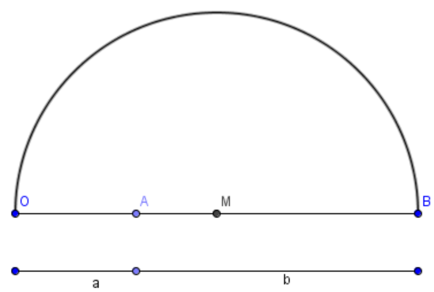
Sobre una recta horizontal se ubican los puntos O, A y B tales que las longitudes de los segmentos OA y AB sean respectivamente a y b.



Gráfica 43

M es el punto medio del segmento OB y OM mide $\frac{a+b}{2}$, lo que genera la media aritmética.

Se traza una semicircunferencia con centro en M y radio OM.

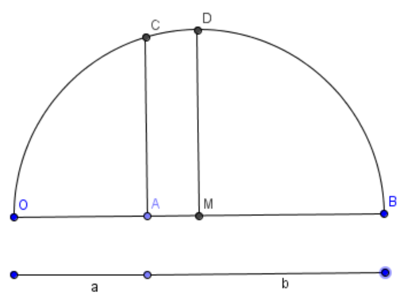


Gráfica 44

Se trazan rectas perpendiculares a la recta OB, por A y M.

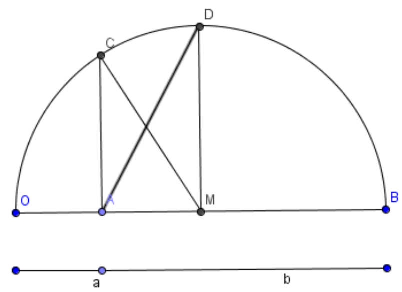
Puntos C y D intersección entre las rectas perpendiculares y la semicircunferencia respectivamente.

Los segmentos CA y DM son perpendiculares al segmento OB. Los puntos C y D son puntos de la semicircunferencia.



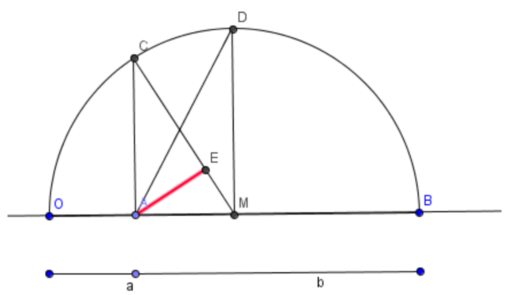
Gráfica 45

Se trazan los segmentos AD Y CM.



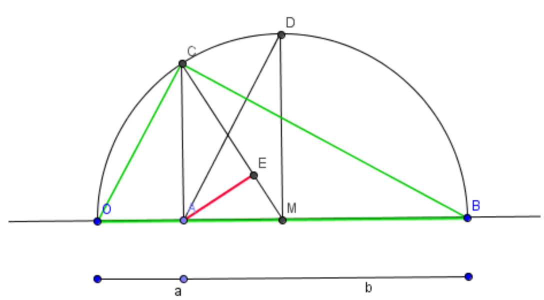
Gráfica 46

El segmento AE es perpendicular al segmento CM.



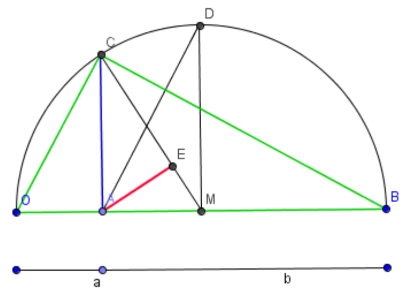
Gráfica 47

Los puntos O, C y B son vértices de un triángulo rectángulo. En un triángulo rectángulo, la altura correspondiente a la hipotenusa es media geométrica de los segmentos en los cuales la hipotenusa es dividida por dicha altura.



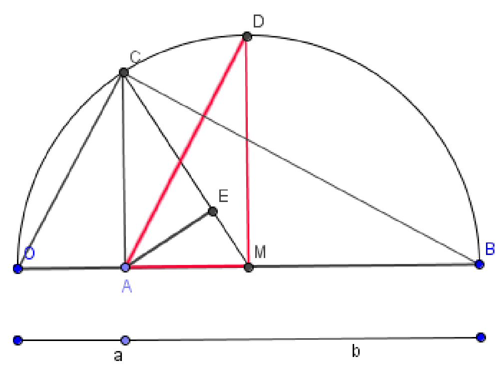
Gráfica 48

Aplicando este teorema, el segmento CA es media geométrica de los segmentos OA y AB, es decir, CA=$\sqrt{ab}$.



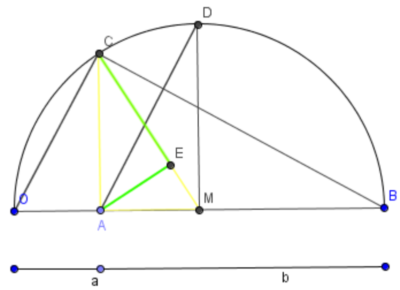
Gráfica 49

El segmento AM mide $\frac{b-a}{2}$ y el segmento DM mide $\frac{a+b}{2}$. Aplicando el teorema de Pitágoras en el triángulo AMD se obtiene que el segmento AD mida $\sqrt{\frac{a^2+b^2}{2}}$ la media cuadrática.



Gráfica 50

El segmento CM mide $\frac{a+b}{2}$ y los triángulos CAM y CEA son semejantes.



Gráfica 51

De la proporción $\frac{CE}{CA}=\frac{CA}{CM}$ resulta $CE=\frac{2ab}{a+b}$ de donde resulta la media armónica.

Hasta aquí se tiene que:

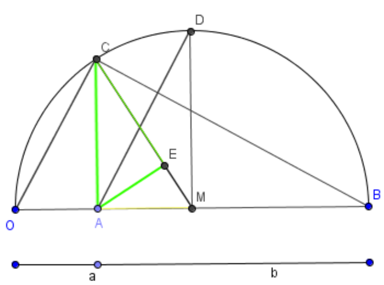
$CE=\frac{2ab}{a+b}=H$ (Media armónica).

$CA=\sqrt{ab}=G$ (Media geométrica).

$CM=DM=\frac{a+b}{2}=\overline{X}$ (Media aritmética).

$AD=\sqrt{\frac{a^2+b^2}{2}}=Q$ (Media cuadrática).

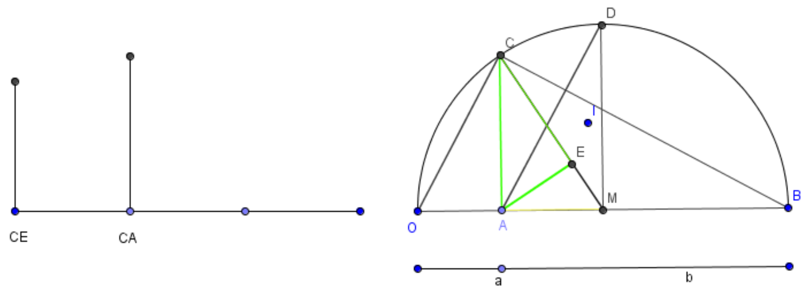
Para demostrar la cadena de desigualdades, que se muestra al inicial la construcción observe que en el triángulo AEC.



Gráfica 52

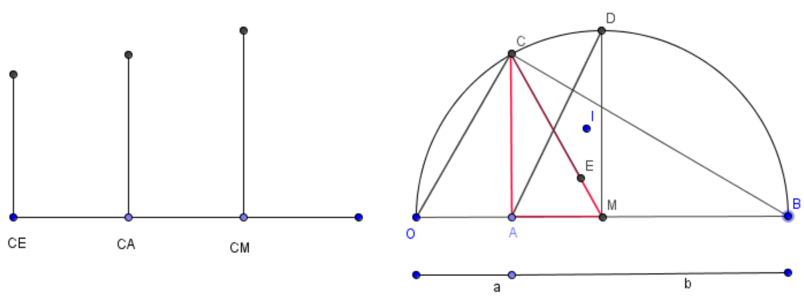
$CE=\dfrac{2ab}{a+b}=H$ es un cateto y $CA=\sqrt{ab}=G$ es la hipotenusa, por lo tanto H<G.

$\dfrac{2ab}{a+b}<\sqrt{ab}$.



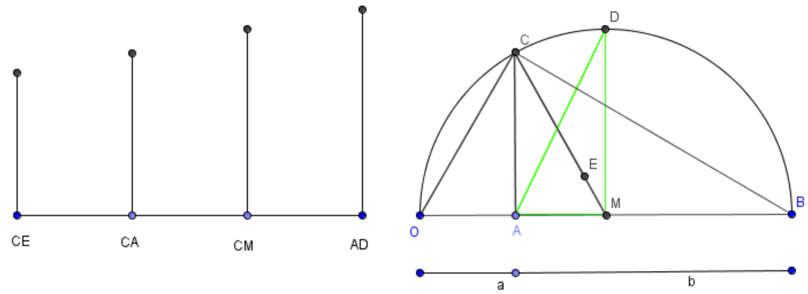
Gráfica 53

En el triángulo MAC, $CA=\sqrt{ab}=G$ es un cateto y $CM=\dfrac{a+b}{2}=\overline{X}$ es la hipotenusa, entonces G< $\overline{X}$.



Gráfica 54

En el triángulo AMD, $DM=\dfrac{a+b}{2}=\overline{X}$ es un cateto y $AD=\sqrt{\dfrac{a^2+b^2}{2}}=Q$ es la hipotenusa, por lo tanto $\overline{X}$<Q.



Gráfica 55

Así queda demostrada la desigualdad entre medias.

$H < G < \bar{X} < Q.$

CE < CA < CM < AD.

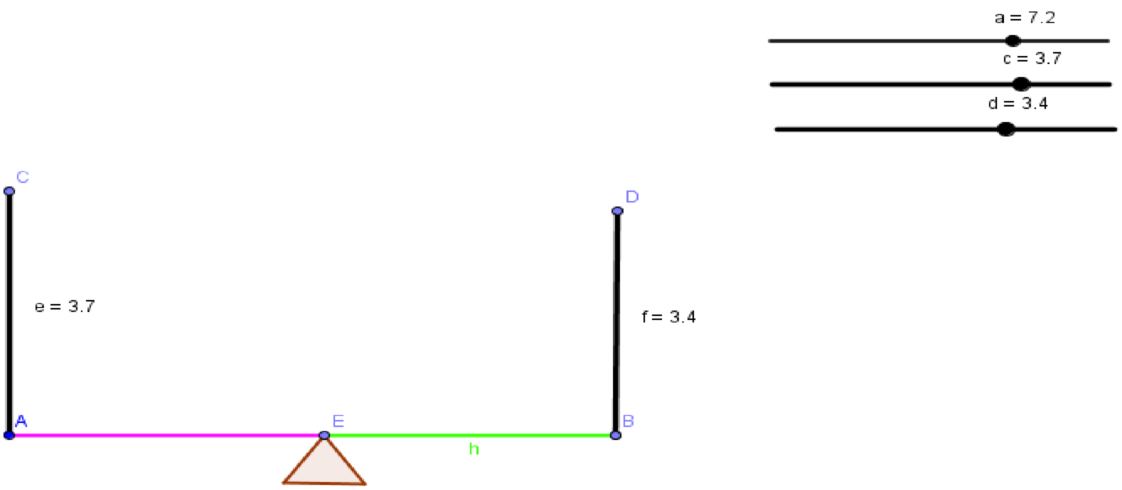
$$\frac{2ab}{a+b} < \sqrt{ab} < \frac{a+b}{2} < \sqrt{\frac{a^2+b^2}{2}}.$$

4.1.1 Modo de uso

Para esta construcción se sugiere al docente llevar el Applet al aula de clase, con la sugerencia se organizar los estudiantes por parejas, seguido de dar a cada pareja la construcción y solicitarles que tomando dos valores a y b hallen las medias y exploren la construcción, uno de los estudiantes trabaje sobre un papel en donde por medio formulas estadísticas llegue hallar el valor de las medias y las compare entre ellas y el otro estudiante tenga la posibilidad de explorar la construcción, para conjeturar las desigualdades que se puedan presentar entre las diferentes medias, para continuar con la actividad se solicita a los estudiantes realizar una comparación sobre los valores encontrados de manera estadística y geométrica, finalmente debe ser el docente quien aclare las dudas presentadas por los estudiantes y explique la desigualdad que se presenta entre los valores de las medias con ayuda del tablero y el Applet.

4.2 Media Ponderada

La media ponderada es el promedio de datos donde existen coeficientes que representan el número de veces que un valor es más importante que otro.; por ejemplo el promedio de un listado de notas donde cada una equivale a un porcentaje distinto de la nota final.



Gráfica 56

En esta representación geométrica se hace del principio de la balanza por Arquímedes, en el cual el punto de equilibrio ilustra el significado de la media ponderada; luego E=AE*AC+EB*BD, donde AE y EB son las distancias a las cuales se ubican los pesos a partir del punto de equilibrio, y EB, BD son los pesos ubicados en cada extremo de la balanza.

El usuario podrá verificar que el punto de equilibrio cambia de acuerdo a como cambien los segmentos que representan el peso de dos objetos, ya que las distancias entre el punto de equilibrio y la ubicación de los pesos también depende de la variación de estos. Se hace uso delos deslizadores a y c para identificar las características de la construcción y las variantes ya planteadas.

4.2.1 Diseño de la construcción:

Se usan dos deslizadores, cada uno para cambiar el tamaño de los segmentos que representan el peso de los “objetos” ubicados a los extremos del segmento c. Se construye un segmento y sobre el otros dos que representan las veces que son el peso que está colgado en su extremo.

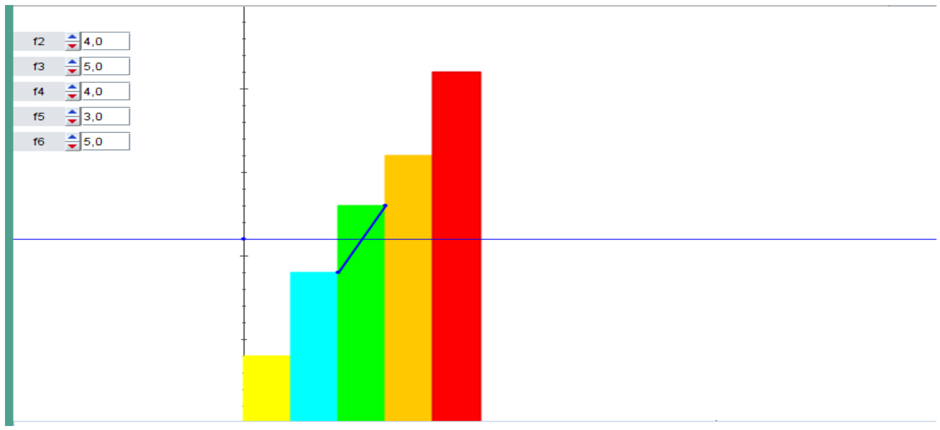
Se construye con la herramienta compas un segmento perpendicular al eje *x* sobre el segmento c de medida a (a es uno de los deslizadores); de igual manera construimos un segmento b sobre el otro extremo.

4.2.2 Modo de uso:

Se sugiere al docente que realice un taller de forma experimental; es decir que se realice un laboratorio, en el cuál sobre una base horizontal se colocan dos pesos; si los pesos son iguales el punto de equilibrio seria el punto medio del segmento que representa la distancia total entre los pesos. Si los pesos son distintos se identificara que se debemos mover la base que nos da el punto de equilibrio; ya que el segmento que representa la distancia entre los pesos, se va a inclinar hacia un lado.

4.3 Mediana

La mediana para datos agrupados es igual a: $\mathrm{Q_j} = \mathrm{L_{i-1}} + \frac{j\frac{\mathrm{n}}{4} - (\mathrm{N_{i-1}})}{\mathrm{n_i}} \mathrm{c_i}$, Es decir al ubicar la clase donde está ubicada la mediana se calcula el valor exacto para esta. La mediana es el valor ubicado en la mitad del conjunto de datos



Gráfica 57

En la representación gráfica se ilustra por medio de un histograma para datos agrupados la mediana el valor de la abscisa del punto correspondiente a la intersección entre, la recta paralela al eje de las abscisas por el punto medio del segmento $LsL\imath$, donde L_S es el límite superior de la clase que antecede a la clase donde se encuentra la mediana o segundo cuartil, y L_i es el límite inferior de esta misma clase.

4.3.1 Diseño de la construcción:

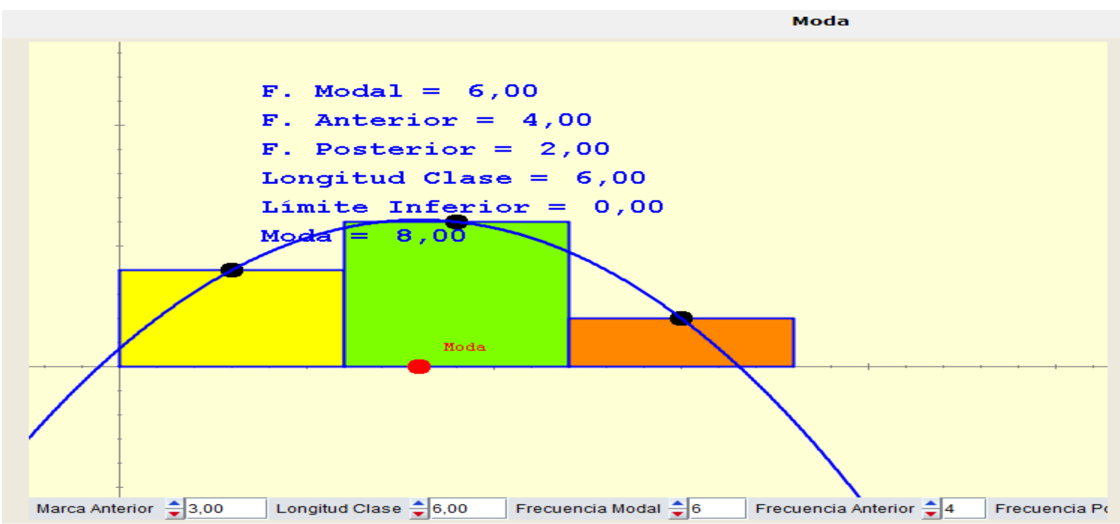
En el programa descartes se realiza el histograma correspondiente para un conjunto de datos agrupados, se toma la marca de clase del intervalo que esta exactamente en la mitad de todos los demás, trazamos un segmento que tendría coordenadas (LS_{X-1}, LS_X); donde LS es el límite superior, en principio para la clase anterior, a la media y la segunda es el límite superior de la clase donde se encuentra la mediana. Luego se traza una recta paralela al eje x por el punto medio de (a, b); donde a y b son los valores de las frecuencias dadas en la clase mediana y su antecesora.

4.3.2 Modo de uso

Se sugiere al docente que realice un taller donde sea posible organizarlos en grupos, y se les plantee que uno de los integrantes deberá diseñar una encuesta que arroje datos de tipo continuo, y que otro integrante deberá hallar la mediana por medio de los cálculos estadísticos, mientras el otro halla la mediana de manera geométrica con ayuda del applet.

4.4 Moda

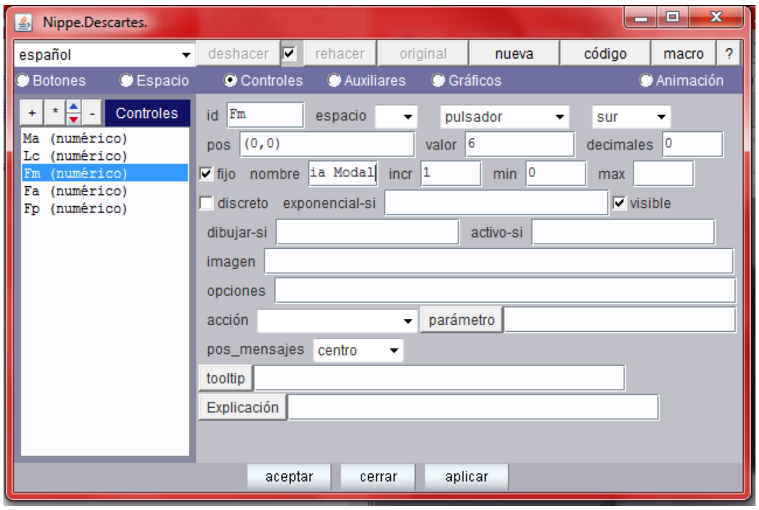
Cuando se tienen distribuciones de frecuencias agrupadas en intervalos y se identifica en la columna de frecuencias el valor de la distribución al que corresponde la mayor frecuencia, el respectivo intervalo se le llamará clase modal.



Gráfica 58

5.4.1 Diseño de la construcción

Se construyen 5 controles uno que sea la marca de clase del intervalo anterior a la clase modal, otro para la longitud de clase, otro para la frecuencia modal, otro para la frecuencia del intervalo anterior y por último uno para la frecuencia del intervalo posterior a la clase modal.



Gráfica 59

Luego se construye el histograma correspondiente, marcamos los puntos (m_i, f_i); donde m_i: es la marca de clase y f_i: la frecuencia de cada intervalo. Y finalmente se construye la ecuación de la parábola, con la cual se identifica la moda de dicha distribución ya que el vértice de la parábola se corresponde con este, ya que es el punto máximo.

5.4.2 Modo de uso

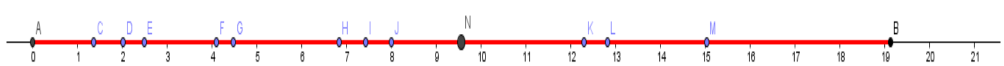
Para esta construcción se sugiere al docente tomar varios ejemplos en el aula de clase, para hallar el valor de la moda y la clase modal, seguido de una breve explicación sobre la parábola donde se explique su importancia y sus partes tales como: vértice, eje focal, directriz y foco, para darle al estudiante las herramientas suficientes para que con ayuda del applet pueda explorar la construcción y conjeturar acerca del vértice de la parábola y el valor de la moda, en este tipo de actividades el docente tiene claro su objetivo y es que por medio de la experiencia el estudiante descubra la relación que existe entre la estadística y la geometría.

5. REPRESENTACIONES GEOMÉTRICAS DE MEDIDAS DE DISPERSIÓN

En este capítulo se definirá geométricamente las medidas de dispersión por medio de construcciones hechas en Geogebra.

5.1 Rango medio

El rango medio es igual a: $R = \frac{v_s + v_i}{2}$, geométricamente podemos representar este valor como el punto medio del segmento formado por la unión del valor mínimo y el valor máximo de los datos recogidos.



Gráfica 60

En esta representación se da una distribución de puntos colineales, donde el rango medio es representado por la coordenada de la abscisa del punto medio del segmento AB.

En esta construcción el usuario puede cambiar la dimensión del segmento e identificar que la correspondencia entre el rango medio y su representación geométrica se mantiene.

5.1.1 Diseño de la construcción:

En esta representación se da una distribución de puntos colineales, donde el rango medio es representado como la coordenada de la abscisa del punto medio del segmento AB.

Para el diseño de esta representación se ubicó sobre el eje x 12 puntos colineales, luego se trazó un segmento AB, siendo A el valor mínimo de los datos y B el valor máximo. Luego con la herramienta punto medio que encontramos en el segundo despliegue a la izquierda, hallamos el punto medio de AB.

5.1.2 Modo de uso

Esta representación se puede utilizar en una clase de exploración del concepto, donde el estudiante defina que es Rango medio en una distribución de datos unidimensionales con base a las características que permanecen constantes identificadas en la exploración de la representación.

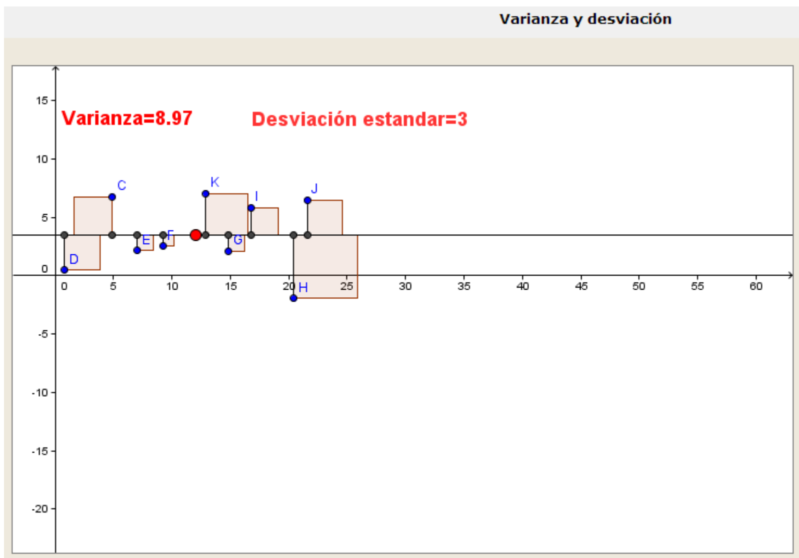
5.2 Desviación estándar y Varianza

La varianza es equivalente a la sumatoria de los cuadrados de $a = x - x$, tal que x; corresponde a los datos recogidos y x, es la media de la distribución.

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - x)^2}{n}$$

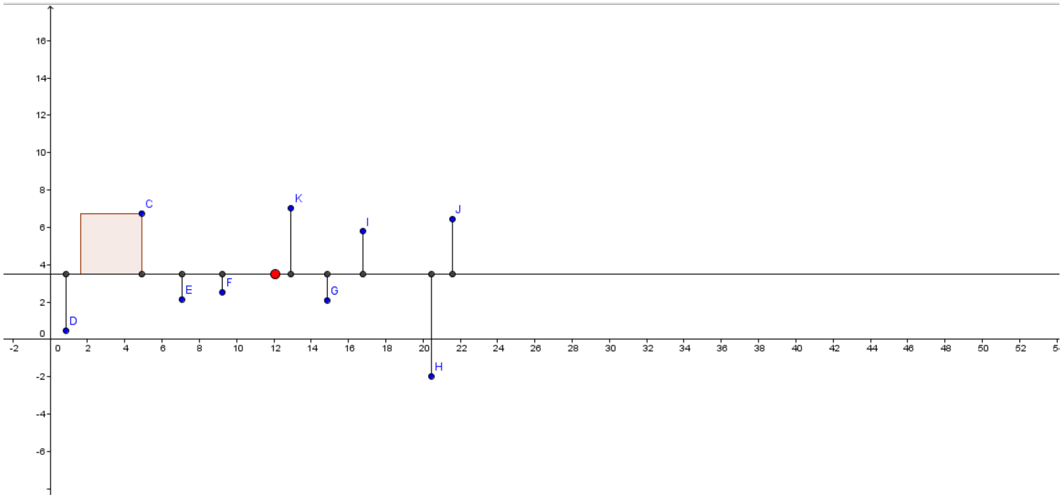
La desviación estándar es igual a la raíz de la varianza

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x)^2}{n}}$$



Gráfica 61

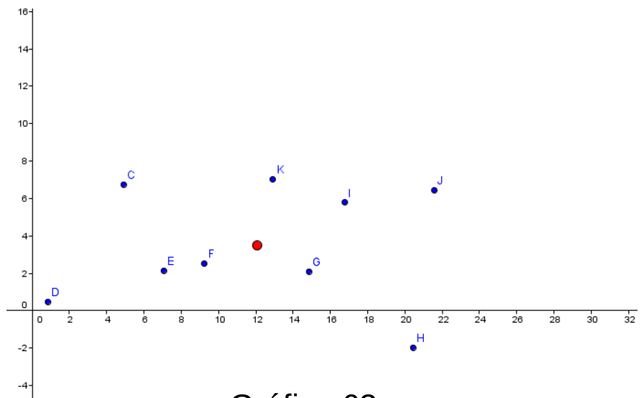
En esta construcción se presenta cada diferencia como el segmento que representa la distancia entre el promedio y cada dato, el cuadrado de cada una de estas diferencias se corresponde con el área de cada cuadrado, con medida del lado igual a cada una de las diferencias.



Gráfica 62

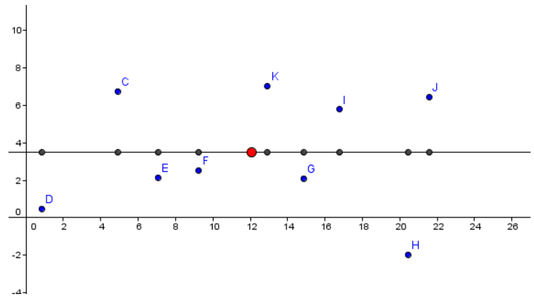
5.2.1 Diseño de la construcción

Dada una distribución de puntos, se halla el punto promedio entre ellos, se trazan rectas paralelas al eje y, una por cada punto; luego una recta paralela al eje x que pase por el punto promedio.



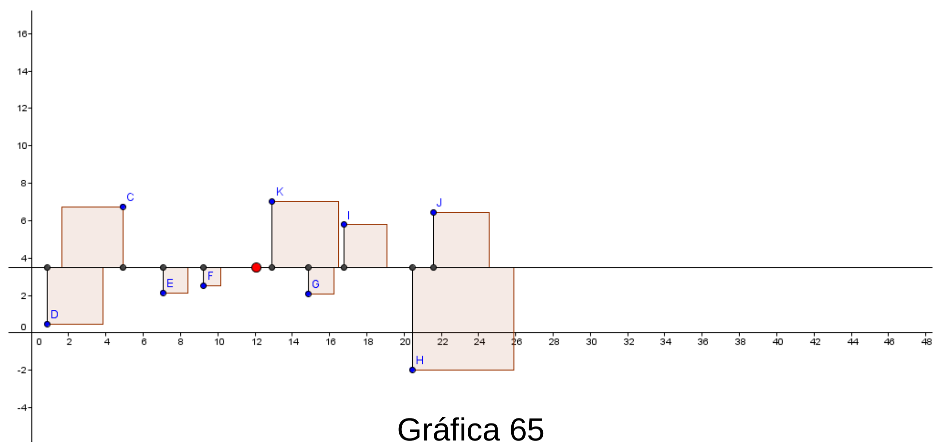
Gráfica 63

Luego con la herramienta punto de intersección, encontramos los puntos de intersección entre las rectas verticales que pasan por cada punto y la que pasa por el punto promedio.



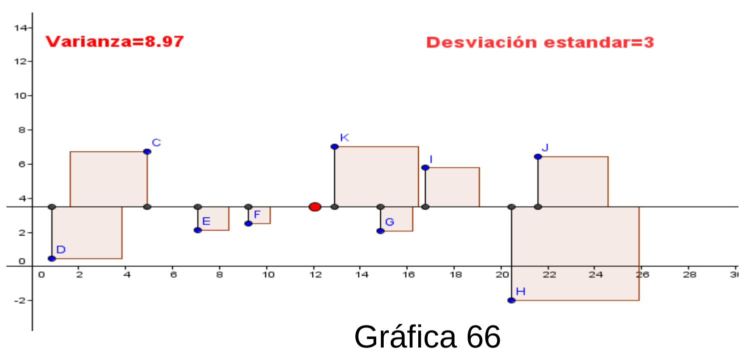
Gráfica 64

Después con la herramienta segmento trazamos los segmentos determinados por los puntos de intersección y los puntos dados, cada segmento de los construidos será el lado generador del cuadrado por medio de la macro hecha inicialmente.



Gráfica 65

Finalmente generamos los cuadrados y hallamos el promedio entre sus áreas, de tal manera que este valor será aproximadamente la varianza y la raíz cuadrada de este se corresponderá con la desviación.



Gráfica 66

6.2.2 Modo de uso

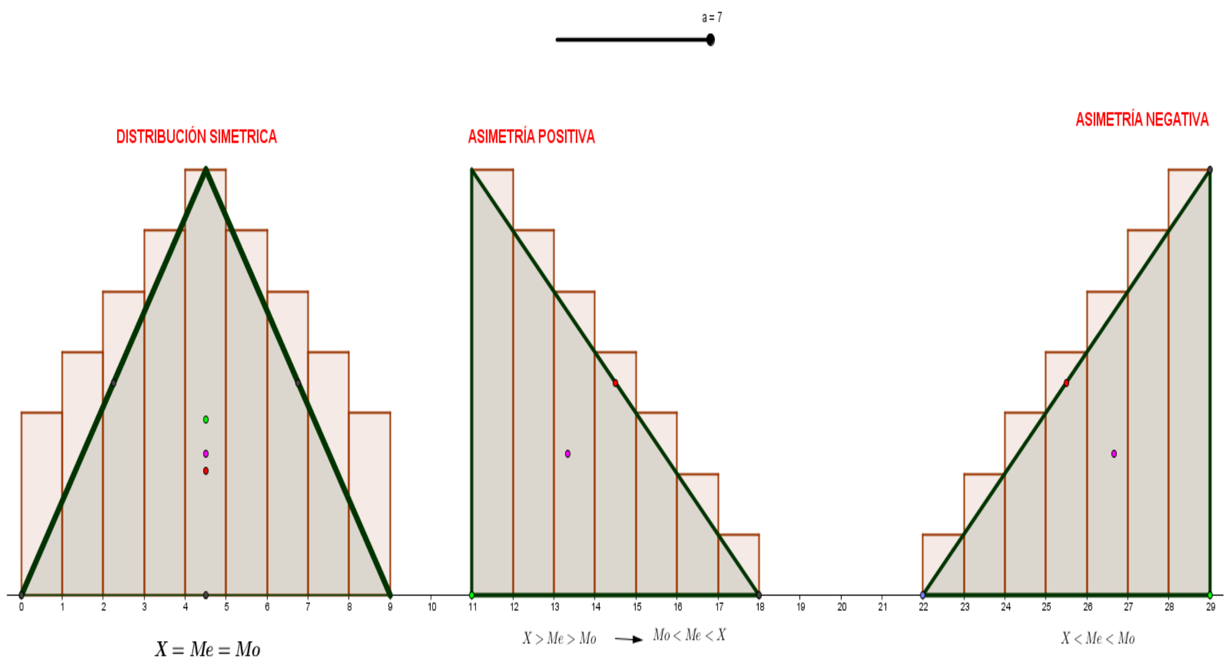
Se sugiere que el docente le proporcione una tabla de datos a los estudiantes y que ellos por medio del applet encuentre el valor aproximado de la varianza, luego hallen el valor por medio de cálculos estadísticos y encuentren el error entre los resultados, es decir que tan aproximado es el método usado. Luego con base a la construcción determine las relaciones entre esta y la definición de varianza.

7. REPRESENTACIÓN DE MEDIDAS DE FORMA

En este capítulo podemos estudiar la relación que se mantiene entre la media, la mediana y la moda en los tipos de asimetría que se dan en una distribución no simétrica, y lo q ue sucede en una distribución simétrica.

7.1. Asimetría

La asimetría mide el grado al cual los valores de la variable equidistan de la media, existen dos clases de asimetría, la asimetría positiva identificada por la relación: $x > Me > Mo$, la asimetría negativa donde se presenta la siguiente relación: $x < Me < Mo$. Cuando la relación dada en una distribución es: $x = Me = Mo$, tenemos una distribución simétrica.



Gráfica 31

7.1.1 Diseño de la construcción

Dado un conjunto de datos se realiza un histograma que ilustre la información suministrada, luego se construye un triángulo que tendrá los siguientes vértices: A, el valor mínimo de la distribución; B, el punto que tiene como coordenadas, la moda y la altura del triángulo; C, el punto máximo para la coordenada de la abscisa y 0 para la ordenada. Cuando ya se ha construido el triángulo se traza sus alturas y en la intersección de estas, diremos que hemos obtenido el valor de la moda y su frecuencia, ya que la moda se corresponde con la abscisa y la frecuencia con la ordenada. Después ubicamos el punto medio de cada lado y trazamos rectas que pasen por un vértice y el punto medio del lado opuesto; por lo tanto en la intersección de estas rectas obtendremos en el valor de la abscisa la mediana. Finalmente trazamos la mediatriz de cada segmento y en su intersección obtendremos el valor de la media, el cual también se corresponde con la coordenada de la abscisa.

7.1.2 Modo de uso

Al construir el histograma podemos por medio de cálculos estadísticos encontrar los valores de la media, la mediana y la moda para dicha distribución; a partir de estos valores podemos determinar que la distribución es asimétrica positiva, negativa o es simétrica y luego compararla con lo obtenido en el applet. Luego el docente podría realizar un taller donde sus estudiantes cambien los valores de las frecuencias del histograma realicen los cálculos estadísticos y comparen lo obtenido con lo que presenta.

8. REPRESENTACIÓN GEOMÉTRICA DE LA RECTA DE REGRESIÓN

Para la recta de regresión lineal se realiza una representación dinámica en la cual se muestra la explicación del tema, el manejo de la construcción y su importancia, con el tipo de variable para la realización de cada simulación.

8.1 Recta de regresión lineal

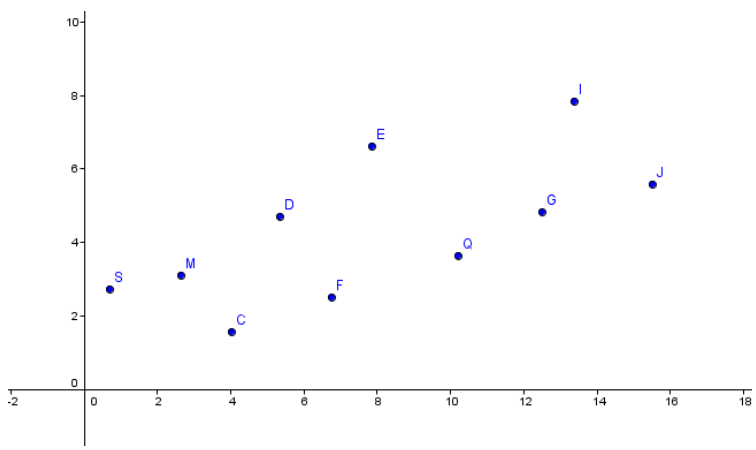
La recta de regresión es la que mejor se ajusta a la nube de puntos ubicados aleatoriamente, la recta pasa por el punto x, y llamado centro de gravedad, el promedio de los X y Y, después de ubicar el centro de gravedad se puede hallar la pendiente de la recta, o el punto de corte con eje y.

Para la construcción se utilizaron las distancias de los puntos a una recta que pasa por x, y , mediante la construcción de cuadrados que tienen como lado la distancia de cada punto a la recta, trazando una recta perpendicular al eje y que pasa por x, y , utilizando el lugar geométrico los puntos, se genera una parábola por la cual se traza una recta paralela al eje y , que pase por el vértice y en ese punto de intersección entre la recta paralela y el eje y , se encuentra el punto de corte de la recta de mayor ajuste a los puntos.

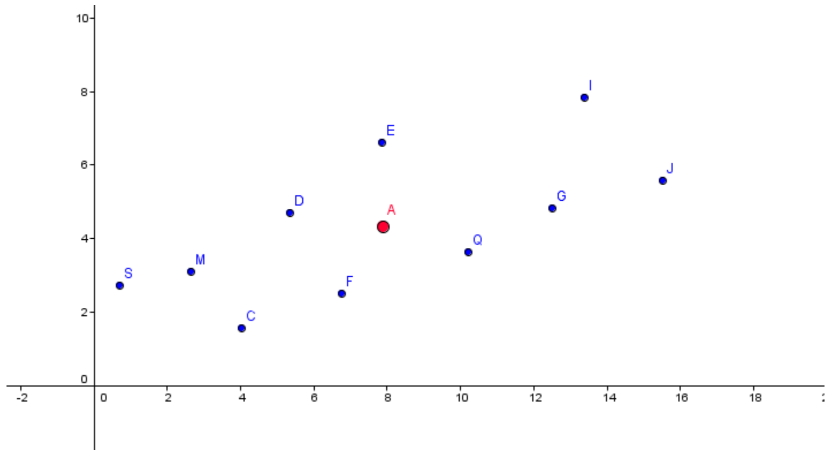
Gráfica 31

8.1.1 Diseño de la construcción

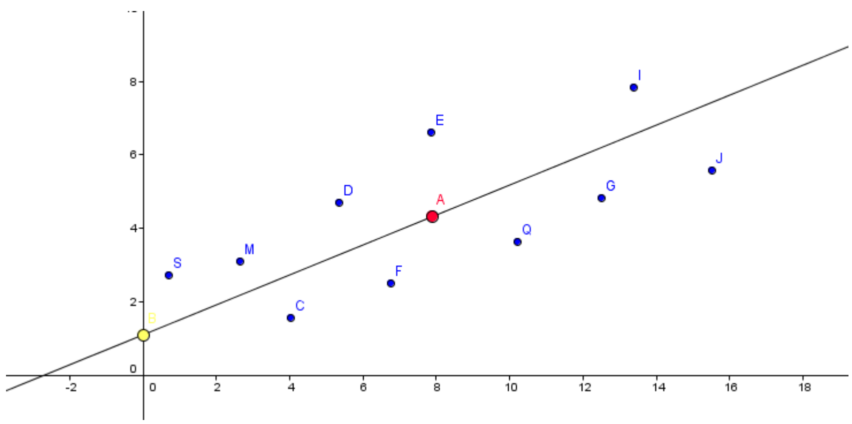
En el plano cartesiano se ubican 10 puntos de manera aleatoria, en el primer cuadrante.



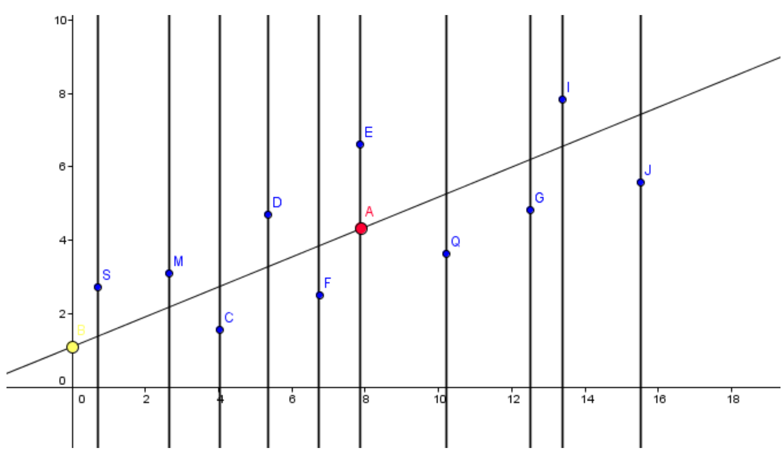
Se ubica el punto A que es promedio coordenadas en x y y o el centro de gravedad.



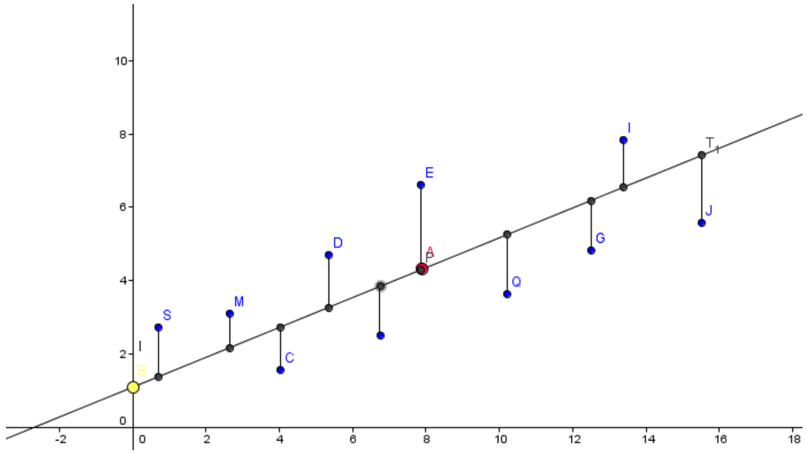
Se traza una recta que pase por el punto A y un punto en el eje y, ya que la recta de regresión pasa por el punto A y el punto de B es de manera aletoria.



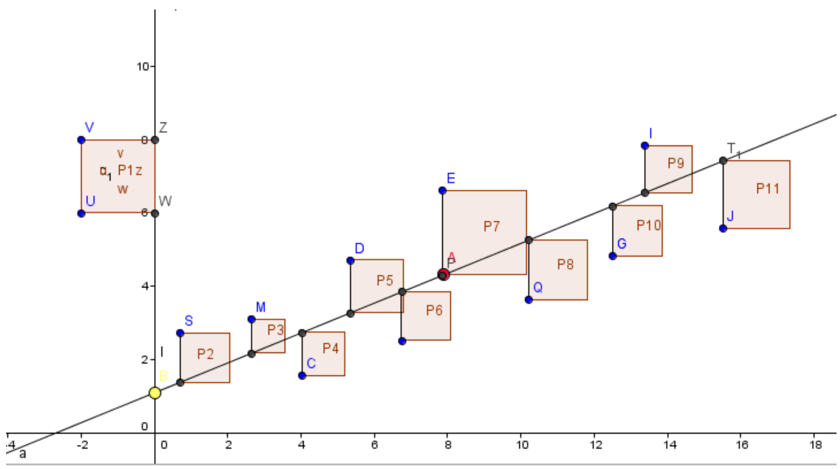
Se trazan rectas perpendiculares al eje x que pasen por cada punto.



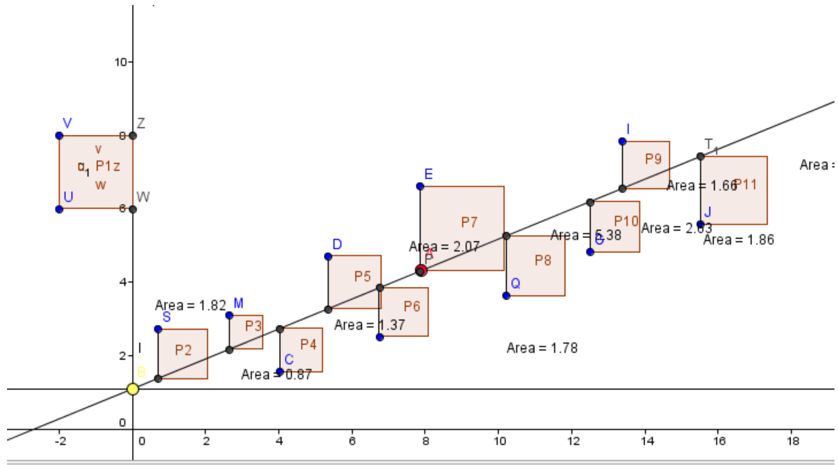
Se encuentran los puntos de intersección entre las rectas perpendiculares y la recta que pasa por el punto A, para trazar los segmentos desde cada punto a la recta.



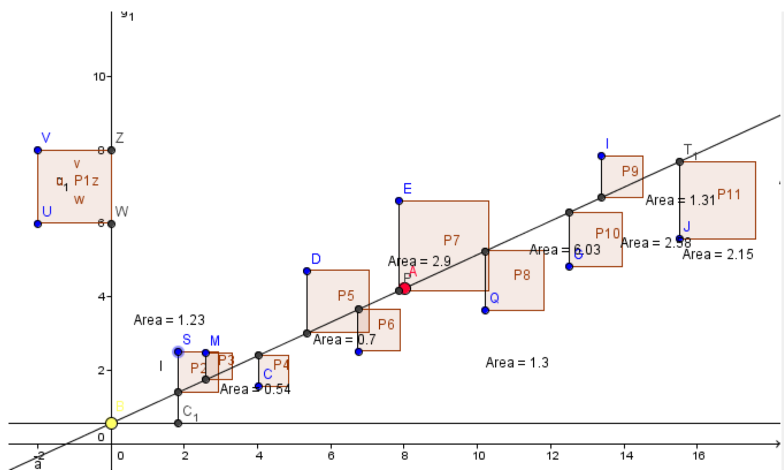
Con cada segmento se construye un cuadrado.



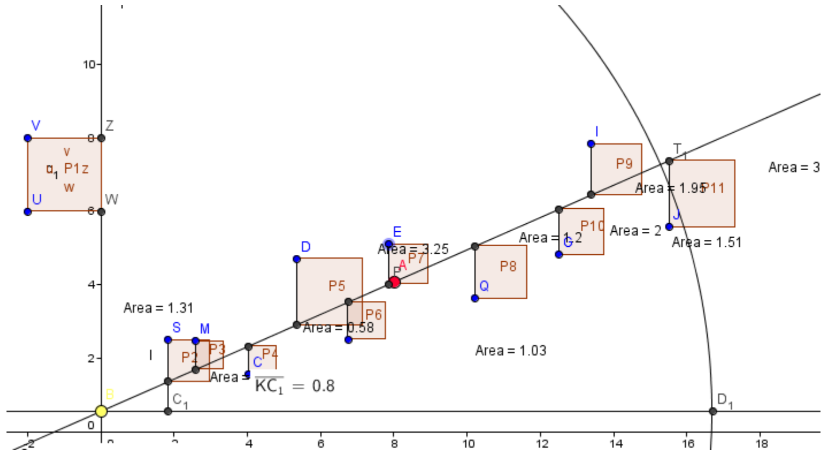
SR la suma las áreas de los cuadrados, se traza una recta perpendicular al eje y que pase por B.



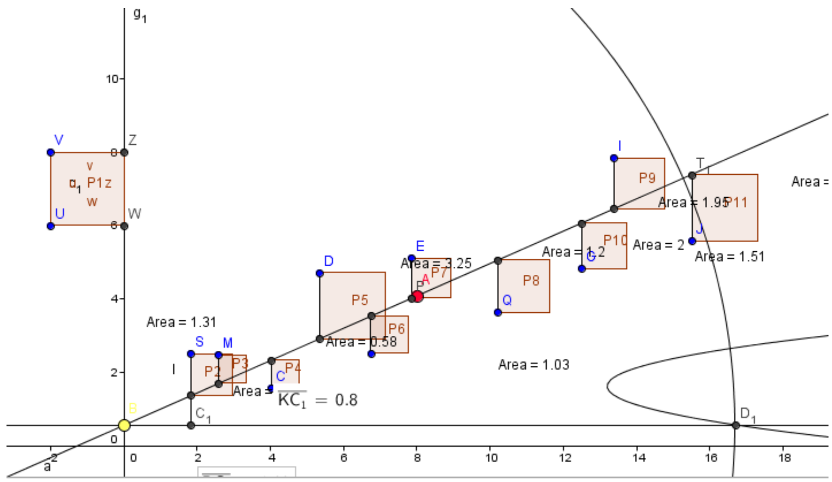
Se encuentra el punto C1 de intersección entre la recta perpendicular al eje y que pasa por B y la recta perpendicular al el eje x que pasa por el punto S.



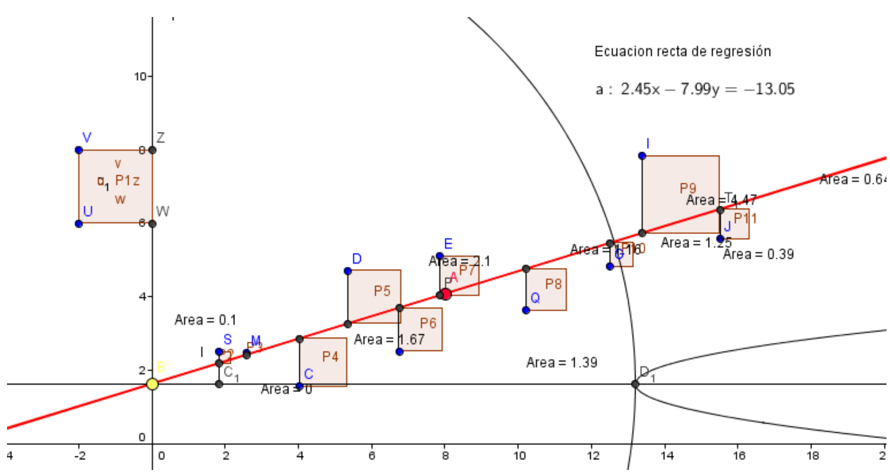
Se traza una circunferencia con centro en B y radio SR, D_1 punto de intersección entre la cricnferencia y la recta perpendicular al eje y que pasa por B.



Se genera el lugar de geométrico (D1, B).



Se genera una parábola cuyo vértice se proyecta al eje y, y da el punto de corte de la recta de regresión.



8.1.2 modo de uso

Se sugiere al docente realizar ejercicios de ajuste por regresión lineal, en el que los estudiantes utilicen este applet como ayuda en clase y verificador de que la recta de regresión lineal tiene la ecuación correcta y es la de mayor ajuste a los datos.

Ejemplo

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	2	3	4	4	5	6	6	7	7	8	10	10
Física	1	3	2	4	4	4	6	4	6	7	9	10

Tabla 13

Hallar las rectas de regresión y representarlas.

X_i	Y_i	$X_i \times Y_i$	${X_i}^2$	${Y_i}^2$
2	1	2	4	1
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
5	4	20	25	16
6	4	24	36	16
6	6	36	36	36
7	4	28	49	16
7	6	42	49	36
8	7	56	64	49
10	9	90	100	81
10	10	100	100	100
72	60	431	504	380

Tabla 13

Se Hallan las medias aritméticas.

$$x = \frac{72}{12} = 6 \qquad y = \frac{60}{12} = 5$$

Se calcula la covarianza.

$$\sigma_{xy} = \frac{431}{12} - 6.5 = 5.92$$

Se Calculamos las varianzas.

$$\sigma_x^2 = \frac{504}{12} - 6^2 = 6$$

$$\sigma_y^2 = \frac{380}{12} - 5^2 = 6.66$$

Recta de regresión de Y sobre X.

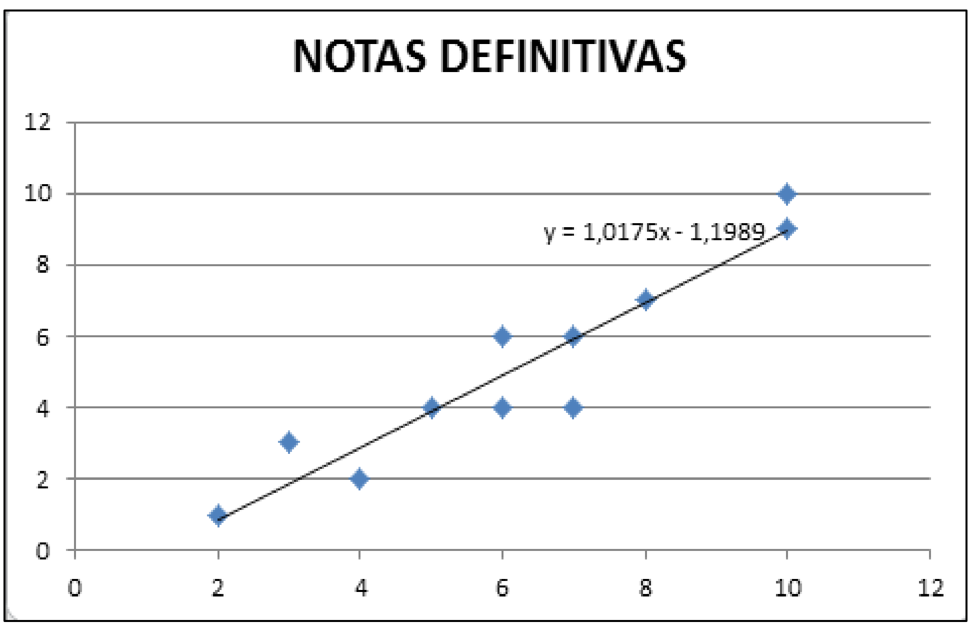
$$y - 5 = \frac{5.92}{6}(x - 6)$$

$$y = 0.987x - 0.922$$

Recta de regresión de X sobre Y.

$$x - 6 = \frac{5.92}{6.66}(y - 5)$$

$$x = 0.889y - 1.556$$



Gráfica 32

CONCLUSIONES

En las conclusiones se presenta una breve descripción de los aportes que obtuvimos en el proceso de construcción de este recurso virtual, y otras que deducimos mediante el desarrollo del trabajo de grado.

- De acuerdo al primer objetivo: *Recopilar construcciones geométricas sobre temas de Estadística Descriptiva.* Se concluye que en los textos se encuentran las ilustraciones de los gráficos estadísticos, aunque están no son dinámicas, y también es posible encontrar la construcción de mediana, moda y comparación de medias para datos agrupados. Estas representaciones muestran cómo llegar a las formulas por medio de semejanza de triángulos.

- *Diseñar representaciones geométricas para ilustrar los diferentes conceptos de Estadística descriptiva.* Se concluye que por medio de diferentes elementos geométricos se pueden realizar construcciones geométricas interactivas para conceptos de estadística descriptiva

- *Elaborar un C.D con la recopilación de las ilustraciones geométricas sobre los conceptos de Estadística descriptiva.* Se concluye que un recurso bibliográfico como este será una herramienta de clase que permitirá corroborarolos procesos hechos en “lápiz y papel”

- La geometría y la estadística son dos ciencias que permiten hacer construcciones que las relacionan y por medio de las propiedades de los diferentes elementos geométricos se hace una lectra del concepto desde otra perspectiva.

- La construcción de un recurso que relacione dos ciencias nos permite identificar la relación entre ellas y así mostramos que no son estudios independientes, si no que por el contrario podemos observar cómo se enlazan para el desarrollo de diferentes temas.

BIBLIOGRAFÍA

- Fernández, F. y Sarmiento B. (2009). Curso básico de estadística, introducción al análisis de datos. Bogotá, Colombia. Editorial CARGRAPHIS S.A
- Triola, M. Estadística. (2004). México. Editorial Pearson. 614p
- Gonzáles, M. Estadística aplicada. (2009). Madrid, España. Editorial Díaz De Santos. 759 p.
- Murray R. Spiegel, John J. Schiller, R. Probabilidad y estadística. (2010). México. Editorial McGraw- Hill. 425 p.
- Ross, M Sheldon. Introducción a la Estadística. (2007). Barcelona- Bogotá- Buenos Aires- Caracas- México. Editorial REVERTÉ S. A.
- Martínez, C. Estadística básica aplicada. (2006). Santa fe de Bogotá: ECOE Ediciones. 386 p.
- MONTGOMERY, Douglas C. & RUNGER, George C. (1997). Probabilidad y Estadística aplicadas a la Ingeniería. México: McGraw Hill.
- Johnson, R. Kuby, P. ESTADÍSTICA ELEMENTAL. (2008). México. Editorial impresiones S.A
- Cuervo, C. Edilberto. Estadística Matemática. (2009). Bogotá, Colombia. Editores Litográficos Ltda.