



Nombre de alumnos: Fernanda
Patricia Hernández Díaz

Nombre del profesor: Rosario Gómez

Nombre del trabajo: Aplicación de la
estadística y representación tabular

Materia: Estadística Descriptiva

Grado: 2^o Cuatrimestre

Grupo: LPS19SSCO9-A

APLICACIÓN DE LA ESTADÍSTICA EN LA PSICOLOGÍA

Introducción

Para empezar es importante destacar que la **estadística** describe, analiza y representa un grupo de datos utilizando **métodos** numéricos y **gráficos** que resumen y presentan la información contenida en ellos. Se puede definirse como aquel **método que contiene la recolección, organización, presentación y resumen de una serie de datos.**

Además podemos agregar que la estadística es una **ciencia** o **método científico** que en la actualidad es considerada como un **poderoso auxiliar** en las investigaciones científicas, que le permite a ésta aprovechar el material cuantitativo. Cabe destacar que **desde los comienzos** de la civilización han existido formas sencillas de estadística, pues ya **se utilizaban representaciones gráficas y otros símbolos en pieles, rocas, palos de madera y paredes de cuevas para contar el número de personas, animales o cosas.** Hacia el año 3000 a.C. los babilonios usaban pequeñas tablillas de arcilla para **recopilar datos** sobre la producción agrícola y sobre las especies vendidas o cambiadas mediante trueque.

Es de importancia saber los conceptos básicos de la estadística, fundamentalmente empezando con el **universo**, en estadística, es la totalidad de **elementos o características** que conforman el **ámbito** de un estudio o investigación. A su vez conocemos que la **población** es la totalidad de los valores posibles (mediciones o conteos) de una característica particular de un grupo especificado de **personas, animales o cosas** que se desean estudiar en un momento determinado. Por lo tanto a la **muestra** en otras palabras es una **parte de la población** que sirve para **representarla**. También el **muestreo** Es el procedimiento mediante el cual se obtiene una o más muestras de una población determinada. Existen dos **tipos de muestreos** a saber: **Parámetros** Son cualquiera característica que se pueda medir y cuya medición se lleve a cabo sobre todos los elementos que integran una población determinada, los mismos suelen representarse con letras griegas. **Dato numéricos** un conjunto de **valores numéricos** que tienen relación significativa entre sí. Los mismos pueden ser comparados, analizados e interpretados en una investigación cualquiera. La **frecuencia** es el número de veces que se repite (aparece) el mismo dato estadístico en un conjunto de observaciones de una investigación determinada, las frecuencias se les designan con las **letras fi**, y por lo general se les llaman frecuencias absolutas. La **distribución de frecuencias** existen investigaciones relacionadas con los precios de los productos de la dieta diaria, la estatura y el peso de un grupo de individuos, los salarios de los empleados, los grados de temperatura del medio ambiente, las calificaciones de los estudiantes, etc., que pueden adquirir diferentes **valores gracias a una unidad apropiada**, que recibe el nombre de **variable**. La representación numérica de las variables se denomina dato estadístico. La distribución de frecuencia es una disposición tabular de datos estadísticos, ordenados ascendente o descendentemente, con la frecuencia (f_i) de cada dato.

Cabe destacar que las distribuciones de frecuencias pueden ser para **datos no agrupados** y para **datos agrupados** o de intervalos de clase.

Distribución de frecuencia para datos no Agrupados: Es aquella distribución que indica las frecuencias con que aparecen los datos estadísticos, desde el menor de ellos hasta el mayor de ese conjunto sin que se haya hecho ninguna modificación al tamaño de las unidades originales. **Distribución de frecuencia de clase o de datos Agrupados:** Es aquella distribución en la que las disposiciones tabulares de los datos estadísticos se encuentran ordenados en clases y con la frecuencia de cada clase; es decir, los datos originales de varios valores adyacentes del conjunto se combinan para formar un intervalo de clase.

Así mismo los **Componentes de una distribución de frecuencia de clase son:**

1.- **Rango o Amplitud total** (recorrido).- El rango es el **tamaño del intervalo** en el cual se ubican todos los valores que pueden tomar los diferentes datos de la serie de valores, desde el menor de ellos hasta el valor mayor estando incluidos ambos extremos. El rango de una distribución de frecuencia se designa con la letra **R**.

2.- **Clase o Intervalo de clase**.- Son divisiones o categorías en las cuales se agrupan un conjunto de datos ordenados con características comunes. En otras palabras, **son fraccionamientos del rango o recorrido de la serie de valores para reunir los datos que presentan valores comprendidos entre dos límites**. Para organizar los valores de la serie de datos hay que determinar un número de clases que sea conveniente. En otras palabras, **que ese número de intervalos no origine un número pequeño de clases ni muy grande**. Un número de clases pequeño puede ocultar la naturaleza natural de los valores y un número muy alto puede provocar demasiados detalles como para observar alguna información de gran utilidad en la investigación.

Tamaño de los **Intervalos de Clase**

Los intervalos de clase pueden ser de tres tipos, según el tamaño que estos presenten en una distribución de frecuencia:

a) Clases de igual tamaño, b) clases desiguales de tamaño y c) clases abiertas.

3.- **Amplitud de Clase, Longitud o Ancho de una Clase** La amplitud o longitud de una clase es el número de valores o variables que concurren a una clase determinada. Existen diversos criterios para determinar la **amplitud de clases**, ante esa diversidad de criterios, se ha considerado que lo más importante es dar un **ancho o longitud** de clase a todos los intervalos de tal manera que respondan a la naturaleza de los datos y al objetivo que se persigue y esto se logra con la práctica.

4.-**Punto medio o Marca de clase**

El centro de la clase, es el **volar de los datos** que se ubica en la posición central de la clase y representa todos los demás valores de esa clase. Este valor se utiliza para el cálculo de la media aritmética.

5.-**Frecuencia de clase**

La frecuencia de clase se le denomina frecuencia absoluta y se le designa con las letras f_i . Es el **número total de valores de las variables** que se encuentran presente en una clase determinada, de una distribución de frecuencia de clase.

6.- **Frecuencia Relativa**

La frecuencia relativa es aquella que resulta de dividir cada uno de los f_i de las clases de una distribución de frecuencia de clase entre el número total de datos (N) de la serie de valores. Estas frecuencias se designan con las letras fr ; si cada fr se multiplica por 100 se obtiene la frecuencia relativa porcentual ($fr \%$).

7.-Frecuencias acumuladas

Las frecuencias acumuladas de una distribución de frecuencias son aquellas que se obtienen de las sumas sucesivas de las f_i que integran cada una de las clases de una distribución de frecuencia de clase, esto se logra cuando la acumulación de las frecuencias se realiza tomando en cuenta la primera clase hasta alcanzar la última. Las frecuencias acumuladas se designan con las letras fa . Las frecuencias acumuladas pueden ser menor que ($fa < que$) y frecuencias acumuladas mayor que ($fa > que$).

8.- Frecuencia acumulada relativa

La frecuencia acumulada relativa es aquella que resulta de dividir cada una de las fa de las diferentes clases que integran una distribución de frecuencia de clase entre el número total de datos (N) de la serie de valores, estas frecuencias se designan con las letras far . Si las far se multiplican por 100 se obtienen las frecuencias acumuladas relativas porcentuales y las mismas se designan así: $far \%$.

La mediana

La mediana (M_d) es una medida de posición que divide a la serie de valores en dos partes iguales, un cincuenta por ciento que es mayor o igual a esta y otro cincuenta por ciento que es menor o igual que ella. Es por lo tanto, un parámetro que está en el medio del ordenamiento o arreglo de los datos organizados, entonces, la mediana divide la distribución en una forma tal que a cada lado de la misma queda un número igual de datos. Para encontrar la mediana en una serie de datos no agrupados, lo primero que se hace es ordenar los datos en una forma creciente o decreciente y luego se ubica la posición que esta ocupa en esa serie de datos; para ello hay que determinar si la serie de datos es par o impar, luego el número que se obtiene indica el lugar o posición que ocupa la mediana en la serie de valores, luego la mediana será el número que ocupe el lugar de la posición encontrada.

La moda

La moda es la medida de posición que indica la magnitud del valor que se presenta con más frecuencia en una serie de datos; es pues, el valor de la variable que más se repite en un conjunto de datos. De las medidas de posición la moda es la que se determina con mayor facilidad, ya que se puede obtener por una simple observación de los datos en estudio, puesto que la moda es el dato que se observa con mayor frecuencia. La moda se designa con las letras Mo .

Desviación típica o estándar

Es la medida de dispersión más utilizada en las investigaciones por ser la más estable de todas, ya que para su cálculo se utilizan todos los desvíos con respecto a la media aritmética de las observaciones, y además, se toman en cuenta los signos de esos desvíos. Se le designa con la letra castellana **S** cuando se trabaja con una muestra y con la letra griega minúscula s (Sigma) cuando se trabaja con una población. Es importante destacar que cuando se hace referencia a la población el número de datos se expresa con N y cuando se refiere a la muestra el número de datos se expresa con n . La desviación típica se define como:

Interpretación de la desviación estándar

La desviación típica como medida absoluta de dispersión, es la que mejor nos proporciona la variación de los datos con respecto a la media aritmética, su valor se encuentra en relación directa con la dispersión de los datos, a mayor dispersión de ellos, mayor desviación típica, a menor dispersión, menor desviación típica.

Varianza

Es otra de las variaciones absolutas y la misma se define como el cuadrado de la desviación típica; viene expresada con las mismas letras de la desviación típica pero elevada al cuadrado, así S^2 y s^2 . Las fórmulas para calcular la varianza son las mismas utilizadas por la desviación típica, exceptuando las respectivas raíces, las cuales desaparecen al estar elevados el primer miembro al cuadrado.

La Estadística dentro del Método Científico

Es importante considerar que la estadística no se puede utilizar como una caja mágica para extraer certezas, donde se introducen datos y se extraen leyes. La estadística, en el contexto de probabilidades y técnicas de inferencia, **es incapaz por sí misma de suplantar al Método Científico, sólo es un gran apoyo.**

Delimitamos el Método Científico como un método o **conjunto sistematizado** de procesos en los que se basa la ciencia para explicar cualquier fenómeno y las leyes que los administran.

La **estadística descriptiva** es la herramienta más útil en la etapa de observación, ya que nos permite extraer información para **realizar nuestras hipótesis fundadas en estos resultados.** También es utilizada para valorar los resultados del experimento.

La **estadística analítica** se utiliza a partir de **la observación**, ya que dependiendo de los datos observados, se utilizará una técnica u otra, y por supuesto en el proceso del experimento, ya que su diseño **dependerá en cierta medida de las técnicas estadísticas más apropiadas,** además, la estadística analítica es el primer y principal razonamiento válido.

Percibimos que la estadística proporciona un gran apoyo al Método Científico en las fases de observación y experimentación, pero en el proceso de hipótesis y en el de la obtención de una ley científica son otras las bases

Objetivos de la Metodología Estadística.

- **Plantear:** en todo tipo de investigación cuantitativa es principal, plantear detalladamente el estudio.
- **Debatir:** se refiere a opinar sobre cualquier tema dichos anteriormente, o sobre otro tema interesante.
- **Solucionar:** proponer una solución específica que ayude a cualquier tema en estudio.
- **Unir:** el fundamento principal de la estadística es la cuantificación de los elementos desde una muestra o cifra de observaciones.

CATEGORÍAS

Se encuentran ciertos conceptos y categorías que forman parte exclusivamente del campo de la estadística descriptiva. Algunos se listan a continuación:

- **Dispersión:** es la diferencia que existe entre los valores incluidos dentro de una misma variable. La dispersión también incluye el promedio de dichos valores.
- **Promedio:** es el valor que resulta de la sumatoria de todos los valores incluidos en una misma variable y la posterior división del resultado por el número de datos incluidos en la sumatoria. Se define como la tendencia central de una variable.
- **Sesgo o curtosis:** es la medida que indica qué tan inclinada es una curva. Es el valor que indica la cantidad de elementos que se encuentran más próximos al promedio. Existen tres tipos diferentes de sesgo (leptocúrtica, mesocúrtica y platicúrtica), cada uno de ellos indica qué tan alta es la concentración de datos alrededor del promedio.
- **Gráficos:** son la representación gráfica de los datos obtenidos del análisis. Usualmente, son utilizados diferentes tipos de gráficos estadísticos, incluidos los de barras, circulares, lineales, poligonales, entre otros.
- **Asimetría:** es el valor que muestra la manera como los valores de una misma variable se encuentran repartidos con relación al promedio. Puede ser negativa, simétrica o positiva

ESTADÍSTICA DESCRIPTIVA E INFERENCIAL

La principal diferencia entre la **estadística descriptiva** y la **inferencial** radica en que la primera busca **ordenar, resumir y clasificar** los datos derivados del análisis de variables. Por su parte, la estadística inferencial, lleva a cabo **deducciones** con base a los datos previamente obtenidos. Por otro lado, la estadística inferencial **depende del trabajo de la estadística descriptiva** para llevar a cabo sus inferencias. De este modo, la estadística descriptiva **constituye la base sobre la que posteriormente la estadística inferencial llevará a cabo su trabajo.**

PARAMETROS ESTADISTICOS

Un parámetro estadístico es un número que se obtiene a partir de los datos de una distribución estadística.

Los parámetros estadísticos sirven para sintetizar la información dada por una tabla o por una gráfica.

Tipos de parámetros estadísticos

Hay tres tipos parámetros estadísticos:

De **centralización**: Nos indican en torno a qué valor (centro) se distribuyen los datos.

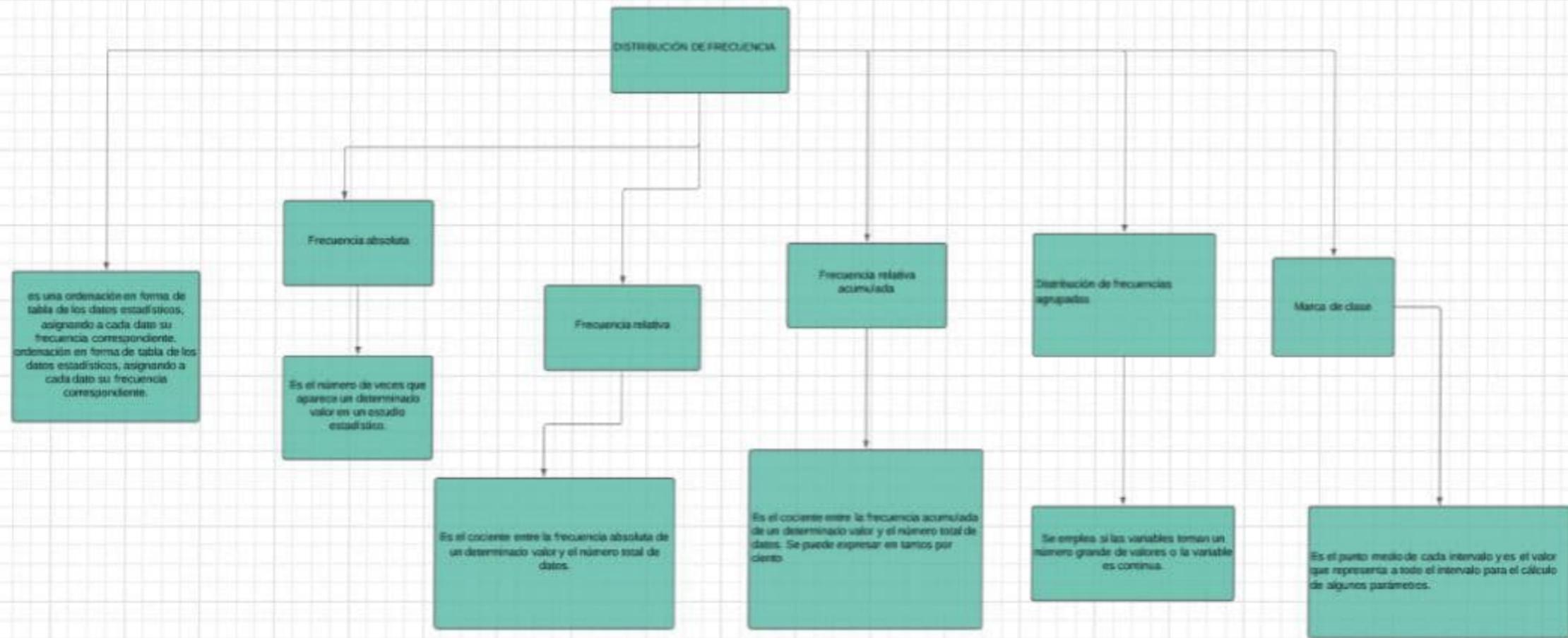
De **posición**: Las medidas de posición dividen un conjunto de datos en grupos con el mismo número de individuos.

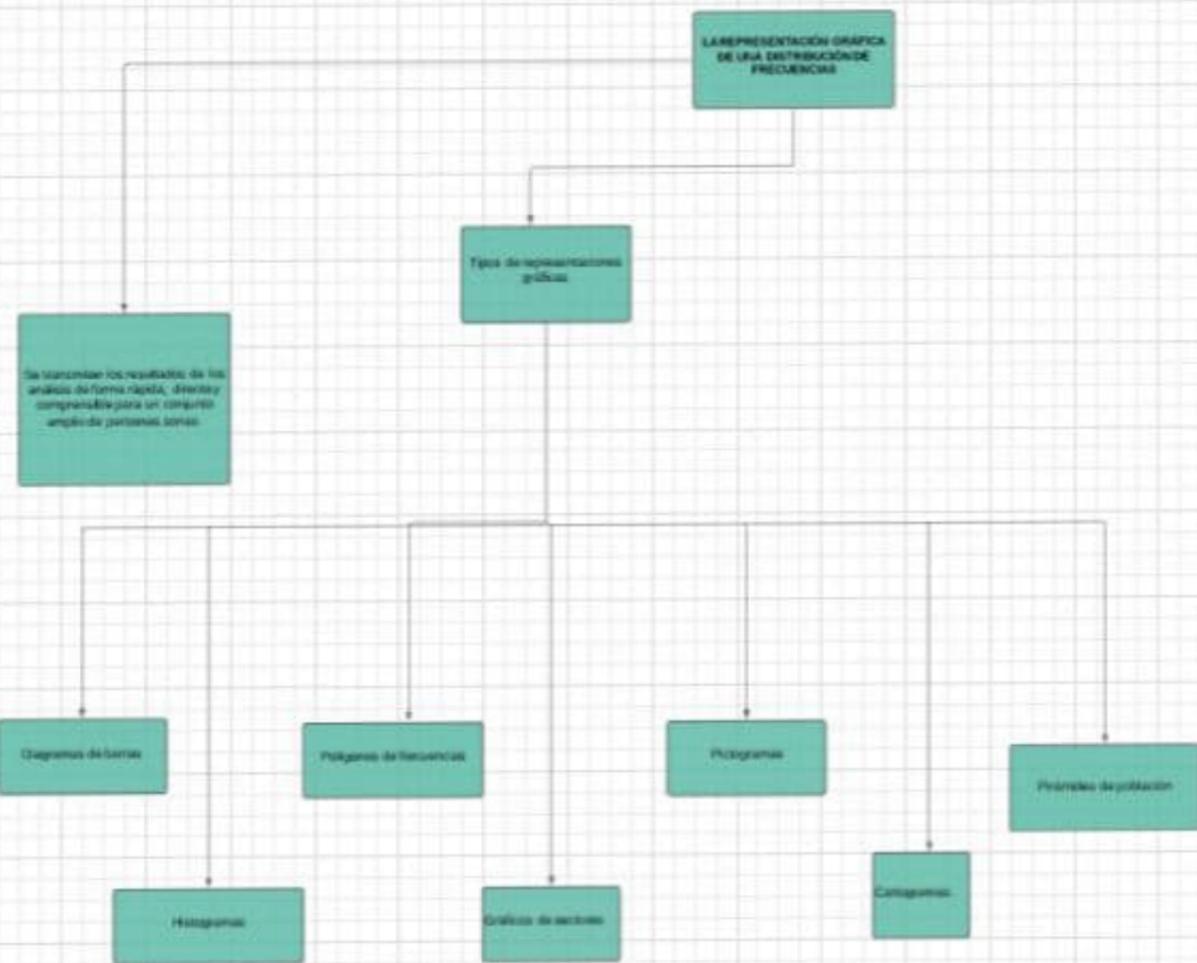
De **dispersión**: Las medidas de dispersión nos informan sobre cuanto se alejan del centro los valores de la distribución.

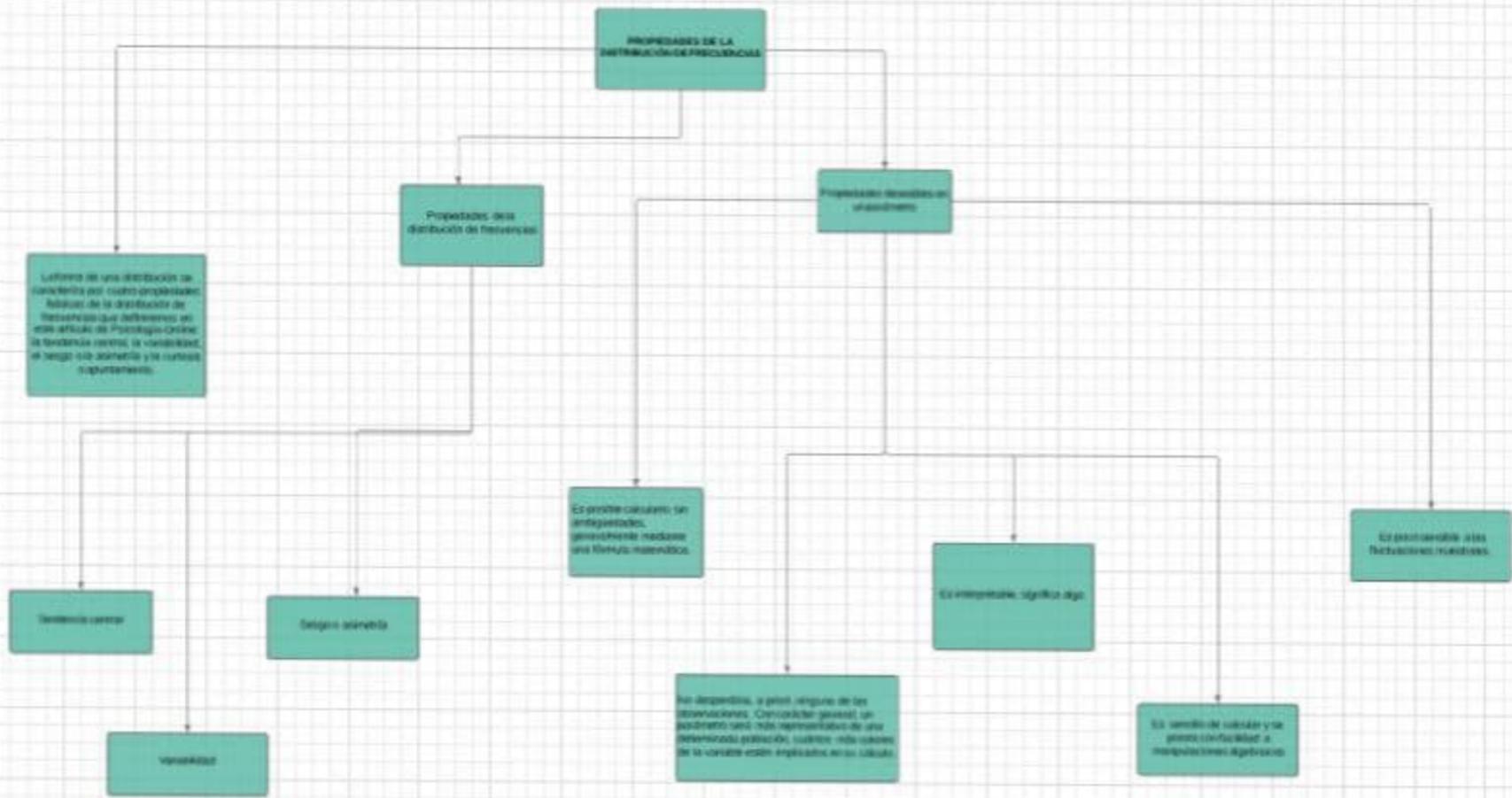
BIBLIOGRAFÍA

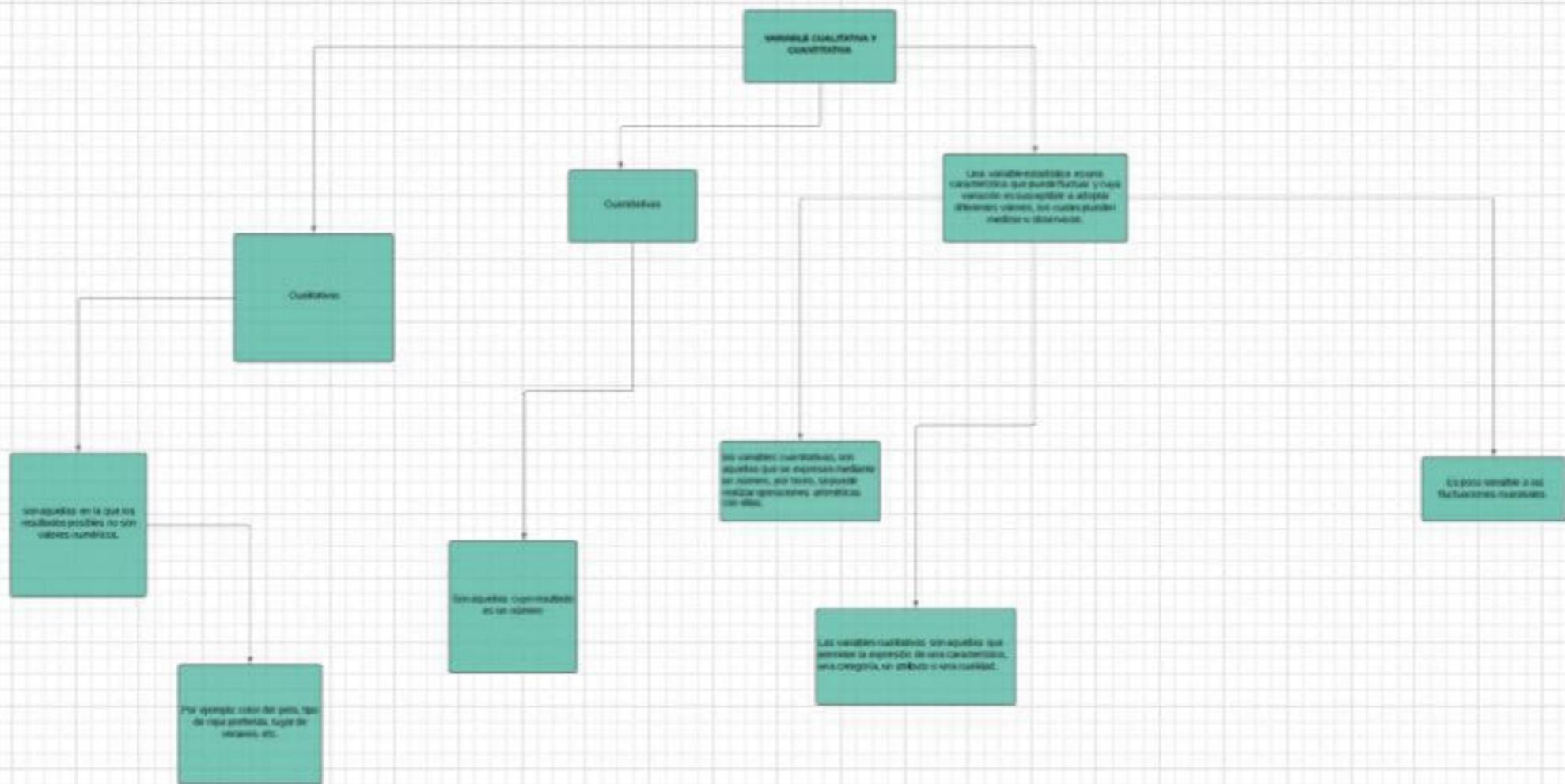
Probabilidad y estadística de George Canavos

Estadística de Murray R. Spiegel









EJERCICIO

Un grupo de investigadores pertenecientes a la secretaría de seguridad pública, tomó una muestra aleatoria de las velocidades (km/h) registradas por 30 vehículos en el trayecto Pichucalco-Villahermosa , con el fin de establecer nuevos límites máximos de velocidad para una carretera. La muestra arrojó los datos siguientes: 90, 99, 104, 99, 119, 98, 95, 112, 95, 120, 100, 90, 116, 96, 114, 108, 98, 118, 100, 106, 114, 100, 112, 106, 100, 115, 111, 105, 114, 97

•**Construye una distribución de frecuencia para velocidades**, que tenga frecuencia absoluta acumulada, frecuencia relativa, frecuencia relativa acumulada y marca de clase así como también un histograma.

Intervalo de clase	f	fa	fr	fra	MC
[90-95)	2	2	0.07	0.07	92.5
[95-100)	8	10	0.27	0.34	97.5
[100-105)	5	15	0.17	0.51	102.5
[105-110)	4	19	0.13	0.64	107.5
[110-115)	6	25	0.2	0.84	112.5
[115-120]	5	30	0.16	1	117.5
Total	30		1		

Operación

$$K = 1 + 3.3 \log(30) = 1 + 3.3 (1.4771212547) = 1 + 4.87 = 5.87 \approx 6$$

$$Fr = f/N$$

$$Fr = 2/30 = 0.07$$

$$Fr = 8/30 = 0.27$$

$$Fr = 5/30 = 0.17$$

$$Fr = 4/30 = 0.13$$

$$Fr = 6/30 = 0.2$$

$$Fr = 5/30 = 0.16$$

Histograma de velocidades

